Comparative analysis of CORINE and Climate Change Initiative Land Cover maps in Europe: implications for wildfire occurrence estimation at regional and local scales

L. Vilar[1]*, J. Garrido[2], P. Echavarría[1], J. Martínez-Vega[1], M.P. Martín[1]

[1]*, *Institute of Economics, Geography and Demography (IEGD-CCHS), Spanish National Research Council (CSIC), Albasanz 26-28, 28037, Madrid, Spain, lara.vilar@cchs.csic.es*
[2] *COTESA-Centre for space observation and remote sensing S.A.U. Antracita 7, 28045, Madrid, Spain*

Abstract: Updated and harmonized land cover (LC) data is essential for wildfire estimation in fire-prone areas as is the case in southern Europe. CORINE Land cover (CLC) and ESA Climate Change Initiative Land Cover (CCI-LC) maps have been analyzed and compared their performance in the estimation of wildfire occurrence in Europe at regional and local scales for the period 2010-2014. LC maps legends were harmonized and similarities and discrepancies were compared. Overall agreement between the two maps for the whole Europe was ~75%. Forest and agriculture showed the largest agreement, while shrubland and grassland the lowest. Quantity and allocation disagreements were calculated including exchange and shift components (Pontius Jr. and Santacruz, 2014) which provided detailed information about the contribution of each class to the overall disagreement. Spatial discrepancies were found in areas where grassland and shrubland were the dominant classes as in United Kingdom or East Turkey. Land Use and Coverage Area frame Survey (LUCAS) was used as ground truth for validation purposes. The agreement with LUCAS was slightly higher for CCI-LC (59%) than for CLC (56%). Generalized Linear Models (GLM), based on presence-absence of wildfires, were used to estimate wildfire occurrence at 3x3 km grid cell resolution from both LC maps at the European scale. LC interfaces and climatic variables (temperature and precipitation) where used as explicative variables while fires from European Forest Fire Information System EFFIS (2010-2014 period) were used as response variable. Wildfire occurrence was also estimated with the two maps at local scale in a test region (Zamora, Spain) using a more precise location of the response variable (x, y fire ignition points). At the European scale models obtained within the two maps showed similar results. CCI-LC model sensitivity was 77.26%, specificity 25.89% and omission error 22.74% while CLC model sensitivity was 75.68%, specificity 29.99% and omission error 24.32%. However, CLC performed slightly better in terms of the percent correct classification (62%). In the test region the models achieved better results in terms of specificity (66.07% and 68.93% for CCI-LC and CLC models respectively) and percent correct classification (~68% for CLC model). At local scale CLC model performed better than CCI-LC model. Wildfire occurrence estimation was more accurate at local scale because of the differences in the spatial accuracy of the response variable used.

1.   Introduction

Wildfires are one of the key natural hazards world-wide, as they impact the environment, destroy property and cause human lives losses. The Euro-Mediterranean region is particularly prone to wildfires with an increase in catastrophic episodes over the past decade (Pausas and Fernández-Muñoz, 2012). Recent events in Portugal (2017) and Greece (2018) caused together over 150 fatalities and near 500,000 ha of area burned (Lekkas et al., 2018). This increase in wildfire frequency and severity (Seidl et al., 2014) combined with progressive development of the wildland-urban interface (WUI) puts more communities at risk resulting in an increase in both the difficulty and cost of fire suppression. On the other hand, forest resources are a vital source of carbon storage, wildlife biodiversity and water conservation. Hence, managing and preserving forest from wildfires is a priority in conservation policies, including accurately monitoring, updating and managing of wildfire risk conditions. The estimation of those conditions needs the identification and mapping of the potential contributing factors at the appropriate spatial and temporal scale, and also their integration in wildfire predictive systems (Chuvieco et al., 2010).

Land use (LU), land cover (LC) and climate changes identified in the last decades are having a strong influence in fire cycles, contributing to the increase in fire size and frequency of large fires. Fuel characteristic (size, moisture, accumulation and arrangement) are changing due to climate and management transformations with a clear trend to foster favorable conditions both for fire ignition and fire propagation (Chuvieco, 2009; Pausas and Fernández-Muñoz, 2012). Accurate and updated information on land use and land cover (LULC) changes are essential for understanding past and future trends in wildfire occurrence, not only because LULC maps provide key data to understand the quantity, arrangement, status and spatial distribution of forest vegetation types, that could derive information on fuels characteristics (Lynch et al., 2015), but also because they can be used as a proxy of socio-economic factors related with human activities that can be considered potential ignition sources. As demonstrated by Gallardo et al. (2015) and Vilar et al. (2016a) the contact areas between specific LC classes or LC interfaces, can be used to represent various anthropogenic activities on the territory that may directly or indirectly relate with wildfire occurrence. LC data source used to derive the interfaces should be consistent in space and time so temporal and spatial variations can be unequivocally assigned to real changes and not to inconsistencies in the LC data sources. (Gallardo et al., 2015)

LC maps have been frequently used as an input for diverse global applications related to e.g. environmental analysis, climate change (Verburg et al., 2011), biodiversity assessment (Tsendbazar et al., 2016), food security (Fritz et al., 2010; Pérez-Hoyos et al., 2017), disaster management or emergency response (Li et al., 2017), etc. However, to our knowledge, no previous studies have compared different LC maps in the context of wildfire estimation and modeling. The CORINE programme was initiated in the European Union in 1985 and currently coordinated by the European Environment Agency (EEA https://www.eea.europa.eu/) in the framework of EU Copernicus programme (https://land.copernicus.eu/pan-european/corine-land-cover) produces LC maps for the whole Europe since 1990. Map´s updates were made available in 2000, 2006, 2012 and 2018 allowing multitemporal analysis of LULC changes at regional scales. For almost three decades, CLC has been widely used for diverse applications including landscape planning (Burkhard et al., 2009) and policy-making (Jokar Arsanjani et al., 2016). CORINE has been also used as LC source for modeling wildfire occurrence. In Martinez et al. (2009) the authors estimated wildfire risk in Spain including agricultural and forest classes as well as interfaces from CLC 1990 and CLC 2000. Rodrigues et al. (2014) included in their wildfire estimation models forest land cover derived from CLC 1990 and 2000 in Spain. Oliveira et al. (2012) used as LC data source the CLC 2000 to analyze fire occurrence in Mediterranean Europe. In Modugno et al. (2016) the variable Wildland Urban Interface (WUI) was obtained from CLC 2006 for mapping large fires patterns in Europe. Vilar et al. (2016a) modeled wildfires in two time points for southern European countries by using LC interfaces obtained from CLC 1990 and 2000.

ESA Climate Change Initiative (CCI) was developed to face climate change challenge at global level. Land cover (CCI-LC) information (Defourny et al., 2016) was derived as part of this initiative including yearly updated global LC maps (1992 to 2015) obtained from MERIS and SPOT satellite images at 300 m spatial resolution. CCI-LC also delivers 3-epoch series of global LC maps (1998-2002, 2003-2007, 2008-2012). Up to now, only few studies have used CCI-LC for wildfire estimation applications. Forkel et al. (2017) proposed a new flexible data-driven fire modeling approach to estimate burned area globally including LC classes from CCI-LC as part of the predictive dataset.

Previous works analyzed and compared CLC or CCI-LC maps with other global LC sets for various purposes as to perform accuracy assessments for other products or to check their usefulness for specific applications related to landscape, biodiversity or cropland monitoring. Pérez-Hoyos et al. (2012) compared global LC products CLC, GLC2000 (Bartholomé and Belward, 2005), MODIS Land Cover (Friedl et al., 2002) and GlobCover (Leroy et al., 2006) in Europe. These products were developed for different purposes and they had different technical characteristics. With the comparison and quality monitoring the authors aimed to assess their usefulness. They compared the LC maps by using Boolean and fuzzy theory approaches and obtaining overall accuracies ranging from 35 to 57% depending on the compared LC map pairs. Tsendbazar el at. (2017) integrated global Globcover 2009, CCI-LC 2010,

MODIS-2010 (Friedl et al., 2010) and Globeland30 (Chen et al., 2015) LC datasets within the aim to derive user-specific maps for characterizing mosaic classes for land system modeling and biodiversity assessment. They obtained an overall correspondence of 80% between the integrated global LC map with the reference LC sample sites. Pérez-Hoyos et al. (2017) compared the global datasets FAO-GLC share (Latham, 2014), Geowiki IIASA-Hybrid (See et al., 2015), Global Land Cover 2000, Global Land Cover by National Mapping Organizations (Tateishi et al., 2014), GlobCover, Globeland30, CCI-LC 2010 and 2015, and MODIS Land Cover product for cropland monitoring by using a Boolean approach. The study was global but the authors focused the analysis in countries with high risk of food insecurity, where crop monitoring is important for early warning. They found that the cropland information varied considerably depending on the LC analyzed source. FAO-GLC share and Globeland30 provided more adequate results. Regardless the large numbers of comparative assessments made for a variety of purposes only few works have compared CLC and CCI-LC datasets. Waldner et al. (2015) identified and collected national to global LC maps (including CLC and CCI-LC) in order to identify the priority areas for cropland mapping. Tsendbazar et al. (2014) found correspondences regarding the forest areas comparing CCI-LC with regional datasets i.e. CLC. Forest and urban areas of regional maps were mostly confused with croplands and mosaic croplands, shrublands, grassland and sparse vegetation classes of the global maps.

In this work CLC and CCI-LC data sources available for Europe were compared and their performance for wildfire occurrence estimation in Europe at two spatial scales was analyzed. The comparison was made using a Boolean approach (Herold et al., 2008) on a per pixel basis. Similarities and differences between both products for the year 2012 were analyzed in terms of quantity and allocation comparison. Land Use and Coverage Area frame Survey (LUCAS) available also for 2012 was used as reference data for validation. Wildfire estimation was accomplished at 3x3km grid cell resolution combining climatic variables and LC interfaces obtained from CLC and CCI-LC data sources by means of General Linear Models (GLM) (see methodological approach in Fig. 1). With this work we try to confirm or discard the hypothesis that differences between LC maps exist and how these differences can affect wildfire occurrence models at different scales

2.   Materials and methods

2.1.   *Study area*

The study area comprises European Economic Area (EEA) 39 countries where CLC information was available for 2012 (CLC12 v.18.5.1). It covers 6,075,176 km2 with ~609,646,000 total inhabitants (https://w3.unece.org/PXWeb/en). ~30% of the territory is covered by forest (woodlands) from which coniferous forest represents 48%, broadleaf forests 34% and mixed forest 17%. More than 30% of the territories are agriculture lands. Urban areas represent ~3%. Woodlands prevail in the northern parts of Europe where topography is dominated by mountains and hilly areas, i.e. woodlands in Finland or Sweden account for more than 60% of the area of these countries (also in Slovenia). On the contrary, in some countries ~50% of their total area is covered by croplands (Denmark or Hungary). Natural and agricultural grasslands dominate the landscape in Ireland (63%) and United Kingdom (>40%), while Malta and the Benelux have large urban areas in relation to their total territory (EEA, 2017). Fig.2 shows the study area and the LC maps CLC and CCI-LC from the year 2012 (CCI-LC12 and CLC12 from now on) used in this work. The area selected for the local analysis is Zamora, a Spanish province (NUTS3 level) located in north-west of Spain (black polygon in Fig. 2).

Wildfires affect mostly to southern Europe. Portugal, Spain, Italy, Greece and southern France are the five most affected countries/regions accounting for ~85% of the total annual burnt area each year (San-Miguel-Ayanz, 2012). In central and northern Europe wildfires are not so frequent but can occur in dry years. For instance, in the summer of 2018 a heat wave affected the northern hemisphere, leading to unusual drought conditions and high temperatures in north Europe. This led to an increase of fire breaks in some countries being e.g. three times more than average in Sweden. Recent studies based on temperature records concluded that human-caused climate change doubled the odds of 2018 summer's European heat wave (https://www.worldweatherattribution.org/). According to the European Fire Database (http://effis.jrc.ec.europa.eu/about-effis/technical-background/european-fire-database/) most fires in southern European countries are related with human activity due to deliberate actions (>50% of the known fires), negligence or accident (~35%) and only ~5 % are due to natural causes (usually dry storm lightning). In northern Europe the percentage of fires caused by natural causes is slightly higher (~8%). and human-caused fires are mostly related with accidental or negligent actions (>70%). In central Europe the causes are similar to southern countries and the incidence of natural fires is very low (<1%).
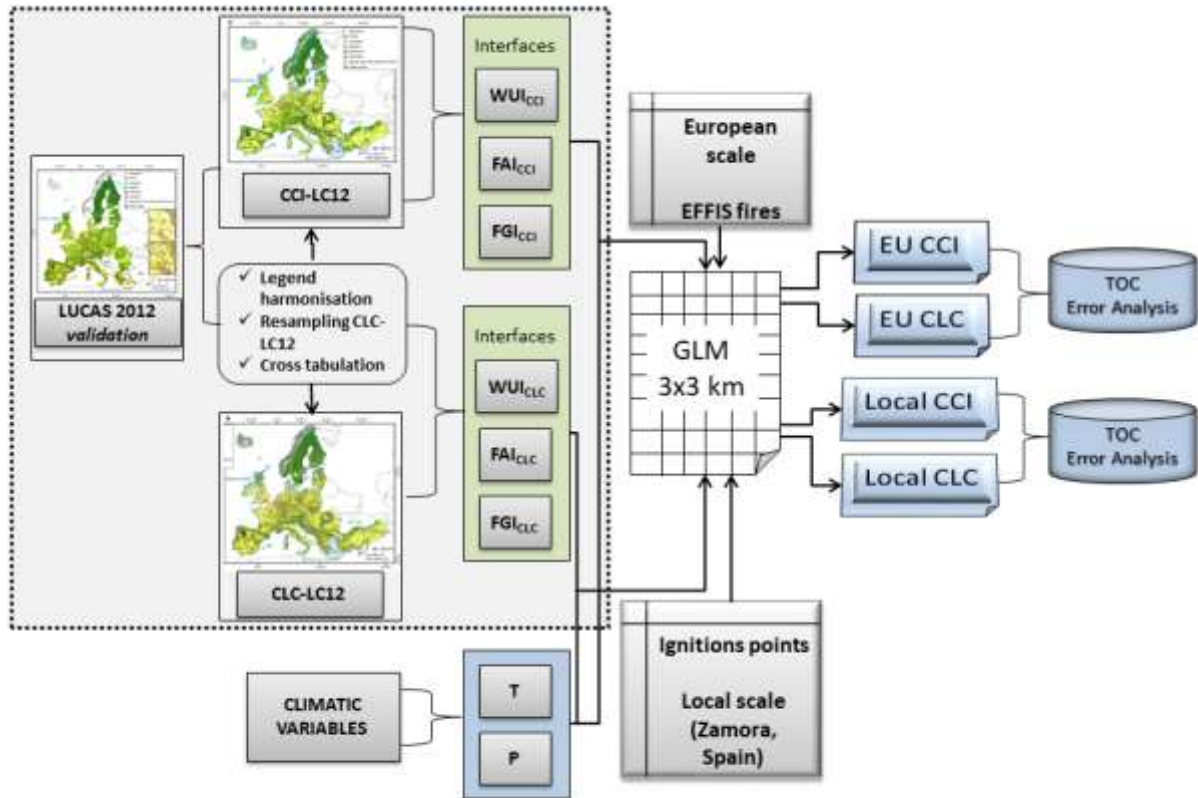
4

Fig. 1. Methodological approach explaining the steps followed in the comparison of CCI-LC12 and CLC12, its validation using LUCAS12 and estimation of wildfire occurrence at European and local scale from LULC interfaces and climatic variables at 3x3 km grid cell resolution using GLM. WUI, FAI and FGI stand for Wildland-Urban, Forest-Agricultural and Forest-Grassland Interfaces. TOC refers to Total Operating Characteristic.
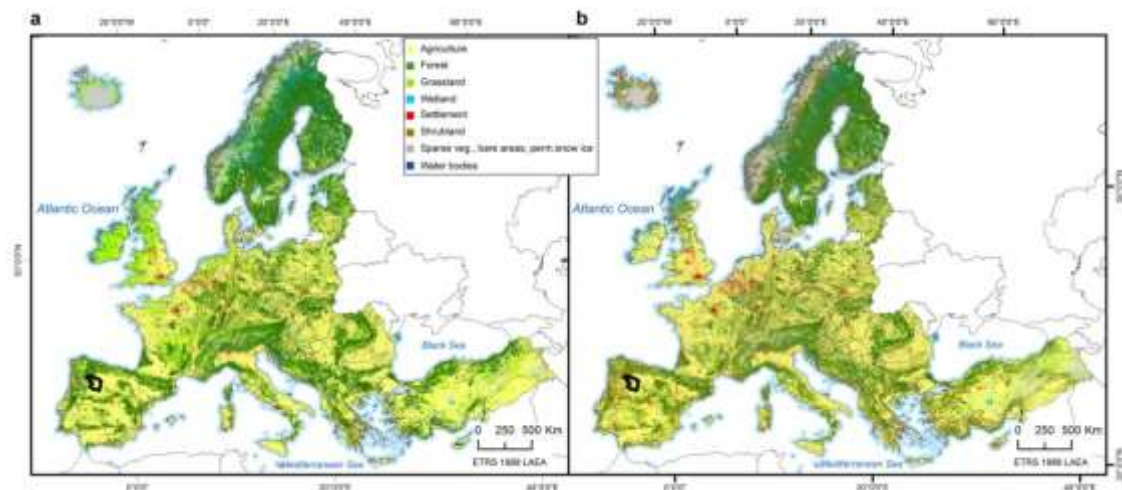

Fig. 2. Study area EEA39 countries. CCI-LC12 (a) and CLC12 (b) maps. Highlighted in black the limits of the local study site (Zamora, Spain).

2.2. Input Data

2.2.1. Land cover datasets

In this study CLC12 and CCI-LC12 maps were compared. Both products are based on Earth Observation (EO) data. CCI-LC12 uses as main input the global archive (2003-2012) of the MERIS-ENVISAT images (Bontemps et al., 2015) at 300 m resolution. Observations acquired by the SPOT Vegetation (SPOT-VGT) were used to extent the temporal coverage of the project over the period 1998-2002 (Defourny et al., 2016). MERIS surface reflectance archive data were pre-processed to correct for radiometric, geometric and atmospheric effects and as for cloud screening (Li et al., 2016). Supervised and unsupervised classification algorithms were combined in an automated procedure to derive LC classes (Defourny et al., 2016). A total of 22 classes were defined using United Nations Land Cover Classification System (UN-LCCS) (Di Gregorio et al., 2016).

CLC12 belongs to Copernicus land cover products (http://land.copernicus.eu/paneuropean/corine-land-cover/clc-2012). For this analysis CLC12 v.18.5.1 was used covering 39 EEA (Fig 2.). This product used high-resolution satellite image coverages (IRS ResourceSat-1/2, SPOT-4/5, RapidEye constellation) acquired between 2011 and 2012 as input data source to update CLC2006 following the "change mapping first approach" (Büttner et al., 2014). Computer Assisted Photointerpretation (CAPI) was the prevailing methodology applied to derive information from the satellite images. National teams started the analysis by identifying all changes larger than 5 ha in their respective countries. During this process the errors detected by the teams in CLC2006 map were corrected providing a CLC2006 revised dataset. Then CLC12 was produced by combining the CLC2006 revised dataset (v. 18.5) and the changes occurred between CLC2006 and CLC12. CLC12 is both available in raster (100 and 250 m resolution) and vector formats. Table 1 shows the characteristics of each LC product.

LUCAS 2012 database (http://ec.europa.eu/eurostat/web/lucas/data/primary-data/2012), based on statistical calculations that interpret observations in the field, was used for the available countries within the study area as the ground truth to compare with the two LC analyzed dataset. LUCAS uses a standardized survey methodology with a sampling plan, defined classifications, data collection processes and statistical approaches to obtain harmonized and unbiased estimates of LULC. The first survey was done in 2001 within the aim to provide early crop estimates for the European Commission. In 2006 the sampling methodology changed to extend the LC surveyed types and it was introduced a three-year survey frequency (2006, 2009, 2012). From 2006 the survey is a two-phase sampling method being the first phase a systematic ground sampling with points spaced 2 km in the four cardinal directions (~1 million points). Each point (circle of a 1.5 m of radius and also 20 m for specified classes) of this first phase sample is photo-interpreted and assigned to one of the seven pre-defined land cover classes (Martino et al., 2009). In LUCAS 2012 ~270,260 points were visited in situ covering EU-27 member states. The main information collected was related to LULC (with a classification comparable to other statistical standards i.e. those used by FAO), with 8 LC categories and 4 LU, (subdivided then in classes and subclasses), soil data and water management.

2.2.2. LC Interfaces

In this work, as in Gallardo et al. (2015) and Vilar et al. (2016a), we considered the interfaces between forest and other land covers, and we used them as indirect drivers of the anthropogenic activity on the territory that may cause a wildfire ignition. It has been reported that human activities common in some LC interfaces can lead to wildfire ignitions either by accident, negligence or deliberate actions (i.e. wildfires caused by using agricultural machinery, pasture burning, recreational activities close to urban areas, etc). For this analysis three LC interfaces were defined from both CCI-LC12 and CLC12 maps: Forest - Agricultural Interface (FAI), Forest - Grassland Interface (FGI) and Wildland - Urban Interface (WUI). As it will be further described in section 2.3, CLC12 was first resampled to 300 m for LC comparison with CCI-LC. From both maps the LC interfaces were obtained, defined as 1 pixel (300x300 m) to each side of the contact among uses that formed each interface type. LC interfaces were then referred to 3x3 km reference cell grid in this study. Interfaces were overlaid with the cell grid and the area occupied by each LC interface by cell was calculated in order to derive final density values by cell.

2.2.3. Climate data

Mean temperature (T) and precipitation (P) over the study region were obtained from the European land-only daily high-resolution gridded data set available in the so-called E-OBS data files platform (https://www.ecad.eu/download/ensembles/download.php). This data set was elaborated under the ENSEMBLES project framework (http://www.ensembles-eu.org/) and contains daily precipitation, minimum, maximum and mean surface temperature since 1950 at 0.25° by 0.25° and 0.5° by 0.5° on a regular latitude-longitude grid and 0.22° by 0.22° and 0.44° by 0.44° on a rotated pole grid (Van den Besselaar et al., 2011). The regular 0.25° grid was chosen for this analysis. Mean temperature and precipitation variables were referred to the 3x3 km reference cell grid as previously describe for the LULC interfaces (see previous section).

Table 1
CCI-LC12 and CLC12 and LUCAS12 main characteristics.

	CCI-LC12	CLC12
Spatial resolution	300 m	100, 250 m
Minimum Mapping Unit	5 ha	25 ha
Data source	MERIS FR[1]/RR[2] global SR composites	IRS P6 LISS III[4], SPOT and RapidEye; CLC2006
Date of satellite data	2012	2011-2012
Classification scheme/legend categories	UN-LCCS[3] based: 22 class	44 class on level 3
LC classification method	Unsupervised spatio-temporal clustering; machine learning classification	Computer Assisted Photointerpretation (CAPI)
Thematic accuracy (data provider)	~74%	$\geq$ 85%
Spatial coverage	Global	39 EEA countries
Reference	(Defourny et al., 2016) http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-PUG-v2.5.pdf	http://land.copernicus.eu/pan-european/corine-land-cover/clc-2012
Update frequency	Annually 1998-2015. 3-epoch series of 5-year periods (1998-2002, 2003-2007, 2008-2012)	1990-2000-2006-2012

[1]FR: Full Resolution (300m)
[2]RR: Reduce Resolution (1000m)
[3]UN-LCCS: United Nations Land Cover Classification System
[4]IRS P6 LISS III: European Spatial Agency (ESA) mission ResourceSat 23.5m spatial resolution (LISS III)
(https://earth.esa.int/web/guest/missions/3rd-party-missions/current-missions/resourcesat-1)

2.2.4. Wildfire data

Fire data used for the analysis at European scale was obtained from the Rapid Damage Assessment (RDA) product from EFFIS (http://effis.jrc.ec.europa.eu/) were fire perimeters of burnt areas for the period 2010-2014 were selected. This period contained the years were the LC maps were elaborated. A 5-year period seemed sufficient for gathering the needed fire data range for modeling. Other works have used varied time-periods e.g. 11 years (Vilar et al., 2016a), 8 years (Oliveira et al., 2012) or 4 years (Padilla and Vega-García, 2011). EFFIS monitors burnt area in Europe since de 90's, producing the first burnt area map in 2000 on the basis of IRS WiFS images (San-Miguel-Ayanz J., 2009). Nowadays the MODIS sensor is used for continuous monitoring and mapping fires of ~40ha or larger (San-Miguel-Ayanz, 2012). Fires are mapped using a semi-automatic procedure including a combination of band thresholds and ancillary information from CLC, the active fire detection product and a fire news application. At local level fire ignitions (x, y coordinates, 2010-2014 period) for the Zamora province (Spain) were used (General Directorate of Environment, Castile and Leon, Spain. https://www.jcyl.es/). Response variables were obtained as the presence-absence of fires by 3x3km grid cell.

Fig. 3 illustrates the response variable (wildfire presence-absence) at European scale from the EFFIS fire perimeters, the period 2010-2014 and referred to 3x3km grid cell. A zoom window in Fig. 3 also shows the response variable used for the local study, the fire ignition points (x, y) from the same time period referred to 3x3km grid cell.
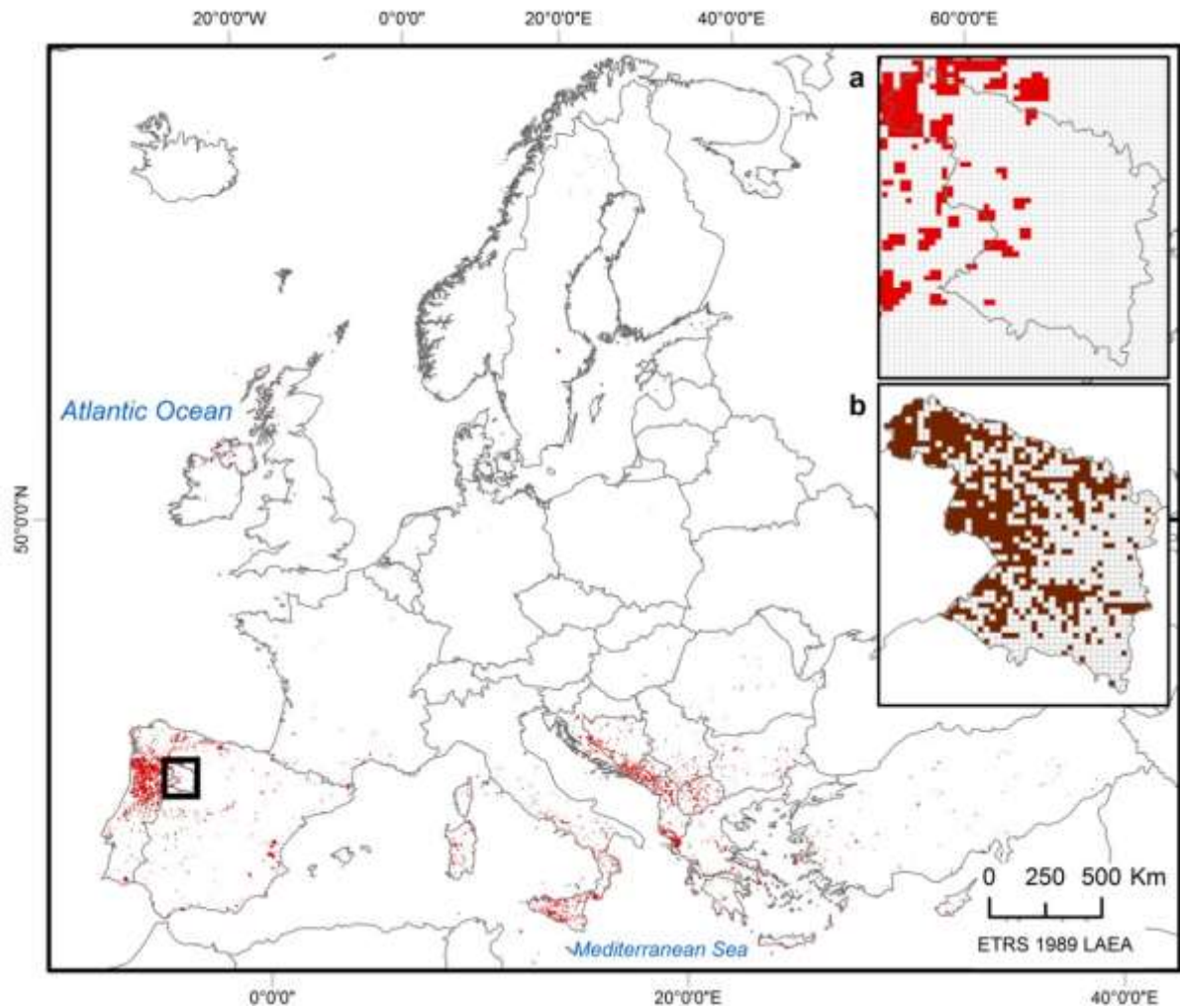
Fig. 3. Wildfire occurrence from the period 2010-2014 (EFFIS). Response variable for the analysis at European scale, presence (red)-absence of wildfire ignitions by 3x3km grid cells. (a) Window detail of the European scale response variable located in North-West Spain, province of Zamora; (b) Response variable at local scale (province of Zamora, Spain), presence (brown)-absence of wildfire ignitions (x, y) by 3x3km grid cells.

2.3. Land cover maps' comparison

As a first step, CCI-LC12 was cropped to the extent of CLC12 and CLC12 was resampled using nearest neighbor to match CCI-LC12 300 m resolution. This method uses the value of the closest pixel to assign to the output pixel value. The legends of the CCI-LC12 and CLC12 maps were harmonized into 8 common classes as shown in table 2. LUCAS12 cover classification was also adapted to the same classes (Table 2).

The legend aggregation and harmonization are crucial steps in map comparison (Pérez-Hoyos et al., 2017). In this study decisions were made both to aggregate and to find equivalences between the analyzed LC datasets and the final common legend.

CCI-LC12 and CLC12 maps were first contrasted using a non-site specific approach, so the total area of each LC category from both maps where compared without regard to their spatial location. After this analysis a site-specific comparison was applied by spatial overlay of the two maps. A cross-tabulation matrix was produced in order to quantify spatial agreement or disagreement between the two maps. CLC was taken as reference map. From this matrix overall spatial agreement (OA) was first derived by dividing the sum of the entries of the main diagonal by the total number of pixels taken (Story and Congalton, 1986). The off-diagonal values represent the assignation errors (Chuvieco, 2002). To determine the individual category agreements omission and commission errors were calculated: (1) omission of *X* errors were defined as the pixels that belong to a category *X* but were

classified in a different category. They were calculated by going down the columns for each category and summing the incorrect classifications and dividing them by the total number of pixels of the reference map for each category. In other words pixels that were not correctly classified as category *X* were omitted from the correct category; (2) commission of *X* errors were defined as the pixels that were predicted to be in a category *X* but did not belong to that category. They were calculated by going across the rows for each category, summing all the incorrect classifications and dividing them by the total number of classified pixels for each category.

Quantity and allocation disagreement, exchange and shift components were also calculated from the cross-tabulation matrix (Pontius Jr., 2019; Pontius Jr. and Millones, 2011). Quantity disagreement informs on the amount of difference between two maps that is due to the less than perfect match in the proportions of the categories (Equation 1). On the other hand, allocation disagreement provides the amount of difference between two maps that is due to the less than optimal match in the spatial allocation of the categories, given the category totals in the maps (Equation 2).

$$q_g = \left|\left(\sum_{i=1}^{J} p_{ig}\right) - \left(\sum_{j=1}^{J} p_{gj}\right)\right| \qquad (1)$$

$$a_g = 2min\left[\left(\sum_{i=1}^{J} p_{ig}\right) - p_{gg}, \left(\sum_{j=1}^{J} p_{gj}\right) - p_{gg}\right] \qquad (2)$$

Where the first summation in the equation 1 is the proportion of the category g in the reference map and the second summation in the proportion of the category g in the comparison map. In equation 2, first argument within the minimum function is the omission of category g and the second argument is the commission of category g. The multiplication by two and the minimum function are needed because allocation disagreement for category g comes in pairs, where commission of g is paired with omission of g, so the pairing is limited by the smaller of commission and omission (Pontius Jr. et al., 2004).

Exchange is a transition from category i to category j in some observations and a transition from j to i in an identical number of other observations. In other words, exchange exists for a pair of pixels when one pixel is classified as category i in the first map and as category j in the second map and, at the same time, the paired pixel is classified as category j in the first map and as a category i in the second map (Pontius Jr. and Santacruz, 2014). The total exchange for a category is the sum of all exchanges that involve that category. If there are more than two categories, shift can happen, being the allocation difference that is not exchange i.e. the difference after subtracting quantity difference and exchange from the overall difference. See Pontius Jr and Millones (2011) for further details. To obtain these quantity and allocation disagreement components diffeR v. 0.0-4 package was used (Pontius Jr. and Santacruz, 2015) for R (RCoreTeam, 2017).

For validation and accuracy assessment purposes LUCAS12 was used as reference ground truth data (Fig.

4). LUCAS12 points were overlaid to CCI-LC12 and CLC12. Then, cross-tabulation matrices between LUCAS12 point sampling and the two maps were produced, obtaining overall agreement, omission and commission errors, as well as quantity and allocation disagreement, exchange and shift components.

2.4. Modeling wildfire occurrence at European and local scales using CCI-LC12 and CLC12 maps

Wildfire occurrence for the year 2012 was modeled by GLM using LULC Interfaces derived from both CCI-LC12 and CLC12 maps plus climatic variables (mean temperature and precipitation) as predictors. EFFIS fires and x, y fire events were used as response variable (presence-absence) for the European and the local model respectively. Response variables were overlaid to LC classes and LC interfaces to explore differences in wildfire frequency per class/interface in both maps. Then models were calculated at 3x3km grid cell resolution. GLM are extensions of linear regression models that support dependent variables with non-normal distributions such as binomials (Guisan et al., 2002). The predictor variables Xj (j=1,…,p) are combined to produce a linear predictor LP which is related to the expected value μ=E(Y) of the response variable Y through a link function g() (Equation 3):

$$g\big(E(Y)\big) = LP = \alpha + X^T \beta \qquad (3)$$

where α is a constant called the intercept, $X=(X_1,…,X_p)$ is a vector of *p* predictors and $\beta=\{\beta_1,…, \beta_p\}$ is the vector of *p* regression coefficients (one by predictor). The distribution of *Y* used in this work was binomial with a logit link function.

As usually occurs in fire risk modeling studies, in this work the number of cells with absence of fire was substantially larger than the cells with presence of fire. To solve this unbalance a sample of the data from absence-fire cells is commonly used for model building (Preisler et al., 2004). As a consequence, a deterministic offset term of –log π is introduced into the model. π denotes the response-specific sampling rate. When π=1, π is also 1, and when π=0, π= π, the decided sampling rate for the non-fire cells. In this work a sample of 1% of the absence-fire cells was considered appropriate to retain enough covariate information on the non-ignitions so this offset was included into the model (further details in Preisler et al. (2004)). After sampling the non-fire cells, 75% of 16517 3x3 km cells randomly selected within the study region was used to calibrate the model and remaining 25% was used for validation purposes. Models were fit using car (Fox, 2011) mgcv package 1.8-22 (Wood, 2017) and ROCR (Sing et al., 2005) packages for R (RCoreTeam, 2017).

Table 2

Correspondence between original CCI-LC12, CLC12 and LUCAS12 legend categories and the harmonized legend.

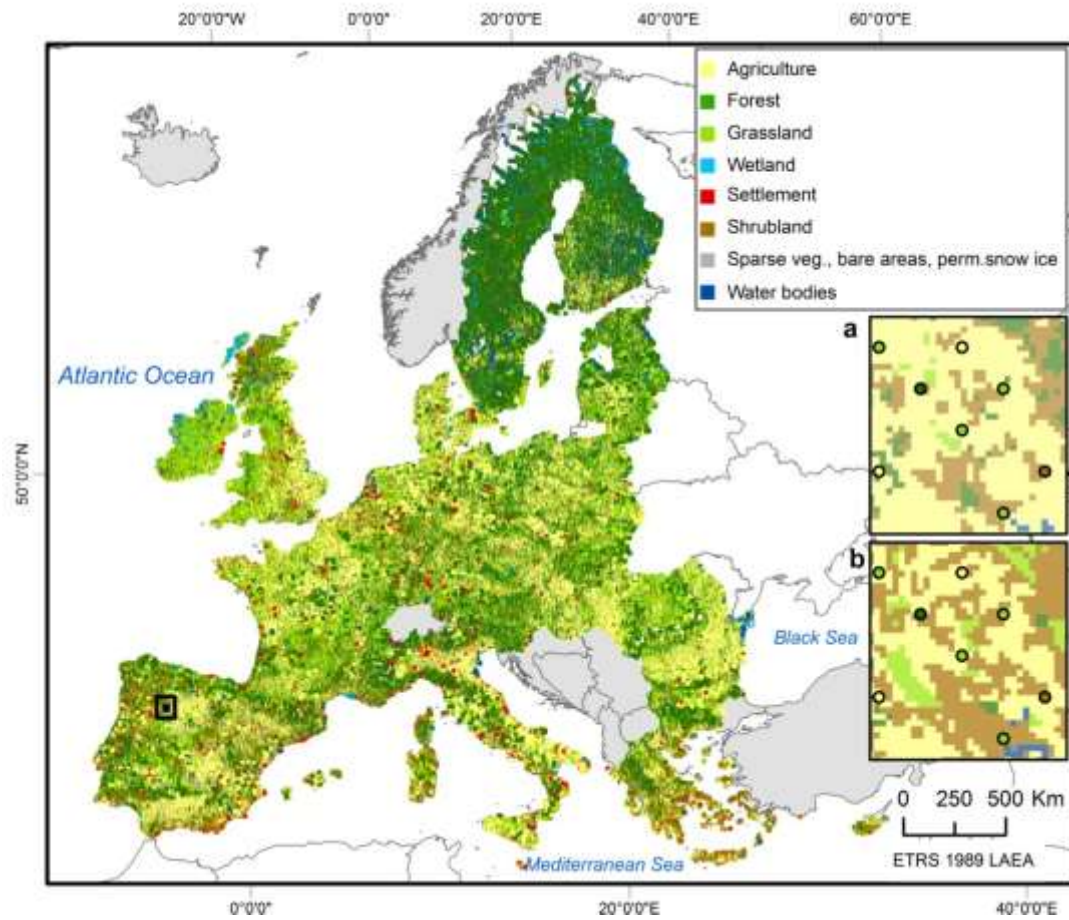| Code	Land cover class	CCI-LC12	CLC12	LUCAS12
1	Agriculture	Cropland, rainfed (10); Herbaceous cover (11), Tree or shrub cover (12)	Agricultural areas (2)	Cropland (B00)
		Cropland, irrigated or post-flooding (20)		
		Mosaic cropland (>50%)/ natural veg. (<50%) (30)		
		Mosaic natural veg. (>50%)/cropland (<50%) (40)		
2	Forest	Tree cover, broadleaved, evergreen, closed to open (>15%) (50)	Forest (3.1)	Woodland (C00)
		Tree cover, broadleaved, deciduous, closed/open (60) (61) (62)		
		Tree cover, needleleaved, evergreen, close/open (70) (71) (72)		
		Tree cover, needleleaved, deciduous, close/open (80) (81) (82)		
		Tree cover, mixed leaf type (90)		
		Mosaic tree and shrub (>50%)/hervaceous cover (<50%) (100)		
		Tree cover, flooded, fresh or brakish water (160)		
		Tree cover, flooded, saline water (170)		
3	Grassland	Mosaic herbaceous cover (>50%)/tree, shrub (<50%) (110)	Natural grassland (3.2.1)	Grassland (E00)
		Grassland (130)		
4	Wetland	Shrub or herbaceous cover, flooded, fresh/saline/brakish water (180)	Wetlands (4)	Wetland (H00)
5	Settlement	Urban areas (190)	Artificial surfaces (1)	Artificial land (A00)
6	Shrubland	Shrubland (120) • Evergreen shrubland (121) • Deciduous shrubland (122)	Moors and heathland (3.2.2)	Shrubland (D00)
			Sclerophyllous vegetation (3.2.3)	
			Transitional woodland shrub (3.2.4)	
7	Sparse vegetation, bare areas, permanent snow and ice	Lichens and mosses (140)	Open spaces with little or no vegetation (3.3)	Bare land/lichens and moss (F00)
		Sparse vegetation (tree, shrub, herb.) (<15%) (150) • Sparse shrub (<15%) (152) • Sparse herbaceous cover (<15%) (153)		
		Bare areas (200) • Consolidated Bare areas (201) • Unconsolidated bare areas (202)		
		Permanent snow and ice (220)		
8	Water bodies	Water bodies (210)	Water bodies (5)	Water areas (G00)

Fig. 4. LUCAS12 map (adapted classification). In light grey countries with no LUCAS data. Window detail of LUCAS sampling points located in North-West of Iberian Peninsula superimposed to (a) CCI-LC12 and (b) CLC12.

The lowest Akaike's Information Criterion (AIC) value selected the best model (Akaike, 1973). Adding independent variables step by step assessed their individual contribution to the model performance. Regression models assume that the predictor variables are not correlated. Therefore, multicollinearity analysis was applied before running GLM (Robinson and Schumacker, 2009) by means of Spearman correlation among variables and by Variance Inflation Factor (VIF). VIF can distinguish the degree of multicollinearity when variables are uncentered (Freund et al., 2003). Spearman correlations higher than 0.7 and/or VIF higher than 10 (Hair et al., 1995) confirmed multicollinearity and so the predictors were not included in the analysis.

To evaluate prediction accuracy of the models, Total Operating Characteristic (TOC) was used (Pontius Jr. and Si, 2014). As Receiver Operating Characteristic (ROC) (Fawcett, 2006) TOC compare the reference variable to the calculated variable by diagnosing each observation as either presence or absence. A threshold of 0.5 resolved the diagnosis in this work. If a predicted observation is greater than or equal to the threshold, then the observation is diagnosed as presence, absence otherwise (Pontius Jr. and Si, 2014). TOC allows having the information of the contingency table for all defined thresholds in a plot. The vertical distance from the horizontal axis to the TOC curve equals to "hits" while the vertical distance from the TOC and the "hits + misses" lines indicates "misses". The horizontal distance from the maximum boundary to the TOC curve indicates "false alarms" while the horizontal distance to the minimum boundary means "correct rejections" (see Pontius Jr. and Si, (2014) for further details). To obtain TOC curves *TOC* v. 0.0-4 package was used (Pontius et al., 2015) for R (RCoreTeam, 2017). The Area Under the Curve (AUC) metric was also calculated, which indicated the probability that a randomly chosen wildfire presence case exceeds the one of randomly choosing an absence. It ranges from 0 to 1, where larger AUCs indicate stronger positive association (Pontius Jr. and Parmentier, 2014). AUC=1.0 means that all wildfire presences have prediction values greater than the prediction values of all wildfire absences (perfectly correct); AUC=0.5 indicates a random association and AUC=0.0 means that all wildfire presences have prediction values that are less than the prediction values of all the wildfire absence observations (perfectly wrong). The classification percentage of the true wildfire presences (sensitivity) and true wildfire absences (specificity) as well as false wildfire absences (omission error) were also calculated.

3.   Results

3.1. Land cover maps' comparison

As it is shown in Fig. 2, agriculture and forest areas are mostly coincident in the two maps. However, some areas in CCI-LC12 are dominated by grassland while in CLC12 by agriculture or shrubland (e.g. UK or Ireland).

Fig. 5 shows the percentage of total surface covered by each of the 8 LC classes in the two maps and percentage of LUCAS sample points.

Agriculture, shrubland, sparse vegetation and settlement classes covered a larger area in CLC12 than in CCI-LC12 while the contrary occurs in forest and grassland. The most noticeable relative differences between maps happened in grassland, ~5 % and in shrubland, ~6 %. LUCAS12 percentage %. LUCAS12 percentage of sample points in agriculture, forest and grassland categories were larger than in the other LUCAS categories. The percentage of LUCAS sampling points in grassland was larger (>20%) than CCI-LC12 or CLC12 percentage in surface. Fig. 6 shows the spatial location of grassland and shrubland classes in the two maps. In CCI-LC12 grasslands were more abundant in United Kingdom and central Europe compared with CLC12. On the contrary, shrubland class was more abundant in CLC12 map, with large areas located in the northern and southern countries

Fig. 7 shows the spatial agreement and disagreement between CCI-LC12 and CLC12 after pixel comparison analysis (non-class specific) while table 3 shows the cross-tabulation matrix with the information per LC classes.
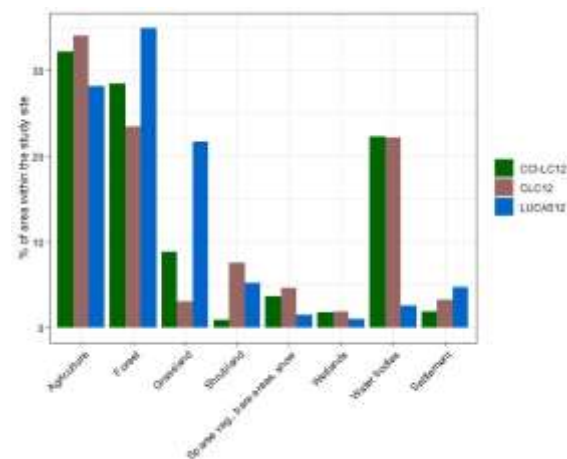

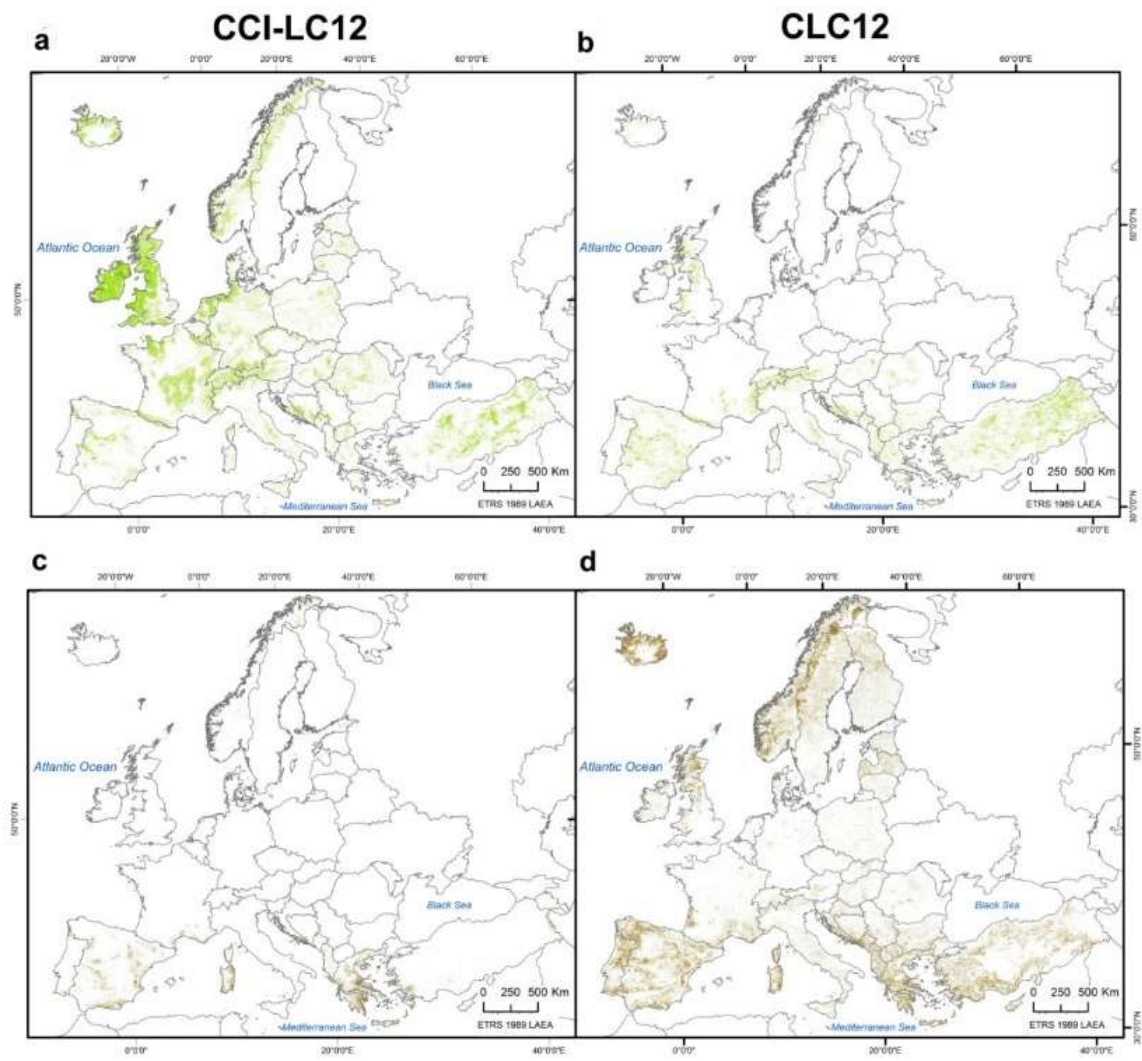**Fig. 5**. Percentage of LC by class for CCI-LC12, CLC12 maps and LUCAS12.

Fig. 6. Grassland and shrubland LC classes in CCI-LC12 and CLC12. (a) CCI-LC12 grassland (b) CLC12 grassland (c) CCI-LC12 shrubland (d) CLC12 shrubland
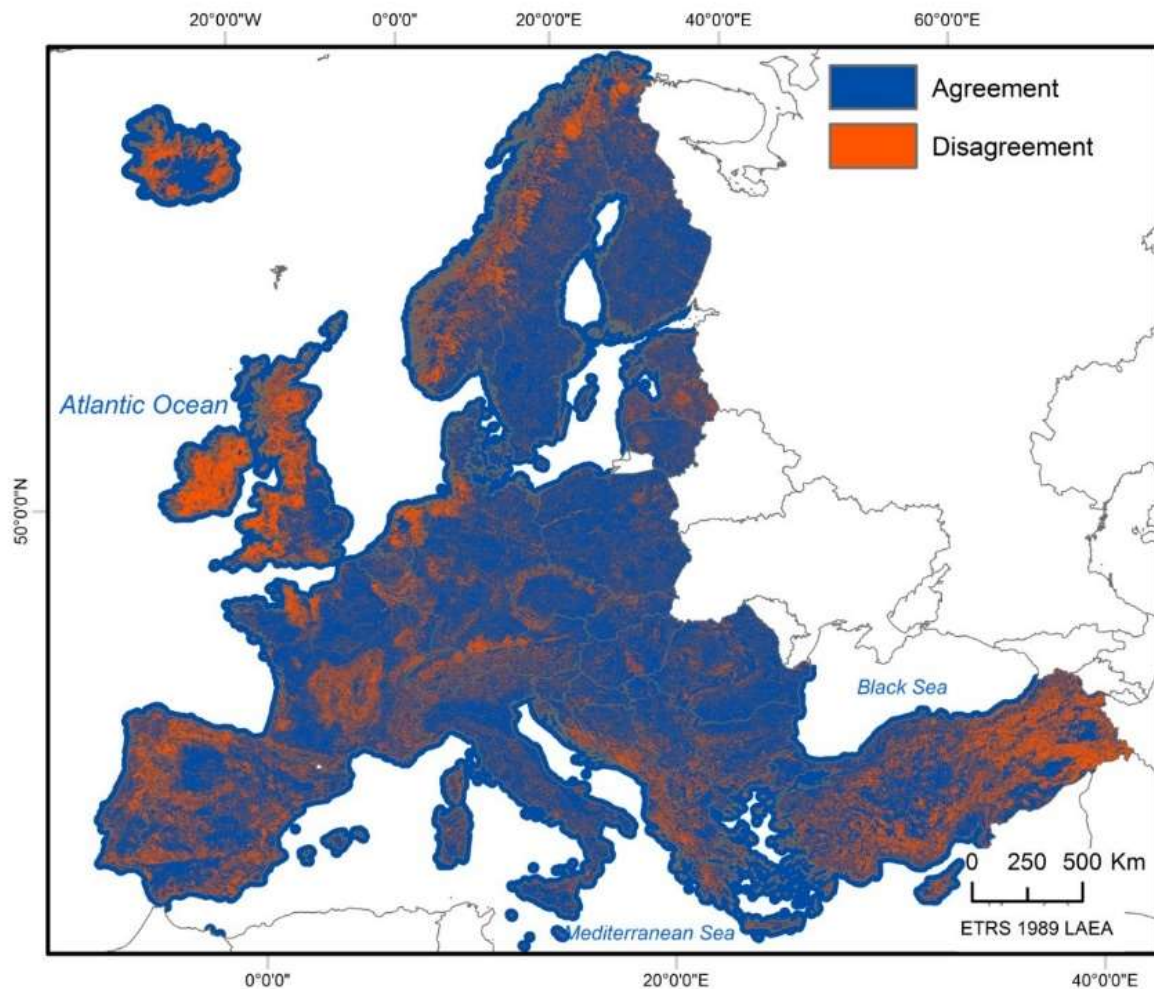
Fig. 7. Agreement and disagreement between CCI-LC12 and CLC12 maps

In Northern Europe disagreements were mainly located in Iceland and Norway, due to the exchange found between the grassland (CCI-LC12) and shrubland (CLC12) classes. Also, in United Kingdom and Ireland the total area of disagreement was high, due to the presence of grasslands in CCI-LC12 classified as agriculture in CLC12. The same happened in the Netherlands and northern France. In East Turkey the disagreement was also important due to the larger presence of agriculture and grassland in CCI-LC12 compared to CLC12 where those areas were classified as shrubland or sparse vegetation categories. In the Iberian Peninsula the disagreements were mostly due to the larger presence of shrubland in CLC12 compared to CCI-LC12.

Cross-tabulation matrix between CCI-LC12 and CLC12 (reference). Diagonal cells (in bold) contain the percentage of pixels (off the total extent) classified as the same category in the two maps. OA shows the general

agreement between the two maps. Commission and omission errors by category are also included. All shown values are in percentage.

The overall agreement (OA) between both land cover sources was 74.72%. By individual categories, agriculture, forest and water bodies presented lower discrepancies. Regarding categories that have been classified in other categories (omission errors) shrubland presented the highest error (>90%), being under estimated related to CLC12. As also seen in Fig. 2 shrublands in CLC12 corresponded to forest or grassland categories in CCI-LC12. On the other hand, agriculture, forest and water bodies presented smaller omission errors. In relation to categories that were classified in a category but did not belong to that category (commission errors) grassland presented the largest error (>80%). In this case, CCI-LC12 grassland mostly corresponded to CLC12 agriculture class

Table 3 Cross-tabulation matrix between CCI-LC12 and CLC12 (reference). Diagonal cells (in bold) contain the percentage of pixels (off the total extent) classified as the same category in the two maps. OA shows the general agreement between the two maps. Commission and omission errors by category are also included. All shown values are in percentage.

|		Agriculture	Forest	Grassland	Wetland	Settlement	Shrubland	Sparse veg.	Water bod.	Commission error
CCI-LC12	Agriculture	**26.28**	1.87	0.95	0.07	1.12	1.39	0.74	0.08	19.14
	Forest	2.66	**20.67**	0.53	0.40	0.24	3.56	0.48	0.18	28.03
	Grassland	4.72	0.41	**1.35**	0.31	0.19	1.24	0.66	0.03	84.82
	Wetland	0.06	0.39	0.03	**0.87**	0.01	0.32	0.08	0.03	51.46
	Settlement	0.26	0.03	0.00	0.00	**1.59**	0.01	0.00	0.01	16.68
	Shrubland	0.15	0.11	0.09	0.01	0.01	**0.55**	0.06	0.00	44.14
	Sparse veg.	0.24	0.07	0.14	0.07	0.07	0.51	**2.56**	0.04	30.93
	Water bod.	0.10	0.19	0.01	0.21	0.03	0.05	0.04	**20.85**	2.95
	Omission error	23.76	12.95	56.34	55.12	51.33	92.80	44.69	1.78	**OA: 74.72**
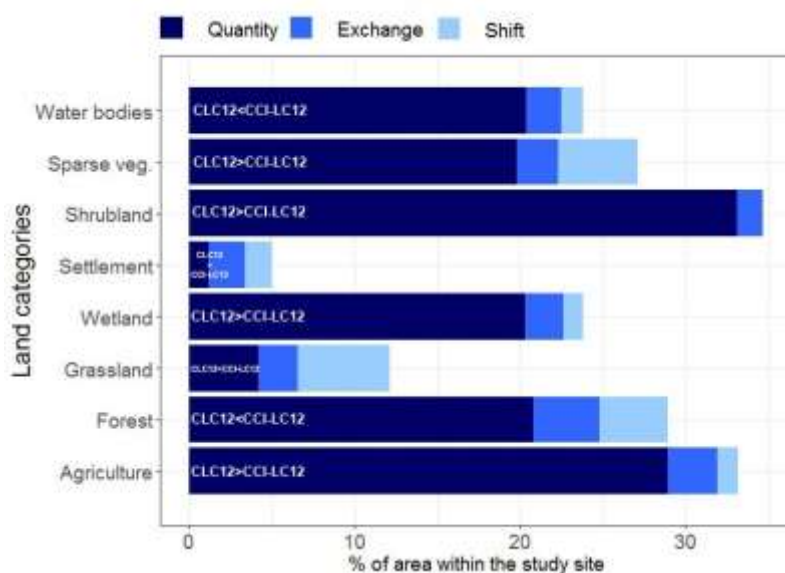
Fig. 8. Quantity, exchange and shift components (%) derived from cross-tabulation matrix between CCI-LC12 and CLC12 by land cover class

(Fig. 2). Agriculture, settlement and water bodies presented lower commission errors.

Fig. 8 shows the results of quantity, shift and exchange components (%) derived from CLC12 and CCI-LC12 cross tabulation matrix.

Overall quantity disagreement, exchange and shift were 74.34%, 10.08% and 9.82% respectively. Quantity disagreement was larger than exchange and shift (allocation difference) in all classes but especially in shrublands with maximum values over ~35 %. This means that the different

number of pixels by category was the main source of disagreement between LC maps. Exchange component was larger in forest and agriculture categories, meaning that pixels transited from different categories between maps but maintaining the same number. Shift component was larger in grassland and sparse vegetation bare areas, permanent snow and ice classes. That means that some pixels from those classes were both in different categories and spatial locations in the two maps. In shrubland category there was no shift, meaning that all of the allocation difference was due to exchange.

Regarding comparison of the two maps with LUCAS12, Tables 4 and 5 show the confusion matrix with the information per LC classes.

Cross-tabulation matrix between CCI-LC12 with LUCAS12 (reference). Diagonal cells (in bold) contain the percentage of pixels (off the total extent) classified as the same category in the two maps. OA shows the general agreement between the two maps. Commission and omission errors by category are also included. All shown values are in percentage.

Cross-tabulation matrix between CLC12 with LUCAS12 (reference). Diagonal cells (in bold) contain the percentage of pixels (off the total extent) classified as the same category in the two maps. OA shows the general agreement between the two maps. Commission and omission errors by category are also included. All shown values are in percentage.

Table 4 Cross-tabulation matrix between CCI-LC12 with LUCAS12 (reference). Diagonal cells (in bold) contain the percentage of pixels (off the total extent) classified as the same category in the two maps. OA shows the general agreement between the two maps. Commission and omission errors by category are also included. All shown values are in percentage

		Agriculture	Forest	Grassland	Wetland	Settlement	Shrubland	Sparse veg.	Water bod.	Commission error	
CCI-LC12	Agriculture	**24.73**	1.36	1.30	0.02	0.44	0.14	0.19	0.06	12.44	
	Forest	6.97	**24.56**	1.61	0.41	0.38	0.33	0.09	0.32	29.18	
	Grassland	11.05	3.20	**5.96**	0.12	0.91	0.12	0.11	0.13	72.41	
	Wetland	0.15	0.41	0.12	**0.30**	0.01	0.00	0.01	0.05	71.17	
	Settlement	2.06	0.64	0.39	0.01	**1.54**	0.03	0.06	0.04	67.83	
	Shrubland	2.08	1.79	0.67	0.09	0.07	**0.63**	0.17	0.05	88.59	
	Sparse veg.	1.03	0.23	0.08	0.01	0.05	0.02	**0.07**	0.03	95.64	
	Water bod.	0.46	0.57	0.12	0.06	0.07	0.00	0.02	**1.30**	50.32	
	Omission error	49.04	25.03	41.83	71.16	55.66	50.26	90.89	34.33	**OA: 59.08**	

Table 5 Cross-tabulation matrix between CLC12 with LUCAS12 (reference). Diagonal cells (in bold) contain the percentage of pixels (off the total extent) classified as the same category in the two maps. OA shows the general agreement between the two maps. Commission and omission errors by category are also included. All shown values are in percentage

		Agriculture	Forest	Grassland	Wetland	Settlement	Shrubland	Sparse veg.	Water bod.	Commission error
CLC12	Agriculture	**26.22**	0.89	0.11	0.02	0.69	0.28	0.01	0.03	7.18
	Forest	6.97	**23.13**	0.29	0.30	0.79	2.93	0.05	0.21	33.30
	Grassland	15.87	2.10	**0.70**	0.17	1.81	0.81	0.05	0.09	96.75
	Wetland	0.17	0.28	0.02	**0.41**	0.01	0.11	0.00	0.04	60.48
	Settlement	1.98	0.45	0.03	0.01	**2.15**	0.12	0.01	0.03	55.03
	Shrubland	1.78	0.93	0.50	0.16	0.13	**1.93**	0.09	0.03	65.17
	Sparse veg.	0.99	0.16	0.04	0.01	0.14	0.13	**0.03**	0.02	97.78
	Water bod.	0.50	0.41	0.01	0.05	0.13	0.05	0.02	**1.43**	45.08
	Omission error	51.88	18.43	58.79	63.64	63.38	69.59	86.66	23.61	**OA: 56.00**

As shown in Tables 4 and 5, the overall agreement between the two LC maps and LUCAS12 reference data was slightly higher for CCI-LC12 (59.08%) compared to CLC12 (56%). By individual categories, agriculture and forest presented lower discrepancies. In both validation assessments sparse vegetation demonstrated the largest omission error followed by wetlands in the CCI-LC12 map and by shrublands in the CLC12 map. As table 2 indicates, LUCAS12 sparse vegetation category contained a mixture of bare land, lichen and mosses, which may explain the high discrepancies found including the large commission error percentage in both maps. The sparse vegetation class resulted again in the largest percentages of commission errors in both maps, followed by shrubland in CCI-LC12 and by grassland in CLC12 map. As seen in table 2, the harmonization process in CCI-LC12 resulted in the shrubland category containing evergreen and deciduous shrubland species, while the transitional woodland shrub category was also included in CLC12, which is very similar to the forest class as it is mixed. On the other hand, CCI-LC12 grassland also included mixed cover type (mosaic herbaceous cover plus tree, shrub, table 2), which led to less omission and commission errors (but still relatively large) compared to CLC12 in the assessment with LUCAS12. In relation to the settlement class, which did not include any mixed categories in the harmonization process, omission and commission errors were >50% in both maps, where commission errors were higher in the CCI-LC12 map while omission errors were higher in the CLC12 map. This result

was contrary as expected, because it was assumed that higher commission and omission errors would occur between thematically close categories such as grassland, shrubland or agricultural land covers.

Fig. 9 illustrates the results of quantity, shift and exchange components (%) between CCI-LC12 and CLC12 versus LUCAS12 reference data. Overall quantity disagreement, exchange and shift were 32.34%, 20.62% and 9.86% in CCI-LC12 and 77.07%, 10.68% and 5.19% in CLC12 respectively. As seen, overall quantity disagreement percentage was notably lower in CCI-LC12 than in CLC12, while exchange and shift percentages were larger, meaning that the disagreements in CCI-LC12 were mostly due to differences in allocation while in CLC12 were due to differences in the quantity of pixels. In CCI-LC12 quantity disagreement was larger for settlement and agriculture classes while greater exchange happened in agriculture, forest and also in grassland classes. Shift happened in agriculture and forest. There was almost no shift in sparse vegetation class. Regarding CLC12 exchange was larger in settlement and forest classes. There was no shift in sparse vegetation and agriculture classes. Comparing both assessments, shrubland class in CCI-LC12 was more accurate in terms of allocation (lower exchange and shift values) and quantity than CLC12, however allocation accuracy was higher in CLC12 for forest and grassland classes.
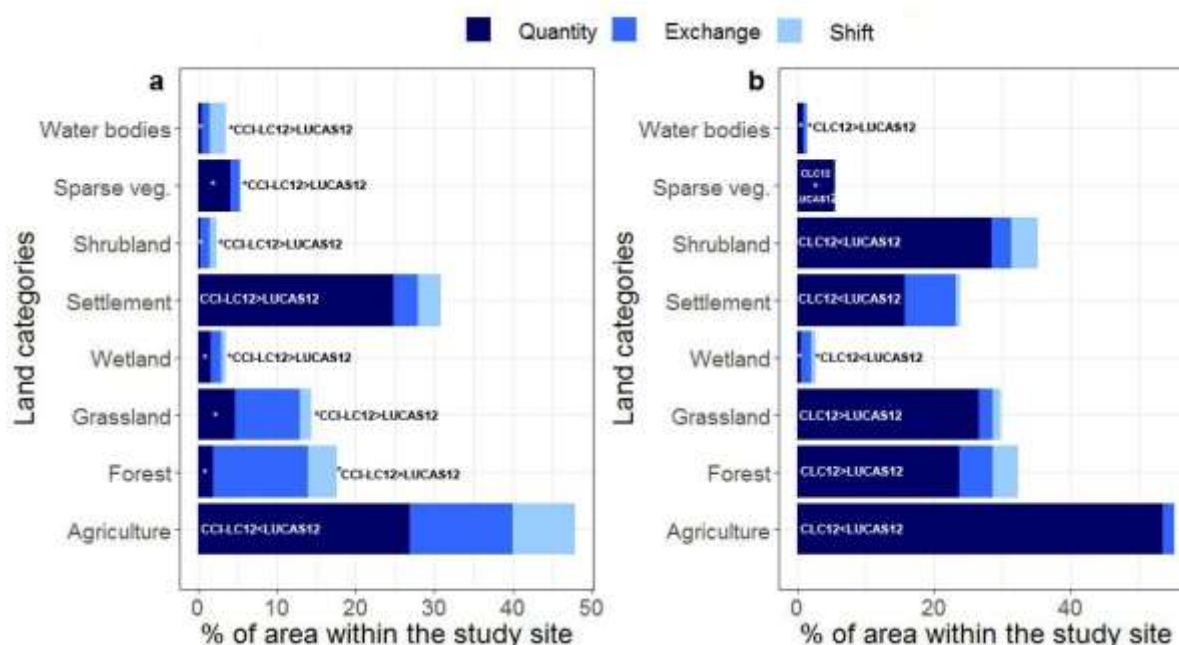

Fig. 9. Quantity, exchange and shift components (%) derived from cross-tabulation matrix between CCI-LC12 and LUCAS 12 (a) and CLC12 and LUCAS12 (b) by land cover class

3.2. Modeling wildfire occurrence at European and local scales using CCI-LC12 and CLC12 maps

Fig. 10 shows LC interfaces derived from CCI-LC12 and CLC12 maps (homogenized legend). Spatial data analysis was performed using ArcGis 10.2.1 (2011). As it can be observed, FGI was more abundant in CCI-LC12 map, being mostly located in northern countries (UK and Ireland) and mountain areas in central and southern Europe. WUI occupied also larger areas in UK and central Europe in the CCI-LC12 map. The FAI has a similar distribution in both maps. The percentage area of each LC interface was 80.95% FAI, 14% FGI and 5.05% WUI for the interfaces derived from CCI-LC12; and 83.67% FAI, 9.16% FGI and 7.17% WUI for the ones obtained from CLC12.

The fire data described in section 2.2.4 was overlaid on the two LC maps and derived interfaces so the frequency of wildfire occurrence by each LC class and LC interfaces CCI-LC12 and CLC12 was obtained (Table 6).

In CCI-LC12 over 70 % of wildfires were located in areas classified as forest and agriculture while in CLC12 the proportion was only 35%. In this map wildfire presence was mainly coincident with shrubland (40%). Wildfire occurrences happened two times more in the forest class in CCI-LC12 than in CLC12 while frequency was four times higher in the shrubland CLC12 class than in CCI-LC12. In relation to the LC interfaces wildfire occurrences were more frequent in CLC12 derived interfaces. In this map higher frequency occurred in the FGI while in CCI-LC12 the higher frequency was observed in FAI.
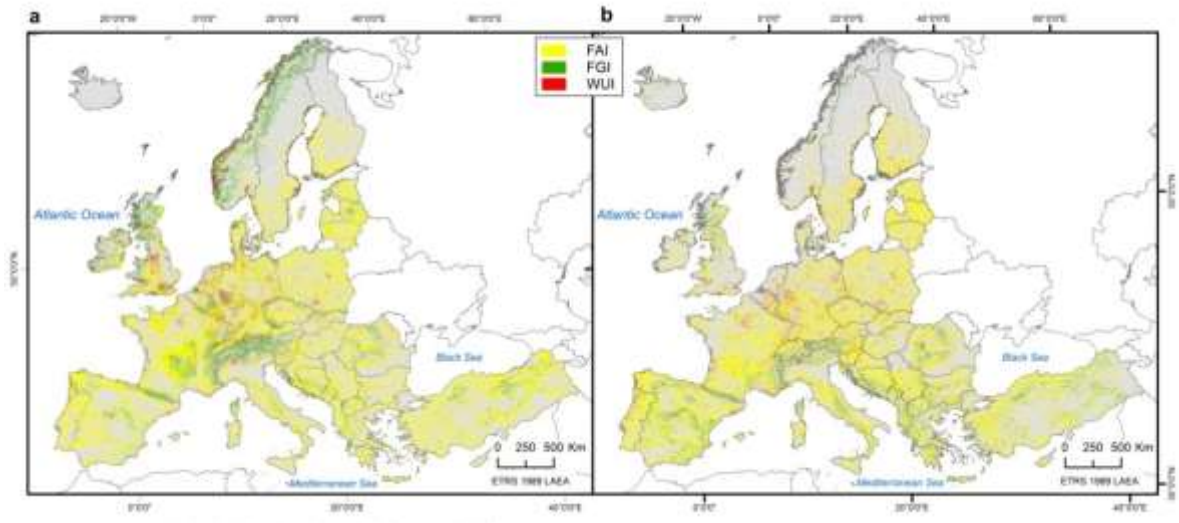

Fig. 10. LULC Interfaces FAI (Forest-Agricultural), FGI (Forest- Grassland), WUI (Wildland Urban Interface) from (a) CCI-LC12 and (b) CLC12 maps.

3.2.1. European wildfire occurrence model

Multicollinearity tests of the predictors (LC interfaces and climatic data) and the responses variables showed an absence of multicollinearity effects, so all predictors were included in the GLM analyses. LC interfaces were standardized for a better comparison of the estimated coefficients (zFAI, zFGI and ZWUI named variables). Table 6 shows the summary of the results of the fitted GLM at the European level.

All terms were statistically significant for both CCI-LC12 and CLC12 based models. Temperature (T) contributed the most in both models, followed by WUI in CCI-LC12 and FGI in CLC12 models. Wildfire occurrence was positively related to all predictors except for WUI. The relationships were as expected except for WUI and precipitation (P). Fig. 11. shows the TOC graph, which was

similar for both models. It indicated that the 0.5 threshold was close to random (the so-called uniform line). Also, the vertical distance from the horizontal axis to the TOC curve (hits) was slightly larger for 0.3 than for 0.5 threshold in both maps, indicating 0.3 threshold provided a better classification fit. The model validation provided also an AUC of 0.52 for the CCI-LC12 and 0.51 for the CLC12. In both cases values were close to random association. Sensitivity and specificity were 77.26 and 25.89% for the CCI-LC12 and 75.68 and 29.99% for the CLC12, respectively. Regarding the omission error it was 22.74% for the CCI-LC12 and 24.32% for the CLC12. The percent correct classification was 61.6% and 62.2% for CCI-LC12 and CLC12 models, respectively.

Table 6 Percentage of wildfire occurrence by LC class and by LC interface for CCI-LC12 and CLC12 maps.

LC class/Interface	CCI-LC12	CLC12
Agriculture	28.52	17.31
Forest	44.33	17.59
Grassland	12.71	14.00
Wetland	2.34	2.53
Settlement	0.08	0.36
Shrubland	10.87	40.31
Sparse veg.	0.98	7.71
Water bod.	0.18	0.20
FAI	1.87	1.87
FGI	1.67	4.19
WUI	0.94	1.60

Table 7 Estimated coefficients in the study site and their in CCI-LC12 and CLC12. Coefficients indicate odds of a wildfire to occur. European level. All p-values are less than 10^{-11}.

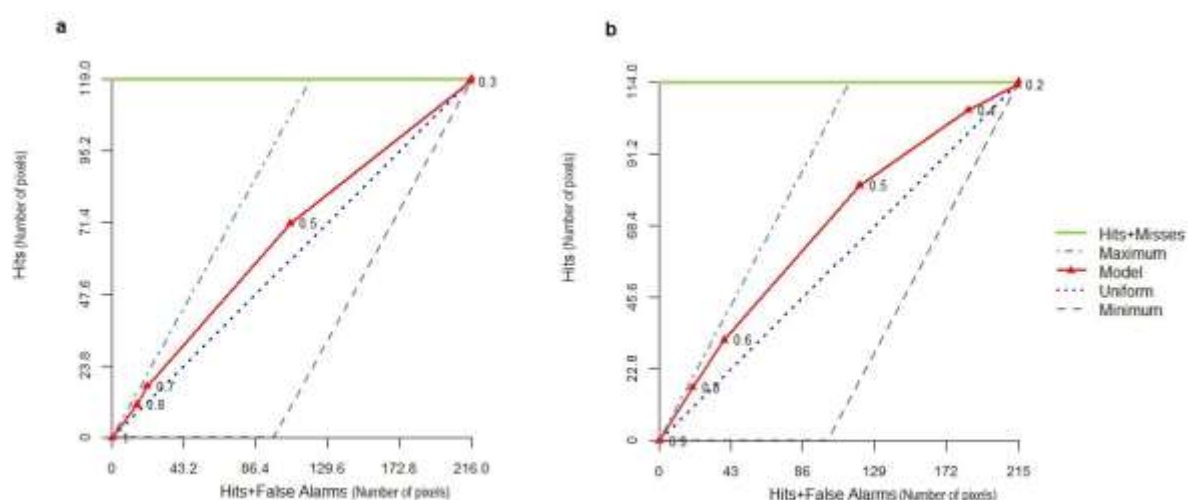|	CCI-LC12	CLC12
	Estimated coefficient	Estimated coefficient
(Intercept)	-15.142	-16.318
zFAI	0.318	0.393
zFGI	0.283	1.292
zWUI	-1.549	-0.388
T	2.205	2.008
P	0.747	0.852
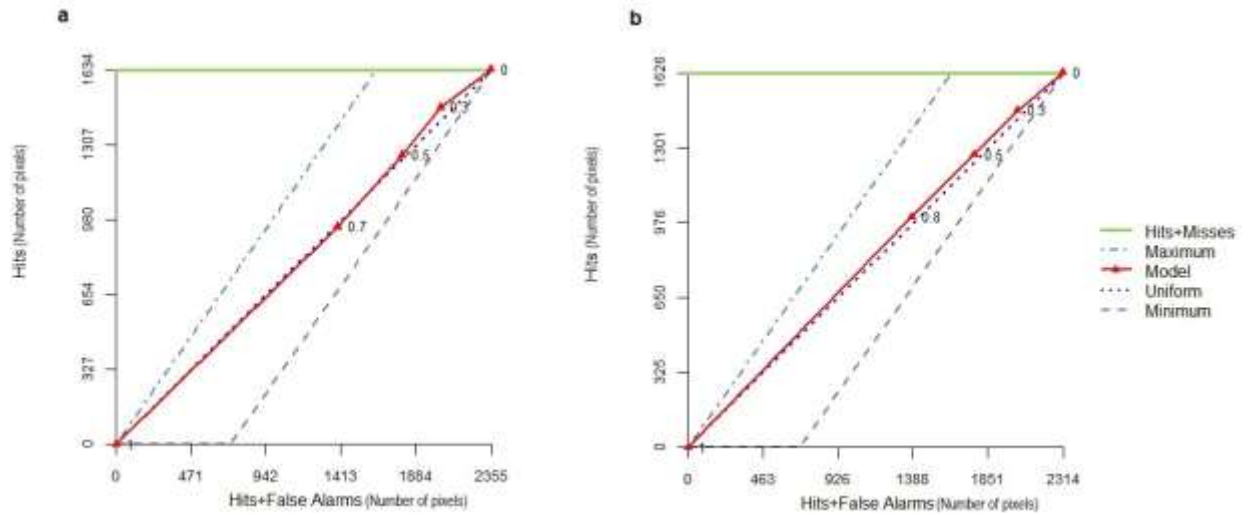
Fig. 11. Total operating characteristic (TOC) for CCI-LC12 (a) and CLC12 (b) models. European scale. Thresholds labeled in red triangles.

3.2.2. *Local wildfire occurrence model*

Table 8 shows the summary of the results at local level.

At local level using a more accurate location of the response variable (fire ignition coordinates) results show for both LC sources that FAI and precipitation (P) were statistically significant and positively related to wildfire occurrence. The positive relationship with P was contrary as expected. Fig. 11 shows the TOC for both models. The vertical distance from the horizontal axis to the TOC curve (hits) was larger for ~0.3 than for the other selected thresholds. However, the distance to the uniform line was larger for 0.5 threshold than the others, describing a less random association, also compared with the European models. The model validation showed an AUC of 0.56 for the CCI-LC12 and 0.57 for the CLC12. In this case sensitivity and specificity were 52.43 and 66.07% for the CCI-LC12 and 66.67 and 68.93% for the CLC12, respectively. Omission errors were 47.57% for the CCI-LC12 and 33.33% for the CLC12. The percent correct classification was 59.5% and 67.8% for CCI-LC12 and CLC12 models, respectively.

Table 8. Estimated coefficients in the study site and their significance in the CCI-LC12 and CLC12 at local level. Coefficients indicate odds of a wildfires to occur.

	CCI-LC12		CLC12	
	Estimated coefficient	Prob(>|z|)	Estimated coefficient	Prob(>|z|)
(Intercept)	-5.74667	0.0372 *	-3.36162	0.207277
zFAI	0.24436	0.0079 **	0.40809	0.000123 ***
zFGI	-0.03436	0.6001	-0.05501	0.489810
zWUI	0.21085	0.1170	0.10698	0.128707
T	0.22835	0.1808	0.05835	0.725337
P	2.06808	1.4e-05 ***	1.84779	6.37e-05 ***

Asterisks indicate significance levels of estimated coefficients: '***' $P < 0$, '**' $P < 0.001$, '*' $P < 0.01$, '.' $P < 0.05$, '' $P < 0.1$, $P < 1$.

Fig. 12. Total operating characteristic (TOC) for CCI-LC12 (a) and CLC12 (b) models. Local scale. Thresholds labeled in red triangles.

4. Discussion

A comprehensive comparison between CCI-LC12 and CLC12 datasets available for Europe and further validation using LUCAS survey has been accomplished and large differences in various LC classes that are critical in wildfire occurrence modeling and also in many other applications has been demonstrated. The year 2012 selected for comparison ensured not time inconsistencies that might affect the assessment. Spatial and thematic homogenization between the two data sources was also thoroughly addressed in order to guarantee the comparability. According to Pérez-Hoyos et al. (2017) the spatial resolution of the original datasets is a key issue for map comparison, as higher it generally implies a better land cover characterization because it is able to better represent smaller features that are mostly in heterogeneous landscapes. In this study the nearest neighbor technique was used to resample CLC (100 m) to CCI-LC pixel size (300 m), which might result in a loss of detail in the CLC map. Nonetheless, following Baboo et al. (2010) this method is appropriate for resampling categorical data such as LC because it does not alter the original values, even if some pixel values may be duplicated or lost.

Legend harmonization required some assumptions to aggregate the original into common LC categories. This was also a crucial step because it implied a simplification, mainly for mixed classes (Pérez-Hoyos et al., 2017). Comparison results between CCI-LC12 and CLC12

showed that agriculture, forest and water bodies were the most coincident classes. Larger disagreements were found in the settlement class with a ~16% of commission and ~50% of omission errors. Tsendbazar et al. (2014) compared global land cover datasets with regional high quality maps for urban and forest classes. In the case of Europe, they compared CLC 2006 and CCI-LC 2005, obtaining a similar correspondence (in relative terms) to our work, ~83.4% and ~51% for forest and urban classes respectively. Largest dissimilarities between CCI-LC12 and CLC12 were found in shrubland and grassland. This result was expected, as larger errors happened between close thematic classes as the above-mentioned, because similar spectral signals might have led to classify the covers in one or another type (CCI-LC derived map). Also, if the covers had comparable characteristics related to color, texture, etc. (CLC derived map using CAPI method to photo-interpret). Overall differences were mainly due to quantity disagreements in the analyzed covers. Grassland class showed larger disagreements in terms of allocation.

LUCAS12 was taken as ground truth for the validation exercise, being CCI-LC12 slightly more accurate than CLC12. In terms of quantity agreement CCI-LC12 was ~45 (in percentage points) more accurate than CLC12. However, CLC12 showed better allocation agreement figures. Forest was the most coincident class for both CCI-LC12 and CLC12. Agricultural class was under estimated in both maps, and also most of the misclassifications happened between other classes and agriculture. Although LUCAS12 intends to survey a current classification of LC types the first aim was to provide early crop estimates (2001 survey). The validation showed that, for both maps, shrubland and grassland classes presented fair to low agreements. In the CLC12 map these categories were under-estimated (larger omission error values than in CCI-LC12). At the same time, shrubland was over-estimated in the CCI-LC12 map. Nonetheless the sparse vegetation, bare areas, permanent snow and ice aggregated class was the least coincident, mainly due to the assumption made when making the harmonization of the legend because LUCAS category include bare areas and lichens and mosses, but being absent the sparse vegetation. Büttner and Maucha (2006) assessed the thematic accuracy of CLC2000 with LUCAS 2001-2002 data. They found ~74% of agreement between the two sources. However, their comparison was based on 22 out of 44 CLC disaggregated categories, finding the highest class-level reliability for the rivers, lakes, two urban categories, agro-forestry and permanently irrigated land. On the contrary sparse vegetation class showed the lowest reliability, indicating the difficulty in interpreting this category. Despite the implications of legends' harmonization, also mosaic or mixed classes tend to generate disagreements due to low spectral separability and mixed vegetation components in the original remote sensing data sources (Fritz et al., 2010; Herold et al., 2008). Tsendbazar et al. (2016) in their comparison of global land cover maps for the year 2005 (Globcover, CCI-LC and MODIS) found high confusion errors for mixed trees, shrubs, grasses and croplands. Herold et al. (2008) showed the limited ability of four global LC maps (IGBP DISCover,

UMD, MODIS 1-km, and GLC2000) to discriminate mixed classes as mosaic of trees, shrubs and herbaceous vegetation. Other authors pointed that savannah and grasslands are commonly interpreted with a low confidence level if the process included visual interpretation (Strahler et al., 2006).

LC interfaces used as driven factors for wildfire occurrence showed also important differences depending on the LC map used to produce them. As a consequence, the frequency of fires located in these areas also changed which eventually affected wildfire occurrence modeling at European and local scales. At European scale, global accuracy of the models using the two different set of interfaces (from CCI-LC12 or CLC12) was very similar with differences in percent correct classification less than 1%, however, the relative importance and relationship of explicative variables changed. In the European model all explicative variables were significant, having the temperature the highest influence in wildfire occurrence followed by WUI (negatively related) in CCI-LC12 model and by FGI in the CLC12 model. According to the validation made with LUCAS12, grassland class from CLC12 presented less allocation disagreements, so CLC12 seems to better spatially locate and explain the relationship of FGI with wildfire occurrence at the EU level. However, larger omission and commission error of grassland in this map made that more pixels were classified as shrubland, while in the case of CCI-LC12 the grassland misclassified pixels were classified as settlement cover instead. In relation to the negative relation of WUI with wildfire occurrence obtained we would have expected nonetheless a positive relationship. This was maybe related with the characteristics of the response variable used at European scale where small fires that could start in this LC interface were not mapped. Other works have included WUI and other LC interfaces derived from CLC for wildfire modeling at European level obtaining diverse results. In Vilar et al. (2016a) the GAM models for the European Mediterranean countries showed an increasing trend of FAI and FGI and an also increasing trend of WUI till it reached a plateau. The authors focused their work on the Mediterranean countries where wildfire occurrence has similar patterns. In the work presented here the models were obtained for the whole Europe, with diverse land cover composition and fire trends. Nevertheless, other authors as Modugno et al. (2016) found positive or negative relationships (depending of the country) between the distance to WUI and large fires, analyzing the regional patterns of large wildfires in WUI in Europe, noticing the strong influence of WUI on wildfires in parts of the Mediterranean regions. Oliveira et al. (2012) found a low and negative relationship between WUI and fire occurrence in their spatial models for the Mediterranean Europe using two methodologies (multiple regression and random forest). Regarding precipitation it was expected to find a negative relationship with the wildfire occurrence, because if the precipitation is high then the humidity of the vegetation would be also higher, which would not facilitate the ignition of a fire. Nonetheless an inverse relationship happened, fact that might be related with having larger

amount of vegetation available to be ignited represented by the LC interfaces.

At the local level, only the precipitation and FAI were significant and positively related to wildfire occurrence in both LC models. This could be related with the different spatial accuracy of the response variable (x, y fire ignitions) compared with the fire perimeters used in the European model. In relation to the positive relationship found at this scale between FAI and wildfire occurrence is in agreement with the works of Rodrigues et al. (2016) and Martínez-Fernández et al. (2013). These authors analyzed wildfire factors (including FAI obtained from CLC) and modeled fire occurrence with different methods (regression and Geographically Weighted Regression) in Spain, where this study region was located and where they also found a positive relationship of FAI with wildfire. When local models using interfaces from CCI-LC or CLC are compared we find the same pattern than in the European scale regarding percent correct classification, with CLC derived model showing highest values, however, a this scale the difference between the two models is higher (8 %) so the election of the LC map is more critical at this scale.

Expected results were found when comparing model accuracy at the two spatial scales with higher percent correct classification and less randomness as a more precise response variable was used. Nonetheless model performance at both scales could be improved by including some other variables that might explain wildfire occurrence e.g. fuel moisture content, other socio-economic factors related to ignition (roads, railways, population, etc.) as analyzed and used in other works in European areas (Vilar et al., 2016b); (Rodrigues, 2014).

5.   Conclusions

Land cover data is essential for wildfire occurrence estimation as for other applications including environmental analysis, global and climate change, food security, land management, etc. The analysis of the suitability of global and regional land cover data sources is needed to obtain the most appropriate estimations. A variety of LC datasets has been used in the literature as one of the inputs for wildfire occurrence modeling at local, regional and global scales but is still unknown which data source might be more appropriate for this specific application. In this study CCI-LC and CLC from 2012 have been compared. Even though both maps refer to the same year each of them derives from different remote sensing data and has been obtained using different methodologies. In this work per-pixel comparison has been applied to spatial and thematic homogenized dataset finding important disagreements for key LC in the context of wildfire analysis as i.e. grassland and shrubland categories. Selecting LC data source will depend on the interest of land and fire managers in specific covers. For instance, location and extension of grasslands would have an effect on, for example, agricultural subsidies policy measurements. On the other hand, shrublands are related to forest clearing policies or forestation measurements, among others. Wildfire occurrence estimators (LC interfaces) have been

derived from the two maps and used to model occurrence at EU and local scales. Differences between models were more meaningful at the local scale. At this scale wildfire estimators adapted better to the actual context of the territory and the performance improved regardless the LC map used. Further work would analyze the influence of the LC inputs in wildfire occurrence models for regions with different characteristics regarding fire incidence as fire size or main causality factors.

References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle, in: B Petran, F.C. (Ed.), International Symposium on Information Theory, Akadeemiai Kiadi: Budapest, Hungary, pp. 267–281.

Baboo, S.S., Devi, M.R., 2010. An analysis of different resmpling methods in Coimbatore district. Global Journal of Computer Science and Technology 10.

Bartholomé, E., Belward, A.S., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. International Journal of Remote Sensing 26, 1959-1977.

Bontemps, S., Boettcher, M., Brockmann, C., Kirches, G., Lamarche, C., Radoux, J., Santoro, M., Van Bogaert, E., Wegmüller, U., Herold, M., Achard, F., Ramoino, F., Arino, O., Defourny, P., 2015. Multi-year global land cover mapping at 300 m and characterization for climate modelling: achievements of the land cover component of the ESA Climate Change Initiative, 36th International Symposium on Remote Sensing of Environment. The International Archives of the

Photogrammetry, Remote Sensing and Spatial Information Sciences.

Burkhard, B., Kroll, F., Müller, F., Windhorst, W., 2009. Landscapes' Capacities to Provide Ecosystem Services- a Concept for Land-Cover Based Assessments Landscape Online 15, 1-22.

Büttner, G., Maucha, G., 2006. The thematic accuracy of CORINE Land Cover 2000. Assessment using LUCAS. European Environment Agency, Copenhagen.

Büttner, G., Soukup, T., Kosztra, B., 2014. CLC2012 Addendum to CLC2006 Technical Guidelines. European Environmental Agency.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., Mills, J., 2015. Global land cover mapping at 30m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing 103, 7-27.

Chuvieco, E., 2002. Teledetección ambiental: la observación de la tierra desde el espacio. Ariel Ciencia.

Chuvieco, E., Aguado, I., Yebra, M., Nieto, H., Salas, J., Martín, M.P., Vilar, L., Martínez, J., Martín, S., Ibarra, P., de la Riva, J., Baeza, J., Rodríguez, F., Molina, J.R., Herrera, M.A., Zamora, R., 2010. Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. Ecological Modelling 221, 46-58.

Chuvieco, E., González, I., Verdu, F., Aguado, I., Yebra, M. , 2009. Prediction of fire occurrence from live fuel moisture content measurements in a Mediterranean ecosystem. International Journal of Wildland Fire 18, 282-287.

Defourny, P., Kirches, G., Brockmann, C., Boettcher, M., Peters, M., Bontemps, S., Lamarche, C., Schlerf, M., M., S., 2016. Land Cover CCI: Product User Guide Version 2. UCL-Geomatics, Louvain-la-Neuve, Belgium.

Di Gregorio, A., Henry, M., Donegan, E., Finegold, Y., Latham, J., Jonkheere, I., Cumani, R., 2016. Land Cover Classification System. Software version 3. Food and Agriculture Organization of the United Nations (FAO), Rome.

EEA, 2017. Landscapes in transition. An account of 25 years of land cover change in Europe. EEA Luxembourg.

ESRI, 2011. ArcGIS Desktop: Release 10, Redlands, Ca: Environmental Systems Resource Institute.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognition Letters 27, 861–874.

Forkel, M., Dorigo, W., Lasslop, G., Teubner, I., Chuvieco, E., Thonicke, K., 2017. A data-driven approach to identify controls on global fire activity from satellite and climate observations (SOFIA V1). Geosci. Model Dev. 10, 4443-4476.

Fox, J., Weisberg, S., 2011. An R Companion to Applied Regression, Thousand Oaks CA.

Freund, R.J., Littell, R.C., Creighton, L., 2003. Regression Using JMP, Cary, NC: SAS Institute, Inc. .

Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcok, C.E., Gopal,

S., Schneider, A., Cooper, A., Baccini, A., Gao, F., Schaaf, C., 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment 83, 287–302.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X.M., 2010. MODIS Collection 5 Global Land Cover: Algorithm Refinements and Characterization of New Datasets. Remote Sensing of Environment 114, 168–182.

Fritz, S., See, L., Rembold, F., 2010. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. International Journal of Remote Sensing 31, 2237-2256.

Gallardo, M., Gómez, I., Vilar, L., Martínez-Vega, J., Martín, M.P., 2015. Impacts of future land use/land cover on wildfire occurrence in the Madrid region (Spain). Regional Environmental Change 16, 1047-1061.

Guisan, A., Edwards, T.C., Hastie, T., 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling 157, 89–100.

Hair, J.F.J., Anderson, R.E., Tatham, R.L., Black, W.C., 1995. Multivariate Data Analysis. Macmillan Publishing Company, New York.

Herold, M., Mayaux, P., Woodcock, C.E., Baccini, A., Schmullius, C., 2008. Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. Remote Sensing of Environment 112, 2538-2556.

Jokar Arsanjani, J., See, L., Tayyebi, A., 2016. Assessing the suitability of GlobeLand30 for mapping land cover in Germany. International Journal of Digital Earth 9, 873-891.

Latham, J., Cumani, R., Rosati, I., Bloise, M., 2014. Global Land Cover SHARE (GLC-SHARE) database. Beta-V1.0 - 2014.

Lekkas, E., Carydis, P., Lagouvardos, K., Mavroulis, S., Diakakis, M., Adreadakis, E., Gogou, M.E., Spyrou, N.I., Athanassiou, M., Kapourani, E., Arianoutsou, M., Vassilakis, M., Parcharidis, P., Kotsi, E., Speis, P.D., Delakouridis, J., Milios, D., Kotroni, V., Giannaros, T., Dafis, S., Kargiannidis, A., Papagiannaki, K., 2018. The July 2018 Attica (Central Greece) Wildfires, p. 51.

Leroy, M., Bicheron, P., Brockmann, C., Krämer, U., Miras, B., Huc, M., Ninô, F., Defourny, P., Vancutsem, C., Petit, D., Amberg, V., Berthelt, B., Arino, O., Ranera, F., 2006. GlobCover: a 300 m global land cover product for 2005 using ENVISAT MERIS time series, ISPRS Commision VII Mid-Term Symposium: Remote Sensing: from Pixels to Processes, Enschede (NL).

Li, S., Cui, Y., Liu, M., He, H., Ravan, S., 2017. Integrating Global Open Geo-Information for Major Disaster Assessment: A Case Study of the Myanmar Flood. ISPRS International Journal of Geo-Information 6, 201.

Li, W., Ciais, P., MacBean, N., Peng, S., Defourny, P., Bontemps, S., 2016. Major forest changes and land cover transitions based on plant functional types derived from the ESA CCI Land Cover product.

International Journal of Applied Earth Observation and Geoinformation 47, 30-39.

Lynch, D., Cuff, N., Russell-Smith, J., 2015. Vegetation fuel type classification for lower rainfall savanna burning abatement projects, in: Murphy, B., Russell-Smith, J., Edwards, A., Meyer, M., Meyer, C.P. (Ed.), Carbon Accounting and Savanna Fire Management. CSIRO Publishing, pp. 73-96

Martínez-Fernández J, C.E., Koutsias N, 2013. Modelling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression. Natural Hazards Earth System Science 13, 311-327.

Martínez, J., Vega-García, C., Chuvieco, E., 2009. Human-caused wildfire risk rating for prevention planning in Spain. Journal of Environmental Management 90, 1241-1252.

Martino, L., Palmieri, A., Gallego, J., 2009. Use of auxiliary information in the sampling strategy of a European area frame agro-environmental survey, First Italian Conference on Survey Methodology (ITACOSM09), Siena (Italy).

Modugno, S., Balzter, H., Cole, B., Borrelli, P., 2016. Mapping regional patterns of large forest fires in Wildland–Urban Interface areas in Europe. Journal of Environmental Management 172, 112-126.

Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A., Pereira, J.M.C., 2012. Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. Forest Ecology and Management 275, 117-129.

Padilla, M., Vega-García, C., 2011. On the comparative importance of fire danger rating indices and their integration with spatial and temporal variables for predicting daily human-caused fire occurrences in Spain. International Journal of Wildland Fire 20, 46-58.

Pausas, J.G., Fernández-Muñoz, S., 2012. Fire regime changes in the Western Mediterranean Basin: from fuel-limited to drought-driven fire regime. Climatic Change 110, 215-226.

Pérez-Hoyos, A., García-Haro, F.J., San-Miguel-Ayanz, J., 2012. Conventional and fuzzy comparisons of large scale land cover products: Application to CORINE, GLC2000, MODIS and GlobCover in Europe. ISPRS Journal of Photogrammetry and Remote Sensing 74, 185-201.

Pérez-Hoyos, A., Rembold, F., Kerdiles, H., Gallego, J., 2017. Comparison of Global Land Cover Datasets for Cropland Monitoring. Remote Sensing 9, 1118.

Pontius Jr., R.G., 2019. Component intensities to relate difference by category with difference overall. International Journal of Applied Earth Observation and Geoinformation 77, 94-99.

Pontius Jr., R.G., Huffaker, D., Denman, K., 2004. Useful techniques of validation for spatially explicit land-change models. Ecological Modelling 179, 445-461.

Pontius Jr., R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. International Journal of Remote Sensing 32, 4407-4429.

Pontius Jr., R.G., Parmentier, B., 2014. Recommendations for using the relative operating characteristic (ROC). Landscape Ecology 29, 367-382.

Pontius Jr., R.G., Santacruz, A., 2014. Quantity, exchange, and shift components of difference in a square contingency table. International Journal of Remote Sensing 35, 7543-7554.

Pontius Jr., R.G., Santacruz, A., 2015. diffeR: Metrics of Difference for Comparing Pairs of Maps, R package version 0.0-4 ed.

Pontius Jr., R.G., Si, K., 2014. The total operating characteristic to measure diagnostic ability for multiple thresholds. International Journal of Geographical Information Science 28, 570-583.

Pontius, R.G., Santacruz, A., Tayyebi, A., Parmentier, B., Si, K., 2015. TOC: Total Operating Characteristic Curve and ROC Curve, 0.0-4 ed.

Preisler, H.K., Brillinger, D.R., Burgan, R.E., Benoit, J.W., 2004. Probability bases models for estimation of wildfire risk. International Journal of Wildland Fire 13, 133-142.

RCoreTeam, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Robinson, C., Schumacker, R.E., 2009. Interaction Effects: Centering, Variance Inflation Factor, and Interpretation Issues Multiple Linear Regression Viewpoints 35.

Rodrigues, M., de la Riva, J., Fotheringham, S., 2014. Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. Applied Geography 48, 52-63.

Rodrigues, M., Jiménez, A., de la Riva, J., 2016. Analysis of recent spatial–temporal evolution of human driving factors of wildfires in Spain. Natural Hazards 84, 2049-2070.

San-Miguel-Ayanz, J., Schulte, E., Schmuck, G., Camia, A., Strobl, P., Liberta, G., Giovando, C., Boca, R., Sedano, F., Kempeneers, P., McInerney, D., Withmore, C., Santos de Oliveira, S., Rodrigues, M., Durrant, T., Corti, P., Oehler, F., Vilar, L., Amatulli, G., 2012. Comprehensive Monitoring of Wildfires in Europe: The European Forest Fire Information System (EFFIS), in: Tiefenbacher, J. (Ed.), Approaches to Managing Disaster - Assessing Hazards, Emergencies and Disaster Impacts. InTech, pp. 87-105.

San-Miguel-Ayanz J., P.J.M.C., Boca R., Strobl P., Kucera J., Pekkarinen A., 2009. Forest Fires in the European Mediterranean Region: Mapping and Analysis of Burned Areas, in: E., C. (Ed.), Earth Observation of Wildland Fires in Mediterranean Ecosystems. Springer-Verlag, Berlin, Heidelberg, pp. 189-203.

See, L., Schepaschenko, D., Lesiv, M., McCallum, I., Fritz, S., Comber, A., Perger, C., Schill, C., Zhao, Y., Maus, V., Siraj, M.A., Albrecht, F., Cipriani, A., Vakolyuk, M.y., Garcia, A., Rabia, A.H., Singha, K., Marcarini, A.A., Kattenborn, T., Hazarika, R., Schepaschenko, M., van der Velde, M., Kraxner, F., Obersteiner, M., 2015. Building a hybrid land cover map with crowdsourcing

and geographically weighted regression. ISPRS Journal of Photogrammetry and Remote Sensing 103, 48-56.

Seidl, R., Schelhaas, M.J., Werner Rammer, W., Johannes Verkerk, P.J., 2014. Increasing forest disturbances in Europe and their impact on carbon storage. Nat Clim Chang 4, 806-810.

Sing, T., Sander, O., Beerenwinkel, N., Lengauer, T., 2005. ROCR: visualizing classifier performance in R, Bioinformatics 21(20).

Story, M., Congalton, R.G., 1986. Accuracy Assessment: A User's Perspective. Photogramm Eng Remote Sens 52, 397-399.

Strahler, A.H., Boschetti, L., Foody, G.M., Friedl, M.A., Hansen, M.C., Herold, M., Mayaux, P., Morisette, J.T., Stehman, S.V., Woodcock, C.E., 2006. Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps. GOFC-GOLD and Office for Official Publications of the European Communities Luxembourg.

Tateishi, R., Hoan, N.T., Kobayashi, T., Alsaaideh, B., Tana, G., Phong, D.X., 2014. Production of Global Land Cover Data – GLCNMO2008. 2014 6.

Tsendbazar, N.-E., de Bruin, S., Herold, M., 2017. Integrating global land cover datasets for deriving user-specific maps. International Journal of Digital Earth 10, 219-237.

Tsendbazar, N.E., de Bruin, S., Mora, B., Schouten, L., Herold, M., 2016. Comparative assessment of thematic accuracy of GLC maps for specific applications using existing reference data. International Journal of Applied Earth Observation and Geoinformation 44, 124-135.

Tsendbazar, N.E., deBruin, S., Defourny, P., Herold, M., 2014. Comparative analsys of global land cover datasets, in: Laboratory of Geo-Information Science and Remote Sensing, W.U.a.R.C. (Ed.), International Conference Global Vegetation Monitoring and Modeling (GV2M), Avignon (France), p. 58.

Van den Besselaar, E.J.M., Haylock, M.R., Van der Schrier, G., Klein Tank, A.M.G., 2011. A European daily high resolution observational gridded data set of sea level pressure. Journal of Geophysical Research 116.

Verburg, P.H., Neumann, K., Nol, L., 2011. Challenges in using land use and land cover data for global change studies. Global Change Biology 17, 974-989.

Vilar, L., Camia, A., San-Miguel-Ayanz, J., Martín, M.P., 2016a. Modeling temporal changes in human-caused wildfires in Mediterranean Europe based on Land Use-Land Cover interfaces. Forest Ecology and Management 378, 68-78.

Vilar, L., Gómez, I., Martínez-Vega, J., Echavarría, P., Riaño, D., Martín, M.P., 2016b. Multitemporal Modelling of Socio-Economic Wildfire Drivers in Central Spain between the 1980s and the 2000s: Comparing Generalized Linear Models to Machine Learning Algorithms. PLOS ONE 11, e0161344.

Waldner, F., Fritz, S., Di Gregorio, A., Defourny, P., 2015. Mapping Priorities to Focus Cropland Mapping Activities: Fitness Assessment of Existing Global, Regional and National Cropland Maps. Remote Sensing 7, 7959.

Wood, S.N., 2017. Generalized Additive Models: An Introduction with R, 2nd ed. Chapman and Hall/CRC.