

# Expert knowledge for translating land cover/use maps to General Habitat Categories (GHC)

Maria Adamo · Cristina Tarantino · Valeria Tomaselli · Vasiliki Kosmidou · Zisis Petrou · Ioannis Manakos · Richard M. Lucas · Caspar A. Mächer · Giuseppe Veronico · Carmela Marangi · Vito De Pasquale · Palma Blonda

Received: 10 October 2013 / Accepted: 2 April 2014 / Published online: 16 April 2014  
© The Author(s) 2014. This article is published with open access at Springerlink.com

**Abstract** Monitoring biodiversity at the level of habitats and landscape is becoming widespread in Europe and elsewhere as countries establish international and national habitat conservation policies and monitoring systems. Earth Observation (EO) data offers a potential solution to long-term biodiversity monitoring through direct mapping of habitats or by integrating Land Cover/Use (LC/LU) maps with contextual spatial information and in situ data. Therefore, it appears necessary to develop an automatic/semi-automatic translation framework of LC/LU classes to habitat classes, but also challenging due to discrepancies in domain definitions. In the context of the FP7 BIO\_SOS ([www.biosos.eu](http://www.biosos.eu)) project, the authors demonstrated the feasibility of the Food and Agricultural Organization Land Cover Classification

System (LCCS) taxonomy to habitat class translation. They also developed a framework to automatically translate LCCS classes into the recently proposed General Habitat Categories classification system, able to provide an exhaustive typology of habitat types, ranging from natural ecosystems to urban areas around the globe. However discrepancies in terminology, plant height criteria and basic principles between the two mapping domains inducing a number of one-to-many and many-to-many relations were identified, revealing the need of additional ecological expert knowledge to resolve the ambiguities. This paper illustrates how class phenology, class topological arrangement in the landscape, class spectral signature from multi-temporal Very High spatial Resolution (VHR) satellite imagery and plant height

---

M. Adamo (✉) · C. Tarantino · P. Blonda  
Institute of Intelligent Systems for Automation (ISSIA),  
National Research Council (CNR), Via Amendola 122/D-  
O, 70126 Bari, Italy  
e-mail: [adamo@ba.issia.cnr.it](mailto:adamo@ba.issia.cnr.it)

V. Tomaselli · G. Veronico  
Institute of Biosciences and BioResources (IBBR),  
National Research Council (CNR), Via Amendola 165/A,  
70126 Bari, Italy

V. Kosmidou · Z. Petrou · I. Manakos  
Information Technologies Institute (ITI), Centre for  
Research & Technology Hellas (CERTH), P.O. Box:  
60361, 6th km Harilaou – Thermi, 57001 Thessaloniki,  
Greece

R. M. Lucas  
Institute of Geography and Earth Sciences, Aberystwyth  
University, Aberystwyth, Ceredigion SY23 3DB, UK

C. A. Mächer  
Alterra, Wageningen UR, Droevendaalsesteeg 3, 6708 PB  
Wageningen, The Netherlands

C. Marangi  
Institute for applied mathematics “Mauro Picone” (IAC),  
National Research Council (CNR), Via G. Amendola 122,  
70126 Bari, Italy

V. De Pasquale  
Planetek Italia (PKI), Via Massaua, 12, 70123 Bari, Italy

measurements can be used to resolve such ambiguities. Concerning plant height, this paper also compares the mapping results obtained by using accurate values extracted from Light Detection And Ranging (LIDAR) data and by exploiting EO data texture features (i.e. entropy) as a proxy of plant height information, when LIDAR data are not available. An application for two Natura 2000 coastal sites in Southern Italy is discussed.

**Keywords** Biodiversity monitoring · General Habitat Categories · VHR satellite imagery

## Introduction

Earth Observation (EO) imagery can provide a continuous synoptic view of land cover/use (LC/LU) patterns and LC/LU changes which have an impact on biodiversity loss (Turner et al. 2003; Townsend et al. 2009). LC/LU classes are not a suitable tool in assessing biodiversity in comparison to habitat classes, which are linked to species, communities and biotopes (Bunce et al. 2012a). According to Lengyel et al. (2008), habitat mapping constitutes one of the possible links among EO data, biodiversity monitoring and ecosystem status assessment, as habitats are linked to species occurrence, and the choice of appropriate taxonomies is a key point in the translation of LC/LU to habitat maps. However, whereas a conversion from LC/LU to habitat classes can be of great help, differences in taxonomies and definitions between the two domains (LC/LU and habitats) have so far limited the establishment of a unified approach for such translation.

Concerning LC/LU, the Food and Agricultural Organisation (FAO) Land Cover Classification System (LCCS) (Di Gregorio and Jansen 2005) was found to be the most useful for translating EO-derived LC/LU classes to habitat categories (Tomaselli et al. 2013), since it allows a better description of natural habitats in comparison to other classification systems (e.g., CORINE, Bossard et al. 2000). Consequently, Tomaselli et al. (2013) proposed a framework to translate LCCS classes into habitat categories, with these described according to Annex I to the Habitats Directive (92/43 EEC). The Habitats Directive is the main European Union (EU) legal instrument concerning biodiversity

**Table 1** List of General Habitat Categories (GHC)

GHC supercategory/primary code acronyms	GHC supercategory/primary code vernacular name
URB/ART	Urban/artificial
URB/NON	Urban/non Vegetated
URB/VEG	Urban/crops
URB/GRA	Urban/herbaceous
URB/TRE	Urban/woody vegetation
URB/ART/ROA	Urban/artificial/road paved
CUL/SPA	Cultivated/bare ground
CUL/CRO	Cultivated/herbaceous crops
CUL/WOC	Cultivated/woody crops
HER/LHE	Herbaceous/leafy hemicryptophytes
HER/CHE	Herbaceous/caespitose hemicryptophytes
HER/THE	Herbaceous/thaerophytes
HER/GEO	Herbaceous/geophytes
HER/HCH	Herbaceous/chamaephytes
HER/CRY	Herbaceous/cryptogams
HER/SHY	Herbaceous wetland/submerged hydrophytes
HER/EHY	Herbaceous wetland/emergent hydrophytes
HER/HEL	Herbaceous wetland/helophytes
HER/SHY/FLO	Herbaceous wetland/submerged hydrophytes/free floating
HER/SHY/LEA	Herbaceous wetland/submerged hydrophytes/leafy
SPV/SEA	Sparsely vegetated/sea
SPV/TID	Sparsely vegetated/tidal
SPV/AQU	Sparsely vegetated/aquatic
SPV/ICE	Sparsely vegetated/ice and snow
SPV/TER	Sparsely vegetated/terrestrial
TER/ROC	Terrestrial/bare rock
TER/BOU	Terrestrial/boulders
TER/STO	Terrestrial/stones
TER/GRV	Terrestrial/gravel
TER/SAN	Terrestrial/sand
TER/EAR	Terrestrial/earth, mud
TRS/DCH	Trees or shrubs/dwarf chamaephytes
TRS/SCH	Trees or shrubs/shrubby chamaephytes
TRS/LPH	Trees or shrubs/low phanerophytes
TRS/MPH	Trees or shrubs/mid phanerophytes
TRS/TPH	Trees or shrubs/tall phanerophytes
TRS/FPH	Trees or shrubs/forest phanerophytes
TRS/GPH	Trees or shrubs/mega forest phanerophytes

**Table 1** continued

GHC supercategory/ primary code acronyms	GHC supercategory/primary code vernacular name
TRS/DCH/EVR	Trees or shrubs/dwarf chamaephytes/ evergreen
TRS/SCH/EVR	Trees or shrubs/shrubby chamaephytes/ evergreen
TRS/LPH/EVR	Trees or shrubs/low phanerophytes/ evergreen
TRS/MPH/EVR	Trees or shrubs/mid phanerophytes/ evergreen
TRS/TPH/EVR	Trees or shrubs/tall phanerophytes/ evergreen
TRS/FPH/EVR	Trees or shrubs/forest phanerophytes/ evergreen
TRS/MPH/ EVR/CON	Trees or shrubs/mid phanerophytes/ evergreen/coniferous
TRS/TPH/EVR/ CON	Trees or shrubs/Tall Phanerophytes/ Evergreen/coniferous
TRS/FPH/EVR/ CON	Trees or shrubs/forest phanerophytes/ evergreen/coniferous
TRS/MPH/DEC	Trees or shrubs/mid phanerophytes/ deciduous
TRS/TPH/DEC	Trees or shrubs/tall phanerophytes/ deciduous
TRS/DCH/NLE	Trees or shrubs/dwarf chamaephytes/non leafy evergreen
TRS/SCH/NLE	Trees or shrubs/shrubby chamaephytes/non leafy evergreen
TRS/LPH/NLE	Trees or shrubs/low phanerophytes/non leafy evergreen

The acronyms include the supercategory. For TRS supercategory additional components are provided, with these including *DEC* Winter deciduous, *EVR* Evergreen, *CON* Conifers, *NLE* Non-leafy evergreen and *SUM* Summer deciduous

and conservation of natural habitats; its Annex I provides a list of natural habitats of Community interest within the EU and it is of central importance for international reporting and Natura 2000 site management (Ladoux et al. 2000; Mehtälä and Vuorisalo 2007). The list does not include anthropogenic and artificial habitats and, although organized into nine groups, it appears that there is no way to identify natural habitats in the field other than by gaining familiarity with the list by local experts who are generally used to interpret the descriptions when answering to national priorities (Bunce et al. 2012b).

In addition, common approaches for monitoring changes in habitats require also definitions and rules that are harmonised continentally and globally. To address such problems, Bunce et al. (2011, 2012b) introduced the General Habitat Categories (GHC) methodology and tested it throughout Europe through in-field campaigns carried out, in different biogeographic regions, on sampling 1 km<sup>2</sup> basic survey areas (0.25 km<sup>2</sup> in complex landscapes), with these being extremely costly. The GHC methodology for habitat classification and monitoring comes as an ecological refinement of the land cover categorization used in LCCS (Bunce et al. 2011) relying on the concept of Life Forms (LFs), which are strictly related to plant structure and morphology. However, LCCS and GHC use different definitions of LFs. In LCCS LFs are defined by physiognomic aspect of plant (Kuechler and Zonneveld 1988), whilst GHC are based on the concept of LFs as proposed by Raunkiaer (1934), defined by the position of buds during unfavourable season. The composition of LFs and the relative abundance in the biological spectra are a direct link among flora, vegetation and environmental (mainly climate) variables. The mapping of GHC, which are reported in Table 1, is firstly based on the identification of LFs and Non-LFs (e.g. artificial areas, rocks) through a set of expert decision rules (Bunce et al. 2011). Then, additional qualifiers on environment, site, management and species composition are collected during in-field campaigns to express variations between elements having the same GHC code and to provide a translation from GHC to Annex I habitat maps (Bunce et al. 2012b). The majority of environmental qualifiers are unlikely to change quickly over time, so that they can be updated during a targeted monitoring process.

LC/LU maps obtained from VHR space imagery, if based on LCCS taxonomy, can be used to provide holistic GHC maps through translation rules (Kosmidou et al. 2014) at lower costs. However, such systematic translation between two different domains (i.e., LC/LU and habitats) has been proven challenging, largely because of discrepancies between class definitions in the related taxonomies. In particular, only few *one-to-one* relationships and several *one-to-many* and *many-to-many* LCCS to GHC class transitions were observed, as systematically identified and described in Figs. 1, 2, 3, and 4 in Kosmidou et al. (2014). These class relationships were crisp and the authors suggested the use of additional information,

**Table 2** Expert rules based on spatial and temporal relations and spectral indices towards advanced LCCS to GHC translation

Rule type	Feature	Condition	Meaning	Rule ID	
Geometric	Border index (BI)	$BI > 1.2$	Segment shape regularity: true	G1	
		$BI < 1.2$	Segment shape non regularity: false	G2	
Temporal	Post biomass peak green/red ratio	$PoBP\_GRR < 1$	Annual vegetation	P1	
		$PoBP\_GRR > 1$	Perennial vegetation	P2	
	Biomass peak blue/NIR ratio	$BP\_BNR > 1$	Perennial water coverage: true	W1	
		$BP\_BNR < 1$	Perennial water coverage: false	W2	
Spatial topological	Rel border to (ranges in [0,1] if it is = 0 no pixels is common with the object labeled as SHY. IF it is $\neq 0$ then there is adjacency)	Rel Border to $\neq 0$	Adjacency to SHY: true	S1	
		Rel Border to = 0	Adjacency to SHY: false	S2	
Spatial non topological	Rel border to adjacency to ART and NON	Rel Border to $\geq 0.5$	Adjacency to (ART or NON) $\geq 50\%$	U1	
		Rel Border to $< 0.5$	Adjacency to (ART or NON) $< 50\%$	U2	
Spatial non topological	Percentage of vegetated (PV) pixels (i.e. pixels with $BP\_GRR \geq 1$ )	$PV > 30\%$	Percentage coverage of vegetation higher than 30 %: true	C1	
		$PV > 30\%$	Percentage coverage of vegetation higher than 30 %: false	C2	
LIDAR derived	LIDAR CHM (canopy height model)	$CHM > 40\text{ m}$	Vegetation height	L1	
		$5 < CHM < 40\text{ m}$		L2	
		$2 < CHM < 5\text{ m}$		L3	
		$0.6 < CHM < 2\text{ m}$		L4	
		$0.3 < CHM < 0.6\text{ m}$		L5	
		$0.05 < CHM < 0.3\text{ m}$		L6	
		$CHM < 0.05\text{ m}$		L7	
		$CHM > 0.6\text{ m}$		L8	
		$CHM < 0.6\text{ m}$		L9	
		Percentage of vegetated (PV) pixels with height $> 0.6\text{ m}$	$PV_{(H,0.6)} > 30\%$	Percentage coverage of vegetation with height higher than 0.6 m $> 30\%$ : true	L10
			$PV_{(H,0.6)} < 30\%$	Percentage Coverage of vegetation with height higher than 0.6 m $> 30\%$ : false	L11
Texture	1st order entropy calculated on the green band of BP image ( $BP\_E_{Green}$ )	$BP\_E_{Green} > 1.6$	Homogeneity degree related to vegetation height	T1	

When not specified the indices are computed for both images

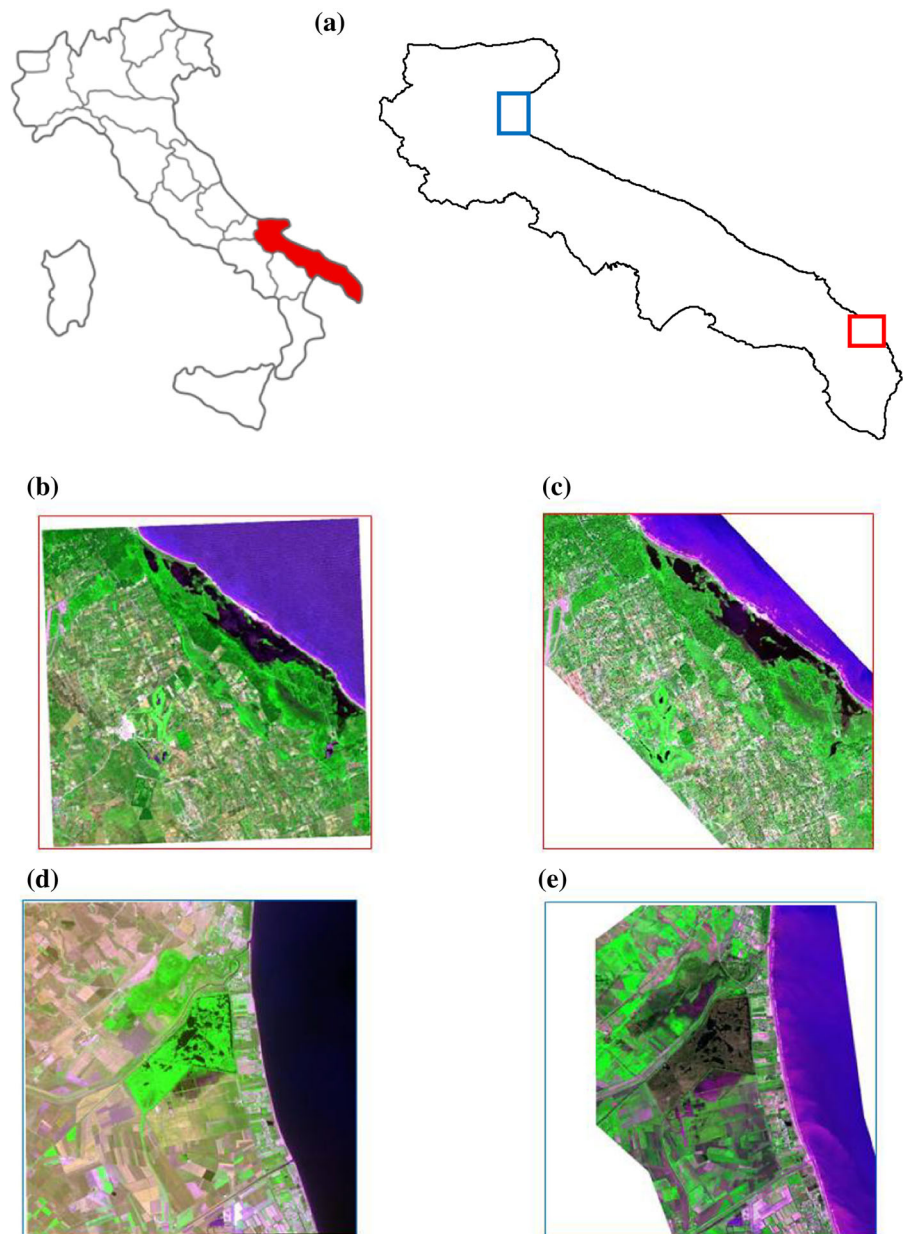
*PoBP* Post biomass peak, *BP* Biomass peak

including vegetation height measurements and land use information, to resolve the ambiguity in the *one-to-many* and *many-to-many* relationships.

There are many studies in the literature, both feature based and pixel based, focusing on the issue of map legend comparison and disambiguation of *one-to-many* and *many-to-many* relations between legends for change detection and validation applications (Herold and Schumliuss 2004; Comber et al. 2004; Herold et al. 2006; Fritz and See 2008; Herold et al. 2009). However

most of them mainly focus on the comparison of map legends concerning the same thematic product (e.g., land cover) produced by using different taxonomies and methods. Fritz and See (2008) compared the Global Land Cover 2000 map with the MODIS land cover data set by firstly reconciling legends through a look-up table and then calculating the fuzzy degree of overlap between legend classes, where there is an overlap, by considering user-driven or external sources-driven importance of different land cover

**Fig. 1** **a** Le Cesine (*red box*) and Lago Salso (*blue box*) test sites location. RGB (Red-NIR-Blue) composition of: **b** Quickbird image acquired on June 2009, **c** Worldview-2 image acquired on October 2010 **c** for Le Cesine site. RGB (Red-NIR-Blue) composition of Worldview-2 image acquired on **d** June 2010 and **e** on February 2011 for Lago Salso site. (Color figure online)



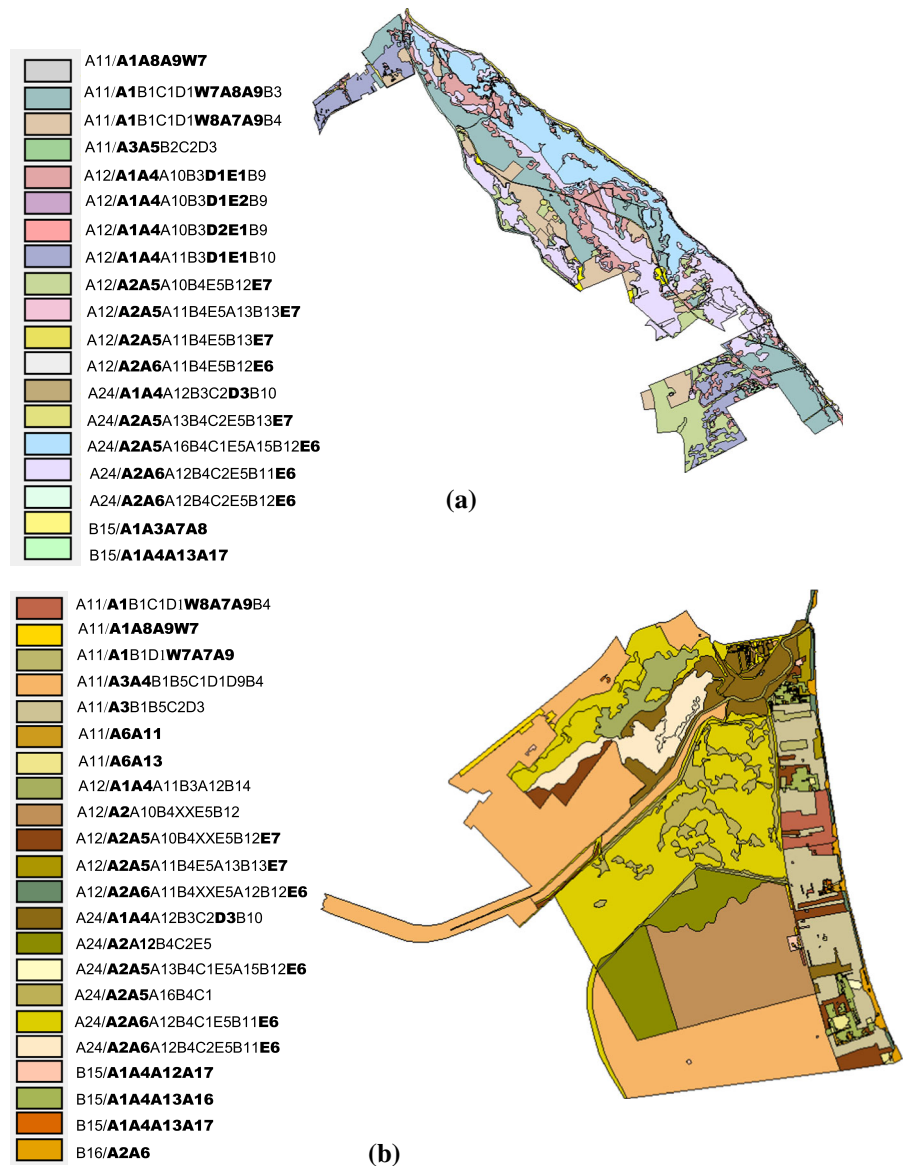
types and finally producing a fuzzy agreement map. Regarding the quantitative comparison of taxonomies, several studies have recently contributed to define a frame where the interoperability between taxonomies is assessed by introducing some semantic similarity measures of the different classification schemes (Feng and Flewelling 2004; Ahlqvist 2004, 2005, 2008; Fritz and See 2008).

The present paper focuses on the translation rules between taxonomies belonging to two different thematic domains (i.e. LC/LU and habitats) and introduces

expert knowledge in the form of crisp rules to resolve the ambiguities in LCCS to GHC translation. LCCS and GHC classes and their relationships are described in terms of spatial (e.g., topological adjacency, enclosure) and temporal relations (i.e., plant phenology and water coverage), as well as class specific spectral signatures in EO data. Applying expert knowledge with reflectance and ancillary data allows better solutions to classification problems (Comber et al. 2004, 2005). In addition, the possibility to formalize spatial–temporal relations (Goodchild et al. 2007; Pierkot et al. 2013) allows for the



**Fig. 2** LC/LU map available and the list of LCCS classes with the alphanumeric code subset considered in LCCS to GHC translation evidenced by *bold* characters for: **a** Le Cesine and **b** Lago Salso

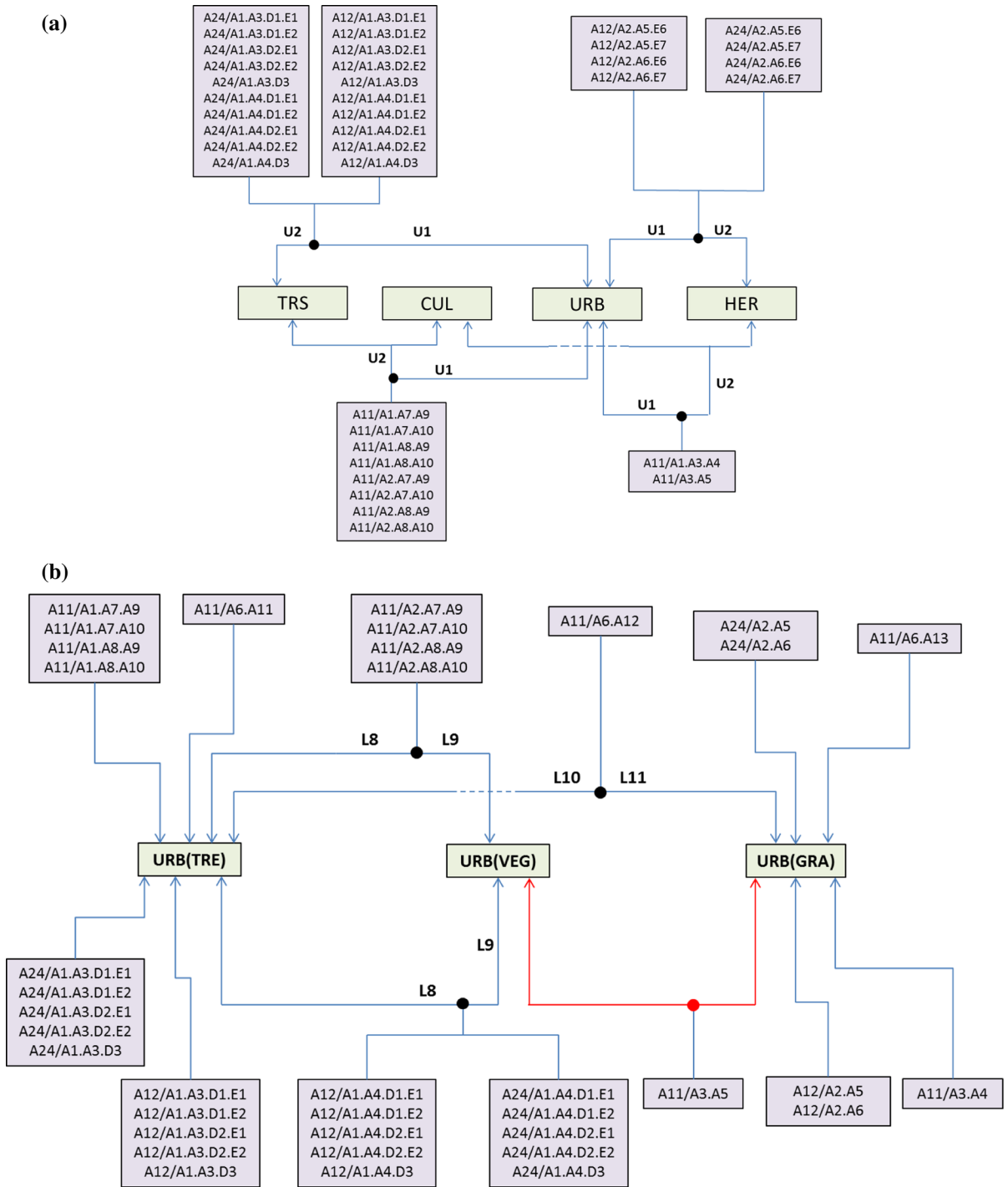


identification of inconsistencies between different datasets.

The purpose of the paper is twofold: (1) to resolve the ambiguities reported through the *one-to-many* and *many-to-many* relations but not dealt with in (Kosmidou et al. 2014); (2) the assessment of texture measurements as surrogate of plant height information in comparison to Light Detection And Ranging (LIDAR) data measurements. Once obtained from LC/LU maps, GHC maps can be translated into Annex I habitats through a specific key (Bunce et al. 2012c), as well as to other habitat LFs based taxonomies outside Europe. The effectiveness of the proposed approach is

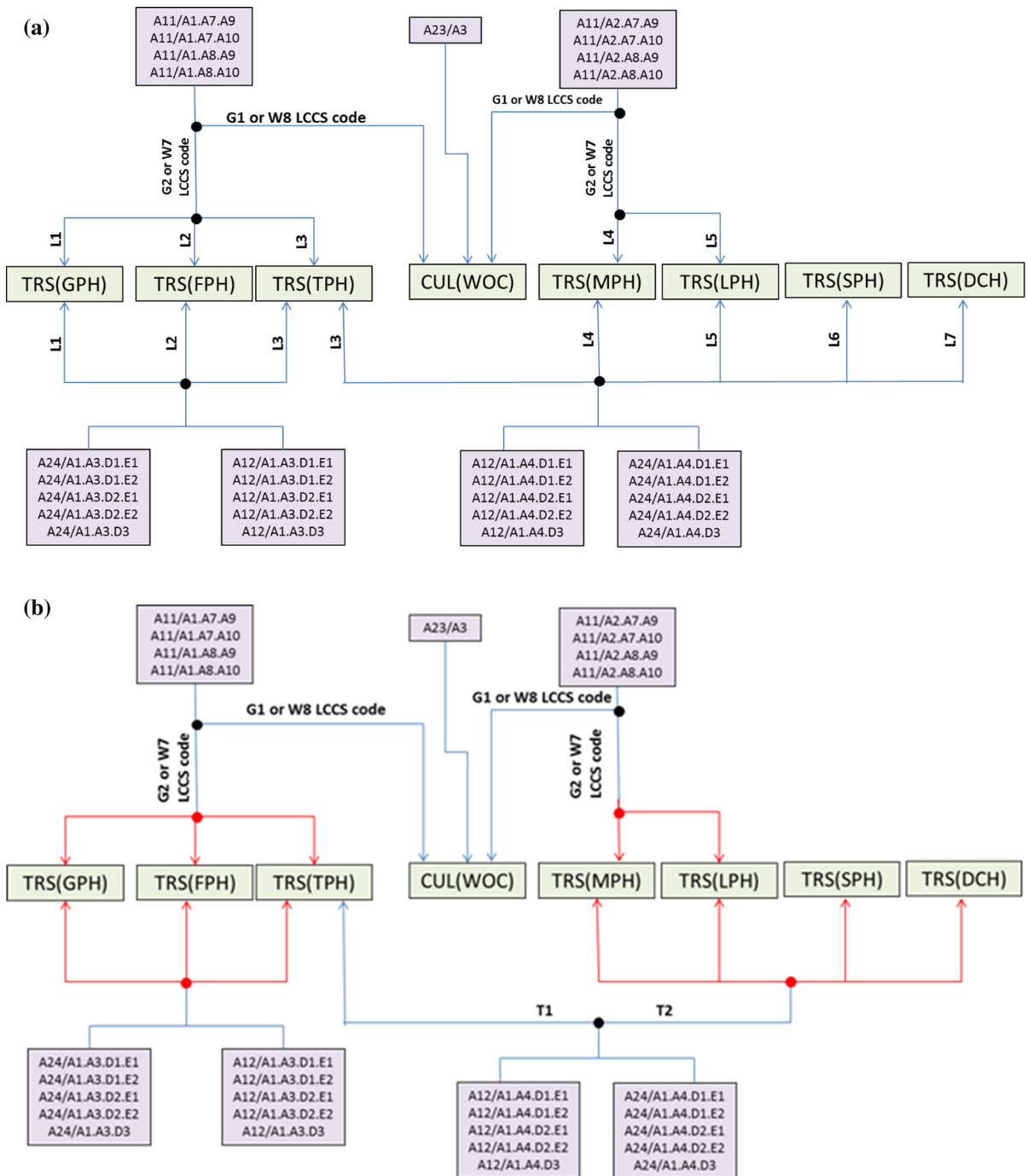
demonstrated for two coastal Natura 2000 sites in Italy as described below but can be extended to other sites.

The proposed approach is considered novel in the sense that it incorporates spatial reasoning (adjacency rules) and temporal relations as well as multi-source data integration (e.g., aerial LIDAR and VHR satellite data) for habitat mapping. To date, in the framework of Mapping and Assessment of Ecosystems and their Services (Maes et al. 2013) there has been little work done in automatic translation of LC/LU to habitat classes based on the use of VHR EO data (fine scale) and expert knowledge.



**Fig. 3** a Diagram for adjacency rule used in the first LCCS to GHC mapping step to discriminate vegetated urban and natural GHC. b Diagram of rules for the identification of vegetated

urban GHC sub-categories. The red arrows represent ambiguities that cannot be resolved with the elements at hand. (Color figure online)



**Fig. 4** Woody GHC categories disambiguation diagram adopted: **a** when LIDAR data are used for plant height measurements; **b** when texture is used as proxy for plant height

information. The *red arrows* represent ambiguities that cannot be resolved without LIDAR data. (Color figure online)



This work was conducted within the three-year BIO\_SOS ([www.biosos.eu](http://www.biosos.eu)) project, funded within the European Union FP7-SPACE third call. The project developed a pre-operational system for cost effective and timely monitoring of changes in land cover and habitats within and along the borders of protected areas in support to policy makers.

### Study sites and EO data

Two study sites, belonging to the Natura 2000 network and located in Southern Italy, Le Cesine (SCI code IT9150032, SPA code IT9150014) and Lago Salso (SCI code IT9110005, SPA code IT9110038) are shown in Fig. 1a.

Le Cesine site is a coastal area, mainly composed by a complex of coastal lagoons and characterized by a system of ponds, marshes and wet meadows. The woody vegetation in this area is a mosaic of *Pinus halepensis* stands and different types of Mediterranean maquis and garrigues. The site covers an area of about 2,148 ha. The wetland is one of the most important sites in southern Italy in terms of biodiversity and is formed by two large coastal lagoons and various channels, marshes and wet grasslands. Reeds and sedges communities are prone to fires spreading from the adjacent agricultural areas and from arsons. Marine erosion has caused a progressive reduction of the sandbank, with frequent ruptures of the dune belt and a progressive salinization of the lagoons and the related environments. Due to water salinization, sedges, rushes and reeds communities have been replaced, over time, by halophytic scrub. In addition, the progressive erosion of the sand bank is determining the reduction and fragmentation of the typical dune habitat types. Finally, agricultural practices taking place both inside and outside the protected area are likely to lead to the disappearance of temporary ponds, which are priority habitats according to the EU Habitats Directive. Along the coastline, the Natura 2000 site is not only exposed to conversion to agricultural practices, but also to illegal urban development and to an increasing tourist pressure.

The Lago Salso site falls in the northern part of the Natura 2000 site “Zone Umide della Capitanata” (IT9110005) and within the Gargano National Park. It consists mainly of a wetland characterized by brackish and freshwater that, at the end of 19th century, covered

an area of more than 4,000 ha, and then dramatically reduced as a result of massive land reclamation. Nowadays, the wetland is diminished to an area of about 540 ha and is characterized by reed, sedge and rush communities. The main pressures are represented by agriculture intensification and farm expansion in the surrounding areas. In the last few years, parts of such agricultural areas have been prone to progressive overflowing, as an action for the recovery of part of the original wetland. Northward, Lago Salso is adjacent to typical salt marshes, characterized by salty soils only periodically flooded and by halophilous annual and shrub vegetation.

Pre-existing validated LC/LU maps available for both sites were used for producing and validating GHC maps. The pre-existing LC/LU map (scale 1:5,000) available for Le Cesine site was produced in CORINE taxonomy and validated by in-field campaigns undertaken in 2008–2009. The map was firstly converted from CORINE taxonomy to LCCS taxonomy based on the report of panel of the Global Observation of Forest and Land Cover Dynamics (Herold and Schullius 2004; Herold et al. 2006, 2009; Tomaselli et al. 2013). The LC/LU map for Lago Salso was produced on the basis of a pre-existing vegetation map and orthophoto interpretation and was based on the LCCS taxonomy (scale 1:5,000). The mapping was validated by in-field campaigns carried out in 2011–2012, to verify the presence and distribution of both artificial and natural and semi-natural habitat types. For natural and semi-natural habitat types, information on vegetation composition and structure was also collected, to validate the GHC translation.

For each coastal wetland site, two multi-temporal VHR images were acquired in two different seasons (e.g., summer and autumn) according to the phenology of the most representative plant communities and the water seasonality. The images were used to extract the spectral indices useful to disambiguate the *one-to-many* LCCS to GHC relations along with spatial and temporal relations, as described in subsequent sections. For Le Cesine site, a QuickBird image acquired on June 2009 and a WorldView-2 image acquired on October 2010 were selected (Fig. 1b, c). For Lago Salso site, two WorldView-2 images were acquired on June 2010 and February 2011 (Fig. 1d, e). All the images were ortho-rectified, co-registered and calibrated in Top of Atmosphere (TOA) Reflectance

values. The spatial resolution (2 m) was compatible with LCCS map scale. Due to differences in the dates related to in-field validation campaigns of LC/LU maps and EO image acquisitions, only reference samples corresponding to no changed classes were selected through visual image inspection to validate the GHC products. LIDAR data were provided by the Italian Ministry of Environment for both sites. The acquisitions were made in spring 2009.

## Methodology

The Figs. 1, 2, 3, and 4 in Kosmidou et al. (2014) evidenced that only 39 out of the total 87 LCCS to GHC mapping relations are *one-to-one*. In this paper, most discrepancies between the two taxonomies were resolved, focusing on the ones related to the manner the land use information is perceived (i.e., a field close to a building changes the categorization in the GHC taxonomy from natural trees and shrubs (TRS) into urban woody (TRE), whereas in LCCS taxonomy it remains as such), and to the differences in the height criteria used to classify the trees/shrubs categories as described hereafter.

### Features used for the translation

Based on the field patches indicated by the input LCCS, an object-based translation approach is adopted. Expert knowledge is used in the proposed framework to describe LCCS and GHC classes and their relationships in terms of spatial and temporal relations and spectral indices extracted from EO data. The description is clearly scale dependent; this means that the methodology can be adapted to any scale if the description of classes is provided at the appropriate required scale. The GHC surveillance and monitoring technique (Bunce et al. 2011) was designed to map GHC through in-field campaigns. Consequently the rules and class descriptions are provided for fine scale mapping. VHR EO data and automatic techniques can help in GHC mapping across the full extent of protected study areas with a significant reduction in the costs related to in-field campaigns. To the authors' knowledge, no fine-scale mapping of GHC has previously been performed from EO data. A

number of key elements of LCCS to GHC map disambiguation are discussed hereafter at fine scale, including:

- (1) *Spatial topological relations among classes.* These were considered to both take into account the spatial onation of wetlands and decide if a vegetated object had to be considered as part of an urban area. More specifically, the *adjacency relation to urban categories* was used.
- (2) *Temporal relations.* To deal with the main phenological (vegetative) stages of both vegetation and water seasonality, two satellite images had to be introduced, corresponding, respectively, to the dry and wet seasons.
- (3) *Geometric attributes* for extracting land use information. The requirement of regularity in the shape of a field (object) was used. This attribute was extracted within the eCognition software using the Border Index feature (Trimble 2011). Based on a rectangular approximation, this feature describes how jagged a segment is. The smallest rectangle enclosing the image object is created and the Border Index is calculated as the ratio between the perimeter length of the object and the smallest enclosing rectangle. These attributes were used to discriminate trees belonging to TRS from WOC GHC categories.
- (4) *Prior spectral knowledge* on class spectral signature. Spectral indices were used as inputs to the translation algorithm, such as the Blue/NIR Ratio (BNR) (Morris and Dupigny-Giroux 2010) and Green/Red Ratio (GRR) (Ritchie et al. 2010) defined as:

$$BNR = \rho(\lambda_{Blue}) / \rho(\lambda_{NIR}), \quad (1)$$

$$GRR = \rho(\lambda_{Green}) / \rho(\lambda_{Red}), \quad (2)$$

where  $\rho$  is the TOA value in the bands considered and it is calculated as the mean value inside the object and  $\lambda_{Blue}$ ,  $\lambda_{Red}$  and  $\lambda_{NIR}$ , indicate the image bands lying in the blue, red and near-infrared part of the electromagnetic spectrum. The bands ratios used in GRR and BNR indices perform well in discriminating vegetated and water covered pixels, respectively.

- (5) *LIDAR derived measurements.* Due to the existing inconsistency in the two taxonomies between the definitions of plant height (Kosmidou et al.

2014), the information of plant height was directly extracted from the available LIDAR dataset and not taken from the height code element (if any) associated to the LCCS label. This information is useful for the disambiguation of Phanerophytes (Raunkiaer's LF including perennial woody plants with buds located more than 30 cm above the soil surface, typically trees and large shrubs) which are divided in 5 GHC (i.e., LPH, MPH, TPH, FPH and GPH) categories on the basis of their height. Up to the date this work was conducted, 43.44 % of the Italian territory (131.118 km<sup>2</sup>) has been covered by LIDAR acquisitions (Costabile et al. 2013). The LIDAR acquisitions were made in May 2009 when plant growth and biomass were seasonally high to minimize laser penetration of the vegetation canopy. The dataset used consists of Digital Terrain Model (DTM), Digital Surface Model (DSM) (from the first and the last pulse) and cloud points. The DTM and DSM are provided with a spatial resolution of 2 m × 2 m and a vertical accuracy of ± 15 cm. The height of each object in the LCCS map was obtained as the difference between the DSM (from the first pulse) and the DTM.

- (6) *Contextual features.* The first order entropy (occurrence measure) texture (Anys et al. 1994), computed in the green band of the biomass peak image, was used to provide plant height information to solve problems when LIDAR data are not available. To compute first-order statistics at a given scale of interest, the pixel values within a moving window were used and the measurement was assigned to the central pixel. Two first-order texture measures (entropy and variance) were selected based on their established ability to characterize vegetation structure (Wuest and Zhang 2009; Wood et al. 2012; Petrou et al. 2012). These were then calculated using imagery acquired at the peak of the biomass. Different size windows were used to account for the heterogeneity of pixel values in relation to vegetation structure and height. The entropy measure, calculated with a kernel window 3 × 3 pixels sized on the image green band was experimentally identified as the most appropriate to discriminate plant height into two main classes in the analysis. The comparison of the

results with the ones from LIDAR data was carried out in terms of the overall accuracy (OA) of the GHC maps.

The translation algorithm receives as inputs: an LCCS map, the multi-seasonal EO images and the set of translation rules. As well known (Di Gregorio and Jansen 2005; Tomaselli et al. 2013), any LCCS class in the map is described by a unique numerical code and a standard name. The alphanumeric elements of the LCCS code correspond to either specific classifiers (i.e., a set of independent diagnostic criteria used to identify each class) or environmental and technical attributes (Di Gregorio and Jansen 2005) useful for class identification. However, some qualifiers in the LCCS code are not related to the GHC semantic class description and consequently are not useful for LCCS to GHC mapping. For this reason, the algorithm considers only a specific sub-set of the alphanumeric elements in the code. More specifically, *life form*, *leaf type and leaf phenology* elements (classifiers) are selected in the code of both natural terrestrial vegetated (i.e., A12 category) and natural aquatic vegetated classes (i.e., A24). Regarding the cultivated classes (i.e., A11), additional codes related to the *technical attributes* (e.g., orchards and plantations) are considered; only the *life form* component is considered for cultivated aquatic classes (i.e., A23); *surface aspect* is considered for both primarily non vegetated artificial surfaces (i.e. B15) and bare areas (i.e. B16), whereas *physical status* is used for artificial water bodies (i.e., B27) and inland waterbodies (i.e. B28).

Figure 2a and b show the input LCCS maps available for Le Cesine and Lago Salso sites, respectively. The Figures include the list of LCCS classes. The elements of the code used for LCCS to GHC translation are evidenced in bold.

#### Expert rules

For translating LCCS classes into GHC, the algorithm applies a set of expert rules listed in Table 3 with the feature used in the text (second column), the feature/spectral index condition (third column) and the description of the rule conditions (second column). Figures 3, 4, and 5 depict in a concise graphical manner the rules for the LCCS to GHC translation. The rules are organized per group of output GHC categories (e.g., URB, TRS, HER), as described hereafter. In

**Table 3** LCCS classes and the corresponding feasible GHC classes located in Le Cesine and Lago Salso sites, based on the work in Kosmidou et al. 2014

Presence (1)/ absence (0) in each site	LCCS dich. code	LCCS hierarchical code and class description	GHC output categories	
10	<b>A11</b> Cultivated and managed terrestrial areas	<b>A3.A5.B2.C2.D3</b> Small sized field of irrigated herbaceous non-graminoid crops.	CUL/CRO or URB/VEG or URB/GRA or HER/LHE or HER/THE or HER/HCH or HER/GEO	
01		<b>A3.A4.B1.B5.C1.D1.D9.B4</b> Rainfed graminoid crops	CUL/CRO or URB/GRA or HER/CHE or HER/THE	
11		<b>A1.B1.C1.D1.W7.A8.A9.B3</b> Monoculture fields of rainfed evergreen needle-leaved tree crops (plantations)	URB/TRE or TRS/TPH/EVR/CON or TRS/ FPH/EVR/CON	
11		<b>A1.B1.C1.D1.W8.A7.A9.B4</b> Monoculture fields of rainfed broad- leaved tree crops orchards (olive groves)	CUL/WOC or URB/TRE	
01		<b>A3.B1.B5.C2.B3</b> Irrigated herbaceous crop	CUL/CRO or URB/VEG or URB/GRA or HER/LHE or HER/THE or HER/HCH or HER/GEO or HER/CHE	
01		<b>A1.B1.D1.W7.A7.A9</b> Monoculture fields of rainfed broadleaved tree crops plantations	CUL/WOC or TRS/TPH/EVR or TRS/FPH/ EVR or URB/TRE	
01		<b>A6.A11</b> Urban vegetated areas. Parks	URB/TRE	
01		<b>A6.A13</b> Urban vegetated areas. Lawns	URB/GRA	
10		<b>A12</b> Natural and semi-natural terrestrial vegetation	<b>A1.A4.A10.B3.D1.E2.B9</b> Broad-leaved deciduous medium/high closed shrubland	URB/TRE or TRS/MPH/DEC or TRS/TPH/ DEC
10			<b>A1.A4.A10.B3.D2.E1.B9</b> Needleleaved evergreen medium/high closed shrubland	URB/TRE or TRS/MPH/EVR/CON or TRS/ TPH/EVR/CON
10			<b>A1.A4.A11.B3.D1.E1.B10</b> Broad-leaved evergreen open dwarf shrubland	URB/VEG or TRS/DCH/EVR or TRS/SCH/ EVR or TRS/LPH/EVR
10			<b>A1.A4.A10.B3.D1.E1.B9</b> Broad-leaved evergreen medium/high closed shrubland	URB/TRE or TRS/MPH/EVR or TRS/TPH/ EVR
11			<b>A2.A6.A11.B4.E5.B12.E6</b> Open perennial medium/ tall grassland	HER/CHE or URB/GRA
11			<b>A2.A5.A11.B4.E5.B13.E7</b> Open annual short forbs	URB/GRA or HER/THE or HER/LHE or HER/HCH or HER/GEO or (weak) TRS/ DCH or (weak) TRS/SCH
11	<b>A2.A5.A10.B4.E5.B12.E7</b> Closed annual medium/tall forbs		URB/GRA or HER/THE or HER/LHE or HER/HCH or HER/GEO	
01	<b>A2.A10.B4.XX.E5.B12</b> Medium tall herbaceous vegetation		HER/LHE or HER/CHE or HER/THE or HER/HCH or HER/GEO or URB/GRA	
01	<b>A1.A4.A11.B3.A12.B14</b> Open medium to high shrubs		TRS/TPH or TRS/MPH or TRS/DCH or TRS/SCH or TRS/LPH or URB/TRE or URB/VEG	

**Table 3** continued

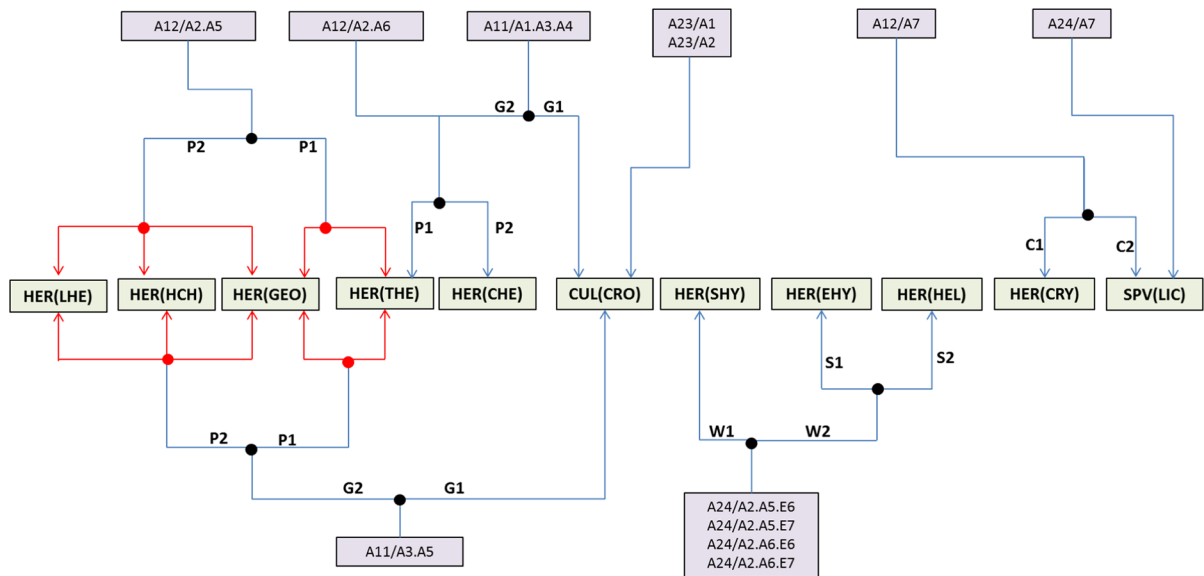
Presence (1)/ absence (0) in each site	LCCS dich. code	LCCS hierarchical code and class description	GHC output categories
10	<b>A24</b>	<b>A2.A5.A13.B4.C2.E5.A8.B13.E7</b>	URB/GRA or HER/EHY or HER/SHY/FLO or HER/HEL or HER/SHY/LEA
10	Natural and semi-natural aquatic or regularly flooded vegetation	Open annual short herbaceous vegetation on temporarily flooded land	
10			
11		<b>A1.A4.A12.B3.C2.D3.B10</b> Aphyllous closed dwarf shrubs on temporarily flooded land	URB/VEG or TRS/SCH/NLE or TRS/LPH/ NLE or TRS/DCH/NLE
10		<b>A2.A6.A12.B4.C2.E5.B12.E6</b> Perennial closed medium-tall grasslands on temporarily flooded land	URB/GRA or HER/EHY or HER/SHY or HER/HEL
11		<b>A2.A5.A16.B4.C1.E5.A15.B12.E6</b> Perennial sparse medium tall herbaceous vegetation on permanently flooded land	URB/GRA or HER/EHY or HER/SHY/FLO or HER/HEL or HER/SHY/LEA
10		<b>A2.A6.A12.B4.C2.E5.B11.E6</b>	URB/GRA or HER/EHY or HER/SHY or HER/HEL
10	Perennial closed tall grasslands on temporarily flooded land		
10			
10			
01		<b>A2.A6.A12.B4.C1.E5.B11.E6</b> Perennial closed tall grassland on permanently flooded land	HER/HEL or HER/EHY or URB/GRA
01		<b>A2.A5.A16.B4.C1</b> Sparse forbs on permanently flooded land	HER/SHY or HER/HEL or HER/EHY or HER/SHY/FLO or URB/GRA
01		<b>A2.A12.B4.C2.E5</b> Mixed closed herbaceous vegetation on temporarily flooded land	HER/EHY or HER/SHY or HER/HEL or HER/SHY/FLO or URB/GRA
10	<b>B15</b>	<b>A1.A3.A7.A8</b>	URB/ART/ROA
	Artificial surfaces	Paved road(s)	
11		<b>A1.A4.A13.A17</b> Scattered urban areas	URB/NON
01		<b>A1.A4.A13.A16</b> Low density urban areas	URB/ART
01		<b>A1.A4.A12.A17</b> Scattered industrial or other areas	URB/NON
01	<b>B16</b>	<b>A2.A6</b>	TER/SAN or TER/GRV or TER/STO
	Bare areas	Loose and shifting sands	

The LCCS code components used for the identification of GHC are evidenced in *bold*

The vernacular name of each GHC category is reported in Table 1

these figures, the LCCS input classes satisfying the same rules are listed in grey rectangular boxes within the diagram of each figure. Such grey boxes are in all figures indicating that LCCS to GHC mapping involve not only *one (LCCS)-to-many (GHC)* relations (only one LCCS class in the grey box) but also *many (LCCS)-*

*to-many (GHC)* relations. To identify the final GHC class among the many GHC outputs (with these evidenced in the central line of each figure) corresponding to the same input LCCS class (or group of LCCS input classes when *many-to-many* LCCS to GHC relations hold), black nodes are introduced in the



**Fig. 5** Herbaceous categories disambiguation. Red links evidence ambiguities still not resolved without in-field data

diagrams to represent alternative decision paths, based on the satisfaction of the rule conditions (see last column of Table 3). The grey arrows represent well defined disambiguation rules, whereas red arrows indicate the presence of ambiguities not resolved with the remote sensing techniques alone.

#### Methodology processing chain

The output GHC categories are identified according to the following sequence:

- (1) Firstly, Artificial Urban (ART) and Urban Non-Vegetated (NON) categories are identified. This step just consists of re-labelling the LCCS objects (classes) originally labelled as buildings, roads, car parks, etc., since these classes correspond to the semantic description of ART category. All non-vegetated and non-artificial objects within an urban area in the LCCS map (identified with the first level LCCS classifiers with codes B15, B16 and B27), including urban water bodies, are re-labelled as NON (Bunce et al. 2011).
- (2) Then, GHC Urban vegetated categories are identified as correspondent to the vegetated classes (objects) in the input LCCS map *adjacent to* ART or NON objects for more than 50 % of their boundaries. In Fig. 3a, the discrimination

between GHC Urban and the remaining GHC vegetated classes (i.e., both natural and cultivated) is shown. The discrimination among the vegetated URB sub-categories, which include Urban vegetables (VEG), Urban herbaceous (GRA) and (TRE) categories, is then based on plant height measurements from either LIDAR data or texture-based information rules reported in Table 3 (last four lines of the 5th column). Fig. 3b provides more details.

- (3) Plant height information extracted from LIDAR data, see Fig. 4a, or alternately from first order entropy texture measurements, see Fig. 4b, is engaged.
- (4) Within the TRS category, Coniferous (CON), Deciduous (DEC), Evergreen (EVR), Non-Leafy Evergreen (NLE) and Summer Deciduous (SUM) are finally discriminated within the output layers of the previous step 3), mainly based on phenological class properties exploiting spectral indices (i.e., P1 and P2 in Table 3) extracted from multi-temporal EO images.
- (5) Finally, Herbaceous (HER), Herbaceous Wetland (HER), Cultivated (CUL) and Sparse Vegetation (SPV) categories are obtained from the remaining vegetated LCCS objects, according to geometric features related to the shape of fields and temporal relations (Fig. 5). In particular, for



Wetland Herbaceous categories (HER), expert knowledge based on spatial topological rules related to habitats spatial pattern has been used. It is well known that in coastal and wetland environments plant communities tend to reside in more or less regular belts, with specific zonation and that water regime is a major determinant of plant zonation patterns in wetlands. Water regime can be described by depth, duration, frequency, timing of flooded and dry phases in a wetland (Casanova and Brock 2000). Studying the two wetland sites, the following rule was identified: (i) submerged hydrophytes (SHY: plants that grow in aquatic conditions with the whole plant in water) are located in the inner part of a lagoon, flooded during the whole year; (ii) emergent hydrophytes (EHY: plants that grow in aquatic conditions and have emergent shoots out of the water) are arranged in an external (and adjacent) belt, flooded during part of the year; and (iii) helophytes (HEL: helophytes, plants that grow in waterlogged conditions) tend to form a further external belt, on soils flooded for short periods or only occasionally and waterlogged in the remaining part of the year. A schematic representation of the zonation rule would appear as: SHY → EHY → HEL.

As an example, the LCCS code “A24.A2.A6.E6” will give as output the GHC class HEL if the considered object simultaneously satisfies the following conditions (see Fig. 5): (i) Adjacency to ART or NON: “less than 50 %”; (ii) BNR of the summer image: “less than 1”; (iii) adjacency to SHY class: “0 %”.

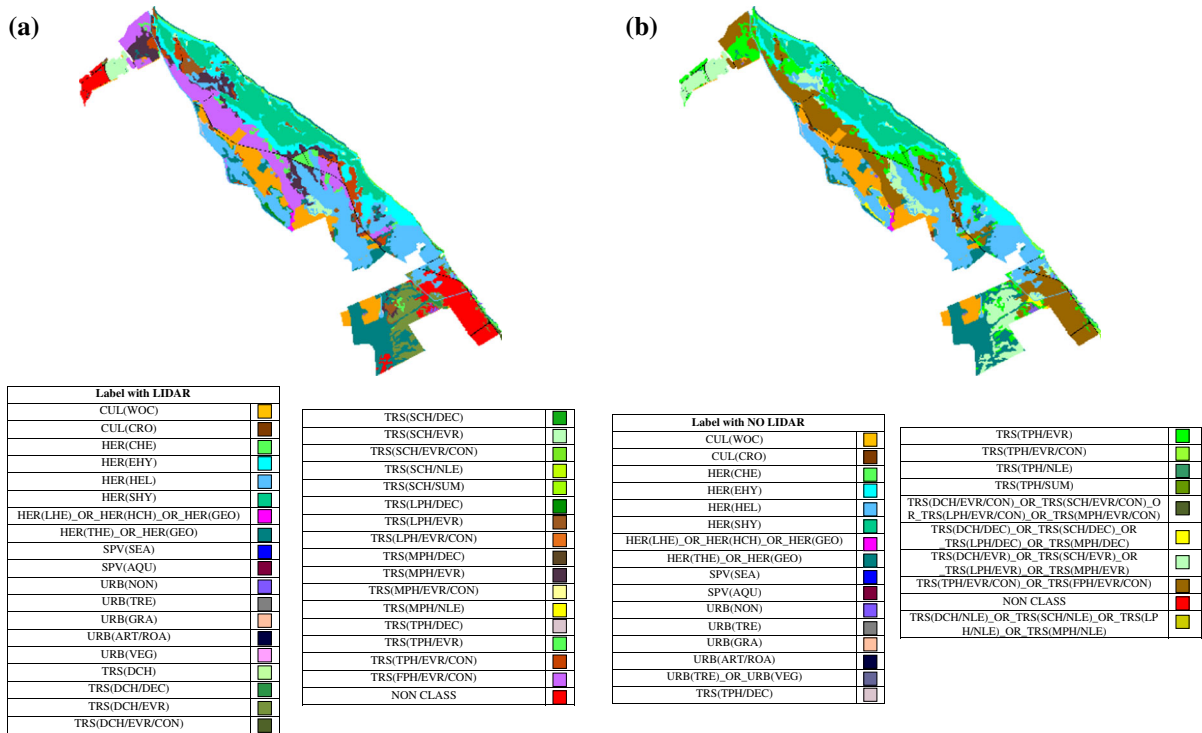
## Results and discussion

The lists of LCCS classes characterizing the two study sites are reported in Table 3 for both Le Cesine site and Lago Salso site. For each LCCS class, the first column of Table 3 includes a binary code of two elements indicating whether the class is present (1) or absent (0) in Le Cesine (first digit) and Lago Salso (second digit) sites; the second and third columns report the dichotomous (e.g. A11, A12, A24, B15, B16) and hierarchical LCCS code components; the last column includes possible GHC output classes

corresponding to the same input LCCS class according to the framework introduced by Kosmidou et al. (2014). As evident in Table 3, the previously proposed framework results in many *one-to-many* and *many-to-many* relations in LC/LU to habitat translation to be resolved through the proposed rules of Table 2.

Figures 6 and 7 show the output GHC maps obtained from the pre-existing input LCCS maps applying the disambiguation rules in Table 2, both with (6a and 7a) and without (6b and 7b) LIDAR data, for Le Cesine and Lago Salso sites, respectively. For Le Cesine site, when comparing Fig. 6a with Fig. 8, which represents the output when applying the generic LCCS to GHC translation framework described in Kosmidou et al. (2014), it becomes evident that the majority of the ambiguity is resolved. Most of the resulting classes after applying the approach proposed in this paper are single GHC classes: an adequate number of multiple classes in Fig. 8 were resolved in Fig. 6a to single ones with the use of the additional features (spatial relations and temporal relations, context features and geometric properties) and expert translation rules embodied in the approach. This constitutes the main contribution of this study to the advancement of automatic complex translation between legends and domains. As an example, LCCS class A12.A1.A4.A10.B3.D1.E2\_B9 in Le Cesine site, representing a “broadleaved deciduous medium–high closed shrubland,” may correspond to either MPH/DEC, TPH/DEC (deciduous medium or tall phanerophytes, respectively) or TRE (urban trees), as seen in Table 2 and in Fig. 8. Inclusion of the ‘*adjacency to ART or NON urban categories*’ rule helped in discriminating the urban TRE classes from the natural vegetation, i.e. MPH/DEC and TPH/DEC. Information on height provided by the LIDAR data made possible the further discrimination between the MPH/DEC and TPH/DEC categories, being 0.6–2 m and 2–5 high, respectively. As a note, the ‘*adjacency to ART or NON*’ rule intensely enhanced the class discrimination task, by particularly resolving ambiguities between artificial or managed and natural vegetated areas.

To validate the maps, GHC reference samples were selected through visual inspection of the multi-seasonal VHR images and in situ campaigns. The samples were collected only in no-changed areas, with these identified as the ones characterized by the same LCCS class both on the ground at the date of survey, in



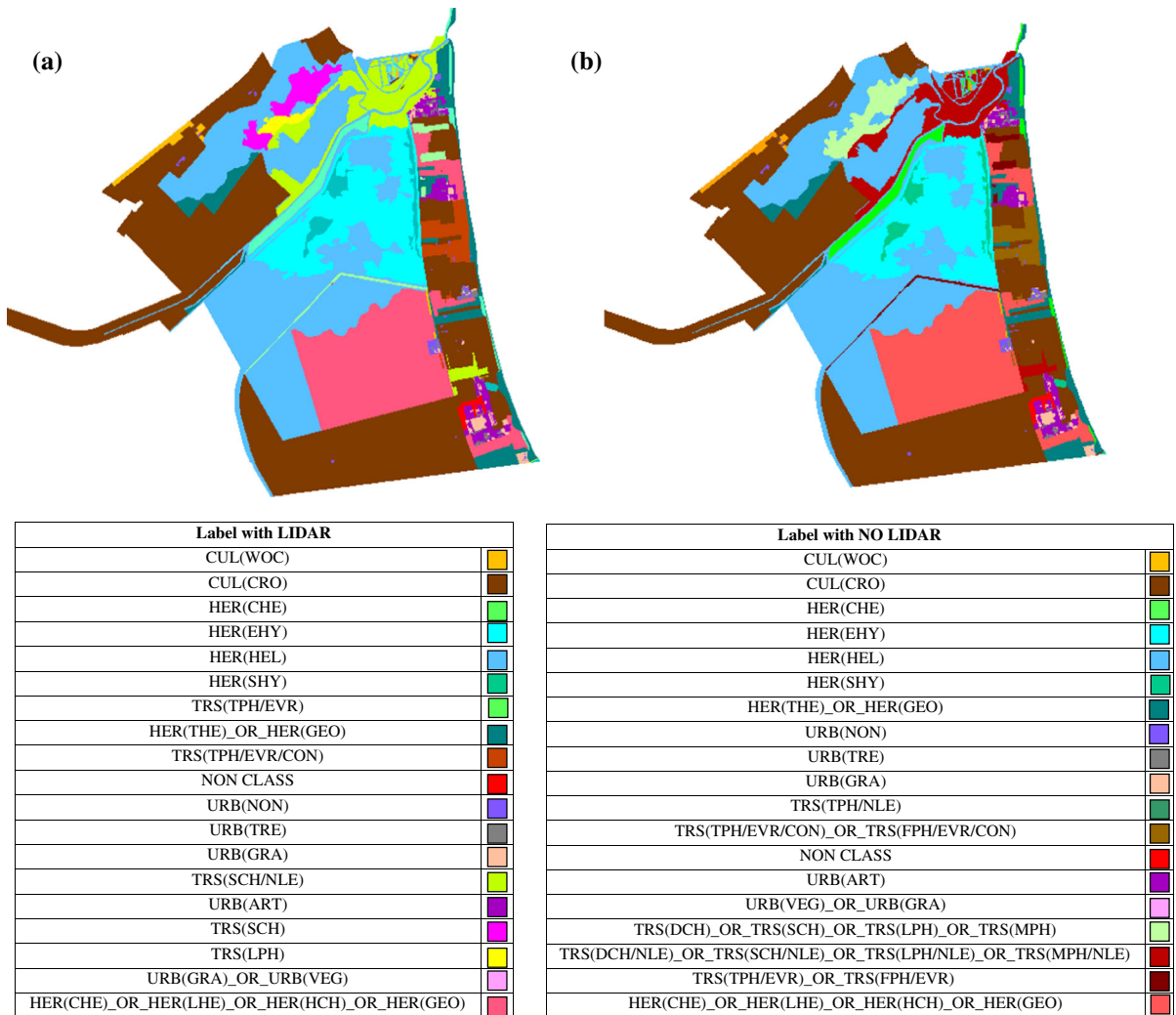
**Fig. 6** a GHC map produced for Le Cesine site: with the use of LIDAR data. b. GHC map produced for Le Cesine site: with the use of texture feature

the pre-existing LCCS map, and in the images used as input to the discrimination procedure. The reason was to avoid classification errors due to differences in the label of reference samples at different dates. For each site, classification confusion matrices were produced for the GHC maps obtained with and without LIDAR. The columns in the matrices represent ground truth classes and the rows the classification results.

Table 4 shows the confusion matrix obtained with LIDAR data for Le Cesine. All results are affected by the quality of the input LCCS map and also by the error in the LIDAR data. In addition, due to the object based classification approach, when trees and shrubs objects (TRS) are considered, the average plant height value from LIDAR includes also the height information of all background vegetated pixels (if any) contained in the object. This might generate an underestimation of the average plants height and, consequently, an incorrect discrimination of the TRS super-category into the sub-categories (e.g., LPH, MPH) associated to different height ranges. As Table 4 reports, an Overall Classification (OA) accuracy equal to 69.9 % with an error of

5.3 % is achieved. An additional problem might be related to the fact that, for HER (herbaceous wetland) object identification, the adjacency rule is effective when habitats are in an ideal conservation status; if such condition is not verified misclassifications may occur. In fact, relationships between flooding duration and plant distribution only partially explain the plant community zonation, because other micro-ecological factors may influence the vegetation pattern. Above all, unwise human activities usually produce habitat reduction, fragmentation and isolation, with deep changes in the spatial pattern of plant communities.

The results highlight how a set of rules, including spatial topological and temporal rules, can disambiguate one-to-many and many-to-many LC/LU to GHC relationships. Concerning wetland spatial patterns, the spatial topological rules used to resolve ambiguities, even though resulting from a generalization of some main ecological gradients, should be carefully evaluated for each site in analysis, in consideration of site-specific ecological conditions and processes. As an example, the salinization occurring in Le Cesine site in

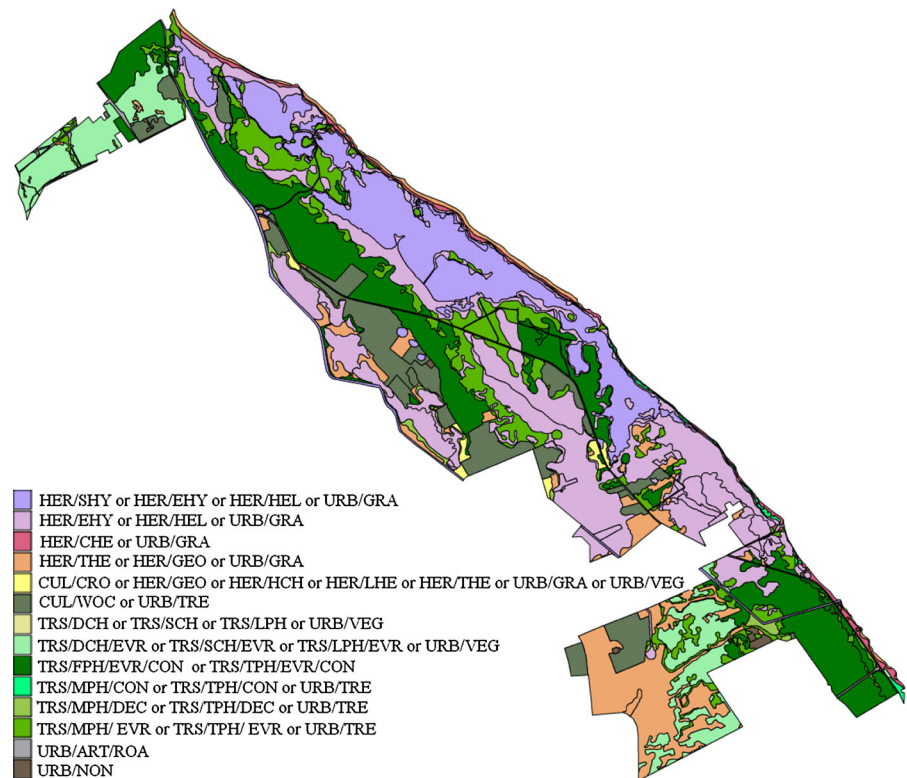


**Fig. 7** **a** GHC map produced for Lago Salso site: with the use of LIDAR data. **b** GHC map produced for Lago Salso site: with the use of texture feature

the last few years is determining a modification of plant communities and their spatial pattern, leading to an alteration of the spatial rules described above. So, new rules should be applied where the spatial pattern has been altered. In fact, plant communities of salt marshes are spatially arranged according to flooding period, soil salinity and moisture (Rogel et al. 2000; Molina et al. 2003). Also in this case, it is possible to generalize: submerged hydrophytes (SHY) in the inner part of the lagoon; therophytes (THE: annual plants; in this case pioneer communities rich in glasswort and other succulent plants of salt marshes, usually appearing during the late summer) forming narrow stripes or belts surrounding the water bodies and only in the

summer period; and shrubby chamaephytes (SCH; in this case, succulent dwarf shrubs typical of salt marshes) arranged in external belts. Then, we could schematize such a zonation: SHY → THE → SCH. Future researches will test these rules in different contexts, in order to verify their effectiveness. Concerning temporal relations, expert knowledge on class water coverage seasonality was used and combined with topological information to solve some many-to-many relations evidenced in Fig. 5. As an example, the four LC/LU aquatic classes (A24) grouped in the second pink box of last line in Fig. 5 (i.e., A24/A2.A6.E6, A24/A2.A6.E7, A24/A2.A5.A6 and A24/A2.A5.E7) can correspond to three possible GHC

**Fig. 8** The GHC map deduced from the pre-existing LCCS map, available for Le Cesine site, without the disambiguation rules proposed in the present paper and published as supplementary material by Kosmidou et al. (2014). The output GHC classes include many ambiguities (multiple GHC classes)



depending on the presence of water, which was measured through spectral indices, in the two bi-seasonal satellite images considered, and also the adjacency to SHY category. According to the rules and legends in Table 3, the final GHC class can be: (i) SHY, if W1 holds or (ii) EHY, if W2 and S1 hold or (iii) HEL, if W2 and S2 hold. Temporal phenological information was used to solve discrepancy in LCCS and GHC terminology. Even though LCCS uses codes to identify natural woody trees broadleaved as evergreen (E1) and deciduous (E2) (i.e., full codes A12/A1.A3.D1 and A12/A1.A3.E2) it is not possible to identify GHC classes corresponding to summer or winter deciduous. To do that, the spectral Green/Red Ratio (GRR) values from the Post Biomass Peak (PoBP) image are considered according to the rule in Table 3. As evidenced in Figs. 3, 4, and 5 by red lines, there are ambiguities that cannot be resolved by using temporal, spatial and spectral information from EO data. As an example, in Fig. 5, the one-to-many relations corresponding to the association of natural herbaceous forbs (A12/A2.A5) to three GHC classes (i.e., LHE, HCH and GEO) cannot be solved from EO data observations because the discrimination is related

to plant morphological and structural properties. In such case in-field inspection is mandatory for final GHC identification. As an additional example, in Fig. 4b it is not possible to resolve the ambiguities related to accurate plant height measurements, because the texture feature was able to discriminate only two coarse groups of plant height. When height information is measured through texture measurements, only two macro height clusters can be detected representing low and high textured vegetation. This means that it is possible to select a threshold in the texture image histogram for separating two macro clusters. In the analysis, to identify the plant height ranges corresponding to such clusters, the reference samples were grouped in two different ways to validate the GHC map and obtain the OA value of the associated confusion matrix. The first grouping merged all reference samples labelled as DCH, SCH, LPH and MPH (respectively, dwarf and shrubby chamaephytes, low and mid phanerophytes), corresponding to trees and shrubs with height <2 m, into one cluster and TPH, FPH (tall and forest phanerophytes) and GPH (Giant phanerophytes over 40 m) corresponding to trees and shrubs with height >2 m







into a second cluster. With the first grouping, the OA was equal to 63 % with an error of 5.5 %. The second grouping merged only DCH, SCH and LPH sub-categories corresponding to trees and shrubs with height  $\leq 0.5$  m in the first cluster and MPH, TPH, FPH and GPH corresponding to height  $> 0.5$  m in the second cluster. The corresponding OA reached 78.4 % with an error of 5 %, as reported in Table 5. The latter result is comparable with the OA value (79.1 %, with an error of 4.7 %) obtained with LIDAR data when the same grouping of reference sample was adopted. As expected, this result is higher than the one in Table 4 (i.e. 69.9 %) when the full set of plant height values was considered and confirms the feasibility to use texture for height discrimination. Based on the experience gained on Le Cesine site, the translation of LCCS to GHC maps was carried out also for the second site, Lago Salso. The OA value for the second site with LIDAR was equal to 66.1 % with an error of 12.7 %. By grouping the reference samples as in the previous case (i.e.,  $\leq 0.5$  and  $> 0.5$  m), the OA value obtained with texture and LIDAR resulted equal to 81.4 % with error 9.9 % and 77.9 % with error 10.6 %, respectively. However, the error values were higher than the ones of Le Cesine. This is due to the different number of reference samples available for Lago Salso and Le Cesine, corresponding to the 22 % and 10 % of the total number of classified objects, respectively.

## Conclusions

The findings of this paper confirm the feasibility of automatic procedures for translating LC/LU maps into GHC maps and resolve the one-to-many and many-to-many relations identified in a previous paper (Kosmidou et al. 2014), through the combined use of EO data and expert knowledge rules to combine temporal, spatial and geometrical features of both LCCS classes and GHC classes. In addition, ambiguities in plant height definition are directly solved by using LIDAR data. However, when LIDAR is considered to be too costly or is not available at all, entropy texture feature from the image green band can help to separate plant heights into two coarse height groups (i.e.,  $\leq 0.5$  and  $\geq 0.5$  m). As a result, based on expert knowledge, a number of multiple classes can be resolved to single ones and the method allows for the mapping of large

areas without the need of any in-field ground truth for training the classifiers. This constitutes the main contribution of this study to the advancement of automatic complex translation between legends and domains and to the mapping of inaccessible areas where the collection of in-field data (ground truth) is impractical.

The methodology can be applied to any LC/LU validated map in LCCS taxonomy or translated in such taxonomy (e.g., from CORINE). Consequently, GHC maps can be produced also in the past if archive EO multi-temporal images (two seasons) close to the date of LC map validation can be found. To the authors' knowledge, no fine-scale GHC mapping has previously been performed using EO data. So far, GHC maps have been only produced by in-field campaigns with the help of in-field computers to register GHC labels for each patch within a specific  $1 \text{ km} \times 1 \text{ km}$  grid. Therefore, the main contribution of the method described in this paper is the possibility to automatically provide GHC maps from validated VHR LCCS maps and EO data (at 2 m spatial resolution) for areas as large as needed and without the class ambiguities discussed but not resolved in the previous work (Kosmidou et al. 2014). Such extended GHC maps can be used to determine areas to which specific in-field campaigns should concentrate for locating the vegetation plots (Bunce et al. 2011) and collecting in situ data on species and environmental attributes. The in-field campaigns can contribute to the refinement of the remaining LCCS to GHC translation ambiguities, such as the one-to-many relations corresponding to the association of natural herbaceous forbs (A12/A2.A5) to three GHC classes (i.e., LHE, HCH and GEO). These often cannot be solved from EO data observations because their discrimination is related to specific plant morphological and structural properties. The fuzzification of the proposed rules can further improve their generalization (Petrou et al. 2013). Within the BIO\_SOS project ([www.biosos.eu](http://www.biosos.eu)), the proposed rule based disambiguation approach has been applied to other Natura 2000 sites in different areas (e.g. Wales, The Netherlands, Greece).

To conclude, the proposed framework is a very promising tool for the automatic translation of LCCS to GHC maps, offering a service for habitat mapping with consistent cost reduction of in-field campaigns in the age of rapid biodiversity decline. The GHC maps

can be regularly updated to detect habitat changes for providing support to decision makers.

**Acknowledgments** The work presented in this paper was carried out as part of the European Union's Seventh Framework Programme FP7/2007–2013, under grant agreement 263435, project *BIOdiversity Multi-Source Monitoring System: from Space To Species* (BIO\_SOS), coordinated by Palma Blonda, CNR-ISSIA, Bari-Italy (<http://www.biosos.eu>). LIDAR data were provided by Dr. S. Costabile from the Geoportale Nazionale - Ministero dell' Ambiente e della Tutela del Territorio e del Mare. The authors would like also to thank Prof. B. Bunce and other BIO\_SOS colleagues (i.e., M. Borges, R. Jongman, E. Padoa-Schioppa, P. Mairota and J. P. Honrado) for fruitful discussions on GHC.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Ahlqvist O (2004) A parameterized representation of uncertain conceptual spaces. *Transactions in GIS* 8:493–514
- Ahlqvist O (2005) Using uncertain conceptual spaces to translate between land cover categories. *International Journal of Geographical Information Science* 19:831–57
- Ahlqvist O (2008) Extending post classification change detection using semantic similarity metrics to overcome class heterogeneity: a study of 1992 and 2001 National Land Cover Database changes. *Remote Sens Environ* 112(3):1226–1241
- Anys H, Bannari A, He DC, Morin D (1994) Texture analysis for the mapping of urban areas using airborne MEIS-II images. *Proceedings of the First International Airborne Remote Sensing Conference and Exhibition*, vol 3, Strasbourg, pp 231–245
- Bossard M, Feranec J, Otahel J (2000) CORINE land cover technical guide: addendum 2000. European Environment Agency, Copenhagen, Tech Rep 40
- Bunce RGH, Bogers MMB, Roche P, Walczak M, Geijzen-dorffer IR, Jongman RHG (2011) Manual for Habitat Surveillance and Monitoring and Vegetation in Temperate, Mediterranean and desert Biomes. Alterra Report 2154. <http://www.ebone.wur.nl/NR/rdonlyres/DADAAB1E-F07C-4AA3-8621-20548A9B7DE6/135332/report2154.pdf>
- Bunce RGH, Bogers MMB, Evans D, Halada L, Jongman RHG, Mücher CA, Bauch B, de Blust G, Parr TW, Olsvig-Whittaker L (2012a) The significance of habitats as indicators of biodiversity and their links to species. *Ecol Ind* 33:19–25
- Bunce RGH, Bogers MMB, Evans D, Jongman RHG (2012b) Field identification of habitats directive Annex I habitats as a major European biodiversity indicator. *Ecol Ind* 33:105–110
- Bunce RGH, Bogers MMB, Evans D, Jongman RHG (2012c) Rule based system for in situ identification of Annex I habitats. Alterra Report 2276. [http://www.wageningenur.nl/upload\\_mm/a/9/b/6a771b7e-dac6-45f0-9b4f-96933cbf636a\\_Report2276Rulebasedsystemforinsituidentificationof.pdf](http://www.wageningenur.nl/upload_mm/a/9/b/6a771b7e-dac6-45f0-9b4f-96933cbf636a_Report2276Rulebasedsystemforinsituidentificationof.pdf)
- Casanova MT, Brock MA (2000) How do depth, duration and frequency of flooding influence the establishment of wetland plant communities? *Plant Ecol* 147:237–250
- Comber AJ, Law ANR, Lishman JR (2004) Application of knowledge for automated land cover change monitoring. *Int J Remote Sens* 25(16):3177–3192
- Comber A, Fisher P, Wadsworth R (2005) Comparing statistical and semantic approaches for identifying change from land cover datasets. *J Environ Manage* 77:47–55
- Costabile S, Martini S, Petriglia L, Paci M, Lopilato G (2013) Proceedings of The Italian National Geoportal: perspectives and innovations, INSPIRE Conference 2013, Florence, 26 June 2013
- Di Gregorio A & Jansen LJM (2005) Land Cover Classification System (LCCS): classification concepts and user manual. Software version 2. Food and Agriculture Organization of the United Nations, Rome
- Feng CC, Flewelling DM (2004) Assessment of semantic similarity between land use/land cover classification systems. *Comput Environ Urban Syst* 28:229–246
- Fritz S, See L (2008) Identifying and quantifying uncertainty and spatial disagreement in the comparison of global land cover for different applications. *Glob Change Biol* 14(5):1057–1075
- Goodchild MF, Yuan M, Cova TJ (2007) Towards a general theory of geographic representation in GIS. *Int J Geogr Inf Sci* 21:239–260
- Herold M, Schmullius C (2004) Report on the harmonization of global and regional land cover products meeting, Rome, 14–16 July 2004. [http://www.fao.org/gtos/gofcgold/docs/GOLD\\_20.pdf](http://www.fao.org/gtos/gofcgold/docs/GOLD_20.pdf)
- Herold M, Woodcock C, Di Gregorio A, Mayaux P, Belward AS, Latham J, Schmullius CC (2006) A joint initiative for harmonization and validation of land cover datasets. *IEEE Trans Geosci Remote Sens* 44(7):1719–1727
- Herold M, Hubald R, Di Gregorio A (2009) Translating and evaluating the land cover legends using the Land Cover Classification System (LCCS). GOF-C-GOLD Report No. 43. [http://nofc.cfs.nrcan.gc.ca/gofc-gold/Report%20Series/GOLD\\_43.pdf](http://nofc.cfs.nrcan.gc.ca/gofc-gold/Report%20Series/GOLD_43.pdf)
- Kosmidou V, Petrou Z, Bunce RGH, Mücher CA, Jongman RHG, Bogers M, Lucas RM, Tomaselli V, Blonda P, Padoa-Schioppa E, Manakos I, Petrou M (2014) Harmonization of the Land Cover Classification System (LCCS) with the General Habitat Categories (GHC) classification system. *Landscape Ecol* 36:290–300
- Kuechler AW, Zonneveld IS (eds) (1988) *Vegetation mapping. Handbook of vegetation science*, vol 10. Kluwer Academic, Dordrecht
- Ladoux L, Crooks S, Jordan A, Turner RK (2000) Implementing EU biodiversity policy: UK experiences. *Land Use Policy* 17:257–268
- Lengyel S, Kobler S, Kutnar L, Framstad E, Henry PY, Babij V, Gruber B, Schmeller D, Henle K (2008) A review and a

- framework for the integration of biodiversity monitoring at the habitat level. *Biodiv Conserv* 17:3341–3356
- Maes J, Teller A, Erhard M, Lique C, Braat L, Berry P, Egoh B, Puydarrieux P, Fiorina C, Santos F, Paracchini ML, Keune H, Wittmer H, Hauck J, Fiala I, Verburg PH, Condé S, Schägner JP, San Miguel J, Estreguil C, Ostermann O, Barredo JJ, Pereira HM, Stott A, Laporte V, Meiner A, Olah B, Royo Gelabert E, Spyropoulou R, Petersen JE, Maguire C, Zal N, Achilleos E, Rubin A, Ledoux L, Brown C, Raes C, Jacobs S, Vandewalle M, Connor D, Bidoglio G (2013). Mapping and assessment of ecosystems and their services. An analytical framework for ecosystem assessments under action 5 of the EU biodiversity strategy to 2020. 1st technical report. Publications office of the European Union, Luxembourg. doi: [10.2779/12398](https://doi.org/10.2779/12398)
- Mehtälä J, Vuorisalo T (2007) Conservation policy and the EU Habitats Directive: favourable conservation status as a measure of conservation success. *Eur Environ* 17:363–375
- Molina JA, Casermeiro MA, Moreno PS (2003) Vegetation composition and soil salinity in a Spanish Mediterranean coastal ecosystem. *Phytocoenologia* 33(2–3):475–494
- Morris B, Dupigny-Giroux L (2010) Using the Nir/blue surface moisture index to explore feature identification at multiple spatial resolutions. In: abstracts of AGU Fall Meeting, vol 1, pp 1298
- Petrou ZI, Tarantino C, Adamo M, Blonda P, Petrou M (2012) Estimation of vegetation height through satellite image texture analysis. Proceedings of the International archives of the photogrammetry, remote sensing and spatial information sciences, XXII ISPRS Congress, Melbourne, 25 Aug–01 Sept 2012, vol XXXIX-B8, pp 321–326
- Petrou ZI, Kosmidou V, Manakos I, Stathaki T, Adamo M, Tarantino C, Tomaselli V, Blonda P, Petrou M (2013) A rule-based classification methodology to handle uncertainty in habitat mapping employing evidential reasoning and fuzzy logic. *Pattern Recogn Lett*. doi:[10.1016/j.patrec.2013.11.002](https://doi.org/10.1016/j.patrec.2013.11.002)
- Pierkot C, Andrés S, Faure JF, Seyler F (2013) Formalizing spatiotemporal knowledge in remote sensing applications to improve image interpretation. *J Spat Inf Sci* 7:77–98
- Raunkiaer C (1934) The life forms of plants and statistical plant geography, being the collected papers of Raunkiaer C. Clarendon, Oxford
- Ritchie GL, Sullivan DG, Vencill WK, Bednarz CW, Hook JE (2010) Sensitivities of normalized difference vegetation index and a green/red ratio index to cotton ground cover fraction. *Crop Sci* 50:1000–1010
- Rogel JA, Ariza FA, Silla RO (2000) Soil salinity and moisture gradients and plant zonation in Mediterranean salt marshes of Southeast Spain. *Wetlands* 20:357–372
- Tomaselli V, Dimopoulos P, Marangi C, Kallimanis AS, Adamo M, Tarantino C, Panitsa M, Terzi M, Veronico G, Lovregina F, Nagendra H, Lucas R, Mairota P, Múcher S, Blonda P (2013) Translating land cover/land use classifications to habitat taxonomies for landscape monitoring: a mediterranean assessment. *Landscape Ecol* 28(5):905–930
- Townsend PA, Lookingbill TR, Kingdon CC, Gardner RH (2009) Spatial pattern analysis for monitoring protected areas. *Remote Sens Environ* 113:1410–1420
- Trimble (2011) *Ecognition Developer 8.7*. Reference Book. <http://www.ecognition.com/>
- Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M (2003) Remote sensing for biodiversity science and conservation. *Trends Ecol Evol* 18:306–314
- Wood EM, Pidgeon AM, Radeloff VC, Keuler NS (2012) Image texture as a remotely sensed measure of vegetation structure. *Remote Sens Environ* 121:516–526
- Wuest B, Zhang Y (2009) Region based segmentation of QuickBird multispectral imagery through band ratios and fuzzy comparison. *ISPRS Journal of Photogrammetry and Remote Sensing* 64(1):55–64