Proc. Estonian Acad. Sci. Biol. Ecol., 2000, **49**, 2, 177–193
https://doi.org/10.3176/biol.ecol.2000.2.02

MAPPING OF WETLAND HABITAT DIVERSITY USING SATELLITE DATA AND GIS: AN EXAMPLE FROM THE ALAM-PEDJA NATURE RESERVE, ESTONIA

Kiira AAVIKSOO[a], Jaanus PAAL[b], and Tiina DIŠLIS[a]

[a] Estonian Environment Information Centre, Ministry of the Environment, Mustamäe tee 33, 10616 Tallinn, Estonia; kiira@ic.envir.ee, tiina_d@ic.envir.ee
[b] Institute of Botany and Ecology, University of Tartu, Lai 40, 51005 Tartu, Estonia; jpaal@ut.ee

Received 15 November 1999, in revised form 15 February 2000

Abstract. The experience of compiling a habitat map of the Alam-Pedja Nature Reserve, Estonia, based on integrating remote sensing (Landsat 5 TM, aerial photos), Geographic Information Systems (GIS), and national biological survey data is reported and discussed. Unsupervised and supervised classification and the maximum likelihood algorithm were used for the classification of satellite image pixels. Training polygons for justification of the classification were established after the repetitive checking of all classification units during the classification procedure. On that basis the mapping of different habitats, specific for the temperate zone and especially for Estonian nature, using a classification scheme with an emphasis on spectral separation of wetland and related forest types on satellite images was developed. The resultant map (1 : 50 000, Transverse Mercator Projection) includes 26 mapping units and fits rather well the Estonian habitat classification scheme. The map can be employed for land use conflict resolution, for monitoring at landscape and habitat level, as well as in biodiversity studies.

Key words: GIS, habitat classification, Landsat TM, mapping, remote sensing, wetlands.

INTRODUCTION

Estonian diverse landscapes have a large proportion of natural and seminatural vegetation types that have been lost in many parts of Europe. For better planning of sustainable usage or protection of sites that are valuable for biodiversity, a multilateral inventory and a well-grounded monitoring system are necessary. The core of that system will be based on large protected areas, constituting a reference set for other territories under direct human impact. An essential part in the inventory of landscapes or habitats is the composing of several thematic maps. After the abandoning of Soviet restrictions there is a possibility to react to the urgent need for accurate cartographic information in Estonia and the other Baltic

States. The fastest way to produce medium-scale habitat maps is by using computer-aided classification of satellite data, supported by available databases, previous maps, and the necessary fieldwork. The level of detail that can be identified and the accuracy of such analyses depend on the nature of the region studied as well as on the spectral, spatial, temporal, and radiometric resolution of the sensor used. The best data for these purposes is provided by the Landsat satellite Thematic Mapper (TM) sensor, where radiometric information is gathered from seven different wavebands of the electromagnetic energy spectrum (Table 1). Landsat TM images have been successfully used for habitat mapping, mainly in North America, e.g. in Kansas (Lauver & Whistler, 1993), in the region of northern lakes (Wolter et al., 1995), and in Maine (Sader et al., 1997).

Table 1. Remote sensing and auxiliary data available for the study area

Data	Scale/resolution
Digital satellite data (band sequential format)	
Landsat Thematic Mapper (TM)	30 m (spatial)
Waveband TM1 – blue	0.45–0.52 µm (spectral)
Waveband TM2 – green	0.52–0.60 µm
Waveband TM3 – red	0.63–0.69 µm
Waveband TM4 – reflective infrared	0.76–0.90 µm
Waveband TM5 – mid-infrared	1.55–1.75 µm
Waveband TM6 – thermal infrared	10.4–12.5 µm
Waveband TM7 – mid-infrared	2.08–2.35 µm
Digital data bases (ARC/INFO coverages)	
Rivers	1 : 50 000
Roads	1 : 50 000
Settlement	1 : 50 000
Nature reserve area	1 : 50 000
Reserve area with a 2 km wide buffer zone	1 : 50 000
Auxiliary data	
Aerial photographs, 1991	1 : 10 000
Forest survey maps, 1987–1993	1 : 20 000
Topographic maps	1 : 50 000
Peatland map	1 : 400 000
Soil maps	1 : 200 000
Field observations, 1996, 1997	

Up to now, visual interpretation of satellite images (similar to the work with aerial photos) has been used for mapping purposes in Estonia. In that way, using SPOT (Système Pour l'Observation de la Terre) data, the Estonian base map at scale 1 : 50 000 (Sagris & Krusberg, 1997) and the CORINE land cover map (1 : 100 000) have been prepared (Sagris et al., 1997; Meiner, 1999). In some cases computer-aided classification of satellite data has been carried out as well (Öberg et al., 1992; Nisell et al., 1993; Aaviksoo, 1995).

For extraction and interpretation of data from a satellite image, a classification scheme of land cover and habitat types is needed. The land cover and land use classification scheme of Anderson et al. (1976) is usually applied worldwide. The first attempt at adopting this scheme for interpretation of satellite images of Estonian mire landscapes was made earlier for the mapping of the Endla Nature Reserve (Aaviksoo, 1995). Another scheme with 33 classification units (Aaviksoo, 1998) was proposed for landscape satellite monitoring of the whole territory of Estonia (Table 2).

Table 2. Land cover classification scheme for Estonia (Aaviksoo, 1998) and its adequacy with clusters of unsupervised and supervised satellite data classification for the Alam-Pedja NR area

No.	Land cover type	Clusters of the unsupervised classification	Clusters of the supervised classification
1	Inland waterbodies	16	1
2	Sea	NP	NP
3	Sand beaches	NP	NP
4	Gravel, shingle, and stone beaches	NP	NP
5	Coastal meadows	NP	NP
6	Floodplain grasslands	16	1
7	Reed beds	7	1
8	Open fens and transitional bogs	8, 13, 22, 24, 30, 32, 34	2, 7
9	Wooded fens and transitional bogs	8, 24, 30, 33	3, 5
10	Minerotrophic swamp forests, transitional bog forests	2, 4, 6, 7, 12	4, 6
11	Open bogs	8, 12, 34	7, 8, 9
12	Open ridge–hollow–pool bogs	7, 8, 12, 16, 18, 19	1, 10
13	Dwarf shrub and wooded bogs	1, 6, 7, 8, 13, 16, 18, 19, 27	11, 12, 13
14	Bog forests	1, 7	14
15	Natural grasslands	9, 15, 23, 31	15
16	Alvar shrublands	NP	NP
17	Shrublands	5, 11	4, 12, 16
18	Clear-cut areas	31	15
19	Natural regeneration of clear-cuts	4, 5, 10, 11, 31, 35	16
20	Cultivated forest stands	10	17
21	Young coniferous forests	10	17
22	Old coniferous forests	1, 7, 25	22
23	Mixed forests	*	*
24	Deciduous forests	2, 3, 4, 5, 22, 37	18, 19, 20, 21
25	Croplands	9, 31	23
26	Vegetable fields	10, 11, 14, 17, 20, 21, 29, 36	24
27	Industrial crops	NP	NP
28	Cultivated grasslands	31	15
29	Milled peat areas	20, 26, 28	25
30	Abandoned peat areas	19, 20, 28	26
31	Sand and gravel pits	**	**
32	Continuous urban areas	**	**
33	Discontinuous urban areas	**	**

* Added when necessary during the processing of satellite data.
** Differentiated on the basis of GIS coverages by cutting off from the satellite map.
NP, not present.

Still, the land cover classification schemes should be revised for any new application. Depending on the level of investigation as well as on landscape character (natural, semi-natural, and rural), an aggregation or splitting of the existing classification units will be required. In all cases, difficulties with the classification accuracy of land cover types are met. As a consequence, the necessity for gathering field data in situ is obvious: this is also a very desirable prerequisite for making land use decisions and management plans (Nedler et al., 1995).

Composition of the current map was initiated by the need to elaborate the management plan for a new area – the Alam-Pedja Nature Reserve (NR) – taken under protection in Estonia. This area, mostly consisting of wetland, was for more than four decades sealed off by the Soviet army and no detailed habitat or vegetation map has ever been made of this area. The vegetation map of Estonia (Laasimer, 1965) at a scale of 1:600 000 dates back to the 1950s, the forest survey map (1:20 000) covers only the corresponding areas.

The main goals of our study were:

(i) to use Landsat TM data for demonstration and analysis of the wetland diversity;

(ii) combining the satellite data with all available relevant auxiliary material and field data, to elaborate a corresponding mapping scheme for Estonian wetland dominated landscapes;

(iii) to compose for the Alam-Pedja NR and its closest surroundings (2 km wide buffer zone) a habitat type map, which could be used for different purposes: for description of biodiversity at corresponding levels, for planning nature protection activities and management, for resolution of land use conflicts and for (retrospective or perspective) change detection.

MATERIAL AND METHODS

Area

The Alam-Pedja NR (260 sq. km^2, centre co-ordinates 58°29′ N and 26°12′ E) is situated in the central part of Estonia (Fig. 1), where the Pedja River joins the Põltsamaa River and then the Suur-Emajõgi River. Orographically the nature reserve lies in the Võrtsjärv Lowland landscape region, and most of its territory was embraced by postglacial Great-Võrtsjärv Lake after the retreat of the glacier about 11 000 BP (Arold, 1993).

The Alam-Pedja NR is one of the wettest areas in Estonia (Lõhmus et al., 1993). There are 12 watercourses within the nature reserve, with a total length of 114.5 km, 55 oxbow lakes (previous meanders of the Suur-Emajõgi River), and brooks with a total length of 50.6 km (Ader & Tammur, 1997). Mires occupy about half of the total area of the nature reserve. Among them prevail raised bogs, separated by paludified and swamp forests of different types. Along bigger rivers extensive floodplain meadows as well as fragments of deciduous floodplain forests can be found. In the surroundings of the nature reserve the topography is more elevated and agricultural lands and forests on automorphic soils are found there.

Data

The Landsat 5 TM image from 12 June 1995 was used as the source material. The Landsat TM sensor records data in seven wavebands: TM1, TM2, and TM3 in the visible region, TM4 in the near infrared, TM5 and TM7 in the middle infrared, and TM6 in the thermal infrared region (Table 1).

Aerial photos, various maps (Table 1), as well as vegetation descriptions made for the purposes of vegetation and floristic inventory were used to support satellite image interpretation.

Pre-processing of data

The satellite image was enhanced for visual display by means of linear contrast stretching, using saturation points at 2.5%. Geometric correction procedure used 60 ground control points for rectifying the satellite image into Transverse Mercator projection by means of the Baltic Map System, created in 1993. Linear mapping function and nearest neighbour resampling were used until the residual mean square was about half of the pixel size – 18 m.

Image processing

The main task in image processing is to extract thematic information from the 7-dimensional feature space. In this study, data of the six reflective TM bands were used for both unsupervised and supervised classification while TM6 was omitted. The TM6 band, especially scanned during the nighttime, is important for indication of unfavourable sites for regeneration of clear cut areas (Nordberg, 1993); in nature reserve areas the usage of this thermal electromagnetic energy band does not add essential information.

For classification of land use and land cover characteristics in regions of forest and emergent wetland habitats, Landsat TM bands 3, 4, and 5 are considered to be superior to other TM spectral regions (Trolier & Philipson, 1986; Sader, 1989; Sader et al., 1995). Combining the data in a red–green–blue (RGB monitor) colour-gun in the order 4–5–3, we get an image which represents the basic Landsat TM satellite information in terms of greenness, brightness, and moisture content. Thus, the spectral band TM4 mainly contains information about the vegetation biomass; TM5 characterizes water content in plant tissues, and TM3 the absorption of chlorophyll (Eastman, 1997).

Classification

Classification of satellite data is based on the physical methods of detecting biotic and abiotic features of the landscape. By this procedure, labels are attached to picture elements (pixels) according to their spectral character. This labelling is implemented by a computer, "taught" beforehand to recognize pixels with spectral similarities (Richards, 1994). Each pixel within a satellite image (matrix of scanned rows and columns) can be clustered using two main approaches: automatic or unsupervised classification, and supervised classification, which identifies clusters on the basis of a priori knowledge, obtained through a combination of fieldwork, analysis of aerial photography and maps, and personal experience (Jensen, 1996).

In the unsupervised classification procedure, reflectance data are automatically clustered by a computer program. By retaining all clusters established by the program, clustering algorithm implies grouping of pixels in the multispectral space according to the histogram peaks technique. In IDRISI, an iterative optimization clustering procedure is realized by the ISODATA algorithm presented by Ball & Hall (1965, cited in Eastman, 1997). In addition to the raw image bands, ISODATA clustering requires a colour composite image (TM453 in our case) for the cluster seeding process and the iterative process thereafter exploits the maximum likelihood method. According to that, for every pixel the product of its probability of belonging to a certain class and the class's integral probability are calculated, and the pixel is assigned to the class having the largest product (Swain, 1978).

At the labelling of each spectral class of the automatically clustered image, we found that several land cover types were included in one cluster. For example, comparison of the colour composite image with clustering results revealed that digital number (DN) 42 in TM453 may represent vast reed-bed areas and ridge–hollow–pool complexes of raised bogs, and even several pine forest habitats. The automatic clustering algorithm aggregated all these pixels into cluster 7 (Table 2) even if the data from all six reflective bands were used in the classification procedure. Because of this misclassification, the task of relating clusters with similar reflectance values to different mapping units was undertaken by using supervised classification.

For supervised classification, a set of "computer training polygons" must be created beforehand. The training polygons were first delineated following the classification scheme of Pan-Estonian land cover types (Table 2, the 1st column) identified in the study area (land cover types 2–5, 16, and 27 are missing in the Alam-Pedja NR and types 31–33 were not differentiated spectrally, but were cut off from the satellite image). For this purpose the aerial photo interpretation results were used. For woodlands and clear-cuts, the training polygons were chosen on the basis of the forest site type maps; cultivated areas were partly determined from land use maps. Also the plant communities inventory data collected in 1996–97 and linked to 1 : 10 000 topographic maps were utilized.

Altogether 78 training polygons were digitized on the screen using the false colour composite image (TM453). The criterion for the minimum polygon area was the size of at least 5 pixels inside the homogenous spectral pattern, i.e. the reference mapping unit had to cover at least 0.45 hectares. Before initiating the maximum likelihood classifier, the spectral signatures of all clusters (six bands – TM1–TM5 and TM7) were analysed. Many habitats with different ecological conditions had very similar signature curves, especially in forests. These training polygons were not used, therefore, in the classification procedure.

After merging summer crops and bare soils into the joint cropland class, and swamp forests with pubescent birch (*Betula pubescens*), black alder (*Alnus glutinosa*), and also those with a scattered regrowth of spruce (*Picea abies*) into the mobile water deciduous swamp forests class, the list of mapping units was compiled. Checking of the mapping units was repeatedly carried out in nature, and when misclassifications were recognized, new training polygons were digitized and the classification procedure performed again.

Finally the study area (nature reserve plus buffer zone around it) was cut out from the satellite image, georectified before into Transverse Mercator Projection with the central meridian 24°00′00″ E, based on the Baltic Map System, 1993. All vector layers and the legend were added as well (Fig. 1). The final map explains diversity and representation of habitat types in landscape mainly according to their nutrient status and life forms.

RESULTS

Unsupervised classification

The unsupervised classification procedure yielded 37 spectral clusters. Table 2 reveals the resultant mismatch between these clusters and the underlying land cover types: most of the clusters (21 of 37, or 57%) correspond to two or more land cover types, and vice versa – numerous land cover types are dispersed between several clusters. Especially splintered are clusters 7, 8, 10, and 4. Cluster 7 represents six land cover types: reed beds, minerotrophic swamp forests and transitional bog forests, open ridge–hollow–pool bogs, dwarf shrub and wooded bogs, bog forests, and old coniferous forests on mineral soil. Hence, it embraces an extremely wide amplitude of ecological conditions – from very wet habitats to relatively dry ones. A similar situation appears for other analogous clusters. For example, cluster 4 is spread over only three land cover types, which, however, are all significantly different: minerotrophic swamp forests and transitional bog forests, natural regenerations of clear-cut areas, and deciduous forests (Table 2). Efforts to divide these clusters into subclusters failed due to a great similarity in dense canopy cover, which does not allow the determination of ecological features of the habitat.

Supervised classification

The supervised classification yielded finally 26 clusters, having a surprisingly low average standard error of reflectance values – 1.65 (Table 3). Variation in

Table 3. The mapping units in terms of habitat types, established by means of supervised classification procedure, and their mean reflectance value (digital number) of Landsat TM wavebands

Unit No.	Habitat type	Habitat code*	Mean reflectance value of Landsat TM waveband ± SE					
			TM1	TM2	TM3	TM4	TM5	TM7
1	Shallow waterbodies, submerged floodplains, reed beds, ridge–pool bogs	6.1.1; 2.2; 2.2.1.1; 3.2.2.3	58±2.0	21±1.2	18±1.3	18±5.12	10±3.7	5±1.4
2	Open fens, transitional bogs with reeds	3.1.1; 3.1.2.1	61±1.0	26±0.6	22±0.8	67±3.1	57±2.2	18±1.1
3	Wooded fens	3.1.1	60±0.9	26±0.6	22±0.8	72±1.2	66±2.1	21±1.0
4	Low swamp birch forests and birch shrublands	1.4.1.1	61±0.8	26±0.5	22±0.9	83±2.0	76±1.6	24±1.1
5	Higher swamp birch forests	1.4.1.1	58±1.2	24±0.7	19±0.7	81±2.8	57±1.7	16±1.0
6	Transitional bog pine forests	1.4.2.1	61±1.2	26±0.6	24±1.0	66±2.8	47±1.8	16±1.1
7	Laggs, transitional quagmires	3.1.2.2	62±1.1	27±0.7	25±1.2	80±3.0	47±3.8	15±1.4
8	Open cotton-grass hummock bogs	3.2.2.1	62±1.1	27±0.7	26±0.8	67±2.5	58±2.0	20±1.1
9	Open hummock bogs	3.2.2.1	65±1.4	29±0.7	30±1.2	69±2.3	66±1.6	24±1.1
10	Open ridge–hollow bogs	3.2.2.2	62±1.3	27±0.7	27±1.5	81±2.7	58±2.1	19±1.1
11	Wooded cotton-grass hummock bogs	3.2.2.2	61±1.2	26±0.6	23±0.0	63±1.4	55±1.5	19±1.0
12	Wooded hummock bogs, floodplain willow shrublands	3.2.2.1; 1.2.2	61±1.3	25±0.7	23±1.3	56±2.3	46±2.4	17±1.3
13	Wooded ridge–hollow bogs	3.2.2.2	62±1.2	26±0.7	25±0.9	55±3.0	57±1.9	21±1.1
14	Bog pine forests	1.4.3.1	58±1.0	22±0.5	18±0.7	45±0.9	34±1.5	11±0.9
15	Natural grasslands, cultivated grasslands, grassy clear-cut areas	2.1; 8.1	62±1.1	27±0.7	21±0.6	118±8.7	77±2.2	22±1.0
16	Willow shrublands, shrubby clear-cut areas	1.2.2	59±1.1	26±0.5	20±0.4	97±3.3	69±2.4	20±1.1
17	Planted young spruce stands	–	61±1.1	27±0.7	23±0.9	95±2.9	86±2.1	27±1.3
18	Paludified birch and alder forests	1.3.1	58±1.0	23±0.7	18±0.5	78±1.2	50±1.0	14±0.7
19	Mobile water deciduous swamp forests	1.4.1.2	58±1.1	23±0.9	18±0.6	69±3.0	48±2.4	14±1.2
20	Alder forests	–	59±1.0	24±0.7	19±0.5	94±3.1	59±1.7	16±0.9
21	Aspen forests	–	60±0.8	22±0.6	18±0.6	81±3.2	49±0.9	13±0.7
22	Coniferous forests	–	57±1.1	21±0.5	17±0.5	51±3.5	30±3.3	9±1.0
23	Croplands	–	62±1.3	27±0.8	22±0.9	108±2.3	66±1.4	19±1.0
24	Arable lands, unvegetated lands	–	64±1.2	29±0.6	22±0.7	93±3.2	67±1.5	20±0.9
25	Milled peat areas	–	64±1.5	28±1.3	37±2.1	57±2.7	101±5.0	50±3.3
26	Abandoned peatlands	–	59±1.1	24±0.9	28±1.1	43±1.5	63±4.4	21±2.1

* Habitat code number is given according to Paal (1997); the first position shows the habitat formation number (forests, grasslands, mires, etc.); the second position, habitat class number; the third position, habitat group number; and the fourth, habitat type number; – not distinguished.

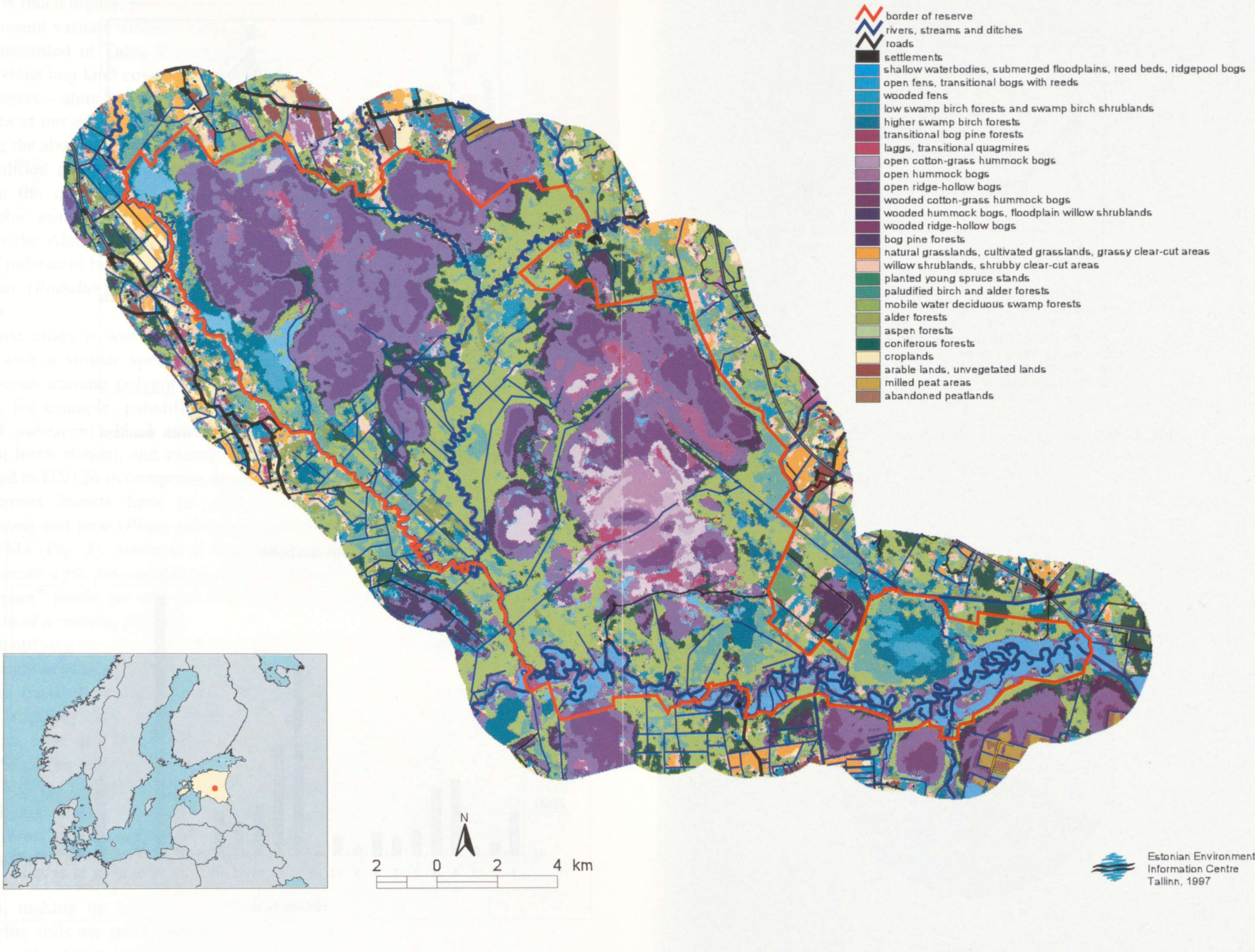

Fig. 1. Location of the study area and the habitat type map of the Alam-Pedja Nature Reserve.

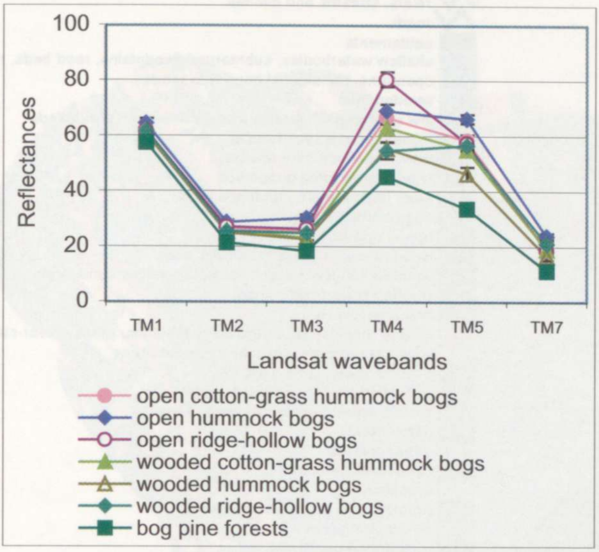

Fig. 2. Spectral signatures of bog habitats with standard deviation values.

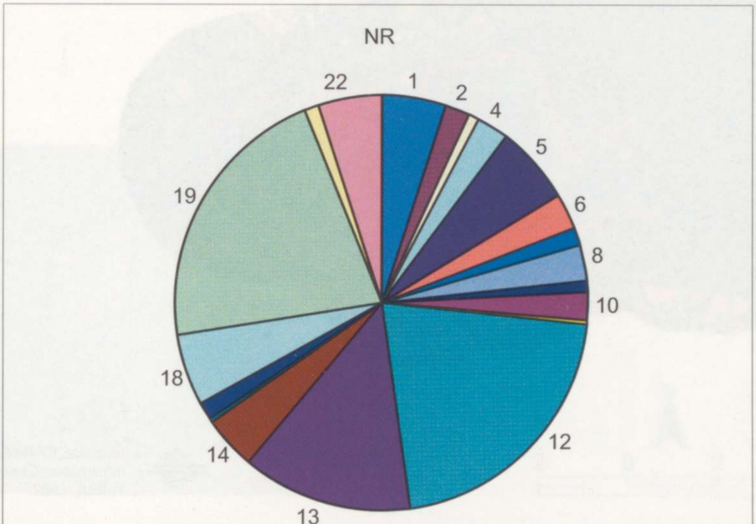

Fig. 4. Occurrence of habitat types in the Alam-Pedja Nature Reserve (25987.6 ha). Habitat type see Table 3.

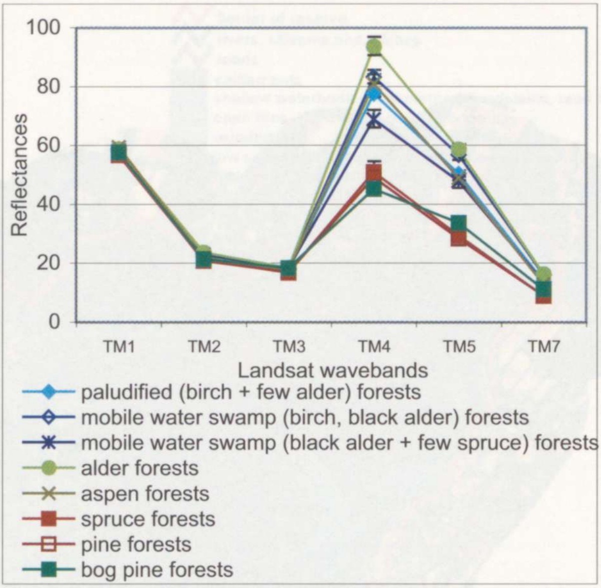

Fig. 3. Spectral signatures of forest habitats with standard deviation values.

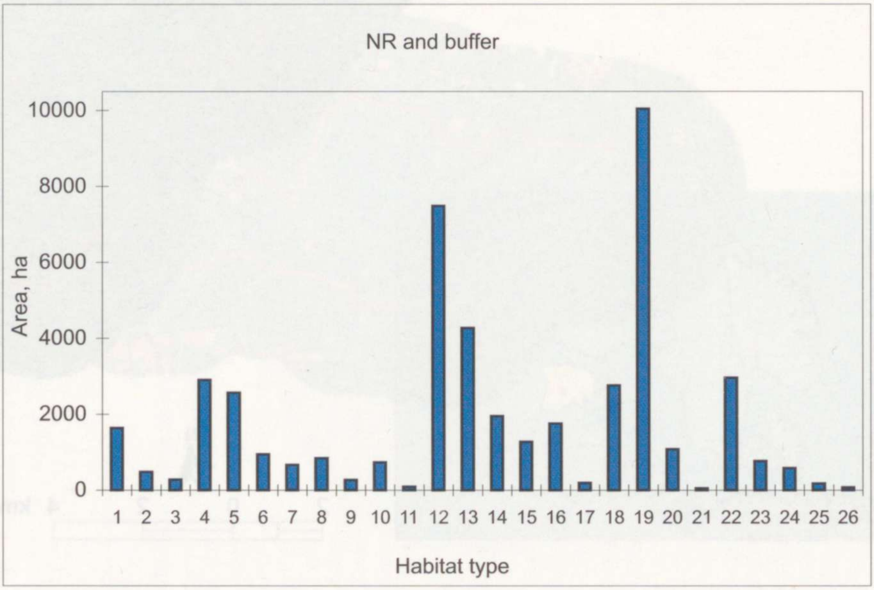

Fig. 5. Occurrence of habitat types in the Alam-Pedja Nature Reserve (NR) and buffer zone (46848.8 ha). Habitat type see Table 3.

cluster 1 is much higher, proving heterogeneity of this cluster. Classification units 1–14 represent various wetland habitat types (Table 3). The number of land cover types represented in Table 2 is notably smaller – only 9 classes (units 6–14); the 4 previous bog land cover types (units 11–14, Table 2), for example, fall now into 7 clusters – units 8–14 (Table 3, Fig. 2).

In spite of our efforts to teach the computer to recognize different forest types, following the above procedure, we did not succeed. Still, it enabled us to discover that paludified deciduous forests and mobile water swamp forests have TM4 values in the interval of 69–78, while this value for deciduous forests on automorphic mineral soils is even higher than 80 (Fig. 3). The wet deciduous forests in the Alam-Pedja NR are represented mainly by mobile water swamp stands of pubescent birch or black alder; in dryer places gray alder (*Alnus incana*) and aspen (*Populus tremula*) forests are characteristic of deciduous stands (Table 3).

In some cases it was possible to distinguish ecologically different forest habitats with a similar spectral response by delineating rather extensive and homogeneous training polygons using forest site type maps or field studies. In that way, for example, paludified birch forest (consisting of *Betula pendula* as well as *B. pubescens*), black alder forest, mobile water deciduous forest (mainly pubescent birch stands), and swamp birch shrubland (*B. pendula*, *B. pubescens*), all referred to DN126 in composite image, were differentiated (Table 3).

Coniferous forests have an almost similar spectral curve for spruce (*Picea abies*) and pine (*Pinus sylvestris*); spruce exhibits a slightly higher value only in TM4 (Fig. 3). Attempts to separate these stands were discarded because spruce forests were not composed of homogeneous enough stands in the study area; 5 "pure" pixels per one site and 60 for the whole set are required for the delineation of a training polygon.

For identifying the real content of clusters 2, 4, 18, and 24 (Table 3) fieldwork or large-scale auxiliary maps were also needed.

On the basis of the established classification, we can analyse the area of different mapping units as well (Table 4, Figs. 4 and 5). Various wetlands – floodplains, fens, and bogs – represented in units 1–14 and partly in 16, cover about 60% of the territory in the Alam-Pedja NR (marked with red line on the map in Fig. 1). Dominant among them are raised bogs (especially wooded hummock bogs and wooded ridge–hollow bogs), which occupy 45% of the area. Of mire forest vegetation, low swamp forests and birch shrublands (unit 4) are very common. Very typical are mobile water deciduous swamp forests (unit 19), which cover 21%, and paludified forest stands – paludified birch and alder forests (unit 18), making up 6% of the nature reserve. Pine and spruce forests on automorphic soils are spread altogether on 5% of the area; gray alder and aspen stands are rather fragmented and cover only 1%.

Table 4. Presentation of the mapping units in the Alam-Pedja NR (Fig. 4) and in the whole territory under investigation (NR and buffer zone, Fig. 5). Habitat type unit see Table 3

Unit	NR		NR and buffer zone	
No.	ha	%	ha	%
1	1322.6	5.1	1640.0	3.5
2	451.5	1.7	483.9	1.0
3	223.0	0.9	275.2	0.6
4	641.1	2.5	2899.8	6.2
5	1601.1	6.2	2563.8	5.5
6	756.0	2.9	949.0	2.0
7	356.6	1.4	664.9	1.4
8	726.2	2.8	842.7	1.8
9	171.3	0.7	267.8	0.6
10	589.4	2.3	738.2	1.6
11	87.2	0.3	90.9	0.2
12	5524.6	21.3	7484.0	16.0
13	3406.2	13.1	4274.3	9.1
14	1087.2	4.2	1946.7	4.2
15	94.5	0.4	1278.6	2.7
16	363.6	1.4	1753.5	3.7
17	13.8	0.1	191.5	0.4
18	1444.8	5.6	2758.1	5.9
19	5532.8	21.3	10041.9	21.5
20	267.5	1.0	1083.2	2.3
21	6.5	<0.1	54.4	0.1
22	1319.4	5.1	2966.4	6.3
23	0.3	<0.1	764.5	1.6
24	0.4	<0.1	578.6	1.2
25	NP	NP	179.2	0.4
26	NP	NP	77.7	0.2
Total	25987.6	100	46848.8	100

NP, not present.

In the 2 km wide buffer zone, habitats under strong human impact – arable lands with or without crops, milled peat areas, and abandoned peatlands – are found (Fig. 1). Clear-cuts, planted young conifer stands, and cultivated grasslands are also present there.

Statistical accuracy of the classifications

For the assessment of the accuracy of both maps, generated by means of unsupervised and supervised classification, information from two sources is necessary: (i) the remote-sensing-derived classification map and (ii) reference test information (Jensen, 1996). Stratified random sampling was used for the selection of pixels. This sampling procedure combines the strong geographic coverage of the systematic approach with the low potential for bias of the random scheme (Eastman, 1997).

Unfortunately, it is not possible to give a statistically very sound calculation of errors. This is connected with the great number of classification units and their complex character. If we take into consideration only the number of classification units (according to the supervised classification 26 habitat/vegetation types) and Congalton's (1991) suggestion to collect a minimum of 50 samples for each cluster in the error matrix, 1300 control samples are needed. Jensen (1996) suggests the use of even 75 or 100 samples per type when the classification has more than 12 categories, i.e. in the current case 1950–2600 samples. It was impossible to obtain such a big massif of data owing to the limited investigation period. Moreover, for gathering the relevant set of field data, a rather detailed vegetation map is necessary. However, once we already have such a map, the fundamentally less precise map based on satellite data will become pointless. Therefore, only a very general assessment of the accuracy of the two maps is given here.

Using stratified random pixels for accuracy assessment, it appeared that by unsupervised classification only up to 12% of the control pixels were determined exactly. These included open cotton-grass hummock bog, grasslands, shrublands, paludified forests, mobile water swamp birch forests, and swamp alder forests. For supervised classification the estimated mapping units were inspected three times in the field. These data could be used as reference information to improve supervised classification results to the extent that 90% of the pixels were determined exactly. Even if this assessment is very rough, it may be concluded that unsupervised clustering can be used only for general determination of land cover types, but it does not satisfy the requirements of classification at habitat or vegetation type level.

DISCUSSION

The current study suggests the following main shortcomings of automatic unsupervised classification:

(i) several clusters correspond simultaneously to numerous land cover types, e.g. pixels of cluster 7 are scattered in six land cover types: reed beds, minerotrophic swamp forests and transitional bog forests, open ridge–hollow–pool bogs, dwarf shrub and wooded bogs, bog forests, and old coniferous forests on automorphic mineral soils (Table 2);

(ii) merging into one cluster various principal types of mires – fens, transitional bogs and raised bogs, each one of which has a specific ecology and represents a different mire development stage;

(iii) grasslands (natural and cultivated) with luxurious vegetation are indistinguishable from winter rye fields, where crop biomass is rather high in the beginning of June;

(iv) abandoned peatlands (in the buffer zone) are almost inseparable from milled peat areas as well as from dwarf shrub bogs and wooded bogs.

Therefore, on the one hand it is obvious that the ecological amplitude (cf. ii above) and vegetation characteristics of these clusters are very variable and inappropriate as mapping units. On the other hand, we can also conclude that the classification scheme of land cover types elaborated for satellite monitoring of Estonian landscapes (Aaviksoo, 1998), which includes altogether 33 types, is not suitable for labelling the coarse clusters derived from unsupervised classification of satellite data. The less time and work-consuming automatic classification is justified only in cases where a quick and/or draft overview of the study area is needed. Even then we recommend the use of no more than 12–15 clusters.

The identification difficulties we met in the process of classification (the appearance of wet and dry habitats in one cluster) could be largely overcome by exploiting the supervised classification method. For receiving a suboptimal solution, with ecologically better-defined mapping units, through this approach, the delineation of training polygons for all actual units as well as a corresponding field inventory are necessary for an iterative justification of the classification scheme. As a matter of fact, even this approach did not solve all the identification problems (the differentiation of forests in particular still remained rather rough), but this is a question of the inadequacy of spectral-based classification and/or remote sensing for detailed mapping purposes rather than of classification methodology.

A considerable amount of information needed for interpretation of Landsat TM data is contained in topographic, forest site type, and land-use maps. These maps enable quite a simple separation of the sites of different ecology intermixed in some mapping units of the final scheme (Table 3). For example, despite the fact that cluster 1 includes sites of four land cover types (Table 2) all having similar reflectance parameters, floodplain grasslands never border on bogs and their separation is not problematic on the map. Difficulties may arise sometimes in distinguishing between shallow water, reed bed, and floodplain grasslands, which may have continuous transitions in nature. Three habitat types of cluster 15 (Table 3) – natural grasslands, cultivated grasslands, and grassy clear-cut areas – all have a large biomass. They are rather easily separable from each other also on the map: cultivated grasslands with clear contours are present only in the buffer zone of the nature reserve, clear-cut areas are confined to forest massifs and do not mix on the map with grasslands either.

In this connection a question arises: what is the real essence of this kind of map if even in the case of supervised classification several clusters are split between two or more land cover types. At the same time, the land cover types themselves do not have very clear vegetational or ecological interpretation. Some of them are defined mainly on the basis of vegetation (reed beds, natural grasslands, shrublands, mixed forests, etc.); for others the landscape features – natural (e.g. beaches, waterbodies) or anthropogenic (exploited peat areas, urban areas) – are determinative (Table 2). Several types are defined using both criteria, e.g. floodplain grasslands, coastal meadows, various bogs, clear-cut areas, etc.

During the composition of the map on the basis of supervised classification, the obtained clusters were compared with existing general classifications of the Estonian vegetation for a more conscious interpretation. It appeared that the vegetation classification units, which have been drawn up on the basis of ecological–phytocoenological principles (Laasimer, 1965; Marvet, 1970; Krall et al., 1980; Lõhmus, 1984), do not agree well with the mapping units established on the basis of satellite data. This is caused mainly by a failure of mutual fitting of the geographical ranges of vegetation units and reflectance classes. Only the spectral characteristics of bogs permit us to get units that are in rather good conformity with the ecological–structural typology of mires traditionally used in Estonia (Masing, 1975, 1984; Masing & Paal, 1998; Paal et al., 1998). Differences in bog spectral signatures are obvious, and allow us to separate seven different types (Fig. 2) that were all well classified by the maximum likelihood algorithm. Where necessary they were specified using training polygons inside well-known (and checked in the field or on aerial photos) communities.

The best possibility for finding general concordance between the classification of vegetation, landscape typology, and Landsat 5 TM data reflectance units seems to be afforded by habitat classification. Habitat includes both biological (plant community) and abiotic features of a certain site (landscape facet or cell). Since the species composition and general physiognomy (structure) of plant communities are the best integrated indicators of ecological conditions in a site (Whittaker, 1965), habitats are usually named and classified according to vegetation. This method has been used, for example, in the CORINE Biotopes project (Devillers et al., 1992) and EUNIS habitat classification (Davies & Moss, 1997) as well as in the Estonian vegetation site types (Paal, 1997).

Indeed, if we interpret mapping units of supervised classification in terms of habitat types, using the vegetation inventory data for that, the ecological features of most established units become rather clear. Many units are equivalent with habitats at their lowest (type) level, for others the correspondence follows habitat group or habitat class level (Table 3). Noteworthy is the fine scale of bog habitat separation by means of the reflectance parameters; some of them are even divided into habitat subunits according to density of the tree layer or other vegetation peculiarities (e.g., open hummock bogs, open cotton-grass hummock bogs, wooded cotton-grass hummock bogs). On that basis we can conclude that the Landsat 5 TM data will enable good mapping of bog habitats; the result concerning other mire types is also quite satisfying. Only the wettest sites have inseparable spectral signatures but, as said before, in most cases these habitats may be distinguished on the map according to their topographic features.

As to forest habitats, their more detailed classification is not feasible in satellite remote sensing, because the stands can at best be distinguished only according to the dominant tree species, the crowns of trees overshadowing the lower layers of vegetation that are usually more informative characteristics of site conditions (Nilson & Olsson, 1995). An essential role in the separation of forest stands belongs to the vertical canopies with inherently low actual cover, whose nadir view includes substantial amounts of soil/water cover (Spanglet et al., 1998). Therefore, discrimination between alder (thick) and aspen (sparse) forest stands, with different lower layers, was possible, even if their spectral signatures were very similar, owing to differences only in TM4 (Fig. 3).

CONCLUSIONS

A versatile analysis of spectral signatures of satellite data enabled the development of a classification scheme of sites that can be interpreted in terms of habitat types. The map created on that basis describes well different wet habitats – floodplains, mires, paludified forests; at the same time the separation of forests, especially on mineral soil, remains rather inexact. Still, a map of this quality is sufficient for general nature management planning in the reserves where wetlands dominate. It is possible to establish areas (zones) where different conservation regimes must be introduced or to identify habitats demanding special care.

The results demonstrate that the present level of remote sensing does not allow discrimination of different tree species. What can be extracted from a single image are really more integral characteristics, such as total amount of chlorophyll or water, canopy cover, some kind of surface roughness, leaf angle (aspen *versus* linden), etc.

In the GIS environment, creation of the map is quickly achieved by overlaying all possible data coverages (roads, settlements, etc.) onto the raw image. The classified image, especially when it has more a priori data taken into account in the supervised classification process, enables one to get an overview of all the mapping units or their aggregated categories and to calculate their areas on the required level of generalization. The resulting map can, in addition to nature reserve management purposes, also be used for natural resource management, for monitoring the protected area and its neighbourhoods, and for the assessment of mutual spatial influence of different landscapes (ecology *versus* economy).

ACKNOWLEDGEMENTS

Satellite data courtesy of the Estonian Environment Information Centre, Ministry of the Environment. Authors are thankful to Prof. J. Aaviksoo and Prof. T. Nilson for critical revision of the manuscript and to Mr. I. Part for help with English.

REFERENCES

Aaviksoo, K. 1995. Vegetation of Endla Nature Reserve classified on the basis of Landsat TM data. In *Consortium Masingii* (Aaviksoo, K., Kull, K., Paal, J. & Trass, H., eds.), pp. 27–36. Tartu University Press, Tartu.

Aaviksoo, K. (ed.) 1998. *Maastike kaugseire aasta-aruanne 1997*. EV Keskkonnaministeerium, Info- ja Tehnokeskus, Tallinn.

Ader, A. & Tammur, E. 1997. *Alam-Pedja looduskaitseala*. EV Keskkonnaministeerium, Info- ja Tehnokeskus, Tallinn.

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. In *U.S. Geological Survey. Professional Paper 964*, pp. 1–28. U.S. Geological Survey, Washington, DC.

Arold, I. 1993. *Eesti maastikud*. Tartu University Press, Tartu.

Ball, G. H. & Hall, D. J. 1965. *A Novel Method of Data Analysis and Pattern Classification*. Stanford Research Institute, Menlo Park, CA.

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

Davies, C. E. & Moss, D. 1997. *EUNIS Habitat Classification. Final Draft. Submitted to EEA by ETC/NC*. Museum National d'Histoire Naturelle, Paris.

Devillers, P., Pinborg, U., Devillers-Terschuren, J. & Ledant, J.-P. 1992. *CORINE-Biotopes. Habitats of the Nordic Countries*. Institut Royal des Sciences Naturelles de Belgique, Dansk miljøministeriet and Skov-og Naturstyrelsen, Brussels.

Eastman, J. R. (ed.) 1997. *User's Guide. Idrisi for Windows Version 2*. Clark Labs for Cartographic Technology and Orographic Analysis, Worcester.

Jensen, J. R. 1996. *Introductory Digital Processing: A Remote Sensing Perspective*. 2nd edn. Prentice Hall, Upper Saddle River, New Jersey.

Krall, H., Pork, K., Aug, H., Püss, E., Rooma, I. & Teras, T. 1980. *Eesti NSV looduslike rohumaade tüübid ja tähtsamad taimekooslused*. ENSV PM IJV, Tallinn.

Laasimer, L. 1965. *Eesti NSV taimkate*. Valgus, Tallinn.

Lauver, C. L. & Whistler, J. L. 1993. A hierarchical classification of Landsat TM imagery to identify natural grassland areas and rare species habitat. *Photogramm. Eng. Remote Sens.*, **59**, 627–634.

Lõhmus, E. 1984. *Eesti metsakasvukohatüübid*. ENSV Agrotööstuskoondise IJV, Tallinn.

Lõhmus, A., Püttsepp, Ü. & Tammur, E. 1993. Projekt "Alam-Pedja looduskaitseala". *Eesti Loodus*, 11/12, 353–356.

Marvet, A. 1970. Eesti taimekoosluste määraja. *Abiks loodusevaatlejale*, **61**, 3–56.

Masing, V. 1975. Mire typology of the Estonian S.S.R. In *Some Aspects of Botanical Research in the Estonian S.S.R.* (Laasimer, L., ed.), pp. 123–138. Acad. Sci. Estonian S.S.R, Tartu.

Masing, V. 1984. Estonian bogs: Plant cover, succession and classification. In *European Mires* (Moore, P. D., ed.), pp. 120–148. Academic Press, London.

Masing, V. & Paal, J. 1998. Estniska våtmarker – klassificering och biodiversitet. *Svensk Bot. Tidskr.*, **92**, 147–161.

Meiner, A. (ed.) 1999. *Eesti maakate. CORINE Land Cover projekti täitmine Eestis*. EV Keskkonnaministeerium, Info- ja Tehnokeskus, Tallinn.

Nedler, V. J., Crossley, D. C. & Cofinas, M. 1995. Using Geographic Information Systems (GIS) to determine the adequacy of sampling in vegetation surveys. *Biol. Conserv.*, **73**, 1–17.

Nilson, T. & Olsson, H. 1995. Effect of thinning cutting on boreal forest reflectance: A comparison of simulations and Landsat TM estimates. *Int. J. Remote Sens.*, **16**, 2963–2968.

Nisell, J., Lindell, T., Öberg, M., Meiner, A. & Saare, L. 1993. *Land-use/land cover classification of the Matsalu Bay drainage basin. Report No. 12*. Centre for Image Analysis, Swedish Space Corporation, Estonian Environment Information Centre, Uppsala.

Nordberg, M.-L. 1993. *Integration of Remote Sensing and Ancillary Data for Geographical Analysis and Risk Assessment for Forest Regeneration Failures*. Doctoral Dissertation. Stockholm University, Department of Physical Geography, Stockholm.

Öberg, M., Nisell, J. & Lindell, T. 1992. Development of open-cast mining areas in Estonia. *Remote Sensing. Information from the Swedish Space Corporation*, 22, 6–9.

Paal, J. 1997. *Eesti taimkatte kasvukohatüüpide klassifikatsioon*. EV Keskkonnaministeerium, Info- ja Tehnokeskus, UNEP, Tallinn.

Paal, J., Ilomets, M., Fremstad, E., Moen, A., Børset, E., Kuusemets, V., Truus, L. & Leibak, E. 1998. *Eesti märgalade inventeerimine 1997.a. Projekti "Eesti märgalade kaitse ja majandamise strateegia" aruanne*. Eesti Loodusfoto, Tartu.

Richards, J. A. 1994. *Remote Sensing Digital Image Analysis. An Introduction*. 2nd edn. Springer-Verlag, Berlin.

Sader, S. A. 1989. Multispectral and seasonal characteristics of northern hardwood and boreal forest types in Maine. In *Image Processing 89*, pp. 109–116. University of Nevada-Reno, ASPRS, Bethesda, MD.

Sader, S. A., Douglas, A. & Liou, W.-S. 1995. Accuracy of Landsat-TM and GIS rule-based methods for forest wetland classification in Maine. *Remote Sens. Environ.*, **53**, 133–144.

Sader, S. A., Hepinstall, J. A., Krohn, W. B., Randall, B. B. & Bartlett, R. I. 1997. Development and testing of a map for Maine. A status report. (*http://wlm13.umenfa.maine.edu/progs/un .../layers/veg/Classification-report.html*).

Sagris, V., Kask, R., Aaviksoo, K. & Meiner, A. 1997. CORINE land cover project: An Estonian approach. In *GIS – Baltic Sea States '97* (Vilu, R. & Arumaa, H., eds.), p. 47. Tallinn Technical University.

Sagris, V. & Krusberg, P. 1997. Estonian base map project and its applications. In *Geographical Information '97. Joint European Conference and Exhibition on Geographical Information. Proceedings*, Vol. 2 (Hodgson, S., Rumor, M. & Harts, J. J., eds.), pp. 937–945. IOS Press, Amsterdam.

Spanglet, H. J., Ustin, S. L. & Rejmankova, E. 1998. Spectral reflectance characteristics of California subalpine marsh plant communities. *Wetlands*, **18**, 307–319.

Swain, P. H. 1978. Fundamentals of pattern recognition in remote sensing. In *Remote Sensing: The Quantitative Approach* (Swain, P. H. & Davis, S. M., eds.), pp. 136–187. McGraw-Hill, New York.

Trolier, L. J. & Philipson, W. R. 1986. Visual analysis of Landsat Thematic Mapper images for hydrologic land use and cover. *Photogramm. Eng. Remote Sens.*, **52**, 1531–1538.

Whittaker, R. H. 1965. Dominance and diversity in land plant communities. *Science*, **147**, 250–259.

Wolter, P. T., Mladenoff, D. J., Host, G. E. & Crow, T. R. 1995. Improved forest classification in the northern Lake States using multi-temporal Landsat imagery. *Photogramm. Eng. Remote Sens.*, **61**, 1129–1143.

MÄRGALADE MITMEKESISUSE KAARDISTAMINE SATELLIITANDMETE JA GEOGRAAFILISTE INFOSÜSTEEMIDE ABIL ALAM-PEDJA LOODUSKAITSEALA NÄITEL

Kiira AAVIKSOO, Jaanus PAAL ja Tiina DIŠLIS

Alam-Pedja looduskaitseala elupaikade kaardistamise näitel on käsitletud kaugseire (Landsat 5 TM digitaalsed satelliitpildid, aerofotod), geograafiliste infosüsteemide (GIS) ja välitööandmete integreeritud kasutamisega seotud

metoodilisi probleeme, samuti satelliitinformatsiooni põhjal saadud kaardistamisüksuste ökoloogilise interpreteeritavuse küsimusi. Satelliitpildi üksuste – pikslite klassifitseerimiseks kasutati nii automaatset protseduuri kui ka suurema tõepära algoritmil põhinevat klassifitseerimist, millele eelnes konkreetsete elupaikade (arvutitöötlusel etteantavate nn. õpetuspiirkondade) uurimine looduses. Algsete õpetuspiirkondade sobivust kontrolliti klassifitseerimisprotseduuri käigus korduvalt välitöödel ja vajaduse korral lisati uusi. Enam tähelepanu pöörati kaardistamisel nendele üksustele, mida on võimalik piiritleda satelliitpildi ruumilist (piksli mõõtmed 900 m^2) ja spektraalset (7 laineala) lahutusvõimet arvestades. Selle põhjal töötati välja Eesti looduslikele tingimustele vastav soode, samuti soostunud metsade ja soometsa-elupaikade täpsustatud klassifikatsiooniskeem, milles kasutatakse kokku 26 kaardistamisüksust. Töö tulemusel valmis Eesti baaskaardi projektsioonis keskmise mõõtkavaga kaart (1 : 50 000), mida kasutati Alam-Pedja looduskaitseala kaitsekorralduskava koostamisel. See kaart võimaldab lahendada maakasutuses tekkivaid konfliktsituatsioone, samuti on see rakendatav elupaikade ja maastike seireks ja bioloogilise mitmekesisuse uurimiseks viimaste tasandil.