Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 August 2016 doi:10.20944/preprints201608.0059.v1

Peer-reviewed version available at *Remote Sens.* **2016**, *8*, 925; doi:10.3390/rs8110925

Article

Europe's Green Arteries—A Continental Dataset of Riparian Zones

Christof J. Weissteiner [1,*], Martin Ickerott [2], Hannes Ott [2], Markus Probeck [2], Gernot Ramminger [2], Nicola Clerici [3], Hans Dufourmont [4] and Ana Maria Ribeiro de Sousa [4]

[1] Independent consultant; Via Milite Ignoto 132, I-21027 Ispra (VA), Italy
[2] GAF AG, Arnulfstrasse 199, D-80634 Munich, Germany; martin.ickerott@gaf.de (M.I.); hannes.ott@gaf.de (H.O.); markus.probeck@gaf.de (M.P.); gernot.ramminger@gaf.de (G.R.)
[3] Universidad del Rosario, Functional and Ecosystem Ecology Unit (EFE) Biology Program, Universidad del Rosario Kr 26 No 63B-48, Bogotá D.C., Colombia; nicola.clerici@urosario.edu.co
[4] European Environment Agency (EEA), Kongens Nytorv 6, 1050 København K, Denmark; hans.dufourmont@eea.europa.eu (H.D.); ana.sousa@eea.europa.eu (A.M.R.S.)
* Correspondence: mail@weissteiner.eu; Tel.: +39-349-660-4261

Abstract: Riparian zones represent ecotones between terrestrial and aquatic ecosystems and are of utmost importance to biodiversity and ecosystem functions. Modelling/mapping of these valuable and fragile areas is needed for an improved ecosystem management, based on an accounting of changes and on monitoring of their functioning in time. In Europe, the main legislative driver behind this goal is the European Commission's Biodiversity Strategy to 2020, on one hand aiming at reducing biodiversity loss, on the other hand enhancing ecosystem services by 2020, and restoring them as far as feasible. A model, based on Earth Observation data, including Digital Elevation Models, hydrological, soil, land cover/land use data, and vegetation indices is employed in a multi-modular and stratified approach, based on fuzzy logic and object based image analysis, to delineate potential, observed and actual riparian zones. The approach is designed in an open modular way, allowing future modifications and repeatability. The results represent a first step of a future monitoring and assessment campaign for European riparian zones and their implications on biodiversity and on ecosystem functions and services. Considering the complexity and the enormous extent of the area, covering 39 European countries, including Turkey, the level of detail is unprecedented. Depending on the accounting modus, 0.95%–1.19% of the study area can be attributed as actual riparian area (considering *Strahler's* stream orders 3-8, based on the Copernicus EU-Hydro dataset), corresponding to 55,558–69.128 km^2. Similarly depending on the accounting approach, the potential riparian zones are accounted for about 3-5 times larger. Land cover/land use in detected riparian areas was mainly of semi-natural characteristics, while the potential riparian areas are predominately covered by agriculture, followed by semi-natural and urban areas.

Keywords: riparian zone; transitional environment; riparian forest buffer; spatial modelling; mapping; spatial ecology; ecosystem functions

---

1. Introduction

Riparian ecotones, transition zones between water and land, are providing a variety of important ecosystem functions and services. The range extends from filtering/buffering of sediment and nutrient load to stream bank stabilization, from water storage/release to aquifer recharge, and from habitat provision to recreational and educational opportunities [1–4]. Riparian zones are considered exceptionally rich in biodiversity [5,6] and extremely fragile at the same time [7]. Despite its high ecological value, a

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

large part of the natural riparian vegetation has already been lost, degraded or fragmented due to human activity [8,9].

To counteract further riparian decline, initiatives at several scales have been put in place. At a global scale, the Millennium Ecosystem Assessment requests measures for systematic assessment of riverine habitats [1]. The EU Biodiversity Strategy to 2020 addresses the issue at European policy level, aiming to halt the loss of biodiversity and the degradation of ecosystem services in the EU by 2020, restoring them as far as feasible [10]. The EU Biodiversity strategy's target 2 focus is on a better protection and restoration of ecosystems and the services they provide, and greater use of green infrastructure, *sensu* Benedict and McMahon [11]. As a consequence, Copernicus (previously known as GMES, Global Monitoring for Environment and Security), the European flagship initiative for Earth Observation and Monitoring, is addressing these ecologically important areas. As part of the Copernicus Land Monitoring Service's local component, "riparian zones" have been mapped during the Copernicus Initial Operations 2011-2013 phase on request of the European Environment Agency (EEA). Local component products are designed to provide specific and more detailed information focusing on specific types of hotspots, in this case riparian zones. Moreover, the local component "riparian zones" is expected to support the MAES initiative (Mapping and Assessment of Ecosystems and their Services, [12]) and link to other European policy areas or initiatives such as the Water Framework Directive, the Habitats and Birds Directives with their centrepiece the Natura 2000 network, the Floods Directive and the European Commission's Green Infrastructure strategy [13–18]. Furthermore, the Blueprint to safeguard Europe's waters calls for *"strengthened measures to help the EU protect its water resources and become more resource (including water) efficient"* [19], urging for measures such as the restoration of wetlands and floodplains to increase the take-up of natural water retention.

Lastly, the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES, www.ipbes.net) would certainly benefit from these new datasets of continental extent (including Turkey) in their assessments on the state of biodiversity and of the ecosystem services it provides to society, namely, for the European assessment.

The aim of Copernicus' local component "riparian zones" is to provide information on spatial extent, distribution and land cover/land use characteristics of riparian zones, as to allow for future systematic assessment of freshwater ecosystems and riverine habitats.

Riparian definitions are conceptual and fuzzy [20] and among scientists sometimes controversial. In the present study we consider as riparian zone, in general terms, transitional areas occurring along land and freshwater ecosystems, characterized by unique soil, hydrology and biotic conditions strongly influenced by the stream water [21,22]. The riparian zone encompasses the stream channel between the low and high water marks and that portion of the terrestrial landscape from the high water mark toward the uplands where vegetation may be influenced by elevated water tables or flooding and by the ability of the soils to hold water [23,24].

European riparian zones, following the definition of Naiman et al. [21], have been mapped in the past by the European Commission's Joint Research Centre (JRC) [25,26]. An enhanced version of this data set has been developed and employed for ecosystem service assessments of riparian buffer capacity for European rivers [27,28]. Both, original and refined versions are JRC products and represent the first pan-European maps of riparian areas. At that time, the best available data sources were used to create a consistent and harmonized European product. However, now, as the technological progress continues, several improved base products have become available, which allow the compilation of a more complete, accurate and detailed product. Moreover, there are new requirements based on policy requests, which call for advanced monitoring, such as change analysis of land use/land cover, ecosystem condition and delivery of ecosystem services, including habitat and biodiversity monitoring. A riparian data set of high quality and detail, based on scientifically sound approaches, is needed to satisfy these requests. The demanding requirements add complexity to such an endeavour, but do certainly guarantee a high utility

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

of the product. In terms of area extent, the requested coverage is bound to the 33 member and the 6 cooperating countries of the EEA. In terms of scale, the product is of high detail. Based on multi-resolution and multi-source satellite imagery, the Minimum Mapping Unit (MMU) is 0.5 ha.

The goal of this work is to design a riparian zones delineation model of high scientific value, being consistent, transparent, open for further input and repeatable in time, and at the same time serving a multitude of environmental needs.

Apart from the JRC riparian zones map [25–28], another similar work has been conducted by the Centre for International Forestry Research (CIFOR) [29], focusing on global tropical wetlands. On a local or regional extent, riparian zones modelling and mapping has been carried out frequently, with a variety of approaches relying on all kind of Earth Observation (EO) data of different scale [30–38]. In all cases, the technique had been based on remote sensing data, often relying on a geomorphic approach, considering the topography, and/or relying on the proximity to water and/or the identification of riparian vegetation and features. Scientific evidence of the drivers determining the width of the riparian zone is reported by Naiman and Décamps [22], which are in general related to the size of the stream, the position of the stream within the drainage network, the hydrological regime, and the local geomorphology. The employed data sources for regional/local analyses are often of high spatial detail, sometimes based on LIDAR (Light Detection and Ranging) data [39]. To date, for a continental endeavour of riparian zone modelling and mapping such spatially very detailed data are not yet available. However, this might change in future.

The general approach applied in this study is modular, whereby each module can be divided into several sub-modules. This way, the approach remains open for modifications, extensions, reductions or exchanges of modules or sub-modules. The design also ensures that riparian zones can still be modelled and mapped in case one or some input data sets are missing in a specific region. The main modules are organised such as to allow consecutive creation of "Potential Riparian Zones" (PRZ), mapping areas with a natural, physio-geographic disposition to host riparian zones; "Observable Riparian Zones" (ORZ) delineating effectively observed riparian zones; and "Actual Riparian Zones" (ARZ), representing the intersection of PRZ and ORZ.

PRZ are modelled based on the input features river network, terrain topography, soil properties, flood zones, modelled topographic wetness and the land cover/land use (LCLU) class "Water". Following a fuzzy logic approach [40], and an object based image analysis approach (OBIA)[41], PRZ-specific membership (MS) values are assigned to segmented objects through a combination of feature-specific membership functions (MSFs). Regionalization/stratification is applied by assigning within each ECRINS river basin (European catchments and Rivers network system, see Table 1) individual regional calibration factors when combining the input feature data.

The ORZ delineation is mainly relying on satellite observations. The approach is analogous to the one of PRZ: MS values with respect to the input features LCLU, Normalized Difference Vegetation Index (NDVI)[42], and Normalized Difference Water Index (NDWI)[43] are combined.

PRZ and ORZ MSs are finally combined to the ARZ MS, expressing a probability to encounter riparian zones on ground. Eventually, applying a hard threshold to the raster based MS value, a vector-based delineation of riparian areas can be derived.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 August 2016 doi:10.20944/preprints201608.0059.v1

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

2. Materials and Methods

Study site

Greater Europe is characterized by a high variety of bio-climatological regions or eco-regions [44,45]. This variety is even more pronounced if geological or soil data is being considered. For example, Europe's most northern parts are considered arctic and boreal which contrast the Mediterranean or Turkey's Anatolian regions. The extent of the study area ranges from 71.2º N to 34.6º S, and from 24.6º W to 44.9º E, comprising 33 member countries of the EEA and 6 cooperating countries (see Figure 1). Besides all 28 EU member countries, also Iceland, Norway, Switzerland, Liechtenstein, Turkey and Albania, Bosnia and Herzegovina, Kosovo under the UN Security Council Resolution 1244/99, Former Yugoslav Republic of Macedonia, Montenegro and Serbia are covered. In addition, the small countries Andorra, San Marino and Vatican City are included.

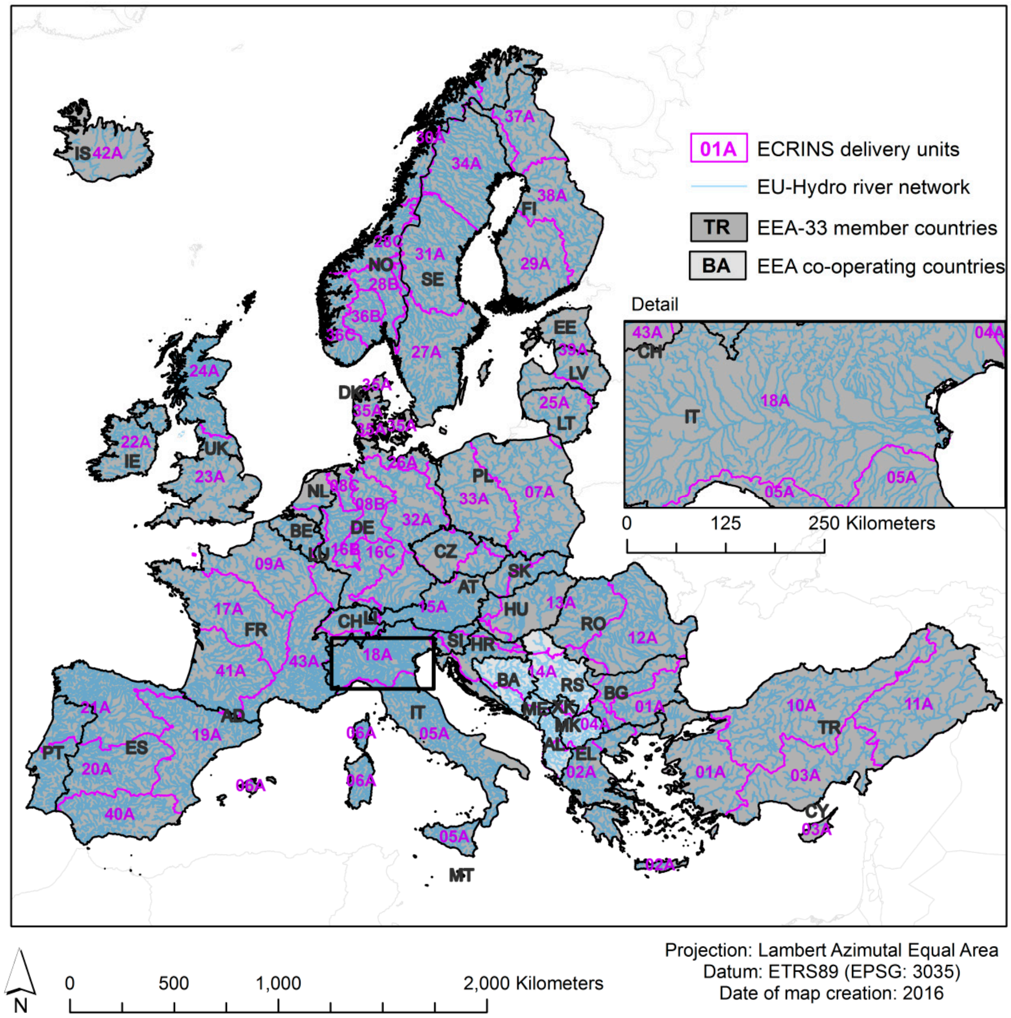

Figure 1: Full study area divided into countries (black polygons and labels) and into hydrologically based Delivery Units (DUs) derived from ECRINS (pink). Country labels are listed in Table 3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 August 2016 doi:10.20944/preprints201608.0059.v1

Peer-reviewed version available at *Remote Sens.* **2016**, *8*, 925; doi:10.3390/rs8110925

The area of interest (AOI) of this study comprises a generously supposed area of water influence around rivers and lakes: for a first delineation, based on the hydrological dataset EU-HYDRO, rivers of *Strahler's* stream order 3-8 [46] were selected, dynamic buffers built around them and then combined with a flood hazard map which delineates the 100-year flood return period [47]. The preliminary AOI extents to approx. 500,000 km^2 and is drained by a total river length of approximately 470,000 km. The river network comprises also lentic water bodies if one of the abovementioned selected rivers runs through them. This preliminary relatively coarse AOI has subsequently been expanded by approximately 10%, through adding additional, modelled PRZ area, as well as several additional relevant riparian areas and features (e.g. oxbow lakes of the relevant river systems) as visually identified in the Copernicus riparian zones LCLU dataset. Thus, the AOI comprises a total of 556,658 km^2.

Data

The riparian zones delineation model relies on data sets of different sources and scales, described hereafter and presented in Table 1. The multi-scale approach makes it almost impossible to assign a common scale to the final product. However, to get a better understanding of this product's scale, it might help to know some core characteristics of two crucial input datasets: the LCLU data set and the Digital Elevation Model (DEM). The Minimum Mapping Unit of the used LCLU data set is 0.5 ha, with an original spatial resolution of the underlying EO-input data ranging between 2 m (Pleiades) and 30 m (Landsat 8), whereas the applied EU-DEM has a spatial resolution of 25 m. Typically, the scales for these raster resolutions lie around 1:5,000 (for 2 m) and 1:75,000 (for 30 m).

Data are generally referring to the reference year 2012, with some exceptions or deviations, which in the case of quasi-static data (e.g. soil data) should not affect the result. All data are provided compliant to the provisions of INSPIRE [48] and projected as ETRS89 Lambert Azimuthal Equal Area (LAEA) projection, conforming to EPSG 3035.

doi:10.20944/preprints201608.0[illegible]v1

Table 1: Input data sets and sources with their main characteristics

Data category	Data set	Source	Data specification/remarks	Spatial resolution	Extent	Acquisition date
Optical remote sensing data	Landsat8	USGS	ToA reflectance, cloud masking,	30 m	AOI	2013-2014
	IRS Liss III	ESA	NDVI,NDWI	23.5 m	AOI	2011-2013
	Pleiades, SPOT5/6, RapidEye	ESA	Used for LCLU classification	1.5 -5 m	AOI	2010-2014
Digital Elevation Model (DEM)	EU-DEM	EEA (http://land.copernicus.eu/pan-european/satellite-derived-products/view) Statistical validation: http://land.copernicus.eu/user-corner/technical-library/eu-dem-2013-report-on-the-results-of-the-statistical-validation	Altitude	25 m	AOI	2000
	GSD elevation data, (Digital Surface Model)	National Land Survey of Sweden (https://www.lantmateriet.se/en/Maps-and-geographic-information/Maps/oppna-data/hamta-oppna-geodata/)	Altitude	50 m	Sweden	2009-2016
	Digital Terrain Mode	The Norwegian Mapping Authority (http://data.kartverket.no/download/content/digital-terrengmodell-10-m-utm-33)		10 m	Norway	2008
	Digital Surface model	National Land Survey of Finland (http://www.maanmittauslaitos.fi/en/digituotteet/elevation-model-10-m)		10 m	Finland	2001-2014
Hydrology	EU-Hydro: River network, inland water	http://land.copernicus.eu/user-corner/publications/eu-hydro-flyer/view http://land.copernicus.eu/pan-european/satellite-derived-products/view	River and waterbody delineation	1:30,000 – 1:50,000	AOI	2006
	ECRINS v1: European catchments and Rivers network system	http://www.eea.europa.eu/data-and-maps/data/european-catchments-and-rivers-network	Catchment delineation	1:250,000	AOI	1990-2006
	European Flood hazard map	JRC [47]	100-year flood return period delineation	100 m	AOI	1990-2010
	Open Street Map: Water	http://openstreetmapdata.com/data/water-polygons	Water body delination	Not defined, varying	AOI	
	HRL permanent water bodies	http://land.copernicus.eu/pan-european/high-resolution-layers/permanent-water-bodies/view	Water delineation	20 m	AOI	2012
Soil data	Harmonized World Soil Data Base, v1.2	FAO and IIASA http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/index.html?sb=1	Various soil properties, see Table 2.	1:1M for Europe; 1 km grid size	AOI	2006
LC/LU data sets	HRL forest layer	http://land.copernicus.eu/pan-european/high-resolution-layers/forests/view	Tree cover density and/or forest type	20 m	AOI	2012
	CORINE land use/land cover (CLC)		LCLU classes	1:100,000	AOI	2006, 2012
	HRL wetlands	http://land.copernicus.eu/pan-european/high-resolution-layers/wetlands/view	Wet areas	20 m	AOI	2012
	RAMSAR sites	http://www.ramsar.org/	Wetlands			
	Natura 2000 sites	http://www.eea.europa.eu/data-and-maps/data/natura-1	Protected areas	1:100,000	AOI	
	Urban Atlas: Land use and land cover data for Large Urban Zones of Europe	http://www.eea.europa.eu/data-and-maps/data/urban-atlas	Urban area delineation	1:10,000	AOI	2006, 2012
	HRL imperviousness	http://land.copernicus.eu/pan-european/high-resolution-layers/imperviousness/view	Imperviousness	20 m	AOI	2012
Other	EEA-39 borders and coastline	EEA	National borders and coastline		AOI	

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

Optical remote sensing data of various sensors and resolution form the backbone of the underlying work (Landsat 8 and LISS III for riparian mapping, and SPOT 5/6 and Pleiades for LCLU classification). Top of atmosphere reflectance (ToA) in the blue, green, red, infrared and shortwave infrared range was calculated and used, after applying cloud masking.

The EU-DEM is a digital surface model (DSM), which is a fusion of the Shuttle Radar Topography Mission (SRTM) and ASTER GDEM data. Its accuracy showed to be lower north of 60°N, which can be explained with the absence of SRTM in that region. In particular, areas of flat topography were found to be heavily affected by bad data values. For areas of Norway, Sweden and Finland north of 60°N, the EU-DEM was therefore, in this study, substituted by freely available national DEMs being resampled to the resolution of the EU-DEM. The few water bodies modelled in Iceland showed to be situated in regions of steep topography. Therefore, the EU-DEM was deemed to be of sufficient quality.

The DSM-character of the EU-DEM has obviously an impact on the riparian area modelling approach in forested and settlement areas. The intrinsic height difference between the DSM surface (i.e. forest canopy, top of buildings) and the terrain would cause significant regional reductions of riparian zones probability and extent, since such "artificially elevated" areas often would act as barriers. To reduce this distortion, an adjustment method has been developed and applied in forested areas. The altitude values of forest areas (based on HRL-Forest) larger than 1 km^2 and within 5 km of the river were masked and interpolated with the surrounding altitude values, rendering the DEM in those areas similar to a DTM, for the purpose of this study.

EU-Hydro is a pan-European river network and water bodies dataset based on Image 2006 [49], which is an EO based data collection with two nearly cloud-free coverages for the EEA-39 countries (20 m spatial resolution). Locations of river courses are therefore highly precise, compared to river networks based on DEM based river extraction modelling approaches. The data set contains rivers, lakes and other hydrological elements as lines and polygons.

Riparian zones are often located in flood zones. A pan-European flood hazard map (FHM) of 100 m spatial resolution, based on a combination of distributed hydrological and hydraulic models has been recently compiled [50]. An updated version of the maximum spatial extents of flood return periods of 20, 50, and 100 years was on purpose produced by the JRC. The FHM is based on the hydrologic LISFLOOD model [51] coupled with a hydraulic model and run as multi-scale process. A discharge model and an 21-year observed meteorological data set are adopted to generate different return periods of flood peaks, which are used for local hydraulic simulations. Note, that the employed DEM is the SRTM data set [52] with 3 arc seconds (approx. 90 m) original resolution, while as river network the CCM2 data set [53] was used. Both data sets are not fully congruent with the here employed data sets but were considered the best available option at this moment. Moreover, the generation of FHM has been restricted to catchments of more than 500 km^2.

The Harmonized World Soil Data Base (HWSD) [54] is a global soil database, composed of different sources, including the European Soil Data Base (1:1,000,000) [55]. The data set consists of raster data of about 1 km grid size (30 arc seconds), aggregated to soil units. The advantage of the HWSD is that the data are organized in a way, that soil attributes are available for all soil types contained within a soil unit, not just for the dominant one. This way, all associated soils of a soil unit can be considered, weighting them by their relative area share via a database operation. The relevant parameters or indicators chosen to model riparian zones are listed in Table 2.

Table 2: Indicators used and MS function applied

Domain	Indicator	Source	Indicator value	MS assigned	MSF type
Geo-morphology	Path Distance (PD)	DEMs		1 at 0, 0 at >=1,000	Sigmoid
	Saga Wetness Index (SagaWI)	DEMs		1 at >=15, 0 at 0	Sigmoid
	Flood Hazard Map (FHM)	JRC	20 yr; 50 yr; 100 yr; >100 yr return period	1; 1; 1 ;0	Discrete
Soil	Soil type	HWSD/ ESDB	Fluvisols (FL); Gleysols (GL); Histosols (HS); all other	1; 0.8; 0.7; 0.1	Discrete
	Available water storage capacity (AWC) [mm/m]	HWSD	150 mm/m; 125; 100; 75; 50; 15; 0	1; 0.83; 0.67; 0.5; 0.33; 0.1; 0	Discrete
	Obstacle to roots between the given depth range [cm]	HWSD	No obstacles 0-80 cm; obstacles 60-80 cm; 40-60 cm; 20-40 cm; 0-80 cm; 0-20 cm	1; 0.8; 0.6; 0.4; 0.2; 0	Discrete
	Impermeable layer (IL) between the given depth [cm]	HWSD	No impermeable within 150 cm; impermeable 80-150 cm; 40-80 cm; <40 cm	1; 0.77; 0.4; 0.27	Discrete
	Soil water regime (WR)	HWSD	Not wet within 80 cm for over 3 months, nor wet within 40 cm for over 1 month; wet within 80 cm for 3 to 6 months, but not wet within 40 cm for over 1 month; wet within 80 cm over 6 months, but not wet within 40 cm for over 11 month; wet within 40 cm depth for over 11 month	0.1; 0.3; 0.6; 1	Discrete
	Top soil and subsoil organic carbon content [%]	HWSD	< 0.2%; 0.2-0.6%; 0.6-1.2%; 1.2-2.0%; >2.0%	0.1; 0.2; 0.45; 0.8; 1	Discrete
Vegetation	NDVI	High resolution EO-data		0 at <=0.2, 1 at >=0.55	Sigmoid
	NDWI			0 at <=0.2, 1 at >=0.35	Sigmoid

Land cover/land use (LCLU) data within the study area had been produced in parallel to this study in very high resolution (with 0.5 ha MMU), by the Copernicus Riparian Zones project, following the MAES ecosystem types and specific nomenclature guidelines [56]. This dataset had been produced in a complex visual delineation and interpretation process, making synergistic use of a variety of EO and in-situ data, such as EU-Hydro, Open Street Map (OSM) and Urban Atlas (UrbA) data, CORINE land cover (CLC) 2006/2012, national LCLU classifications, the High Resolution Layers (HRL) "Imperviousness", "Forest", "Water" and "Wetlands", RAMSAR sites, NATURA 2000, EU-DEM and the Potential Riparian Zones. More details and sources of the mentioned data sets are reported in Table 1. Furthermore, Landsat time series were considered for differentiation of irrigated vs. rain fed cropland or for assignment of complex classes such as managed grasslands and other croplands. Urban and forest classes were differentiated according to density of imperviousness and tree cover. The full range of LCLU classes is reported in Table S1. The thematic accuracy of the LCLU product has been validated so far in 23 out of 43 Delivery Units (DU), covering 52% of the area but the full variety of bio-geographic regions, achieving an overall thematic accuracy of 85% with values ranging from 77% to 94% [57].

Model setup

Riparian area modelling of this extent and scale requires a rich data pool and a model with a sufficient degree of detail. The model was on the one hand required to be able of coping with the variety of

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

conditions encountered and on the other hand to be not overly complex in order to allow operational product generation. This requires the model to run in a robust manner even with incomplete data coverage. Therefore, it was designed in such a way to allow addition or removal of data sets, according to availability or appropriateness. Spatially independent Delivery Units (DU), based on the ECRINS river basins, enable to run the model on subsets, enabling easier data handling and, in addition, local calibration.

Figure 2 depicts the organization of the model and the process flows. Overall, three output layers are created for riparian area delineation: The Potential Riparian Zone (PRZ), the Observable Riparian Zone (ORZ) and the Actual Riparian Zone (ARZ). The key input variables and indicators, the strategy to combine them, and the output layers are explained hereafter.

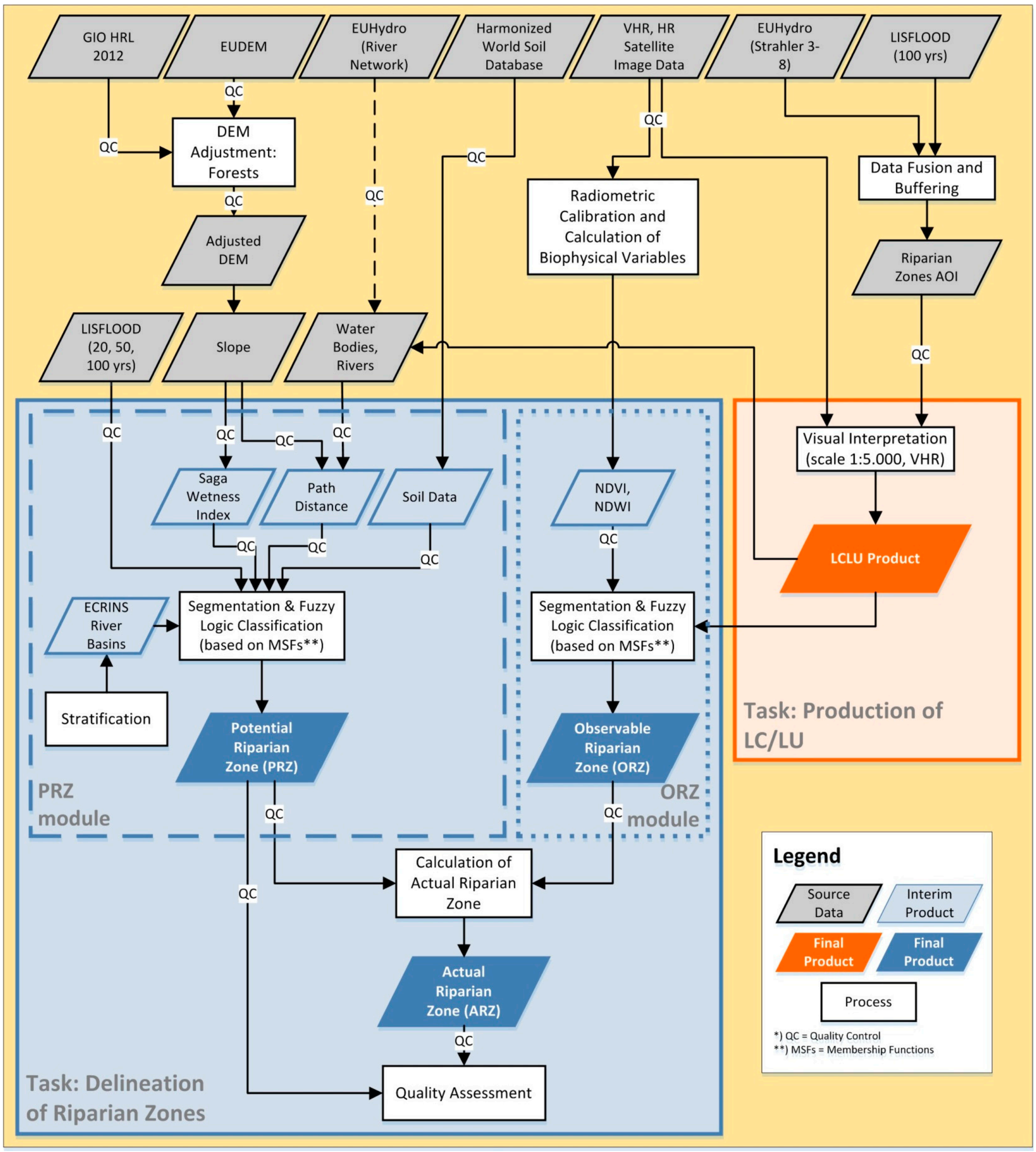

Figure 2: Schematic overview of the riparian zones delineation model (simplified).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 August 2016 doi:10.20944/preprints201608.0059.v1

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

Fuzzy logic based classification scheme and fuzzy set fusion

Within this study, the fuzzy set theory (Zadeh, 1965) is applied for classifying riparian zones from a set of physio-geographic factors/datasets. Depending on MSFs, each of the individual input data sets (see Figure 2) is assigned MS degrees between 0 and 1, expressing how much the data set value satisfies the defined concept (e.g. belonging to riparian zone). An example MSF is depicted in Figure 3 (black circle symbols). Such MS values belong to a fuzzy set, describing "soft" relations with un-sharp or "fuzzy" boundaries with respect to input factor's values, in contrast to dichotomic (binary) MS degrees of "hard" classification approaches. MSFs associate MS degrees to the full value range of the input factor data sets.

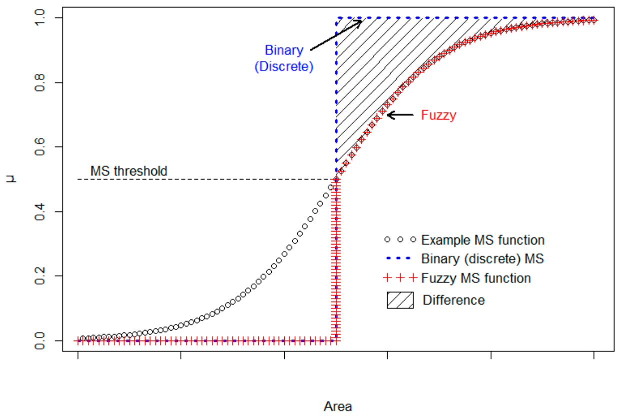

Figure 3: Fuzzy and discrete area accounting applied to an exemplary membership function.

As a number of factors are used, being grouped to modules, fuzzy sets need to be fused (combined). Finding appropriate connectives for the logical combination of fuzzy sets has turned out to be an important issue [58]. There are several techniques to combine or group these fuzzy sets, and results can differ significantly depending on the way the data are fused.

The right choice of such a data fusion framework depends on the type and characteristics of data to be managed [59]. For environmental data of high complexity, often containing imprecise/vague information, a so called hard fusion *(Boolean AND/OR)* would not be adequate, since it can easily produce false positives/negatives. Instead, soft aggregation operators such as *AND like/OR like* allow a modelling towards AND/OR, and can cope with extreme situations, where e.g. only one/all factor/s is/are satisfied [60]. In the frame of this study, Generalized Conjunction Disjunction (GCD) aggregation operators [61,62] have been applied. GCD allows a soft aggregation *(soft OR/soft AND),* which not only can deal with an unlimited number of factors, but also with weights assigned to them. A parameter r defines the degree of simultaneity or conjunction degree (for *soft AND*) and replaceability or disjunction degree (for *soft OR*) of the contributing factors, reflecting the required level of satisfaction of these two fundamental logic connectives. The GCD (parameterized in r) of m values v_k in [0,1] with importance degrees i_k which sum to 1 is defined as follows:

$$GCD\ (v,i) = \left(\sum_{k=1}^{m} i_k * (v_k)^r\right)^{\frac{1}{r}} \tag{1}$$

Where v_k are the values to be aggregated, i_k are their relative importance weights, r is a parameter ranging from $-\infty$ to $+\infty$ and $r \neq 0$ and $\Sigma i_k = 1$. For our purpose we chose a weak simultaneity/replaceability, for

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

which r values are recommended as r=0.26 *(soft AND)* and r=2.018 *(soft OR)* [62]. A comparable approach has been applied to model rural land abandonment with multi-source data [63].

Input features

Input features are usually generated from original model input data, such as remote sensing data or a DEM. After assigning a MS to the generated data sets, based on an individual MSF (detailed in Table 2), the data are integrated in the model as depicted in Figure 2. The characteristics of the applied MSFs, their derivation and MS fusion principles are described in Table 2 and later in this section.

The *Path Distance (PD)* is a frequently used measure to determine the minimum accumulative travel cost of fluids from a source to each location on a raster surface. It was successfully used in previous riparian area mapping [25]. The PD can be expressed as

$$PD = D * F_{sl} \quad (2)$$

with D as surface distance from source and F_{sl} as friction factor determined by the local slope. For longer distances from the source (river) and/or higher slopes, the PD results higher. This general rule is inverted by the applied MSF assigning higher MS degrees to areas closed to the source.

Topography determines largely the gravitational flow of water within the landscape. Local landforms control the hydraulic head, water flow and water distribution [29]. The *Topographic Wetness Index (TWI)* [64] is the most widely used indicator to model the local hydrological behaviour of a varying topography and helps to identify wet or dry sites depending on topography. It is defined as

$$TWI = ln(\alpha / tan\ \beta) \quad (3)$$

where α is the specific catchment area, or drainage area per unit contour width, defined as A/b; A is the upstream catchment area [m^2], while b is the contour width [m]. β is the local slope steepness in degrees. The index takes on higher values, if the upstream catchment area increases or if the slope angle flattens, indicating a stronger wetness. Similarly, the adopted MSF assigns high MS degrees to high TWI values.

The "Saga wetness index" (SagaWI) [65,66] differs from the classical TWI through the use of a "modified catchment area" (α_m), which aims to alleviate the relatively strong effects of slight terrain variability on the catchment area in zones of typically orohydrologic homogeneous conditions, such as valley floors. This is achieved by iterative modification of each grid cell's catchment area in dependence of neighbouring maximum values, using a slope-dependent equation until the result remains unchanged by additional iterations. As result it predicts for cells situated in valley floors with a small vertical distance to a channel a more realistic, higher potential soil moisture compared to the standard TWI calculation. For these reasons the SagaWI was adopted in this study. The SagaWI has been calculated with the software module Saga wetness index in QGIS Desktop 2.5.1 (http://www.qgis.org).

The flood extent zones of the FHM at flood return periods 20, 50, and 100 years were directly assigned a MS degree as outlined in Table 2.

Soil data: Since an ideal soil parameter such as the soil transmissivity, computed of average saturated hydraulic conductivity and the depth to restrictive layer was not available, an alternative had to be computed. The HWSD provides soil attributes for dominant and associated soils within a soil mapping unit (SMU), providing the share of all composing soils. The following soil attributes were selected: area share of SMU, soil type, available water storage capacity (AWC), obstacle to roots, presence of an

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 August 2016 doi:10.20944/preprints201608.0059.v1

Peer-reviewed version available at *Remote Sens.* **2016**, *8*, 925; doi:10.3390/rs8110925

impermeable layer, soil water regime, and top soil and subsoil organic carbon content. For all attributes, dedicated MSFs have been associated, while the MS degree for the attribute "share of SMU" is represented by the share itself (Table 2). All single soil MS were combined by *soft OR* aggregation to a single soil MS.

Indicators for vegetation vigour (Normalized Difference Vegetation Index, NDVI) [42] and leaf water content (Normalized Difference Water Index, NDWI) [43] provide evidence of riparian features. This is particularly evident in drier areas, where the riparian corridor creates a strong contrast with the remaining land cover. Sigmoid MSFs were applied for both indicators (Table 2).

Land use/land cover: LCLU classes have been assigned a MS degree according to their probability of being of riparian nature. Typical riparian features, such as "Riparian broadleaved forest" have been assigned a MS of 1, while typical non-riparian features such as built-up areas or agricultural fields have been assigned a MS of 0. MSs of all classes are detailed in Table S1. LCLU classes which can be of both, riparian and non-riparian nature, have been assigned a so called neutral value, i.e. MS 0.5, in order to keep them in the system. In such cases, the combined final MS is determined mostly by the MS of other input layers.

MS functions were determined by expert knowledge, as in the case of LCLU MSs, or empirically, based on previously mapped riparian zones [25,26], by extracting the histogram of the variable of interest and approximating a MSF from it.

For the modelling of riparian zones it is considered important to take regional differences into account, compensating for soil, geomorphic or climatic particularities. Therefore, a stratification approach was applied through individual adjustments to each of the 43 DUs, applying modifications to the weights (potentially ranging between 0 and 1) of the input layers (i) Path Distance, (ii) TWI, (iii) soil data and (iv) the FHM for PRZ delineation. The standard weights for these layers were set to 0.8, 0.3, 0.1 and 0.6, respectively. Adjustments of up to +/- 0.1 were applied, based on expert judgement by comparing the PRZ extent to high resolution optical satellite data as a reference.

Model output

By applying a segmentation together with the above described fuzzy logic based classification approach to the available input features, the following riparian zone delineation products are derived:

1. Potential Riparian Zones (PRZ)
2. Observable Riparian Zones (ORZ)
3. Actual Riparian Zones (ARZ)

A major advantage of the chosen model design is its ability to create these complementary riparian zone delineation products, which represent different aspects of riparian zone occurence (such as the PRZ and the ARZ), thus enabling further analyses. The ARZ is created as intersecting product of PRZ and ORZ, assuring a higher degree of product reliability, since both data sets are based on independent sources. The ratio ARZ/PRZ provides a first assessment of *riparian extent saturation* or, if inverted, the *riparian extent deficiency*, relating the actual riparian extent to the potential one within a defined area. High *riparian extent saturation* values indicate areas where ARZ is very close to its potential extent, or, in terms of area extent, almost "saturated". Low values in turn express a low degree of saturation, indicating an extent deficiency. The ratio is deemed a richer indicator as the ARZ and/or PRZ by its own. Assuming that PRZ provides the historic or pristine-like extent of riparian coverage, the inverted ARZ/PRZ ratio provides a dimension for a rough assessment of historic riparian area losses.

Peer-reviewed version available at *Remote Sens.* **2016**, *8*, 925; doi:10.3390/rs8110925

In the module PRZ, areas are mapped for their natural disposition to host riparian features, which are not necessarily represented (any more) in today's actual LCLU class. The disposition is computed for several input layers, which can be added or removed in a modular way. The resulting MSs are combined to a preliminary PRZ, applying a *soft OR* aggregation (see Equation 1).

In a last step, standardization is applied to the preliminary PRZ MS raster. That is to ensure that the vector based delineation of the PRZ (MS_PRZ > 0.5) is harmonized with the ones of neighbouring DUs. Based on a threshold (e.g. 0.6) applied to the preliminary PRZ raster, which best fits the PRZ extent at MS $\mu = 0.5$ of neighbouring DUs, a linear rescaling of values lower and higher than this threshold is done, in order to standardize the meaning of the PRZ MS values. The resulting raster is the PRZ MS called MS_PRZ.

In a parallel module, the ORZ is derived, representing the actually observed (often recent) riparian zones, based on different input layers. Also in this case, input layers can be considered modular and are treated and combined the same way as for PRZ. ORZ is based on the observation of certain vegetation classes, as derived from VHR EO imagery and subsequent classification. Additionally, NDVI and NDWI provide further evidence of riparian features. The resulting single MSs of all layers are combined *(soft OR)* to a single MS expressing the probability to encounter riparian features on ground (MS_ORZ).

MS_PRZ and MS_ORZ are combined via soft aggregation to MS_ARZ (*soft AND,* with r=0.26; see Equation 1).

National or supranational riparian datasets can, if available, be overlaid at this point, leaving the door open for assimilation of such further independent data sources. MS_ARZ (or MS_PRZ) are finalized (i) as vector data sets, by offsetting the MS raster at μ=0.5, and (ii) left as raster with the according MS values.

3. Results

All results are grouped into DUs (see Figure 1), and can be consulted or downloaded from the Copernicus site: http://land.copernicus.eu/local/riparian-zones. An example, depicting vector and raster based PRZ results in full resolution for one DU (Weser catchment) is shown in Figure 4.

To facilitate the (statistical) analysis of the output products on a pan-European extent, they were summarized within 1 km^2 grid cells. Although leading to less detailed data, this processing step avoided or minimized information losses. Besides the high spatial detail, also the fuzzy character of the products required particular attention for statistical analysis. In this regard it is important to keep in mind the concepts of fuzzy and dichotomous (binary) MS (explained earlier) for reading the results. Each of these concepts delivers different results for the areas of ARZ and PRZ. The concepts of fuzziness and dichotomy for area measurements are not applied on the full MS range, but only above μ=0.5, as Figure 3 depicts. Area measurements in binary mode assume a sharp cut-off at the MS threshold of μ=0.5, below which no riparian area is being kept or accounted for, while above it the area is fully considered. Equally not accounted for are riparian areas of $\mu < 0.5$ when applying the fuzzy area assessment, while above of this threshold, area measurements are weighted according to the resulting raster based MS. The difference between binary and the here applied fuzzy area assessment can be quantified by the hatched area in Figure 3 and depends obviously on the effective shape of the MSF.

Application of fuzziness on the full MS range can instead be seen from the detailed mapping example in Figure 5 B, where MSs ranging from 0 to 1 are overlaid by a vector based delineation (cut off at μ=0.5). Furthermore, the detailed example shows the ARZ, ORZ and PRZ delineation with a VHR and the LCLU classification background (Figure 5 A, C).

Country-wise and global statistical key numbers of ARZ, PRZ (both fuzzy and binary, named ARZ_{fuz}/ARZ_{bin} and PRZ_{fuz}/PRZ_{bin}) and other indicators, such as LCLU shares, are reported in Table 3. Within the study area, a total of 55,558 km^2 have been delineated as ARZ, and 182,488 km^2 as PRZ (fuzzy approach), which represent 0.95% and 3.13% of the total study area (EEA-39), respectively. If accounted in

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

binary way, ARZ amount at 69,128 km² and PRZ at 341,215 km² or 1.19% and 5.86% of the total study area, respectively.

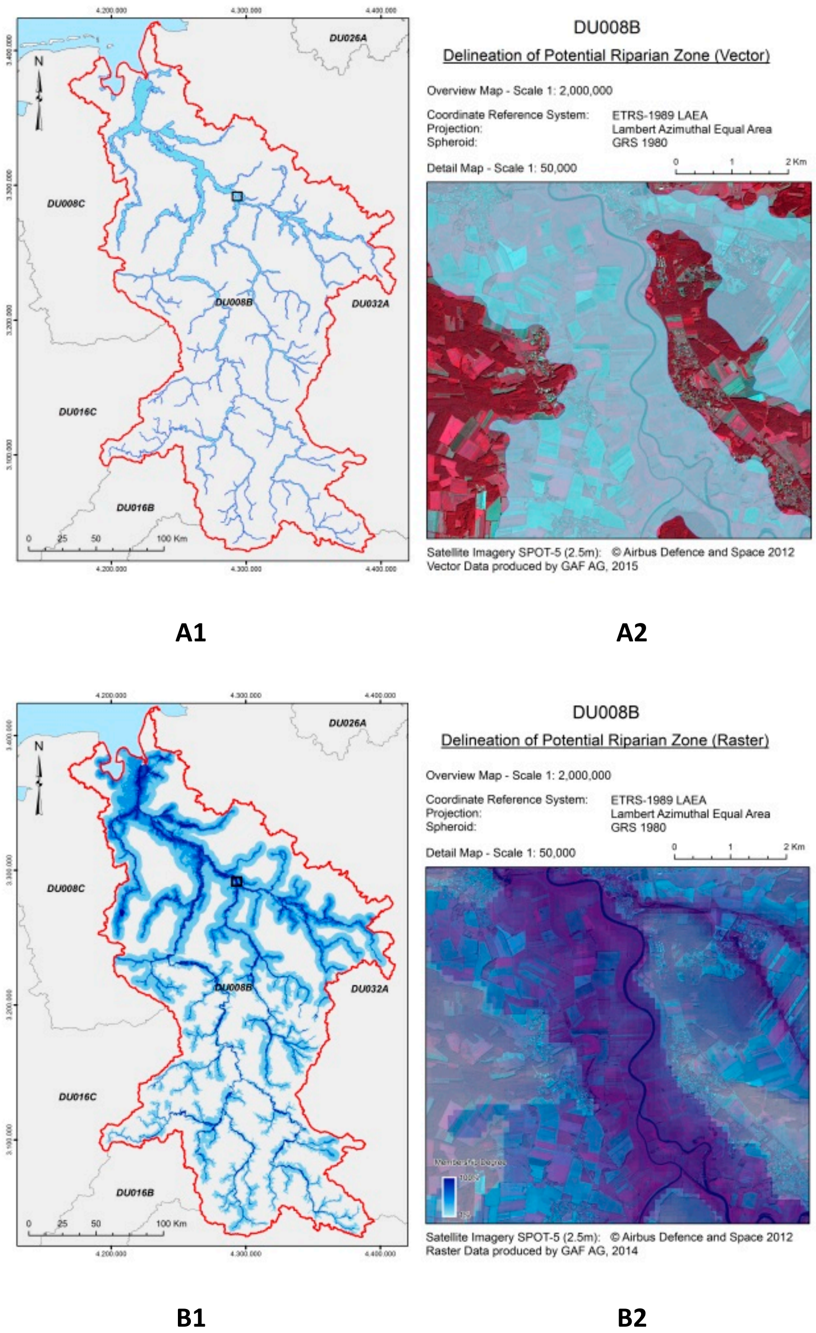

Figure 4: Examples of the Potential Riparian Zone (PRZ) products for Delivery Unit 008B (*Weser* catchment, Germany). A1: PRZ as vector, *Weser* catchment, A2: subset of A1, B1: PRZ as raster, *Weser* catchment, B2: subset of B1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 August 2016 doi:10.20944/preprints201608.0059.v1

Peer-reviewed version available at *Remote Sens.* **2016**, *8*, 925; doi:10.3390/rs8110925

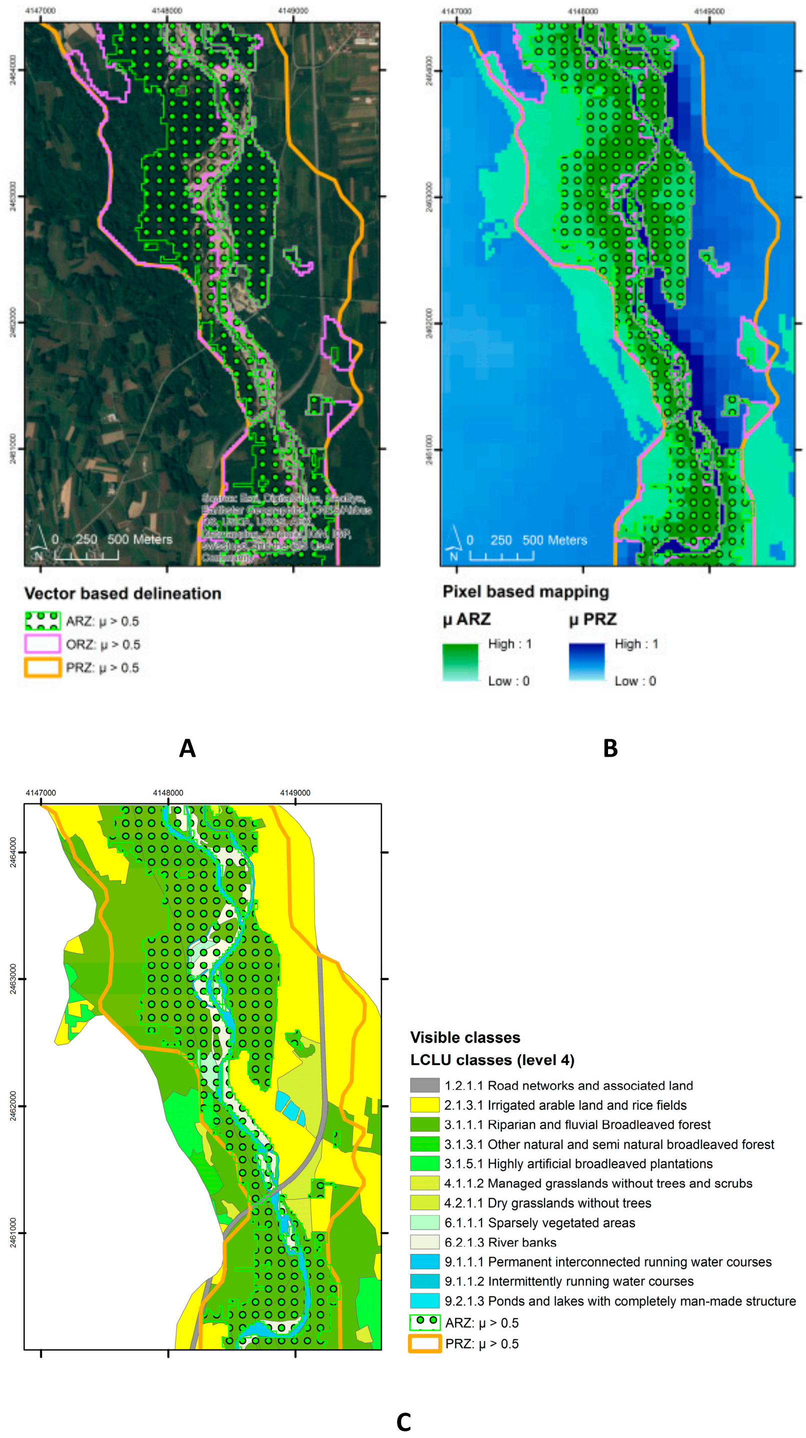

Figure 5: Delineation example: *Orco* River, Piedmont, Italy. A: Vector based delineation and VHR background; B: Pixel based results overlaid by vector based delineation; C: Vector based delineation and LCLU as background.

 doi:10.20944/preprints201608.0 v1

Table 3: Country-wise statistics

Code	Zone	Country_area	ARZ_{fuz}	PRZ_{fuz}	ARZ_{fuz}	PRZ_{fuz}	Ratio ARZ_{fuz}/PRZ_{fuz}	ARZ_{bin}	PRZ_{bin}	ARZ_{bin}	PRZ_{bin}	Urban, agricultural, semi-natural share					
												ARZ_{bin} based			PRZ_{bin} based		
		[km²]	[km²]	[km²]	[%]	[%]	[-]	[km²]	[km²]	[%]	[%]	[-]			[-]		
AD	Andorra	465	4	10	0.76	2.12	0.449	4	11	0.92	2.45	0.02	0.05	0.93	0.03	0.34	0.62
AL	Albania	28,488	318	977	1.12	3.43	0.435	389	1,745	1.37	6.12	0.05	0.06	0.89	0.35	0.38	0.27
AT	Austria	83,928	787	3,105	0.94	3.70	0.397	950	4,543	1.13	5.41	0.01	0.09	0.9	0.28	0.51	0.21
BA	Bosnia and Herzegovina	51,157	465	1,589	0.91	3.11	0.402	581	2,321	1.14	4.54	0.05	0.12	0.83	0.3	0.45	0.25
BE	Belgium	30,668	200	1,050	0.65	3.42	0.300	249	1,564	0.81	5.10	0.	0.16	0.84	0.23	0.59	0.18
BG	Bulgaria	110,979	989	3,266	0.89	2.94	0.451	1,176	4,964	1.06	4.47	0.05	0.14	0.81	0.36	0.46	0.18
CH	Switzerland	41,280	276	1,081	0.67	2.62	0.416	322	2,768	0.78	6.70	0.02	0.15	0.83	0.28	0.51	0.2
CY	Cyprus	9,246	1	41	0.01	0.44	0.133	2	56	0.02	0.61	0.01	0.01	0.98	0.4	0.56	0.04
CZ	Czech Republic	78,870	443	2,158	0.56	2.74	0.326	540	3,197	0.69	4.05	0.01	0.17	0.82	0.35	0.51	0.13
DE	Germany	357,766	2,464	14,922	0.69	4.17	0.258	3,101	23,146	0.87	6.47	0.01	0.18	0.81	0.29	0.59	0.12
DK	Denmark	43,065	53	327	0.12	0.76	0.259	65	508	0.15	1.18	0.01	0.31	0.68	0.39	0.53	0.08
EE	Estonia	45,372	454	896	1.00	1.98	0.545	565	3,081	1.24	6.79	0.01	0.24	0.75	0.13	0.44	0.42
EL	Greece	132,028	830	3,787	0.63	2.87	0.401	1,005	6,191	0.76	4.69	0.07	0.09	0.84	0.42	0.45	0.14
ES	Spain	498,537	4,202	15,635	0.84	3.14	0.406	5,218	23,706	1.05	4.76	0.13	0.15	0.72	0.38	0.44	0.18
FI	Finland	337,838	7,738	11,250	2.29	3.33	0.704	9,673	36,589	2.86	10.83	0.01	0.01	0.98	0.14	0.15	0.71
FR	France	549,061	4,898	19,820	0.89	3.61	0.373	5,814	29,457	1.06	5.36	0.05	0.1	0.86	0.31	0.51	0.18
HR	Croatia	56,536	849	2,625	1.50	4.64	0.392	1,084	3,973	1.92	7.03	0.04	0.12	0.84	0.33	0.44	0.24
HU	Hungary	93,013	1,393	6,959	1.50	7.48	0.262	1,735	11,119	1.86	11.95	0.04	0.25	0.71	0.37	0.52	0.11
IE	Ireland	69,946	440	1,913	0.63	2.73	0.295	554	3,459	0.79	4.95	0.	0.33	0.67	0.05	0.67	0.28
IS	Iceland	102,688	371	1,360	0.36	1.32	0.357	851	2,777	0.83	2.70	0.	0.02	0.98	0.05	0.11	0.84
IT	Italy	300,645	3,281	13,188	1.09	4.39	0.416	3,856	20,023	1.28	6.66	0.05	0.1	0.85	0.4	0.45	0.15
LI	Liechtenstein	160	2	16	1.49	10.25	0.143	3	23	1.66	14.27	0.01	0.02	0.97	0.35	0.53	0.12
LT	Lithuania	64,899	621	1,676	0.96	2.58	0.449	733	2,569	1.13	3.96	0.01	0.34	0.66	0.18	0.57	0.25
LU	Luxembourg	2,596	12	50	0.47	1.92	0.343	15	70	0.56	2.69	0.01	0.05	0.94	0.11	0.59	0.29
LV	Latvia	64,586	654	1,773	1.01	2.75	0.425	800	2,844	1.24	4.40	0.01	0.2	0.79	0.2	0.49	0.32
ME	Montenegro	13,879	141	289	1.02	2.08	0.575	177	645	1.28	4.65	0.05	0.09	0.86	0.12	0.37	0.51
MK	Former Yugoslav Rep. of Macedonia	25,439	236	885	0.93	3.48	0.445	282	1,694	1.11	6.66	0.12	0.14	0.75	0.39	0.44	0.17
NL	Netherlands	37,357	178	1,873	0.48	5.01	0.158	221	3,256	0.59	8.72	0.01	0.31	0.69	0.26	0.68	0.07
NO	Norway	323,383	2,940	5,272	0.91	1.63	0.588	3,875	13,880	1.20	4.29	0.01	0.02	0.97	0.16	0.18	0.67
PL	Poland	311,928	2,616	12,200	0.84	3.91	0.294	3,231	19,115	1.04	6.13	0.01	0.36	0.63	0.25	0.56	0.19
PT	Portugal	88,843	880	2,501	0.99	2.81	0.472	1,113	4,020	1.25	4.52	0.16	0.16	0.68	0.36	0.41	0.23
RO	Romania	238,392	2,966	11,709	1.24	4.91	0.365	3,700	18,664	1.55	7.83	0.05	0.19	0.76	0.36	0.48	0.16
RS	Serbia	77,542	971	3,156	1.25	4.07	0.412	1,202	4,949	1.55	6.38	0.14	0.19	0.67	0.39	0.44	0.17
SE	Sweden	449,718	8,593	13,683	1.91	3.04	0.667	10,651	43,605	2.37	9.70	0.	0.01	0.99	0.14	0.17	0.69
SI	Slovenia	20,277	159	688	0.78	3.39	0.367	194	982	0.96	4.84	0.02	0.06	0.92	0.36	0.47	0.18
SK	Slovakia	49,026	452	2,593	0.92	5.29	0.277	533	3,805	1.09	7.76	0.04	0.24	0.73	0.39	0.51	0.09
TR	Turkey	780,102	1,727	10,779	0.22	1.38	0.279	2,300	24,356	0.29	3.12	0.07	0.24	0.69	0.37	0.48	0.15
UK	United Kingdom	244,574	1,872	7,053	0.77	2.88	0.420	2,269	11,070	0.93	4.53	0.	0.31	0.69	0.23	0.6	0.17
XK	Kosovo UNSCR 1244/99	10,907	81	335	0.75	3.07	0.364	95	472	0.87	4.32	0.14	0.15	0.71	0.41	0.45	0.14
		5,825,185	55,558	182,488	0.95	3.13	0.474	69,128	341,215	1.19	5.86	0.03	0.12	0.85	0.27	0.41	0.32

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

The ratio ARZ/PRZ, calculated on a fuzzy approach, is depicted in Figure 6. The figure provides quickly an impression of where the riparian areas are close to the full potential, such as in large parts of Scandinavia. For spatially identifying *riparian extent deficiencies* (i.e. areas with a low ARZ/PRZ ratio) the situation appears spatially complex. Anyway, some areas such as the Netherlands, large parts of Germany and Eastern European countries, and the Po-valley (shown in detail) can be identified as hot spots.

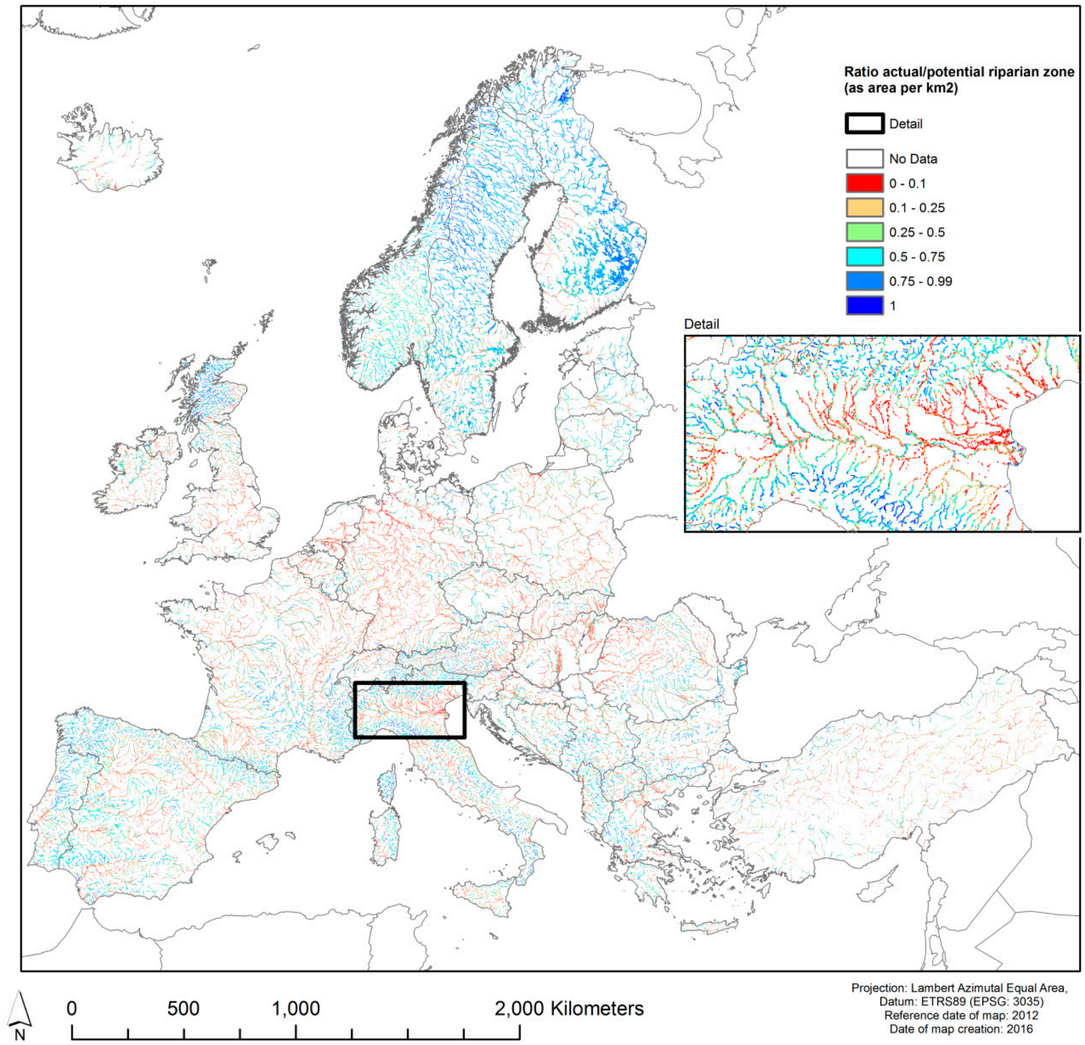

Figure 6: Spatial distribution of the ARZ/PRZ ratio.

Apart from the ratio values, it is interesting to get an idea about its components, the ARZ and PRZ areas. Figure 7 provides a graphical overview of the individual ARZ/PRZ relations for all EEA-39 countries. Figure S1 in the section *Supplementary data* depicts the analogous results per DU. The three Scandinavian countries Norway, Finland and Sweden (NO, FI, SE) exhibit particularly high proportions of ARZ. The correlation between ARZ and PRZ of the individual countries is high, especially without the exceptional Scandinavian contributors (R=0.81 and 0.97, respectively). The linear function reveals a gain of 0.30 and 0.23 (with and without NO, FI, SE, respectively), which means that through such country-wise averaging approach, the actual ARZ amounts to about 30% (with NO, FI, SE) and 22% (without them), respectively, of its potential size. These values, however, if calculated grid cell-wise, result clearly higher: the EEA-39 wide mean and median amount to 0.44 and 0.40 with and to 0.37 and 0.29 without Norway, Finland and Sweden, respectively. The large area contribution (in terms of absolute km²) of the Scandinavian countries with generally higher ARZ/PRZ ratios increase the EEA-39-wide means in this pixel-based calculation, as compared to a country-wise approach.

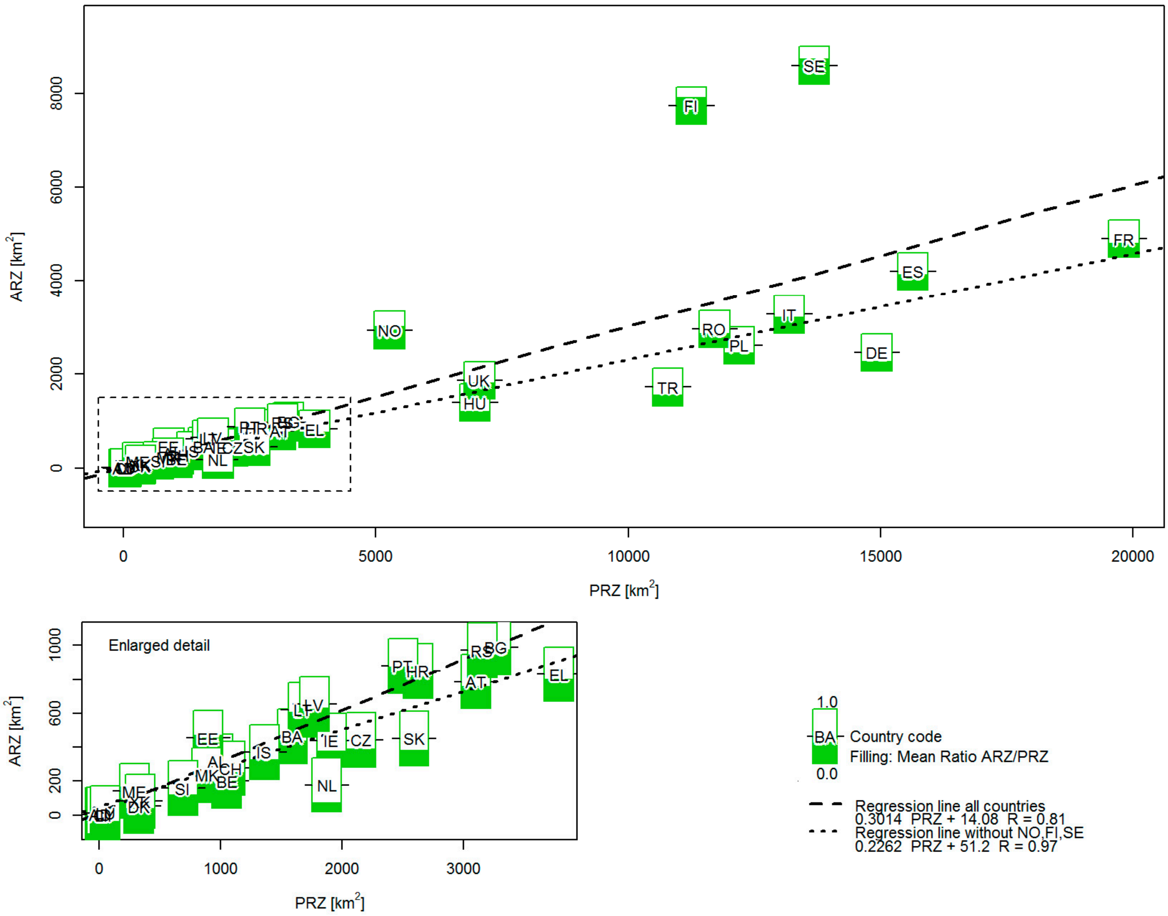

Figure 7: Country-wise comparison of ARZ/PRZ ratios in dependence of the individual overall extents. Due to extra-ordinary high ARZ abundance for three Nordic countries (NO, FI, SE), regression lines were calculated with/without them. The small graph displays the indicated detail.

At this point the pressures or reasons for the *riparian extent deficiencies* become of interest. In terms of LCLU this can be answered by Figure 8, where the riparian extent deficiency (as colour intensity) is combined with the three main LCLU categories "Urban" (in red), "Agriculture" (green), and "Semi-Natural" (blue). The green colour, expressing agricultural land use, is clearly dominating this image for most of Europe, whereas in Scandinavia this changes in favour of semi-natural LCLU and locally, especially in highly urbanized areas, towards urban land use. Colour intensity is higher for areas with high riparian extent deficiency.

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

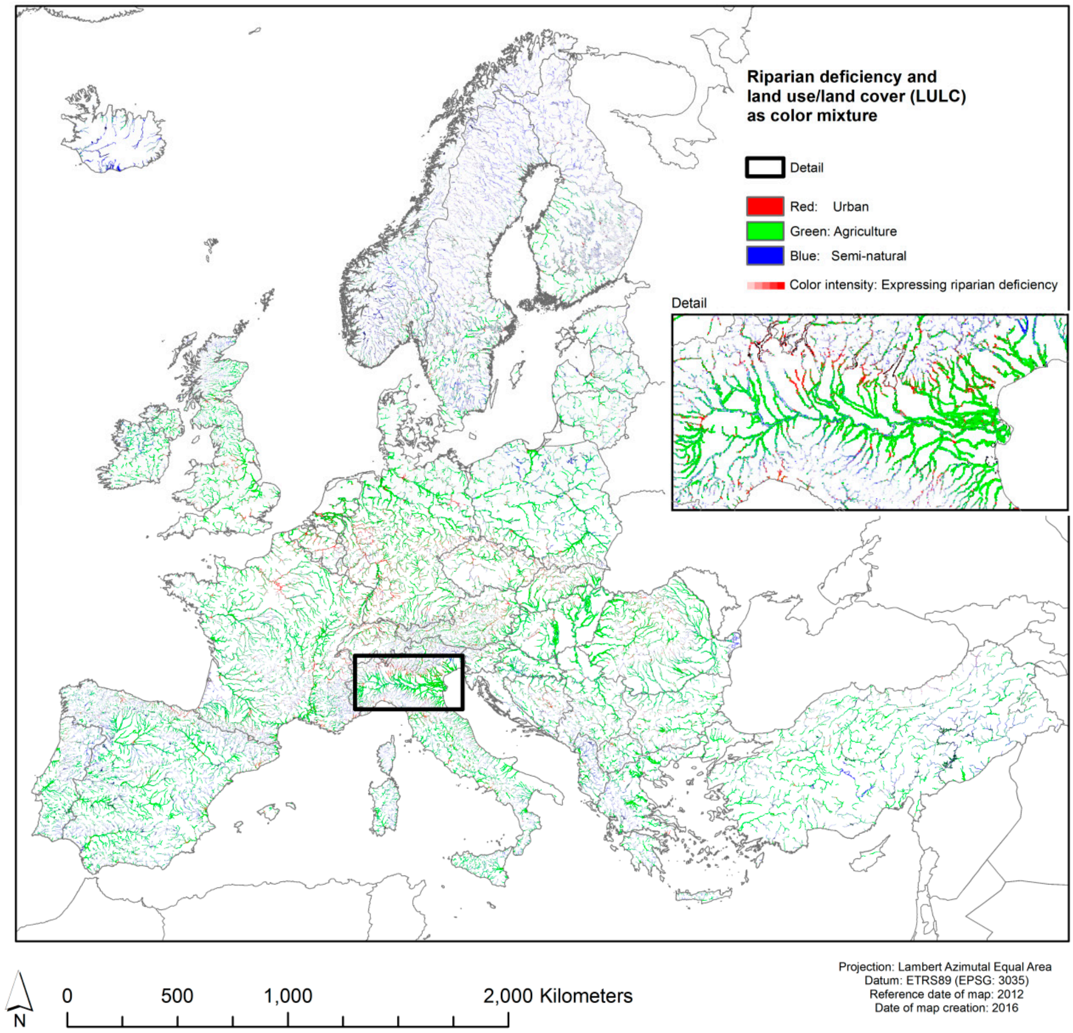

Figure 8: Dominant land cover/land use (as colour) combined with *riparian extent deficiency* (as colour intensity).

Table 3 and Figure 9 provide more detailed numbers for countries and the full study area in relation to LCLU. ARZ_{bin} comprise 3% urban, 12% agricultural and 85% semi-natural land. The high area share for semi-natural land is obviously expected, while the urban and agricultural shares can be explained by the presence of green urban areas and in particular agricultural grasslands, respectively.

PRZ_{bin} instead comprise 27% urban, 41% agricultural and 32% semi-natural land, indicating that there would be a significant potential for riparian expansion, especially within the land category agriculture. In urban areas, an expansion may be more difficult due to constraints such as built-up areas.

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

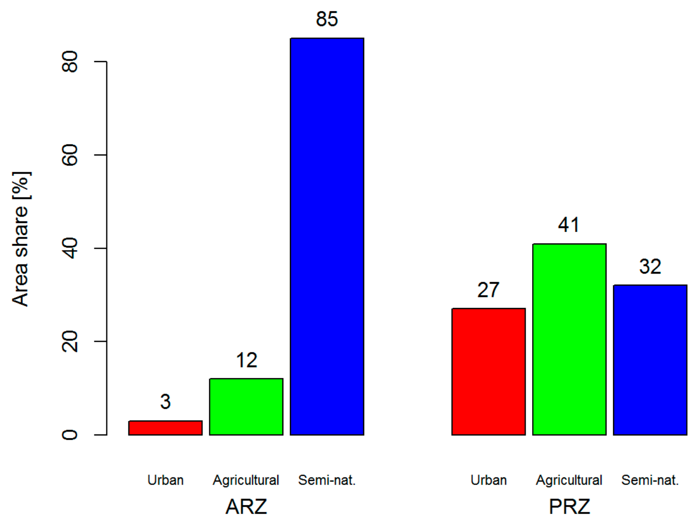

Figure 9: LCLU share of ARZ and PRZ

Accuracy and reliability

Due to excessive costs and time constraints, an ad-hoc field validation for accuracy assessment of the riparian zones dataset on the EEA-39 scale was not considered feasible. Alternatively, two strategies were followed to derive indications of the reliability of the ARZ dataset: (i) a qualitative assessment examining and discussing classification errors identified through an extensive visual analysis of the ARZ dataset with Google Earth Pro© observation viewer [67] and additional VHR imagery; (ii) a quantitative assessment of user and producer accuracy using (a) visual validation points from Google Earth Pro© and (b) independent datasets. Only the ARZ_{bin} dataset has been validated, due to lack of appropriate validation data comparable to the PRZ.

The qualitative reliability evaluation was performed using a set of 11 multispectral VHR images of recent RapidEye, Quickbird and SPOT5 acquisitions, together with Google Earth Pro© imagery. The ARZ were overlaid to the imagery, and three randomly chosen areas (each of approximately 10 – 100 ha) for each of the 47 DUs were visually analysed, summing up to a total of 141 areas. Overall, ARZ were well represented, appearing to be in general terms a reliable product. Overestimations (false positives) are mostly found: (i) in landscapes with complex agricultural/urban land use/land cover, with classification errors more common in green leisure areas and urban gardens, or in agricultural patches within an urban context; (ii) nearby or within flooded areas; (iii) along tree lines of cemented river channels (Figure 10, A and B). Underestimation errors (false negatives) are mostly identified near or within large floodplains of anastomosed rivers and river deltas, and in some areas where the river network appears underestimated (Figure 10, C and D). Generally, regions with less frequent urban or agricultural land cover reveal higher delineation accuracy.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 August 2016 doi:10.20944/preprints201608.0059.v1

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

Figure 10: Examples of riparian misclassification errors (red vector line: riparian zone). Overestimations: (A) Sport field; (B) urban gardens. Underestimations: (C),(D) river meanders.

Quantitative measures of ARZ accuracy were derived on a first instance using visual validation points (VISVAL) from Google Earth Pro© imagery, as applied in previous studies [25,68]. Confidence level δ of classification accuracy estimate p for large sample sets was calculated following Spiegel [69]. User's accuracy (UA) was derived by randomly extracting 200 points (25 x 25 m pixels) from the ARZ dataset, and checking on-screen the number of matches. ARZ UA resulted as 82.0 ± 5.3% at 95% confidence level. Similarly, producer's accuracy (PA) was calculated by identifying randomly distributed riparian points in Google Earth Pro©, using a 25 x 25 m pixel geometry, and limiting the VISVAL points to the hydrological buffer network used in the ARZ processing. After applying filters for *Strahler's* order 3-8 we assessed 191 points and their matches (spatial intersection) with the ARZ dataset. PA resulted as 77.5 ± 5.9%.

As a complementary indication of PA we also exploited three different and independent datasets (i)-(iii) for absence/presence of riparian zones. The datasets are based on: (i) ecological survey of riparian forest ('alluvial and riparian woodlands and galleries close to main European river channels'), as part of the LUCAS 2009 data [70]; (ii) River Habitat Survey data [RHS, 71], a field method for the broad characterization of river streams, from (a) Technical University of Lisbon and (b) various institutions in the context of the MARCE project (http://marce.ihcantabria.es), coordinated by the University of Cantabria. The dataset was limited to points along streams of *Strahler's* order 3-8 and to the hydrological buffer network considered in the ARZ model, resulting in 402 points distributed predominantly over Western Europe. To meaningfully represent the extension of each field survey we assigned a 50 m buffer to the riparian point location, and checked spatial intersection with the ARZ dataset. PA resulted in this case as

 doi:10.20944/preprints201608.0059.v1

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

83.1 ± 3.66%. Overall, the quantitative accuracy assessment was performed with a total of 793 sampling points. Figure S2 shows the spatial distribution of all validation points.

4. Discussion

The achieved results within this work are of high importance to a wide range of disciplines, among them policy makers and land or water (restoration) planners, researchers, engineers, water-basin managers, ecologists and economists. The pan-European extent and the consistent way of data derivation facilitate actions on a large scale, as for the case of EU-policy. With respect to existing data sets, the gap for a uniformly derived dataset covering the whole EEA-39 area whilst being of high spatial detail is filled. A multitude of input data, including topographic, hydrological, soil and EO-based data is employed. The yielded results are not comparable to the previously derived JRC riparian zones but represent a data set of higher value, providing more detail, accuracy and reliability, considering the applied approaches and the input data sets employed. For example, the hydrological network, based on EU-Hydro in combination with the VHR LCLU dataset of riparian zones, has been derived based on optical remote sensing data, from which the geo-location of rivers was precisely captured, in contrast to the hydrological network of the JRC data set, where river locations are mainly based on a drainage network based on the SRTM-DEM. Another example is the LCLU MMU, which has significantly improved from CORINE's 25 ha to 0.5 ha. Accuracy improvements have been achieved also for the applied EU-DEM, in Nordic countries substituted by more reliable national data sets. Apart from data input quality improvements, the multitude of input data sets is expected to significantly increase the reliability, since, similar to a weight-of-evidence approach, a repeated affirmation of an attribute leads to more robust results. A crucial improvement of the riparian zones delineation model compared to previous works is its completeness, covering not only semi-natural areas, but all kind of land cover/land use categories. Moreover, the applied concept with multiple outputs ARZ, ORZ and PRZ enables a much wider use of results. A number of novel applications should become possible, e.g. in ecological fields such as ecosystem service and function analyses, through the additional availability of the PRZ, the ORZ, the ratio ARZ/PRZ and the parallelly derived LCLU data set.

The implementation of this work required an effort from policy side, who is accomplishing with present and future ecological needs, such as ecosystem services and biodiversity monitoring. It also required significant investment from a technical-scientific body carrying out a vast and complex work. The present article describes indeed not one but several output data sets within the riparian context, i.e. mainly the ARZ and the PRZ. Synergistically, a detailed VHR LCLU dataset is available in parallel, covering the wider riverine zone of all rivers of *Strahler's* stream order >=3. Each data set and specifically their combination, provide an excellent starting point for more detailed analyses in the riverine and riparian domain. Data downloads are organized in Delivery Units (DU), which are based on an aggregation of ECRINS river (sub-)catchments. This way, a separate and independent processing of each DU can be performed without running into technical issues of exceeding computer capacities.

Methodologically, the riparian area delineation approach can be divided into two work tasks: the derivation of (i) ARZ and (ii) PRZ (and intermediate result ORZ), while the LCLU classification is considered a parallel task, which in other cases might not be required to be computed. Only the foreseen independency of the modules ARZ and PRZ enables finally the contextual interpretation of their resulting end-products, i.e. by means of the ARZ/PRZ ratio. PRZ is based on the hydrological data set (river network), terrain topography (slope, height above water level), soil properties, flood zone delineation and LCLU class water. Following a fuzzy MS approach, MSs are assigned to each contributing layer, following defined MS functions. In GIS-terms, an OBIA approach has been applied, where the atomic unit is an object, defined as a polygon with relatively homogeneous conditions, created by a segmentation process. The segmentation and the fuzzy MS approach was applied analogously to the ORZ, where the input data

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 August 2016 doi:10.20944/preprints201608.0059.v1

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

sets are constituted by LCLU data and the EO-data derived vegetation indices NDVI and NDWI. ORZ and PRZ are combined to the ARZ. The multitude of input data sets requires strong efforts in terms of processing and labour, but contributes to higher robustness and accuracy reducing the effect of single bad data values. The model design keeps the door open for further input datasets, potentially becoming available in the future (e.g. by expanding the model to all *Strahler's* stream orders).

The accuracy analysis of ARZ_{bin} was performed by different approaches, also employing independent data sets. The quantitative assessment revealed similar accuracies for both UA and PA, and for visual analysis and independent sources, ranging in all cases around 80%, which can be considered in line with good mapping practices [72]. The achieved accuracy is deemed a good result, considering the continental-scale extent of the study area and the complexity of the task.

A basic analysis on European/national and DU-scale provides statistical figures and graphs of the features considered most important. A key indicator is considered the ratio ARZ/PRZ, providing the relation of actual and potential riparian zones and this way the degree of *riparian extent saturation*, or, if inverted, the degree of *riparian extent deficiency*. The combination of the ratio ARZ/PRZ and the LCLU data allows an assessment of the drivers of long-term or historic riparian area loss, which was discovered to be predominantly conversion to agricultural area.

However, it is crucial to note that statistics depend significantly on the riparian area inventory approach applied. Here, two different approaches, a fuzzy and a binary one, have been adopted. At a European extent, i.e. if considering the whole study area (EEA-39), the fuzzy and binary approach reveal for ARZ 55,558 km^2 (0.95% of total study area) and 69,128 km^2 (1.19%), respectively, hence the binary approach delivering a number approximately 24% higher. For PRZ, the fuzzy (182,488 km^2 or 3.13%) and binary based figures (341,215 km^2 or 5.86%) differ by about 87%, the higher discrepancy indicating that PRZ_{fuz} exhibits generally lower MS degrees compared to ARZ_{fuz}. This is expected, since the uncertainty of ARZ, due to a higher number of input layers (sum of ORZ and PRZ input layers) compared to PRZ, is reduced. Also, the PRZ intrinsic uncertainty can be expected higher, linked to higher difficulties to determine its area compared to ARZ. If looking at the shares of ARZ_{bin} in between urban, agricultural and semi-natural areas, only a small part of ARZ_{bin} is found in urban (3%) and in agricultural areas (12%), the majority is dominated by semi-natural areas (85%). PRZ_{bin} instead, is predominantly covered by agriculture (41%), followed by semi-natural areas (32%) and urban areas (27%). These numbers leave space for policy action (e.g. conversion, co-existence, protection), in particular the revealed high share of agricultural land. Urban areas might be more difficult to handle, since constrained by built-up areas.

The country-wise relation of ARZ/PRZ ratios in dependence of the individual ARZ and PRZ absolute extents has been quantified by linear regression analysis, revealing regression coefficients (gains) of 0.22 to 0.3, depending on application of either a country-wise or basin-wise approach. Scandinavian areas (NO, SE, FI) clearly contribute to an increase of the regression coefficient, due to their high abundance of ARZ. This becomes even more evident when looking at the global pixel based mean and median for the ratio, which amount at 0.44 and 0.40 with and at 0.37 and 0.29 without Norway, Sweden and Finland. Here, due to their large extent, the Scandinavian areas tend to raise the values even more than on a country- or basin-wise analysis (regression). However, the pattern of the ARZ/PRZ ratio is not just higher in Nordic areas, but is highly variable also across micro-regions. The quite clear meaning of this ratio and its ease of calculation makes this indicator a precious tool for further planning.

The potential use of the data set is wide-spread: on top of all applications stands certainly an analysis which goes beyond the one presented in this paper, which helps to understand the roots of a variety of underlying drivers, helps geo-localizing them and provides a base for restoration. For land/water planners the data could provide important insights for large scale river restoration and filter (buffer) design, as for example with respect to placement or design of riparian filter strips or how to alleviate nutrient/pesticide emissions to waters. Policy makers may understand better how to allocate funds for effective counter-measures to pollution and plan future actions. Flood mitigation planners may extract useful

Peer-reviewed version available at Remote Sens. 2016, 8, 925; doi:10.3390/rs8110925

information to plan (bio-engineered) flood retention basins. Ecologists or related professions may analyse the status of riparian areas, their longitudinal/lateral connectivity and fragmentation, their quality e.g. regarding biodiversity and find the hotspots for interventions. A number of other applications not listed here do obviously exist. For most applications, the pan-European extent provides a tool for action targeting, or in other words, in which places/regions to invest first in order to reduce the highest risks or to get out the highest revenue. Obviously, for detailed scale planning more accurate data sets will need to be considered.

A limitation of the data set is certainly the exclusion of the headwaters or rivers of *Strahler's* stream order 1-2, being justified by cost considerations. The missing headwaters can be of particular relevance when assessing ecosystem services such as filter or retention capacity, generally known to be higher in lower stream orders, due to preferential non-concentrated flow in these areas. However, this part can be integrated in a second step.

Other limitations concern the variety and quality of input layers. While the quality of the Digital Elevation Model and the hydrological network has been improved compared to the version used for the JRC riparian map, other data sets lack still detail and consequently accuracy, such as the soil maps. The effect of such a rough scale (1:1M), although gridded into 1km^2 pixels, was alleviated with a lower weight (< 0.2) for soil data when combining the data within the processing chain.

5. Conclusions

An important step towards a consistent and harmonized European-wide assessment of riparian areas has been achieved. The presented results provide a very detailed and solid delineation of riparian zones and procure at the same time a baseline for further analyses within the legal requirements such as the European Water Framework Directive, the European Floods Directive or the European Mapping and Assessment of Ecosystems and their Services (MAES) initiative.

Apart from legal obligations, researchers, land/water planners and managers in the environmental and other fields can base their work on the created data sets, which are particularly suited for larger scale analyses.

The gap for a structured and detailed riparian data set, which can be monitored and repeatedly derived with comparable methods in future years, has been filled. To accommodate expected advances in spatial data availability in the next years, the methodological design leaves the door open for new and more advanced input data sets.

A statistical analysis provides a basic assessment of the riparian zones across 39 countries in terms of riparian extent. Riparian areas appear heavily diminished compared to their potential extent, with regionally varying figures. Major losses can be attributed in the first place to historic conversion into agricultural area, followed by extension of urban land use. These results might serve as a starting point for future riparian land restoration endeavours and related identification of hot spots.

The presented results show a snapshot of the situation centred in 2012 which is repeatable at any later stage and its structure should allow monitoring of riparian areas in future, even with relatively dense frequency. Both, as a one-off assessment and as future time-series the data set should provide the basis for further insights into the dynamics of a highly valuable and sensitive ecotone.

Supplementary Materials: Figure S1: DU-wise comparison between ARZ and PRZ. Due to extra-ordinary high ARZ abundance for some Nordic DUs (27A, 29A, 31A, 37A, 38A, 34A), regression lines were calculated with/without them. The small graph displays the indicated detail., **Figure S2:** Spatial distribution of the visual validation points (VISVAL) used in the accuracy assessment., **Table S1:** Nomenclature of the Land Cover and Land Use product, including the highest level of class discrimination at level 4 (level 1 is compatible with the MAES ecosystem types), and MS degree assigned to each class.

Acknowledgments: This study has been conducted using Copernicus data and information funded by the European Union. The work has been prevalently carried out within the Copernicus Initial Operations 2011-2013, Land Monitoring Service Local Component: Riparian Zones, Framework Service Contract EEA/MDI/14/001, awarded by the European Environment Agency to a consortium composed of GAF AG (lead), INDRA Sistemas S. A., METRIA AB, and GeoVille GmbH. Data for validation were kindly provided by Dr. J. Barquin (Cantabria University, Spain), Prof. T. Ferreira and Dr. F. Aguiar (University of Lisbon). We acknowledge the Spanish Ministry of Science and Innovation in relation to project MARCE and all involved institutions for making available RHS data. JRC's input of Flood Hazard maps is greatly acknowledged. The contribution of David Herrmann (GAF AG) in the early stage of model implementation is highly appreciated. We acknowledge the contribution of Karin Larsson, METRIA AB, for her support in accessing *in-situ* data.

Author Contributions: Christof J. Weissteiner is the principal author of this manuscript. He wrote and revised the manuscript, conceived the model design and analysed the results (indicators and statistics). Martin Ickerott, Hannes Ott, Markus Probeck and Gernot Ramminger were the main responsibles for the technical implementation and contributed to model design and writing. Nicola Clerici contributed to the validation of results and writing the article. Ana Maria Ribeiro de Sousa contributed to manuscript writing, Hans Dufourmont contributed to the conceptual model.

Conflicts of Interest: The authors declare no conflict of interest. The opinions expressed in this article are those of the authors and do not reflect official positions of the European Union or the European Environment Agency.

References

1. MA *Millenium Ecosystem Assessment. Ecosystems and Human Well-being: Synthesis*; Island Press: Washington, D.C., 2005.
2. Malanson, G. P. *Riparian Landscapes*; Cambridge Studies in Ecology; Cambridge University Press, 1996.
3. Naiman, R. J.; Decamps, H.; McClain, M. *Riparia - Ecology, Conservation, and Management of Streamside Communities*; Academic Press, 2005.
4. Welsch, D. J. *Riparian Forest Buffers - Function for Protection and Enhancement of Water Resources*; [Broomall, PA:] U.S. Dept. of Agriculture, Forest Service, Northern Area State & Private Forestry, 1991.
5. Naiman and, R. J.; Décamps, H. THE ECOLOGY OF INTERFACES:Riparian Zones. *Annu. Rev. Ecol. Syst.* **1997**, *28*, 621–658.
6. Naiman, R. J.; Decamps, H.; Pollock, M. The Role of Riparian Corridors in Maintaining Regional Biodiversity. *Ecol. Appl.* **1993**, *3*, 209–212.
7. Nilsson, C.; Grelsson, G. The fragility of ecosystems: A review. *J. Appl. Ecol.* **1995**, *32*, 677–692.
8. Doppelt, B.; Scurlock, M.; Frissell, C.; Karr, J. *Entering the Watershed: A New Approach to Save American River Ecosystems*; Island Press: Washington, D.C., 1993.
9. Tockner, K.; Stanford, J. A. Riverine flood plains: Present state and future trends. *Environ. Conserv.* **2002**, *29*, 308–330.
10. European Commission Our life insurance, our natural capital: an EU biodiversity strategy to 2020. COM (2011) 244. Brussels. 2011.
11. Benedict, M. A.; McMahon, E. T. *Green Infrastructure: Linking Landscapes and Communities*; IslandPress: Washington, D.C., 2006.
12. Maes, J.; Liquete, C.; Teller, A.; Erhard, M.; Paracchini, M. L.; Barredo, J. I.; Grizzetti, B.; Cardoso, A.; Somma, F.; Petersen, J.-E.; Meiner, A.; Gelabert, E. R.; Zal, N.; Kristensen, P.; Bastrup-Birk, A.; Biala, K.; Piroddi, C.; Egoh, B.; Degeorges, P.; Fiorina, C.; Santos-Martín, F.; Naruševičius, V.; Verboven, J.; Pereira, H. M.; Bengtsson, J.; Gocheva, K.; Marta-Pedroso, C.; Snäll, T.; Estreguil, C.; San-Miguel-Ayanz, J.; Pérez-Soba, M.; Grêt-Regamey, A.; Lillebø, A. I.; Malak, D. A.; Condé, S.; Moen, J.; Czúcz, B.; Drakou, E. G.; Zulian, G.; Lavalle, C. An indicator framework for assessing ecosystem services in support of the EU Biodiversity Strategy to 2020. *Ecosyst. Serv.* **2016**, *17*, 14–23.

13. European Commission Directive 2014/89/EU of the European Parliament and of the Council of 23 July 2014 establishing a framework for maritime spatial planning 2014.
14. European Commission Green Infrastructure (GI) — Enhancing Europe's Natural Capital 2013.
15. European Commission Directive 2009/147/EC of the European Parliament and of the Council of 30 November 2009 on the conservation of wild birds 2010.
16. European Commission Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks 2007.
17. European Commission Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy 2000.
18. European Commission Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora 1992.
19. European Commission A Blueprint to Safeguard Europe's Water Resources 2012.
20. Verry, E. S.; Dolloff, C. A.; Manning, M. E. Riparian ecotone: A functional definition and delineation for resource assessment. *Water Air Soil Poll* **2004**, *4*, 67–94.
21. Naiman, R. J.; Decamps, H.; McClain, M. *Riparia - Ecology, Conservation, and Management of Streamside Communities*; Academic Press, 2005.
22. Naiman, R. J.; Décamps, H. The ecology of interfaces: Riparian zones. *Annu. Rev. Ecol. Syst.* **1997**, *28*, 621–658.
23. Naiman, R. J.; Fetherston, K. L.; McKay, S.; Chen, J. Riparian forests. In *River Ecology and Management. Lessons from the Pacific Coastal Region*; Springer: New York, 1997.
24. Naiman, R. J.; Décamps, H.; Pollok, M. The role of riparian corridors in maintaining regional biodiversity. *Ecol. Appl.* **1993**, *3*, 309–313.
25. Clerici, N.; Weissteiner, C. J.; Paracchini, M. L.; Boschetti, L.; Baraldi, A.; Strobl, P. Pan-European distribution modelling of stream riparian zones based on multi-source Earth Observation data. *Ecol. Indic.* **2013**, *24*, 211–223.
26. Clerici, N.; Weissteiner, C. J.; Paracchini, M. L.; Strobl, P. *Riparian zones: Where green and blue networks meet*; European Community, D. G. Joint Research Centre: Luxembourg, 2011; p. 60.
27. Weissteiner, C. J.; Pistocchi, A.; Marinov, D.; Bouraoui, F.; Sala, S. An indicator to map diffuse chemical river pollution considering buffer capacity of riparian vegetation — A pan-European case study on pesticides. *Sci. Total Environ.* **2014**, *484*, 64–73.
28. Weissteiner, C. J.; Bouraoui, F.; Aloe, A. Reduction of nitrogen and phosphorus loads to European rivers by riparian buffer zones. *Knowl Managt Aquat. Ecosyst* **2013**, 08.
29. Gumbricht, T. *Mapping global tropical wetlands from earth observing satellite imagery*; CIFOR: Bogor, Indonesia, 2012.
30. Abood, S.; Maclean, A. Modeling riparian zones utilizing DEMs, flood height data, digital soil data and National Wetland Inventory VIA GIS. In *ASPRS Annual Conference Proceedings*; Milwaukee, Wisconsin, 2011; pp. 1–11.
31. Akasheh, O. Z.; Neale, C. M. U.; Jayanthi, H. Detailed mapping of riparian vegetation in the middle Rio Grande River using high resolution multi-spectral airborne remote sensing. *J. Arid Environ.* **2008**, *72*, 1734–1744.
32. Aksoy, S.; Akcay, H. G.; Wassenaar, T. Automatic Mapping of Linear Woody Vegetation Features in Agricultural Landscapes Using Very High Resolution Imagery. *Geosci. Remote Sens. IEEE Trans. On* **2010**, *48*, 511–522.
33. Collins, J. N.; Sutula, M.; Stein, E.; Odaya, M.; Zhang, E.; Larned, K. *Comparison of Methods to Map California Riparian Areas*; San Francisco Estuary Institute: San Francisco, CA, USA, 2006.
34. Johansen, K.; Phinn, S.; Witte, C. Mapping of riparian zone attributes using discrete return LiDAR, QuickBird and SPOT-5 imagery: Assessing accuracy and costs. *Remote Sens. Environ.* **2010**, *114*, 2679–2691.

35. Muller, E. Mapping riparian vegetation along rivers: old concepts and new methods. *Aquat. Bot.* **1997**, *58*, 411–437.
36. Sutula, M.; Stein, E. D.; Inlander, E. *Evaluation of a Method to cost-Effectively Map Riparian Areas in Southern California Coastal Watersheds*; Southern California Coastal Water Research Project: Westminster, CA, USA, 2006.
37. Verry, E. S.; Dolloff, C. A.; Manning, M. E. Riparian ecotone: A functional definition and delineation for resource assessment. *Water Air Soil Poll* **2004**, *4*, 67–94.
38. X. Yang Integrated use of remote sensing and geographic information systems in riparian vegetation delineation and mapping. *Int. J. Remote Sens.* **2007**, *28*, 353–370.
39. Goetz, S. J. Remote sensing of riparian buffers: Past progress and future prospects. *J. Am. Water Resour. Assoc.* **2006**, *42*, 133–143.
40. Zadeh, L. A. Fuzzy Sets. *Inf. Control* **1965**, *8*, 338–353.
41. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16.
42. Rouse, J. W.; Haas, R. H.; Schell, J. A.; Deering, D. W.; Harlan, J. C. *Monitoring the vernal advancement of retrogradation of natural vegetation*; NASA/GSFC: Greenbelt, MD, 1974.
43. Gao, B. C. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote SensEnv* **1996**, *58*, 257–266.
44. Metzger, A.; Bunce, B.; Jongman, R. H. G.; Mucher, C. A.; Watkins, J. W. A climatic stratification of the environment of Europe. *Glob. Ecol. Biogeogr.* **2005**, *14*, 549–563.
45. Olson, D. M.; Dinerstein, E.; Wikramanayake, E. D.; Burgess, N. D.; Powell, G. V. N.; Underwood, E. C.; D'Amico, J. A.; Itoua, I.; Strand, H. E.; Morrison, J. C.; Loucks, C. J.; Allnutt, T. F.; Ricketts, T. H.; Kura, Y.; Lamoreux, J. F.; Wettengel, W. W.; Hedao, P.; Kassem, K. R. Terrestrial ecoregions of the world: A new map of life on Earth. *BioScience* **2001**, *51*, 933–938.
46. Strahler, A. N. Hypsometric (area-altitude) analysis of erosional topology. *Geol. Soc. Am. Bull.* **1952**, *63*, 1117–1142.
47. Alfieri, L.; Salamon, P.; Bianchi, A.; Neal, J.; Bates, P.; Feyen, L. Advances in pan-European flood hazard mapping. *Hydrol. Process.* **2014**, *28*, 4067–4077.
48. European Commission Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) 2007.
49. Soille, P. (Ed) *The Image-2006 mosaic project*; JRC Scientific and Technical Reports; European Commission: Luxembourg, 2008.
50. Alfieri, L.; Salamon, P.; Bianchi, A.; Neal, J.; Bates, P.; Feyen, L. Advances in pan-European flood hazard mapping. *Hydrol. Process.* **2014**, *28*, 4067–4077.
51. Van Der Knijff, J. M.; Younis, J.; De Roo, A. P. J. LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation. *Int. J. Geogr. Inf. Sci.* **2008**, *24*, 189–212.
52. Jarvis, A.; Reuter, H. I.; Nelson, A.; Guevara, E. Hole-filled seamless SRTM data V4 2008.
53. Vogt, J.; Foisneau, S. *CCM River and Catchment Database version 2.0. Analysis Tools*; European Commission, D. G. Joint Research Centre: Luxembourg, 2007; p. 22 p.
54. FAO/IIASA/ISRIC/ISS-CAS/JRC *Harmonized World Soil Database Version 1.0*; FAO; IIASA: Rome, Italy; Laxenburg, Austria, 2008; p. 37 p.
55. European Commission European Soil Database (vs 2.0) 2004.
56. EEA, E. E. A. *CS-3/17 (D2.0) Final nomenclature guideline*; European Environment Agency (EEA): Copenhagen, 2015.
57. EEA, E. E. A. *Copernicus Land Monitoring Services Validation of Riparian Zones Land Cover / Land Use (LCLU) product - Summary of preliminary results*; EEA: Copenhagen, 2015.
58. Dubois, D.; Prade, H. A review of fuzzy set aggregation connectives. *Inf. Sci.* **1985**, *36*, 85–121.

Peer-reviewed version available at *Remote Sens.* **2016**, *8*, 925; doi:10.3390/rs8110925

59. Stroppiana, D.; Boschetti, M.; Brivio, P. A.; Carrara, P.; Bordogna, G. A fuzzy anomaly indicator for environmental monitoring at continental scale. *Ecol. Indic.* **2009**, *9*, 92–106.
60. Yager, R. R. A framework for multi-source data fusion. *Soft Comput. Data Min.* **2004**, *163*, 175–200.
61. Dujmović, J. J.; Larsen, H. L. Generalized conjunction/disjunction. *Spec. Sect. Aggreg. Oper.* **2007**, *46*, 423–446.
62. Dujmović, J. J.; Nagashima, H. LSP method and its use for evaluation of Java IDEs. *Aggreg. Oper. Decis. Model.* **2006**, *41*, 3–22.
63. Weissteiner, C. J.; Boschetti, M.; Böttcher, K.; Carrara, P.; Bordogna, G.; Brivio, P. A. Spatial explicit assessment of rural land abandonment in the Mediterranean area. *Glob. Planet. Change* **2011**, *79*, 20–36.
64. Beven, K. J.; Kirkby, J. J.; Seibert, J. A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* **1979**, *24*, 43–69.
65. Boehner, J.; Koethe, R.; Conrad, O.; Gross, J.; Ringeler, A.; Selige, T. *Soil Regionalisation by Means of Terrain Analysis and Process Parameterisation*; In: Micheli, E., Nachtergaele, F., Montanarella, L. (Eds.); European Soil Bureau: Luxembourg (Luxembourg), 2001; pp. pp. 313–222.
66. Boehner, J.; Selige, T. Spatial prediction of soil attributes using terrain analysis and climate regionalisation. In *In: Boehner, J., McCloy, K. R., Strobl, J. (Eds.): SAGA - Analysis and Modelling Applications*; Goettinger Geographische Abhandlungen; Goettingen, Germany, 2006; pp. 13–28.
67. *Google Earth*; Google, 2016.
68. Pekkarinen, A.; Reithmaier, L.; Strobl, P. Pan-European forest/non-forest mapping with Landsat ETM+ and CORINE Land Cover 2000 data. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 171–183.
69. Spiegel, M. R. *Statistics*; McGraw-Hill.; New York, 1961.
70. Eurostat *LUCAS 2009 (land Use/Cover Area Frame Survey). Land Use and Land Cover: Nomenclature. Version 16/03/2009.*; Eurostat: Luxembourg (Luxembourg), 2009.
71. Raven, P. J.; Holmes, N. T. H.; Dawson, F. H.; Everard, M. Quality assessment using River Habitat Survey data. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **1998**, *8*, 477–499.
72. Thomlinson, J. R.; Bolstad, P. V.; Cohen, W. B. Coordinating Methodologies for Scaling Landcover Classifications from Site-Specific to Global: Steps toward Validating Global Map Products. *Remote Sens. Environ.* **1999**, *70*, 16–28.

Supplementary data

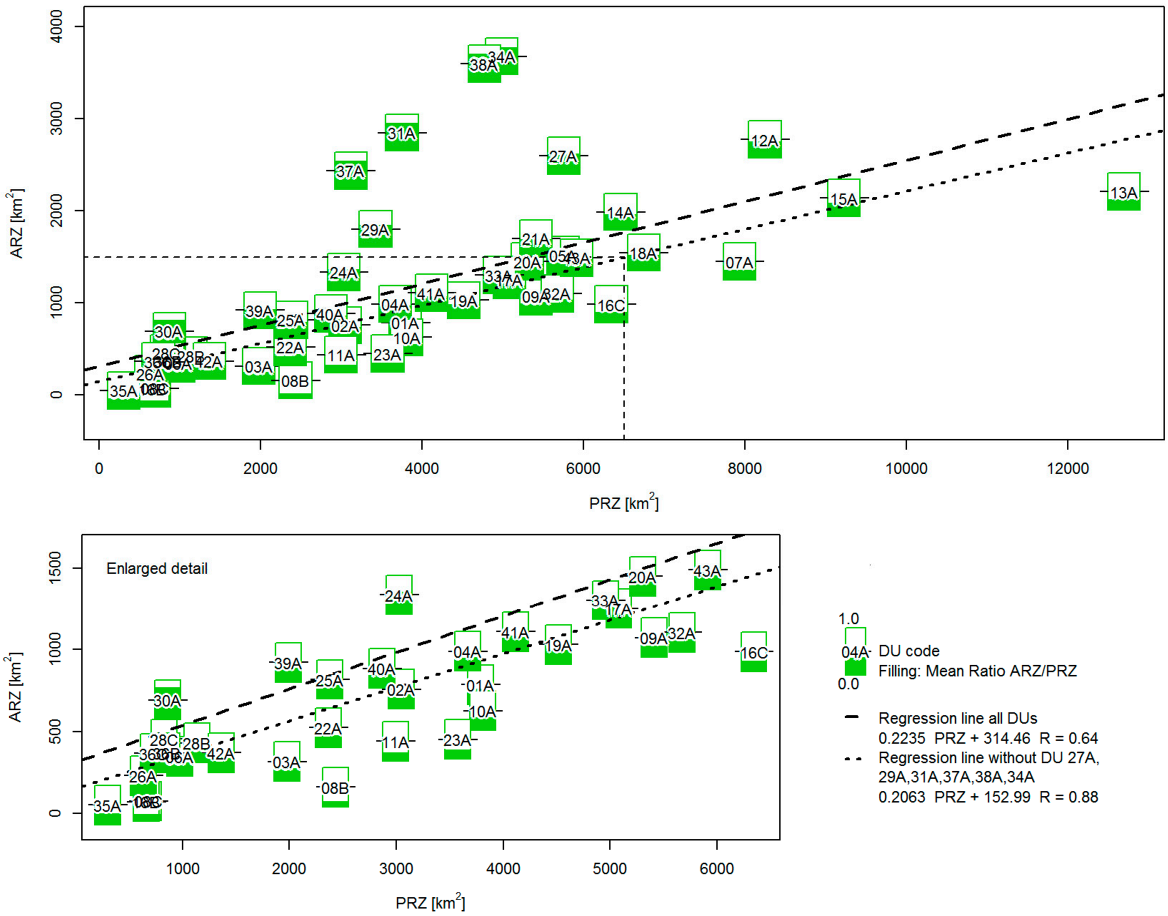

Figure S1: DU-wise comparison between ARZ and PRZ. Due to extra-ordinary high ARZ abundance for some Nordic DUs (27A, 29A, 31A, 37A, 38A, 34A), regression lines were calculated with/without them. The small graph displays the indicated detail.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 August 2016 doi:10.20944/preprints201608.0059.v1

Peer-reviewed version available at *Remote Sens.* **2016**, *8*, 925; doi:10.3390/rs8110925

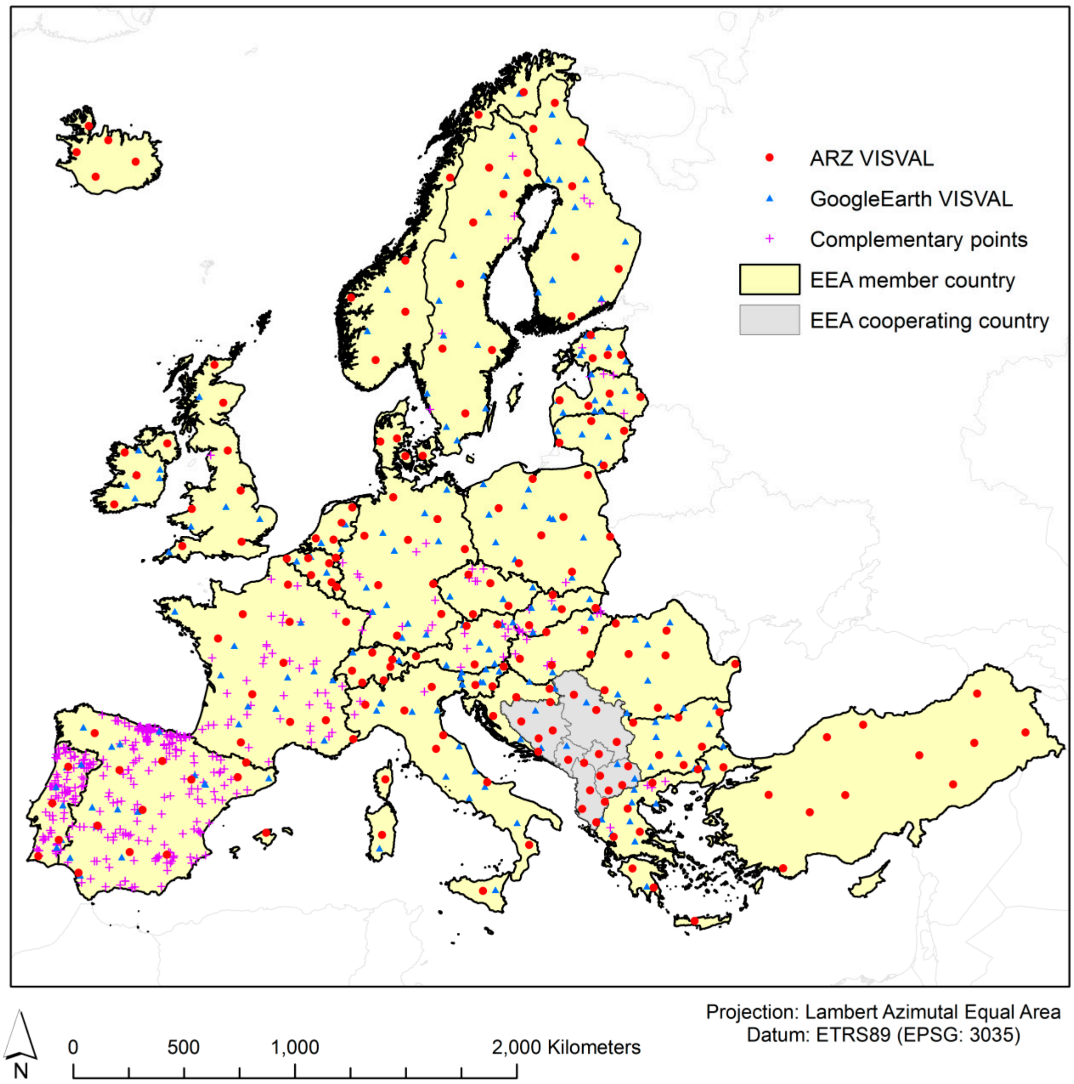

Figure S2: Spatial distribution of the visual validation points (VISVAL) used in the accuracy assessment.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 August 2016 doi:10.20944/preprints201608.0059.v1

Peer-reviewed version available at *Remote Sens.* **2016**, *8*, 925; doi:10.3390/rs8110925

Table S1: Nomenclature of the Land Cover and Land Use product, including the highest level of class discrimination at level 4 (level 1 is compatible with the MAES ecosystem types), and MS degree assigned to each class.

Level 1	Level 2	Level 3	Level 4	MS assigned
1 Urban	1.1 Urban fabric, industrial, commercial, public, military and private units	1.1.1 Dense to medium dense urban fabric (IM.D. >30-100% + industrial, commercial, public, military and private units)	1.1.1.1 Continuous urban fabric (in-situ based or IM.D. >80-100%	0
			1.1.1.2 Dense urban fabric (IM.D. >30-80% + industrial, commercial, public, military and private units)	0
			1.1.1.3 Industrial or commercial units	0
		1.1.2 Low density urban fabric (IM.D. 0-30%)	1.1.2.1 Low density urban fabric (IM.D. 0-30%)	0
	1.2 Transport infrastructure	1.2.1 Transport infrastructure	1.2.1.1 Road networks and associated land	0
			1.2.1.2 Railways and associated land	0
			1.2.1.3 Port areas	0
			1.2.1.4 Airports	0
	1.3 Mineral extraction, dump and construction sites, land without current use	1.3.1 Mineral extraction, dump and construction sites	1.3.1.1 Mineral extraction, dump and construction sites	0
		1.3.2 Land without current use	1.3.2.1 Land without current use	0
	1.4 Green urban, sports and leisure facilities	1.4.1 Green urban areas	1.4.1.1 Green urban areas T.C.D. ≥ 30%	0.5
			1.4.1.2 Green urban areas T.C.D. < 30%	0.5
		1.4.2 Sports and leisure facilities	1.4.2.1 Sports and leisure facilities T.C.D. ≥ 30%	0.5
			1.4.2.2 Sports and leisure facilities T.C.D. < 30%	0.5
2 Croplands	2.1 Arable land	2.1.1 Non-irrigated arable land	2.1.1.1 Non-irrigated arable land	0
		2.1.2 Greenhouses	2.1.2.1 Greenhouses	0
		2.1.3 Irrigated arable land and rice fields	2.1.3.1 Irrigated arable land and rice fields	0
		2.1.4 Complex patterns of irrigated and non-irrigated arable land	2.1.4.1 Complex patterns of irrigated and non-irrigated arable land	0
	2.2 Permanent crops	2.2.1 Vineyards	2.2.1.1 Vineyards	0
		2.2.2 Fruit trees and berry plantations	2.2.2.1 High stem fruit trees (extensively managed)	0
			2.2.2.2 Low stem fruit trees and berry plantations	0
		2.2.3 Olive groves	2.2.3.1 Olive groves	0
	2.3 Heterogeneous agricultural area	2.3.1 Annual crops associated with permanent crops	2.3.1.1 Annual crops associated with permanent crops	0.5
		2.3.2 Complex cultivation patterns	2.3.2.1 Complex cultivation patterns	0
		2.3.3 Land principally occupied by agriculture with significant areas of natural vegetation	2.3.3.1 Land principally occupied by agriculture with significant areas of natural vegetation	0.5
		2.3.4 Agro-forestry T.C.D. ≥ 30%	2.3.4.1 Agro-forestry T.C.D. ≥ 30%	0.5
		2.3.5 Agro-forestry T.C.D. < 30%	2.3.5.1 Agro-forestry T.C.D. < 30%	0.5
3 Woodland and forest	3.1 Broadleaved forest	3.1.1 Riparian and fluvial Broadleaved forest	3.1.1.1 Riparian and fluvial broadleaved forest	1
		3.1.2 Broadleaved swamp forest	3.1.2.1 Broadleaved swamp forest	1
		3.1.3 Other natural & semi natural broadleaved forest	3.1.3.1 Other natural & semi natural broadleaved forest	0.5
		3.1.4 Broadleaved evergreen forest	3.1.4.1 Broadleaved evergreen forest	0.5
		3.1.5 Highly artificial broadleaved plantations	3.1.5.1 Highly artificial broadleaved plantations	0.5
	3.2 Coniferous forest	3.2.1. Riparian and fluvial coniferous forest	3.2.1.1 Riparian and fluvial coniferous forest	1
		3.2.2 Coniferous swamp forest	3.2.2.1 Coniferous swamp forest	1
		3.2.3 Other natural & semi natural coniferous forest	3.2.3.1 Other natural & semi natural coniferous forest	0.5
		3.2.4 Highly artificial coniferous plantations	3.2.4.1 Highly artificial coniferous plantations	0.5
	3.3 Mixed forest	3.3.1.Riparian and fluvial mixed forest	3.3.1.1 Riparian and fluvial mixed forest	1
		3.3.2 Mixed swamp forest	3.3.2.1 Mixed swamp forest	1
		3.3.3 Other natural & semi natural mixed forest	3.3.3.1 Other natural & semi natural mixed forest	0.5
		3.3.4 Highly artificial mixed plantations	3.3.4.1 Highly artificial mixed plantations	0.5
	3.4 Transitional woodland scrub	3.4.1 Transitional woodland scrub	3.4.1.1 Transitional woodland and scrub	0.5
			3.4.1.2 Lines of trees and scrub	0.5
	3.5 Damaged forest	3.5.1 Damaged forest	3.5.1.1 Forest damaged by fire	0.5
			3.5.1.2 Other damaged forest	0.5

Level 1	Level 2	Level 3	Level 4	MS assigned
4 Grassland	4.1 Managed grassland	4.1.1 Managed grassland	4.1.1.1 Managed grasslands with trees and scrubs (T.C.D. ≥ 30%)	0
			4.1.1.2 Managed grasslands without trees and scrubs (T.C.D. < 30%)	0
	4.2 Natural grasslands	4.2.1 Natural grasslands prevailingly with trees and scrubs	4.2.1.1 Dry grasslands with trees (T.C.D. > 30%)	0
			4.2.1.2 Mesic grasslands with trees (T.C.D. > 30%)	1
		4.2.2 Natural grasslands without trees and scrubs	4.2.2.1 Dry grasslands without trees (T.C.D. < 30%)	0
			4.2.2.2 Mesic grasslands without trees (T.C.D. < 30%)	1
			4.2.2.3 Alpine and subalpine grasslands without trees (T.C.D. < 30%)	0.5
5 Heathland and scrub	5.1 Moors and heathland	5.1.1 Moors and heathland	5.1.1.1 Heathlands and Moorlands	0.5
			5.1.1.2 Other scrub land	0.5
	5.2 Sclerophyllous vegetation	5.2.1 Sclerophyllous vegetation	5.2.1.1 Sclerophyllous vegetation	0.5
6 Sparsely vegetated land	6.1 Sparsely vegetated areas	6.1.1 Sparsely vegetated areas	6.1.1.1 Sparsely vegetated areas	0.5
	6.2 Bare soil, rock, perennial snow & ice	6.2.1 Beaches, dunes, sands	6.2.1.1 Beaches	0.5
			6.2.1.2 Dunes	0
			6.2.1.3 River banks	1
		6.2.2 Bare rocks, burnt areas, glaciers and perpetual snow	6.2.2.1 Bare rocks and rock debris	0
			6.2.2.2 Burnt areas (except burnt forest)	0.5
			6.2.2.3 Glaciers and perpetual snow	0
7 Wetland	7.1 Inland marshes	7.1.1 Inland freshwater marshes	7.1.1.1 Inland freshwater marshes	1
		7.1.2 Inland saline marshes	7.1.2.1 Inland saline marshes	0.5
	7.2 Peat bogs	7.2.1 Peat bogs	7.2.1.1 Exploited peat bog	0.5
			7.2.1.2 Unexploited peat bog	0.5
8 Lagoons, coastal wetlands and estuaries	8.1 Maritime wetlands	8.1.1 Salt marshes & salines	8.1.1.1 Salt marshes	0.5
			8.1.1.2 Salines	0
		8.1.2 Intertidal flats	8.1.2.1 Intertidal flats	0.5
	8.2 Marine waters	8.2.1 Coastal lagoons	8.2.1.1 Coastal lagoons	0
		8.2.2 Estuaries	8.2.2.1 Estuaries	0
9 Rivers and lakes	9.1 Water courses	9.1.1 Interconnected running water courses	9.1.1.1 Permanent interconnected running water courses	0
			9.1.1.2 Intermittently running water courses	0
			9.1.1.3 Highly modified natural water courses and canals	0
		9.1.2 Separated water bodies belonging to the river system (dead side-arms, flood ponds)	9.1.2.1 Separated water bodies belonging to the river system (dead side-arms, flood ponds)	0
	9.2 Lakes and reservoirs	9.2.1 Lakes and reservoirs	9.2.1.1 Natural water bodies	0
			9.2.1.2 Ponds and lakes with completely man-made structure	0
			9.2.1.3 Intensively managed fish ponds	0
			9.2.1.4 Standing water bodies of extractive industrial sites	0
10 Marine (other)	10.1 Marine (other)	10.1.1 Marine (other)	10.1.1.1 Marine (other)	0

IM.D = Imperviousness Degree, T.C.D. = Tree Crown Density