Research Article

Mateja Breg Valjavec*, Rok Ciglič, Krištof Oštir, and Daniela Ribeiro

Modelling habitats in karst landscape by integrating remote sensing and topography data

<https://doi.org/10.1515/geo-2018-0011>

Received Aug 11, 2017; accepted Mar 13, 2018

Abstract: Field mapping is an accurate but also time consuming method of detailed mapping of habitat types. Levels of habitat types are usually hierarchically nested at several levels. Our main research question therefore is: ‘How detailed can be modelling of habitat types with decision trees and digital data in karst landscape?’ Similar to studies in other (non-karst) environments we explored the basic properties of the habitats in Dinaric Karst study region (Classical Karst in Southwest Slovenia) and tested modelling of habitat types at three different levels of detail. To seek for the best set of predictor variables we used RapidEye satellite images, airborne images and digital elevation model. We prepared more than 60 explanatory variables and divided habitat polygons into training and testing samples to validate the results. The results proved that modelling with decision trees in Dinaric Karst landscape does not result in high accuracy at high detailed levels. Due to the presence of mine fields in the large area of Dinaric Karst (e.g. in Croatia and Bosnia and Herzegovina) the field mapping in this area is difficult therefore the findings from this study can be used for further development of mapping through remote sensing.

Keywords: habitat types, land cover, Dinaric Karst, multi-spectral satellite images, decision trees, RapidEye, digital elevation model

1 Introduction

Typologies of the environment, including those presenting habitat types, have different purposes such as making inventory, assessing, monitoring, managing, planning, predicting, sampling, modelling, presenting, and analysing the environment [1–7]. Classifications help us to reduce the complexity of the natural world [8] and help us to understand it better. An overview of the environment and its characteristics is necessary [9].

One of the most extensive habitat monitoring scheme in Europe is arguably the Natura 2000 programme, where each European Union (EU) member state develops its own guidelines for field-based conservation status monitoring. These guidelines require detailed surveying of many ecologically relevant variables, most of these classically considered out of reach for remote sensing [10]. Therefore, detailed field mapping of habitats is still the most established and still the most accurate method for mapping, classifying and monitoring Natura 2000 habitats. This method is also in use for karst landscapes. Remote sensing is a well-established method in global land cover mapping [e.g. 3, 11–13]. Due to constant technological development, it must be constantly tested also for mapping of habitat types [e.g. 14–16]. Besides, remote sensing plays an increasing role in ecology, biodiversity, and conservation research [17].

Beginning with simple, often subjective, hand-drawn maps, habitat mapping now involves complex species distribution models using mapped predictor variables derived from remotely sensed data [18]. In addition, hyperspectral imaging enables an accurate mapping of vegetation classes up to a certain detail, but the availability and spatial resolution of even the most advanced airborne hyperspectral sensors limits the differentiation of certain fine scale vegetation characteristics, needed to assess the local conservation status of habitats in a spatially explicit manner [19]. Therefore, we can add additional datasets to improve differentiating. For example, remote sensing data and topography data are widely used in modelling of land cover [20] as well as in detailed mapping of habitats and species distribution [21]. Very-high spatial resolution air-

*Corresponding Author: **Mateja Breg Valjavec:** Anton Melik Geographical Institute, Research Centre of the Slovenian Academy of Sciences and Arts, Gosposka ulica 13, SI – 1000 Ljubljana, Slovenia; Email: mateja.breg@zrc-sazu.si

Rok Ciglič, Daniela Ribeiro: Anton Melik Geographical Institute, Research Centre of the Slovenian Academy of Sciences and Arts, Gosposka ulica 13, SI – 1000 Ljubljana, Slovenia

Krištof Oštir: Faculty of Civil and Geodetic Engineering, University of Ljubljana, Jamova cesta 2, SI – 1000 Ljubljana, Slovenia

borne imaging spectrometry has shown its ability to map species composition with extremely high accuracy. This technology is still only available for dedicated, small-scale surveys, but wherever complemented with ground data it delivers maps of unprecedented thematic (and especially taxonomic) detail, highly relevant for habitat quality mapping [22].

Several authors have shown that modelling with high resolution multispectral satellite and airborne images in combination with topographic data can deliver up-to-date land-cover classifications for land resource planning, biodiversity studies, regional or global biogeochemical cycle, environmental change [23–26], and detailed habitat monitoring in remote tropical [18], desert [19] and in dangerous military or war landscapes [27]. The inclusion of topographic data for assessing vegetation diversity has already been performed in past researches [28].

An examination of above mentioned research cases shows that different modelling techniques may be used for classification of habitats and land cover types. A list of different methods for land cover classification was presented also by [29], nevertheless decision trees are one of the most frequently used methods for habitat modelling. The method has already been used for modelling of potential habitat for black poplar (*Populus nigra* L.) [30], modelling of heathland vegetation [19], different shrub and tree species [31, 32], forest cover [33], invasive or exotic plant species [34, 35], grassland use intensities [36], habitat types [14], marine biotopes [37], plant distribution in marshes [15], abundance of taxonomic group of organisms [38], and wetland habitats [16]. Although decision trees are robust to variables (correlation, noise etc.) there are few researches that also used variable evaluation before the modelling [34].

It has been proven that the classification of satellite images is an efficient tool when determining the land cover in the Classical Karst region [11], which is a part of broader Dinaric Karst region. However, according to our knowledge there is no further research on more detailed typological level of habitat types when dealing with Dinaric Karst landscapes of South East Europe, beside that only Slovenia has established national system for habitat mapping and monitoring. Due to remaining mine fields, some parts of southern Dinaric landscapes, especially in Croatia and Bosnia and Herzegovina, are extremely dangerous for field research and economic activities [39]. Some projects of habitat mapping are already in progress and researchers have already faced with this threat (*“Terrestrial Habitat Mapping of the Republic of Croatia, World bank project ID: P111205”*). Therefore the use of remote sensing and geographic information system (GIS) methods is unavoidable.

Modelling of habitats in Dinaric Karst and Mediterranean is in preliminary phase, few papers are focused on animal habitat modelling [40, 41] or land cover [12], but no comparable research on modelling mosaicked karst habitats has been published so far. Therefore, established methods reported above are not completely transferable to highly mosaicked and biodiversity rich karst landscape due to high land (karst landscape with dolines) and consequential habitat fragmentation, small land plots, land cover changes in short distance and underground hydrology.

Regarding this, the main objective of the paper is to analyse and compare habitat type modelling with decision trees, at different scales in Dinaric Karst landscape. The main research question is: »How detailed can modelling of habitat types be?«. Our aim is to explore the possibility and accuracy of mapping habitats in Dinaric Karst by using multispectral and topographic Earth observation data combined with field data. We used two established typology schemes: detailed Classification of Slovenian habitat types – HTS (si. *habitatni tipi Slovenije*) and general Land cover typology by Food and Agricultural Organisation of United Nations Land Cover Classification Scheme – FAO-LCCS. We made a translation from more detailed habitat typology to more general land cover typology and created series of typologies at three hierarchical levels. Then we modelled typology at each level.

2 Study area and data

2.1 Dinaric Karst study area and habitat characteristics

Dinaric Karst landscapes are extending over two EU member countries (Slovenia and Croatia) and non-member countries of Western Balkans that were included in past war conflicts on Balkan in 1990s. Consequently, many unknown mine fields unable safe and professional field mapping of habitats that should be in the near future included in EU Natura 2000 protected area system. This is the largest continuous karst landscape in Europe [42].

We explored the basic properties of the existent karst habitats by studying Slovene Dinaric Karst (Figure 1). The study area is spreading over a 24 km² of Slovenian region Kras. The area is characterized by its karst phenomena (dolines, caves, and underground hydrology) which contribute to the formation of rough stony terrain, high biodiversity, high land cover diversity and habitat fragmentation [11]. Karst habitats are characteristic for landscapes that are formed in carbonate rocks. The absence of sur-

Figure 1: Delimitation of the Dinaric Karst [43] and the location of the study area.

face water due to predominant groundwater discharge, the presence of specific concave karst landforms (dolines) and distribution of fertile soils in dolines caused traditional land use patterns where cultivated dolines and grasslands prevail.

The study area is sparsely populated and overgrowing process is intensive, which makes it harder to distinguish between bushy and forested habitat types in different succession stages. In only 250 years an almost treeless stony grassland landscape was converted to a forest-dominated landscape [44].

In the past, the dry stony pastures with low sedge and rock knapweed (*Carici humilis-Centaureetum rupestris*) were formed on shallow soils due to grazing.

In slightly deeper soils with less pronounced stoniness and with more humus, dry and semi-dry grasslands have been formed through mowing and been used as meadows and pastures alternatively. Plant communities of the alpine oatgrass and villous viper's grass (*Danthonio-Scorzoneretum villosae*) were formed in this area, and are considered as the most diverse grassland communities.

These grasslands are being overgrown in many places, however they are still better preserved than pastures. Xeromesophile meadows where tall oatgrass (*Arrhenatherum elatius*) and hay meadows are common on sites with deeper soils.

In valleys with deepest and most fertile soils wet meadows with tall oatgrass can be found, which are regularly fertilized and mowed three times a year. This association is mainly formed on the surfaces that have been arable land in the past. Arable lands are nowadays declining. Around the villages there are still vegetable gardens and fields. In the villages, there are still common tall-tree orchards, while vineyards are less present. The development of human infrastructure has changed a large area of these cultural landscapes. The highway has fragmented the landscape. The habitats in its immediate vicinity have been changed, and are dominated by ruderal communities. Unrestrained expansion of craft zones around settlements [45] has destroyed the natural habitats in the bottom of the cultivated dolines by infilling dolines with waste, be-

Table 1: List of source and derived derivative data layers.

Source data layer (resolution)	Layer
Habitat types defined by Kaligarič <i>et al.</i> [45] (polygon vector data)	HTS classification detailed level with 51 habitat types [45]
	HTS classification level with 20 habitat types [45]
	HTS classification/FAO comparative level with 7 land cover types
	FAO LCCS level with 2 land cover types
Digital elevation model (5 m)	Elevation (meters)
	Difference between DEM and smoothed DEM (low filter)
	Difference between DEM and smoothed DEM (focal, rectangle, 9)
	Slope (degrees)
	Aspect categories (8 categories)
	Compound topographic index [48–50]
	Heat load index [50, 51]
	Integrated moisture index [50, 52]
RapidEye images (6.5 m)	Site exposure index [50, 53]
	RapidEye April, bands 1–5
	RapidEye August, bands 1–5
	RapidEye April NDVI
	RapidEye April Red Edge NDVI
	RapidEye August NDVI
	RapidEye August Red Edge NDVI
	NDVI Difference between August and April
Aerial images (0.5 m)	Red Edge NDVI Difference between August and April
	Infrared aerial image, RGB bands
	Aerial image, RGB bands

ing this one of the major environmental issues within the study area.

2.2 Source data and derived variables

The study was based on several vector and raster data sources (Table 1). We used habitat type maps produced by field mapping during vegetation period in 2011. Colour aerial images in the scale of 1:2,000 and 1:4,500 were used as a base map for field mapping [45]. Mapping was based on the HTS classification which was adapted from “A classification of Palaearctic habitats” (catalogue of habitat types organized according to CORINE methodology, intending to increase knowledge of the diversity of Europe’s habitats [46]).

We collected raster data with high spatial resolution: 6.5 m RapidEye satellite images (provided by Blackbridge, now Planet), 0.5 m aerial images and 5 m digital elevation model (both provided by The Surveying and Mapping Authority of the Republic of Slovenia). Digital elevation model vertical accuracy is between 1 (open landscape) and 3 m (forested area). They were used for the derivation of

different explanatory variables – different vegetation and relief indexes, such as site exposure index (see Table 1 for complete list).

Aerial images and digital elevation model were already prepared for the analysis. However, RapidEye images needed additional processing (geometric and radiometric correction) before their usage. The study area was covered by two RapidEye images, taken in spring and summer of two consecutive years (2011-04-20 and 2012-08-18). The processing of satellite images has been performed with STORM—a fully automatic image processing chain developed by our group [47]. The images were first geometrically corrected by automatic extraction of ground control points (GCPs) and matching images onto reference roads, followed by sensor modelling with orthorectification. In the next step the images were radiometrically corrected – atmospheric corrections have been applied. We used AT-COR 2 (Atmospheric & Topographic Correction, version 2) to eliminate the effects of the atmosphere such as scattering and haze, and to define the extent of clouds and cloud shades. Several products have been calculated from the images once the pre-processing was completed. We computed the normalized differential vegetation index (NDVI)

and the Red Edge NDVI (normalized difference of near infrared (NIR) and red edge bands) in order to get information on vegetation health and level of greenness. Near infrared band was also determined as very useful by previous research [26].

All of the explanatory variables were available in raster format with different resolutions varying from 0.5 to 6.5 m. Data on habitat types was the only vector layer (the dependent variable). Each patch was represented by individual polygon and encompassed at least few raster cells. Therefore we assigned information on mean and standard deviation value of specific part of each raster layer to each patch (polygon of habitat type). This resulted in a large database with 61 independent variables (30 different variables with 2 different statistics – mean and standard deviation). Aspect was the only variable with one statistical value (majority). In the case of aspect, the numeric values for mean and standard deviation would be misleading (1° is similar to 360° , but has completely different value). Therefore, we divided aspect into 8 categories (N, NW ...) and calculate the majority class for each polygon. In this way we obtained the information about the prevailing aspect direction for each unit. The polygon area was also a variable. Table 3 shows a complete list of variables.

3 Determination of classification levels and decision trees methodology

3.1 Translation from habitat to land cover classification scheme

While the conversion from land cover classification to habitat classification or vice-versa can be of great utility, differences in definitions and criteria have so far limited the establishment of a unified approach [54]. In this chapter we explain the translation from detailed HTS habitat typology to more general FAO-LCCS land cover typology for the study area. Land cover data at the European and global levels (e.g. Corine Land cover and FAO land cover maps) are the closest approximation of habitat type maps for Dinaric Karst region.

HTS habitat classification was adapted to the diversity of habitats existent in Slovenia [55]. Habitats mapped in the field were divided into 51 types and then grouped into 20 types.

FAO developed a **Land Cover Classification System (LCCS)** [56]. It is based on a set of diagnostic criteria for

a parametric classification approach. The system has two stages: the dichotomous phase (where a dichotomous key identifies 8 major land cover categories) and the modular hierarchical phase (which is open ended and where a set of environmental classifiers allows the definition of more detailed land cover classes) [57]. FAO-LCCS was considered by Tomaselli *et al.* [54] as an appropriate classification scheme for long-term monitoring of the conservation status of habitats.

The translation from HTS habitat types to FAO-LCCS land cover classes generated comparative level with 7 land cover types. Both original level and comparative level are presented in Table 2A.

We limited our research to vegetated areas. Therefore, we selected suitable (vegetated) types at each level. All the categories/types used in the analysis are separately listed in Table 2B for each level (detailed habitat level, comparative level, and general land cover typology).

It is evident that the habitat typology of HTS arises from the need for evaluation and conservation, while in FAO-LCCS categorization of land cover the anthropogenic aspect is at the forefront. As an example of discrepancies we selected four habitat types on the level of HTS. *Urban parks and large gardens* (type 85), *Lines of trees, hedgerows, small anthropogenic woodlands, mosaic landscapes* (type 84), *Orchards, olive groves, tree plantations* (type 83), *Improved grasslands* (type 81) are included in the habitat division of the same class – *agricultural and cultural landscape* (types starting with 8). On the other hand, *Mesophile grasslands* (type 38) are placed in a completely different type – *scrub and grasslands* (types starting with 3). In terms of land cover the mentioned type 81 is much closer to the type 38. Hence it is very difficult to distinguish between these types using remote sensing.

According to this and in order to harmonize with the FAO-LCCS we placed these two types, on the comparative level (HTS/FAO-LCCS), into one category *Artificial/harvested grasslands* (type 1.2.4), the remaining three types (83, 84 and 85) were placed in *Artificial trees and dry grasslands* (type 1.2.5). Similarly, we included the *Humid grassland and tall herb communities* (type 37) and the *Water-fringe vegetation* (type 53), which in HTS habitat typology are classified in two different types, in the same type – *Wetland* (type 1.1.2). For all the other categories the grouping in the transition from one to another categorization was less demanding.

The problematic category division resulting in the great variability of land use is one of the main characteristics when classifying satellite imagery of the Karst region [11]. The problem, which is a result of the fragmented land division, different land use types and rapid over-

Table 2A: The translation from HTS habitat types to FAO-LCCS at the case study area.

DETAILED HABITAT LEVEL		COMPARATIVE LEVEL (FROM HABITAT TO LAND COVER TYPES)		GENERAL LAND COVER TYPOLOGY LEVEL		
HTS	Habitat name	Habitat code	translation from HTS to FAO-LCCS Land cover typology	Comparative code	Land cover type	
	Arable land	82	Agriculture	1.2.1	A11	Cultivated and managed terrestrial areas
	Humid grassland and tall herb communities	37	Wetland	1.1.2	A12	Natural and semi-natural vegetation
	Water-fringe vegetation	53				
	Steppes and dry calcareous grasslands	34	Natural dry grassland	1.1.3		
	Temperate heath and scrub	31	Heathland	1.1.31		
	Sclerophyllous scrub	32				
	Mesophile grasslands	38	Artificial/harvested	1.2.4	A11	Cultivated and managed
	Improved grasslands	81	grasslands			terrestrial areas
	Urban parks and large gardens	85	Artificial trees	1.2.5		
	Lines of trees, hedgerows, small anthropogenic woodlands, mosaic landscapes	84	and grasslands			
	Orchards, olive groves, tree plantations	83				
	Broad-leaved deciduous forests	41	Deciduous forest	1.1.6	A12	Natural and semi-natural vegetation
	Inland cliffs and exposed rocks	62	Natural bare land	2.1.7	B16	Bare areas
	Structures from stacked stones.	84.S6				
	Unpaved roads and paths	86.S72	Nonvegetated	2.2.7	B15	Artificial surfaces and associated areas
	Fallow or recently abandoned arable land	87	artificial land			
	Asphalt roads	86.S712	Urban	2.2.8		
	Built-up areas (cities, villages, industrial areas)	86				
	Villages, suburbs and individual buildings	86.2				
	Surface standing waters	22	Water bodies	2.9	B27/B28	Waterbodies

Table 2B: List of habitat types/land cover types with number of polygons and area covered in the study area.

Level	Type	Area (m ²)	Share (%)	Number of polygons	Share (%)
Level B (HTS)	Lines of trees, hedgerows, small anthropogenic woodlands, mosaic landscapes	745364.6	8.7	821	17.8
	Improved grasslands	94913.4	1.1	63	1.4
	Broad-leaved deciduous forests	8901.0	0.1	9	0.2
	Urban parks and large gardens	66272.4	0.8	144	3.1
	Mesophile grasslands	1116685.1	13.1	619	13.5
	Water-fringe vegetation	6047.3	0.1	1	<0.1
	Humid grassland and tall herb communities	3058.7	<0.1	5	0.1
	Steppes and dry calcareous grasslands	1251858.4	14.6	772	16.8
	Arable land	78439.0	0.9	91	2.0
	Temperate heath and scrub	4985416.6	58.3	1787	38.8
	Orchards, olive groves, tree plantations	191185.5	2.2	289	6.3
	Sclerophyllous scrub	360.1	<0.1	1	<0.1
	Total	8548502.1	100.0	4602	100.0
Level C (HTS/FAO-LCCS)	Agriculture	78439.0	0.9	91	2.0
	Artificial trees and grassland	1002822.4	11.7	1254	27.2
	Artificial/harvested grassland	1211598.5	14.2	682	14.8
	Deciduous forest	8901.0	0.1	9	0.2
	Heathland	4985776.8	58.3	1788	38.9
	Natural dry grassland	1251858.4	14.6	772	16.8
	Wetland (bogs/mires)	9106.0	0.1	6	0.1
	Total	8548502.1	100.0	4602	100.0
Level D (FAO-LCCS)	Cultivated and managed terrestrial areas	2292860.0	26.8	2027	44.0
	Natural and semi-natural vegetation	6255642.1	73.2	2575	56.0
	Total	8548502.1	100.0	4602	100.0

growing of the Karst region, is especially evident, because meadows can be detected through thin bushes that create a similar spectral signature. Kokalj and Oštir [11] pointed out the problem in classification and distinguishing between field, bushes or meadows.

3.2 Supervised modelling with decision tree method

A decision tree is a machine learning method. The main purpose of modelling with decision tree is to clearly describe the relationship between dependent or objective variable (variable we predict) and one or more independent explanatory variables. Explanatory variables are selected at each modelling level according to their information value or impurity measure (e.g. Gini coefficient) and then used to define subsets. A decision tree consists of internal (parent) nodes – attributes, edges – corresponding to subsets of attribute values, and terminal nodes (child

nodes, leaves) – class labels. We can classify certain sample simply by moving from the top of the tree to final terminal node according to rules. With this method we can explain existing data and predict future values [58, 59]. Decision trees do not have specific requirements regarding probability distribution of variables. They can handle nominal (e.g. aspect class) and numeric (e.g. slope) variables, and are not influenced by redundant variables or noise. They do not demand high computation resources and results are easy to interpret. High accuracy can be achieved by using different combination of trees [60].

In this research we used decision trees with algorithm CRT (Classification and Regression Trees; [61]) with Gini index as an impurity measure [59]. There are many possible settings (e.g. minimum number of units in nodes, number of levels, type of impurity measure, inclusion of pruning), thus we were forced into modelling with trial and error. 90% of polygons were used for modelling, and 10% were used for validation.

3.3 The selection of explanatory variables

Identifying the most useful variables for modelling land-cover classes is a critical step [62]. If irrelevant data is removed from analysis, one can reduce costs and improve the implementation and understanding of models [63, 64]. However the selection of variables usually receives less attention than the choice of methods [65]. Decision trees automatically select the most informative variable at given node and thus are not affected by unnecessary less informative variables. Despite that fact we would like to maximize modelling success and to remove variables that are not prominent. Thus all data layers were analysed according to their usefulness for modelling habitats/land cover types. We assessed them based on the:

- relationship between each variable and habitat types,
- correlation between variables.

To define a numeric objective relationship between each explanatory variable and habitat types classification we calculated information gain and gain ratio [58, 59, 66]. Information gain ($InfoGain(A)$) is defined as the amount of information, obtained from the explanatory variable, for determining the type of objective variable [59]. Gain-ratio is defined as information gain, normalized with the variable entropy [58, 59]. Both methods require nominal variables, thus our numeric variables were discretized (for details of discretization used in Weka software see [58, 67] and [68]).

Pearson's and Spearman's coefficients were used to calculate correlation between the most prominent explanatory variables.

3.4 Selection of polygons

We examined the size of the polygons. Mapped habitat types (field mapping data) sized between 100 m² and 100,000 m² were included in the modelling. Polygons larger than 100,000 m² are mainly forested habitats, which have not been detailed mapped and therefore are highly heterogeneous. Polygons less than 100 m² are mostly very small areas of habitat types, which encompass only few cells and thus cannot represent individual habitat type.

3.5 Complete methodology overview

After defining proper classification levels, selecting suitable modelling technique and evaluating appropriate explanatory variables, we modelled habitat types (objective

variable) with several explanatory variables. The whole process is schematically presented in Figure 2.

4 Results and discussion

4.1 Variable selection

We calculated information gain and gain ratio for each explanatory variable at each of three levels. Then we calculated indexed values of the information gain and gain ratio for each variable according to the average value of all variables at given level. Finally, we derived general average value for each explanatory variable (Table 3).

On the basis of the highest values of information gain and gain ratio we can conclude that mean RapidEye band 5 (April), mean NDVI (April; bands 5, 4), and mean difference between August and April NDVI (bands 5, 4) are the most useful explanatory variables for classification of habitats and land cover. Relief indexes (*e.g.* heat load index and integrated moisture index) are less useful in our case. Among these the most prominent variable is mean altitude, and surprisingly aspect is one of the least important for habitat differentiation in karst landscape where a gently sloping hillside is prevailing.

In regard to the results of correlation analysis, both tests, Pearson and Spearman coefficient, gave similar results and we were able to drop large number of variables that were marked as redundant. After the complete evaluation of variables, we selected a group of variables with low correlation.

The assessment of variables resulted in the selection of 9 explanatory variables:

- mean altitude,
- mean slope,
- mean Site exposure index,
- mean RapidEye band 2 (green),
- mean RapidEye band 5 (near infra-red),
- mean Red Edge NDVI in August,
- mean Red Edge NDVI difference (August - April),
- mean aerial infrared color image (band 1), and
- mean visual color aerial image band 1 (red).

There were 4 pairs with moderate correlation (>0.5 or <-0.5), however the rest of 36 pairs of variables had low correlation (Table 4). Nevertheless, mean RapidEye band 5 (April), mean NDVI (August; bands 5, 4), and mean difference between August and April NDVI (bands 5, 4) were kept in the database due to their high informative value. As some authors [34] use collinear variables when modelling

Table 3: Ranking of explanatory variables according to their information values.

Variable	Average value	Variable	Average value
mean RapidEye band 5 (April)	3.294	mean NDVI (August; bands 5, 3)	0.804
mean NDVI (April; bands 5, 4)	2.727	mean RapidEye band 2 (April)	0.772
mean difference between August and April NDVI (bands 5, 4)	2.503	mean NDVI (August; bands 5, 4)	0.673
mean difference between August and April NDVI (bands 5, 3)	2.471	std. deviation of RapidEye band 3 (April)	0.665
mean NDVI (April; bands 5, 3)	2.211	std. deviation of difference between original and smooth relief 2	0.651
mean infrared digital orthophoto band 1	1.861	std. deviation of integrated moisture index	0.645
mean infrared digital orthophoto band 3	1.768	mean RapidEye band 1 (April)	0.617
std. deviation of NDVI (April; bands 5, 4)	1.613	mean heat load index	0.617
mean digital orthophoto band 1	1.514	std. deviation of compound topographic index	0.611
mean altitude	1.509	mean integrated moisture index	0.607
std. deviation of digital orthophoto band 2	1.503	std. deviation of difference between August and April NDVI (bands 5, 3)	0.582
std. deviation of infrared digital orthophoto band 1	1.498	std. deviation of RapidEye band 1 (April)	0.575
mean digital orthophoto band 2	1.487	mean difference between original and smooth relief 2	0.544
mean infrared digital orthophoto band 2	1.412	mean site exposure index	0.515
std. deviation of NDVI (April; bands 5, 3)	1.410	std. deviation of slope	0.509
mean digital orthophoto band 3	1.379	std. deviation of difference between August and April NDVI (bands 5, 4)	0.484
mean RapidEye band 4 (August)	1.365	std. deviation of heat load index	0.482
std. deviation of infrared digital orthophoto band 3	1.338	std. deviation of RapidEye band 1 (August)	0.479
std. deviation of digital orthophoto band 3	1.266	std. deviation of RapidEye band 2 (April)	0.452
mean RapidEye band 2 (August)	1.258	mean RapidEye band 5 (August)	0.446
mean RapidEye band 1 (August)	1.221	std. deviation of RapidEye band 4 (April)	0.442
mean RapidEye band 3 (August)	1.162	std. deviation of site exposure index	0.436
std. deviation of infrared digital orthophoto band 2	1.148	std. deviation of RapidEye band 3 (August)	0.390
mean slope	1.102	std. deviation of RapidEye band 2 (August)	0.381
std. deviation of altitude	1.077	std. deviation of RapidEye band 4 (August)	0.373
mean RapidEye band 4 (April)	1.015	mean difference between original and smooth relief 1	0.363
std. deviation of digital orthophoto band 1	0.987	std. deviation of difference between original and smooth relief 1	0.289
std. deviation of RapidEye band 5 (April)	0.951	aspect (majority class)	0.285
mean compound topographic index	0.922	std. deviation of RapidEye band 5 (August)	0.224
mean RapidEye band 3 (April)	0.907	std. deviation of NDVI (August; bands 5, 4)	0.160
area	0.901	std. deviation of NDVI (August; bands 5, 3)	0.148

Figure 2: Complete methodology of the habitat modelling. Different raster data layers were derived on the basis of DEM, Rapideye images, orthophotos, and infrared orthophotos. After the variables evaluation 9 data variables were selected, which were the explanatory (independent) variables. Original habitat types classification (with 51 types; level A) was hierarchically reclassified into more generalized types (levels B, C and D). Models of habitat types were produced for levels B, C, and D on the basis of training sets. Validation set was used to evaluate modelling results. At the end comparison between the levels was performed.

Table 5: Producer's and user's accuracies for general land cover level (all polygons).

Land cover type	Producer's accuracy	User's accuracy
Cultivated and managed terrestrial areas	80.9%	83.5%
Natural and semi-natural vegetation	87.4%	85.3%

Table 6: Producer's accuracy for general land cover level (training and test polygons separately).

Land cover type	Training sample	Test sample
Cultivated and managed terrestrial areas	81.8%	73.3%
Natural and semi-natural vegetation	88.1%	81.2%

with decision trees and the method can handle correlated variables we assume that the use of low and moderate correlated explanatory variables in our study is not problematic.

4.2 Accuracy assessment

4.2.1 Modelling land cover at general level (FAO-LCCS; level D)

At the general level of classification of land cover (FAO-LCCS level) we modelled two types: *Natural and semi-natural vegetation (A12)* and *Cultivated and managed terrestrial areas (A11)*.

In the differentiation between these two types producer's and user's accuracies were between 75 and almost 90% (Tables 5, 6, and 7). The agricultural landscape in the Dinaric Karst is being overgrown by natural vegetation. Therefore, it is difficult to distinguish between natural or anthropogenic grassland, which is overgrowing abandoned fields. This partly explains that the success of predictive models for *Cultivated and managed terrestrial areas* are always lower than the *Natural and semi-natural vegetation*.

Table 7: User's accuracy for general land cover level (training and test polygons separately).

Land cover type	Training sample	Test sample
Cultivated and managed terrestrial areas	84.4%	75.5%
Natural and semi-natural vegetation	86.0%	79.3%

4.2.2 Modelling at comparative land cover level (level C)

At the comparative level we modelled the land cover in eight categories (Tables 8, 9, and 10). With extremely low achievement the categories *Deciduous Forest* and *Wetland* stand out. This is due to the reason that the two categories together are represented by only 15 polygons from a total of approximately 4,600 polygons, which are included in the modelling. From the remaining categories, the categories *Agriculture* and *Natural dry grassland* present the lowest success rate. The category *Agriculture* is often modelled as *Artificial trees and grassland*, *Artificial/harvested grassland*, *Heathland*, and *Natural dry grassland*. *Natural dry grassland* is frequently modelled as *Heathland*, *Artificial trees and grassland*, and *Artificial/harvested grassland*. The distinction between *Artificial trees and grassland*, and *Artificial/harvested grassland* is hard, especially, if the snapshot is made during the period when the heights of the grass are approximately the same. Moreover, the outstanding value of individual trees of the type *Artificial trees and grassland* loose in average a much larger representation from the surrounding pasture (grassland).

4.2.3 Modelling at HTS habitats level (level B)

At this level we modelled 12 habitat types.

The classification reached its maximum performance with the type *Temperate heath and scrub*, where 86.7% from training polygons and 70.2% from the testing polygons were placed in the correct habitat type (Tables 11, 12 and 13). This habitat type includes mixed grasslands and shrubs, therefore it was expected that the maximum of incorrectly classified polygons were indexed in the similar types, which represent grasslands or trees or the mixture of both. A higher modelling performance was achieved in the classification of *Mesophile grasslands* (67.9% for training sample and 57.0% for the test sample), which refer to anthropogenic grasslands. Higher modelling performance was also obtained for the classification of *Lines of trees*,

Table 8: Producer's and user's accuracies for comparative land cover level (all polygons).

Basic land cover level (FAO-LCCS, level D)	Comparative land cover (level C)	Producer's accuracy	User's accuracy
Cultivated and managed terrestrial areas	Agriculture	29.7%	60.0%
	Artificial trees and grassland	71.3%	73.6%
	Artificial/harvested grassland	67.7%	67.2%
Natural and semi-natural vegetation	Deciduous forest	0.0%	0.0%
	Heathland	82.9%	70.8%
	Natural dry grassland	45.9%	62.9%
	Wetland (bogs/mires)	0.0%	0.0%

Table 9: Producer's accuracy for Comparative land cover level (training and test polygons separately).

Basic land cover level (FAO-LCCS, level D)	Comparative land cover (level C)	Training sample	Test sample
Cultivated and managed terrestrial areas	Agriculture	31.3%	12.5%
	Artificial trees and grassland	72.9%	56.7%
	Artificial/harvested grassland	68.5%	60.6%
Natural and semi-natural vegetation	Deciduous forest	0.0%	0.0%
	Heathland	83.6%	75.9%
	Natural dry grassland	47.7%	27.5%
	Wetland (bogs/mires)	0.0%	0.0%

Table 10: User's accuracy for Comparative land cover level (training and test polygons separately).

Basic land cover level (FAO-LCCS, level D)	Comparative land cover (level C)	Training sample	Test sample
Cultivated and managed terrestrial areas	Agriculture	60.5%	50.0%
	Artificial trees and grassland	75.6%	56.7%
	Artificial/harvested grassland	68.6%	55.6%
Natural and semi-natural vegetation	Deciduous forest	0.0%	0.0%
	Heathland	71.3%	66.5%
	Natural dry grassland	64.7%	42.2%
	Wetland (bogs/mires)	0.0%	0.0%

hedgerows, small anthropogenic woodlands, mosaic landscapes (68.1 % for the training sample and 70.5% for the testing sample). The modelling results show a mixture between the above mentioned land cover types.

From the observation of tables 11, 12, and 13 it is evident that some habitats, represented by a very small number of polygons (from 1 to 9 polygons), cannot be adequately modelled. These habitat types, such as wet habitat types, are rare in karst regions.

As we did not determine the importance of the land cover categories, those categories represented by a small number of patches are automatically deprived. Since there are several categories highly represented and these were at

great extent correctly classified, the overall assessment of the classification has higher accuracy rate.

4.3 Comparison between levels B, C and D

Results from the comparison of the classification accuracies at different levels revealed that model accuracy decreases with the number of habitat type categories (Table 14). This is not surprising, and confirms previous findings [13]. The result was expected as the more detailed the classification is the more similar the categories are.

The comparison of the classification success at different scales for each polygon (Figure 3–5) shows that there

Figure 3: Classification success of modelling land cover at general level – FAO-LCCS. Only patches included in the analysis are presented.

Table 11: Producer’s and user’s accuracies for HTS habitat level (all polygons).

	Producer’s accuracy	User’s accuracy
Lines of trees, hedgerows, small anthropogenic wood-lands, mosaic landscapes	68.1%	70.5%
Improved grasslands	23.8%	65.2%
Broad-leaved deciduous forests	0.0%	0.0%
Urban parks and large gardens	33.3%	55.2%
Mesophile grasslands	67.9%	57.0%
Water-fringe vegetation	0.0%	0.0%
Humid grassland and tall herb communities	0.0%	0.0%
Steppes and dry calcareous grasslands	45.7%	67.4%
Arable land	24.2%	43.1%
Temperate heath and scrub	86.7%	70.2%
Orchards, olive groves, tree plantations	32.2%	52.0%
Sclerophyllous scrub	0.0%	0.0%

Figure 4: Classification success of modelling at comparative land cover level. Only patches included in the analysis are presented.

Table 12: Producer’s accuracy for HTS habitat level (training and test polygons separately).

	Training sample	Test sample
Lines of trees, hedgerows, small anthropogenic wood-lands, mosaic landscapes	68.7%	62.7%
Improved grasslands	23.2%	28.6%
Broad-leaved deciduous forests	0.0%	0.0%
Urban parks and large gardens	35.8%	0.0%
Mesophile grasslands	69.0%	52.4%
Water-fringe vegetation	0.0%	0.0%
Humid grassland and tall herb communities	0.0%	0.0%
Steppes and dry calcareous grasslands	46.8%	35.2%
Arable land	24.1%	25.0%
Temperate heath and scrub	87.7%	79.8%
Orchards, olive groves, tree plantations	33.5%	22.9%
Sclerophyllous scrub	0.0%	0.0%

Figure 5: Classification success of modelling at HTS habitats level. Only patches included in the analysis are presented.

Table 13: User’s accuracy for HTS habitat level (training and test polygons separately).

	Training sample	Test sample
Lines of trees, hedgerows, small anthropogenic wood-lands, mosaic landscapes	70.8%	67.5%
Improved grasslands	68.4%	50.0%
Broad-leaved deciduous forests	0.0%	0.0%
Urban parks and large gardens	57.1%	0.0%
Mesophile grasslands	59.2%	33.8%
Water-fringe vegetation	0.0%	0.0%
Humid grassland and tall herb communities	0.0%	0.0%
Steppes and dry calcareous grasslands	69.1%	51.0%
Arable land	45.7%	20.0%
Temperate heath and scrub	70.3%	69.7%
Orchards, olive groves, tree plantations	53.5%	40.0%
Sclerophyllous scrub	0.0%	0.0%

Table 14: Classification accuracies at all levels.

Level	Number of habitat type categories	Overall accuracy (%)	Training accuracy (%)	Test accuracy (%)
HTS (level B)	20	63.4	67.2	59.9
Comparative land cover (level C)	7	69.2	71.1	59.3
FAO-LCCS (level D)	2	82.6	85.3	77.7

are 51.6% of polygons correctly classified at each level. Regarding to the overall classification accuracies at each level (between 63.4% and 82.6%) this percentage is relatively high. There is a relatively high number of polygons which are falsely classified at least at two levels (48.4%). 7.1% of polygons are falsely classified at all three levels.

The methods and data used in our research cannot assure highly accurate mapping of habitats in Dinaric Karst. It is possible to distinguish relatively well habitat types at very general level, where similar accuracies were achieved also in other environments with similar number (2 or 3) of categories [e.g. 19, 31–33], but the success of correct classification at more detailed level does not give confident results.

The overgrowing of landscape is one of the major problems of the study area, resulting in the loss of the mosaic of habitats. Overgrowing also diminishes differences between certain types. Therefore, it is hard to define clear delineations even in the field.

Another reason for low accuracies is also the spatial resolution of data. Hofer [28] reported that topographic (relief) variability increases plant variability and Rapinel *et al.* [16] achieved better modelling results with lidar data. Therefore, we believe that higher resolution data should be used and tested also in karst areas.

Since this was one of the rare attempts of using remote sensing data in Dinaric Karst region, the research results can be presented as a basis for testing with other datasets (hyperspectral airborne images, lidar data) or in other countries within the region.

5 Conclusions

We can conclude that modelling of habitat and land cover types at several hierarchical levels for the study area on the basis of digital elevation model (5 m resolution), satellite images RapidEye (6.5 m resolution), and aerial images (0.5 resolution) did not achieve high accuracies. The highest accuracy (80%) was achieved at the general level with only two habitat types. At lower classification levels the

accuracies were less than 70%. As expected, it is harder to model classifications with numerous types. Namely, the landscape is very diverse and larger number of habitat types is more difficult to model.

Producer's and user's accuracies for certain types rarely exceed 80%, as a result of similarity of habitat types. With the overgrowing of cultivated areas the differences between types are becoming even smaller. Therefore, the need to differentiate new types of habitats and land cover inside vegetated areas will arise.

In our case relief data with resolution of 5 m does not contribute to the modelling process and can be eliminated in further models of karst areas, especially when dealing with karst plateaus. Further improvements regarding relief data, could be achieved by using very high resolution laser scanning data, especially in highly karstified landscapes. Additionally, laser scanning would also provide data on vegetation height. With the increasing coverage of laser scanning data from national-scale surveys we expect lidar-based habitat quality assessment to allow a major step forward in regional-scale conservation.

Acknowledgement: The authors acknowledge the financial support from the Slovenian Research Agency (research core funding No. P6-0101, L1-7542, L6-6852, J2-6777, and P2-0227).

References

- [1] Bailey R.G., *Ecosystem geography*. Springer, New York, 1996
- [2] Bunce R.G.H, Barr C.J., Clarke R.T., Howard D.C., Lane A.M.J., Land classification for strategic ecological survey. *J. Environ Manage.*, 1996, 47, 1, 37–60. doi: 10.1006/jema.1996.0034
- [3] Varga K., Szabó S., Szabó G., Dévai G., Tóthmérész B., Improved land cover mapping using aerial photographs and satellite images. *Open Geosci.*, 2014, 7, 1. doi:10.1515/geo-2015-0002
- [4] Bastian O., Landscape classification in Saxony (Germany) – a tool for holistic regional planning. *Landsc. Urban Plan.*, 2000, 50, 1-3, 145–155, doi: 10.1016/S0169-2046(00)2900086-4
- [5] Loveland T. R., Merchant, J. M., 2004, Ecoregions and Ecoregionalization: Geographical and Ecological Perspectives. *Environ Manage.*, 2004, 34, S1-S13, doi: 10.1007/s00267-003-5181-x

- [6] Ribeiro D., Somodi I., Čarni A., Transferability of a predictive *Robinia pseudacacia* distribution model in northeast Slovenia. *Acta geogr. Slov.*, 2016, 56, 1, 25–43, doi: 10.3986/AGS.772
- [7] Marković V.N., Vasiljević D.A., Jovanović T., Lukić T., Vujičić M.D., Kovačević M., *et al.* 2017, The effect of natural and human-induced habitat conditions on number of roe deer: case study of Vojvodina, Serbia. *Acta geogr. Slov.*, 57, 2, 57–69, doi: 10.3986/AGS.903
- [8] Torres M., Qiu G., Automatic habitat classification using image analysis and random forest. *Ecol. Inform.*, 2014, 23, 126–136
- [9] Múcher C.A., Bunce R.G.H., Jongman R.H.G., Klijn J.A., Koomen A.J.M., Metzger M.J., *et al.*, Identification and Characterisation of Environments and Landscapes in Europe. Alterra rapport 832. Alterra, Wageningen, 2003
- [10] Ichter J., Evans D., Richard D., Terrestrial Habitat Mapping in Europe: An Overview. European Environmental Agency, Luxembourg, 2014
- [11] Kokalj Ž., Oštir K., Land cover mapping using Landsat satellite image classification in the Classical Karst – Kras region. *Acta Carsologica*, 2007, 36, 3, 433–440
- [12] Rodríguez-Galiano V.F., Chica-Olmo M., Abarca-Hernandez F., Atkinson P.M., Jeganathan C., Random forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sens. Environ.*, 2012, 121, 93–107
- [13] Haest B., Thoonen G., Vanden Borre J., Spanhove T., Delalieux S., Bertels L., Kooistra, L., Múcher C.A., Scheunders P., An object-based approach to quantity and quality assessment of heathland habitats in the framework of Natura 2000 using hyperspectral airborne AHS images. *ISPRS Archives*, 2010, 38, 4/C7
- [14] Erasmi S., Fuchs H., Westphal C., Kleinn C., Mapping habitat type with Terrasar-X, Radarsat-2 and Rapideye data. *Proceedings of the 5. TerraSAR-X Science Team Meeting, Oberpfaffenhofen*, 2013 European Commission (EC): Environment, 2015
- [15] Hladik C., Schalles J., Alber M., Salt marsh elevation and habitat mapping using hyperspectral and LIDAR data. *Remote Sens. Environ.*, 2013, 139, 318–330, doi: 10.1016/j.rse.2013.08.003
- [16] Rapinel S., Hubert-Moy L., Clément B., Combined use of LiDAR data and multispectral earth observation imagery for wetland habitat mapping. *Int. J. Appl. Earth Obs. Geoinf.*, 2015, 37, 56–64, doi: 10.1016/j.jag.2014.09.002
- [17] Wang K., Franklin S.E., Guo X.L., Cattet M., Remote sensing of ecology, biodiversity and conservation: A review from the perspective of remote sensing specialists. *Sensors*, 2010, 10, 9647–9667
- [18] Ackers S.H., Davis R.J., Olsen K.A., Dugger K.M., The evolution of mapping habitat for northern spotted owls (*Strix occidentalis caurina*): A comparison of photo-interpreted, Landsat-based, and lidar-based habitat maps. *Remote Sens. Environ.*, 2015, 35, 2, 277–283
- [19] Delalieux S., Somers B., Haest B., Spanhove T., Vanden Borre J., Múcher C.A., Heathland conservation status mapping through integration of hyperspectral mixture analysis and decision tree classifiers. *Remote Sens. Environ.*, 2012, 126, 222–231, doi: 10.1016/j.rse.2012.08.029
- [20] Zhai L., Sun J., Sang H., Yang G., Jia Y., Large area land cover classification with Landsat ETM+ images based on decision tree. *ISPRS J. Photogramm. Remote Sens.*, 2012, 421–426
- [21] Somodi, I., Čarni, A., Ribeiro, D., Podobnikar, T., Recognition of the invasive species *Robinia pseudacacia* from combined remote sensing and GIS sources, *Biological Conservation*, 2012, 150, 1, 59–67.
- [22] Zlinszky A., Schroiff A., Kania A., Deák B., Mücke W., Vári Á., *et al.*, Categorizing Grassland Vegetation with Full-Waveform Airborne Laser Scanning: A Feasibility Study for Detecting Natura 2000 Habitat Types. *Remote Sens.*, 2014, 6, 8056–8087
- [23] Mokarram M., Sathyamoorthy D., Relationship between landform classification and vegetation (case study: southwest of Fars province, Iran). *Open Geosci.*, 2016, 8, 1, 302–309. doi:10.1515/geo-2016-0027
- [24] Friedl M.A., Sulla-Menashe D., Tan B., Schneider A., Ramankutty N., Sibley A., *et al.*, MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.*, 2006, 100, 223–236
- [25] Chehbouni A., Escadafal R., Duchemin B., Boulet G., Simonneaux V., Dedieu G., *et al.*, An integrated modelling and remote sensing approach for hydrological study in arid and semi-arid regions: the SUDMED Programme. *Int. J. Remote Sens.*, 2008, 29, 5161–5181.
- [26] Schuster C., Förster M., Kleinschmit B., Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. *Int. J. Remote Sens.*, 2012, 33, 17, 5583–5599
- [27] Neumann C., Weiss G., Schmidlein S., Itzerott S., Lausch A., Doktor D., Brell M., Gradient-Based Assessment of Habitat Quality for Spectral Ecosystem Monitoring. *Remote Sens.*, 2015, 7, 2871–2898
- [28] Hofer G., Bunce R.G.H., Edwards P.J., Szerencits E., Wagner H.H., Herzog F., Use of topographic variability for assessing plant diversity in agricultural landscapes. *Agric. Ecosyst. Environ.*, 2011, 142, 144–148
- [29] Li C., Wang J., Wang L., Hu L., Gong P., Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery. *Remote Sens.*, 2014, 6, 964–983, doi: 10.3390/rs6020964
- [30] Debeljak M., Ficko A., Brus R., The use of habitat and dispersal models in protecting European black poplar (*Populus nigra* L.) from genetic introgression in Slovenia. *Biol. Conserv.*, 2015, 184, 310–319, doi: 10.1016/j.biocon.2015.02.004
- [31] Coops N.C., Waring R.H., Beier C., Roy-Jauvin R., Wang T., Modeling the occurrence of 15 coniferous tree species throughout the Pacific Northwest of North America using a hybrid approach of a generic process-based growth model and decision tree analysis. *Appl. Veg. Sci.*, 2011, 14, 402–414, doi: 10.1111/j.1654-109X.2011.01125.x
- [32] Hellesen T., Matikainen L., An object-based approach for mapping shrub and tree cover on grassland habitats by use of LiDAR and CIR orthoimages. *Remote Sens.*, 2013, 5, 558–583, doi: 10.3390/rs5020558
- [33] Hansen M.C., Roy D.P., Lindquist E., Adusei B., Justice C.O., Altstatt A., A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin. *Remote Sens. Environ.*, 2008, 112, 5, 2495–2513, doi: 10.1016/j.rse.2007.11.012
- [34] Sadeghi R., Zarkami R., Sabetriftar K., Van Damme P., Application of genetic algorithm and greedy stepwise to select input variables in classification tree models for the prediction of habitat requirements of *Azolla filiculoides* (Lam.)

- in Anzali wetland, Iran. *Ecol. Model.*, 2013, 251, 44-53, doi: 10.1016/j.ecolmodel.2012.12.010
- [35] Ghulam A., Porton I., Freeman K., Detecting subcanopy invasive plant species in tropical rainforest by integrating optical and microwave (InSAR/PolInSAR) remote sensing data, and a decision tree algorithm. *ISPRS J. Photogramm. Remote Sens.*, 2014, 88, 174-192, doi: 10.1016/j.isprsjprs.2013.12.007
- [36] Asam S., Klein D., Dech S., Estimation of grassland use intensities based on high spatial resolution LAI time series. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 2015, XL-7-W3, 285-291
- [37] Gonzales-Mirelis G., Lindegarth M., Predicting the distribution of out-of-reach biotopes with decision trees in a Swedish marine protected area. *Ecol. Appl.*, 2012, 22, 8, 2248-2264, doi: 10.1890/11-1608.1
- [38] Kocev D., Džeroski S., Habitat modeling with single- and multi-target trees and ensembles. *Ecol. Inform.*, 2013, 18, 79-92, doi: 10.1016/j.ecoinf.2013.06.003
- [39] Šiljković Ž., Čuka A., Pejdo A., Rural area transformation: From cropland to mine fields – Zemunik Donji municipality (Croatia) case study. *Društvena istraživanja*, 2011, 20, 4, 1163-1181, doi: 10.5559/di.20.4.12
- [40] Kobler A., Adamič M., Identifying brown bear habitat by a combined GIS and machine learning method. *Ecol. Model.*, 2000, 135, 291-300, doi: 10.1016/S0304-3800(00)00384-7
- [41] Narce M., Meloni R., Beroud T., Pléney A., Ricci J.C., Landscape ecology and wild rabbit (*Oryctolagus cuniculus*) habitat modeling in the Mediterranean region. *Anim. Biodivers. Conserv.*, 2012, 35, 2, 277-283
- [42] Mihevc A., Prelovšek M., Geographical position and general overview. In: Mihevc A., Prelovšek M., Zupan Hajna N. (Eds.), *Introduction to the Dinaric Karst*, Inštitut za raziskovanje krasa ZRC SAZU, Postojna, 2010, 6-8
- [43] Gams I., *Kras v Sloveniji v prostoru in času*. Založba ZRC, Ljubljana, 2004
- [44] Kaligarič M., Ivajnsi D., Vanishing landscape of the “classic” Karst: changed landscape identity and projections for the future. *Landscape Urban Plan.*, 2014, 132, 148-158, doi: 10.1016/j.landurbplan.2014.09.004
- [45] Kaligarič M., Škornik S., Šajna N., Otopal J., Pipenbaher N., Ivajnsi D., *Kartiranje negozdnih habitatnih tipov Slovenije / Območje KRAS – Kozina in KRAS – Vremščica*. Final Report (in Slovene), Faculty of natural sciences and mathematics, University Maribor, 2011
- [46] Devillers P., Devillers-Terschuren J., A classification of Palaearctic habitats. *Nature and environment*, 1996, no. 78, Strasbourg, Council of Europe
- [47] Pehani P., Čotar K., Marsetič A., Zaletelj J., Oštir K., Automatic Geometric Processing for Very High Resolution Optical Satellite Data Based on Vector Roads and Orthophotos. *Remote Sens.-Basel*, 2016, 8, 4, 1-26. doi: 10.3390/rs8040343
- [48] Moore I.D., Gessler P.E., Nielsen G.A., Petersen G.A., Terrain attributes: estimation methods and scale effects. In: Jakeman A.J., Beck M.B., McAleer M. (Eds.), *Modeling Change in Environmental Systems*. Wiley, London, 1993, 189-214
- [49] Gessler P.E., Moore I.D., McKenzie N.J., Ryan P.J., Soil-landscape modeling and spatial prediction of soil attributes. *Int. J. Geogr. Inf. Syst.*, 1995, 9, 4, 421-432
- [50] Evans J.S., Oakleaf J., Cushman S.A., Theobald D., An ArcGIS Toolbox for Surface Gradient and Geomorphometric Modeling, version 2.0-0, 2014, <http://evansmurphy.wix.com/evansspatial>
- [51] McCune B., Keon D., Equations for potential annual direct incident radiation and heat load. *J. Veg. Sci.*, 2002, 13, 603-606, doi: 10.1111/j.1654-1103.2002.tb02087.x
- [52] Iverson L.R., Dale M.E., Scott C.T., Prasad A., A GIS-derived integrated moisture index to predict forest composition and productivity in Ohio forests. *Landsc. Ecol.*, 1997, 12, 331-348
- [53] Balice R.G., Miller J.D., Oswald B.P., Edminister C., Yool S.R., *Forest surveys and wildfire assessment in the Los Alamos*; 1998–1999. Los Alamos National Laboratory, Los Alamos, New Mexico, 2000
- [54] Tomaselli V., Dimopoulos P., Marangi C., Kallimanis A. S., Adamo M., Tarantino C., et al., Translating land cover/land use classifications to habitat taxonomies for landscape monitoring: a Mediterranean assessment. *Landsc. Ecol.*, 2013, 28, 5, 905-930
- [55] Jogan N., Kaligarič M., Leskovar I., Seliškar A., Dobravec J., *Habitatni tipi Slovenije HTS 2004 – tipologija*. Ministrstvo za okolje, prostor in energijo – Agencija RS za okolje, Ljubljana, 2004
- [56] Di Gregorio A., Jansen L.J.M., *Land Cover Classification System (LCCS): classification concepts and user manual*. Food and Agriculture Organization of the United Nations, 2005, Rome
- [57] Kobler A., Adamič M., *Brown Bears in Slovenia: Identifying Locations for Construction of Wildlife Bridges Across Highways*. Proceedings of the International conference on wildlife ecology & transportation, Missoula, Montana, 1999
- [58] Witten I.H., Frank E., *Data mining: practical machine learning tools and techniques*. Morgan Kaufman Publishers, Amsterdam, 2005
- [59] Kononenko I., Kukar M., *Machine learning and data mining: Introduction to principles and algorithms*. Hordwood Publishing, Chichester, 2007
- [60] Dimitrovski I., Kocev D., Loskovska S., Džeroski S., Hierarchical classification of diatom images using ensembles of predictive clustering trees. *Ecol. Inform.*, 2012, 7, 1, 19-29, doi: 10.1016/j.ecoinf.2011.09.001
- [61] Lin N., Noe D., He X., *Tree-based methods and their applications*. V: Phoam H. (Ed.), *Springer handbook of engineering statistics*. Springer, London, 2006, 551-570
- [62] Lu D., Weng Q., A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.*, 2007, 28, 5, 823-870, doi: 10.1080/01431160600746456
- [63] Jiang B., Ding X., Ma L., He Y., Wang T., Xie W., A hybrid feature selection algorithm: combination of symmetrical uncertainty and genetic algorithms. In: *Proceedings of the second international symposium Optimization and systems biology*, Lijiang, 2008, 152-157
- [64] Tirelli T., Pessani D., Importance of feature selection in decision-tree and artificial neural- network ecological applications. *Alburnus alburnus alborella: A practical example*. *Ecol. Inform.*, 2011, 6, 5, 309-315, doi: 10.1016/j.ecoinf.2010.11.001
- [65] Williams K.J., Belbin L., Austin M.P., Stein J.L., Ferrier S., Which environmental variables should I use in my biodiversity model? *Int. J. Geogr. Inf. Sci.*, 2012, 26, 2009-2047, doi: 10.1080/13658816.2012.698015
- [66] Ciglić R., Information values of absolute elevation and elevation difference for –illustration of thermal belt. *Acta Geogr. Slo.*, 2010, 50, 2, 177-200, doi: 10.3986/AGS50201

- [67] Fayya, U.M., Irani, K.B., Multi-interval discretization of continuous-valued attributes for classification learning. In: Bajcsy R. (Ed.), Proceedings of the Thirteenth international joint conference on artificial intelligence 2, San Mateo, 1993, 1022-1027
- [68] Hall M., Frank E., Holmes G., Pfahringer B., Reutemann P., Witten I.H., The WEKA data mining software: an update. SIGKDD Explor., 2009, 11, 1, 10-18