International Journal of Applied Earth Observations and Geoinformation 94 (2021) 102221

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Comparison of ESA climate change initiative land cover to CORINE land cover over Eastern Europe and the Baltic States from a regional climate modeling perspective

V. Reinhart[a,b,*], C.C. Fonte[c], P. Hoffmann[a,b], B. Bechtel[d], D. Rechid[a], J. Boehner[b]

[a] Helmholtz Zentrum Geesthacht, Climate Service Center Germany (GERICS), Germany
[b] Universität Hamburg, Institute of Geography, Germany
[c] University of Coimbra, Portugal
[d] Ruhr-Universität Bochum, Germany

ARTICLE INFO

Keywords:
Land Use
Land Cover
Quality Assessment
Reference Data
Proportional Area Comparison
Regional Climate Modelling
Eastern Europe

ABSTRACT

High-quality land use and land cover (LULC) information is of crucial importance for the performance of regional climate models (RCMs), in particular at high spatial resolutions down to convection permitting scales below 4 km. Several satellite-based high-resolution products are currently available for implementation into RCMs. One of the most recent products is the European Space Agency Climate Change Initiative Land Cover (ESA CCI LC) dataset. While the ESA CCI LC has been assessed globally, an evaluation against regional, independent LULC datasets is necessary to identify LULC inaccuracies in the respective region of interest and to give regional climate modelers estimates for the uncertainty in the land use forcing. In the present work the ESA CCI LC dataset is compared to the COoRdination and INformation on the Environment (CORINE) Land Cover (CLC). Agreement between the datasets is assessed by proportional area comparison (PAC). The resulting agreement measures are compared to the results of a majority approach (MA) to explore possible differences between the methods. Three timesteps of ESA CCI LC matching the timesteps of CLC are assessed to take a change in agreement over time into account. In addition to the quantification of agreement, spatial patterns of possible issues with ESA CCI LC are identified through utilization of geospatial information systems (GIS). Using the PAC, the agreement of ESA CCI LC with CLC is found to be ~76 % for the research area (RA). Although the agreement decreases slightly using the PAC, no substantial differences in agreement measures were found compared to the results of the MA. Dominant LULC categories agriculture and forest show an agreement of over 80 % with CLC. A few major issues were found for grassland, wetlands, settlements, and water bodies in the RA of which some might influence RCM performance if the dataset is implemented without adjustment. We highly recommend to apply the PAC to other regions in Europe and further globally to investigate if the found issues are also found elsewhere. The use of more independent regional and specified datasets for validation but also for possible improvement of the ESA CCI LC dataset is suggested.

1. Introduction

Production of high-quality and high-resolution land use and land cover (LULC) information has received increased attention in the last decades due to the importance of land cover representation for numerous fields of research. One important application is in regional climate modelling, which is moving towards higher resolution and therefore needs high-quality land cover information with high spatial and temporal resolution and coverage. For regional climate models (RCMs) a realistic LULC representation is crucial to realistically model subsurface and near-surface energy and moisture fluxes as well as to investigate feedback mechanisms and coupling effects between LULC and regional climate (Chu et al., 2011; Verburg et al., 2011; Houghton et al., 2012; Bontemps et al., 2013; Brovkin et al., 2013; Davin et al., 2019; Georgievski and Hagemann, 2019). Several RCM studies focusing on the quantification of uncertainties in near-surface climate parameters caused by LULC showed the benefits of using more precise LULC information in RCMs (Gao et al., 2015; Santos-Alamillos et al., 2015; Sertel et al., 2010).

Continuous LULC information is nowadays mostly produced by a

* Corresponding author at: Helmholtz Zentrum Geesthacht, Climate Service Center Germany (GERICS), Germany.

https://doi.org/10.1016/j.jag.2020.102221
Received 16 April 2020; Received in revised form 10 August 2020; Accepted 17 August 2020
Available online 9 September 2020

combination of manual and automatic interpretation of satellite imagery with inclusion of existing ground truth data and regional or country based LULC information (Loveland and Belward, 1997; Bontemps et al., 2011). For the reconstruction of past LULC and the projection of future LULC, supplementary data such as population development and distribution or changes in climate are used as proxies to assimilate LULC development with a model approach (Pongratz et al., 2008; Hurtt et al., 2011). Historical "ground truth" documents such as cadastral maps are an important input source or a validation instrument when reconstructing historical LULC with a model approach (Ramankutty and Foley, 1999; Petit and Lambin, 2002; Fuchs et al., 2013).

Uncertainties in a final LULC dataset can arise from various well-known sources of error during the development process. Different classification procedures, atmospheric disturbances as well as changing satellite sensors and algorithms can contribute to uncertainties between datasets but also within multiannual datasets (Castilla and Hay, 2007; Verburg et al., 2011; Fuchs et al., 2013). Therefore, throughout the development process of a LULC dataset, a comprehensive assessment is required to gain information about the quality of the product. However, there is no uniformly applied method and the procedure itself is widely discussed (Foody, 2002; Olofsson et al., 2014; Sarmento et al., 2015).

One of the most recent and detailed LULC products is the European Space Agency Climate Change Initiative Land Cover (ESA CCI LC) dataset, a continuous global dataset with 23 annual time steps (1992–2015) at 300 m grid resolution (ESA, 2017). Previous work found the ESA CCI LC epoch time steps (2000, 2005, 2010 and 2015) to be relatively accurate on a global scale (75.1 % for 2015, (Achard et al., 2017). Regional quality assessments gave consistent results when comparing ESA CCI LC to other high-resolution LC products in the investigated regions respectively (Pérez-Hoyos et al., 2017; Yang et al., 2017; Samasse et al., 2018; Koubodana et al., 2019). Hua et al. (2018) indicated a high consistency with other global state of the art land cover datasets for ESA CCI LC over Europe (over 60 % agreement with GLC2000 and GLOBCOVER globally and slightly higher when only Europe was compared). Due to its continuous global and annual availability and the detail in LULC description ESA CCI LC is most promising to be implemented in RCMs for various regional domains. However, a comparison of ESA CCI LC with independent reference data for Eastern Europe on a regional scale is missing.

In this study, a detailed comparison of the global ESA CCI LC product over Eastern Europe and the Baltic States with the COoRdination and INformation on the Environment (CORINE) Land Cover (CLC) dataset is carried out. We consider CLC as reference, since this dataset is more detailed, has a higher resolution and is independent from ESA CCI LC. Further, the quality of CLC is well known (Jaffrain et al., 2017). However, due to the different characteristics of the ESA CCI LC and the reference data, resampling and nomenclature harmonization techniques are necessary to adjust data sets of different resolution and classification for comparison. As a consequence of the application of these techniques at least one of the products is modified majorly during the process which can bias agreement measures (Foody, 2002; Tchuenté et al., 2011; Yang et al., 2017).

When applying a majority resampling approach (MA) to adjust the spatial resolution of two or more products, LULC class areas are changed. However, a recent approach in the region of Coimbra (Portugal) showed the bias in agreement measures due to use of the majority resampling is considerably high (Fonte et al., 2020). Against this background, an alternative approach is required, to reduce the bias due to resampling of LULC products and to improve the information value of LULC map comparisons.

The present work uses an innovative approach to assess the quality of high resolution LULC products in a way that is not restrained by the static grid structure using proportional area comparison (PAC) (Sarmento et al., 2015). By applying PAC, the quantity of correctly classified cells is transformed to a quantity of area that is correctly classified and that is independent from the grid structure. The method can be applied to compare gridded datasets with different resolutions (Fritz et al., 2010, 2011). In addition, the method can provide a spatial pattern of inaccuracies for the individual LULC classes.

In the present work, the method is applied to compare ESA CCI LC to CLC for all time steps that were available for both datasets respectively. In section 2, methods and data used in this study as well as the classification harmonization method are described. The agreement measures for the PAC and the MA for all time steps respectively, followed by maps of spatial disagreement patterns of the LULC categories are presented in section 3. Ways to deal with the identified issues in the individual categories and their possible implications for regional climate modelling are discussed in section 4. Finally, section 5 closes with conclusions and an outlook for further research.

2. Methods and data

2.1. ESA CCI land cover

ESA CCI LC is a continuous global land cover product with 300 m grid resolution. The product has been available online since 2017 and was developed over nine years by the ESA Climate Change Initiative (CCI) program. It provides global annual LC maps for 23 consecutive years from 1992 to 2015 (ESA, 2017). ESA CCI LC is a combination product of global surface reflectance from different satellite missions (Table 1). One of the major purposes of ESA CCI LC was to create a land cover product that meets the requirements of the climate modeling community (Li et al., 2017).

Validation is achieved through use of a dedicated tool provided by ESA and Google Earth images as background data. (Achard et al., 2017). Globally regularly distributed two-stage stratified random sampling including primary and secondary point sample units are validated by a network of experts for the respective region. Satellite based Google Earth data covering the respective time period is utilized as reference and validation is still ongoing. When validated using the GlobCover 2009 validation database, on which the validation tool is based, it is found that the overall accuracy for the ESA CCI LC 2015 map is 75.4 %. (Achard et al., 2017). Global consistency with other existing high resolution satellite based global land cover datasets was found to be relatively high for the European Continent (Hua et al., 2018). In addition, the ESA CCI LC is compared to existing, validated LULC products over the African Continent (Koubodana et al., 2019; Pérez-Hoyos et al., 2017; Samasse et al., 2018) and China (Yang et al., 2017) but up to now, there are no comprehensive assessment activities for ESA CCI LC over Central and Eastern Europe published. Extension of the ESA CCI LC map series until 2018 was provided in October 2019 (ESA, 2019).

Table 1
ESA CCI Product information (ESA, 2017).

Time period	Satellite products
Baseline Production 2003–2012	MERIS FR/RR global SR composites
1992–1999	Baseline 10-year global map AVHRR global SR composites for back-dating baseline
1999–2013	Baseline 10-year global map SPOT-VGT global SR composites for up and back-dating the baseline MERIS FR global SR composites to delineate the identified changes at 300 m spatial resolution PROBA-V global SR composites at 300 m for year 2013 to delineate the identified changes at 300 m spatial resolution
2013–2015	Baseline 10-year global map PROBA-V global SR composites at 1 km for years 2014 and 2015 for updating the baseline PROBA-V time series at 300 m for 2014 and 2015 to delineate the identified changes at 300 m spatial resolution

2.2. CORINE land cover (CLC)

The CLC database was initiated by a European program and includes land cover information for all EU member states in 1985 (Heymann, 1994). The first dataset from 1990 therefore covers 27 countries, while the most recent CLC map from 2018 covers 39 countries (GISAT, 2019).

The CLC development process includes satellite image interpretation (LANDSAT, SPOT, TM and MSS) and regional land cover information such as aerial photography, local knowledge and statistics. Table 2 shows the technical specifications for each available CLC dataset (Copernicus Land Monitoring Service., 2020).

Validation of CLC 2000 was done by reinterpretation based on IMAGE2000 data and by comparison to LUCAS LULC data. Reliability of CLC2000 was found to be 87.0 ± 0.8 % and agreement with LUCAS LULC data to be 74.8 ± 0.6 % respectively (Büttner and Maucha, 2006).

For CLC2006, no individual validation was done but a change reliability study from 2000 to 2006 was carried out (Büttner et al., 2011). The changes were found to be small (1.25 % of total CLC area) and it was concluded, that for CLC 2006 a similar accuracy can be expected. The CLC change product (2000–2006) was validated using stratified random sampling and including a weighted proportion of all occurring change types. Accuracy of changes was found to be 87.8 ± 3.3 % which confirms the assumption that CLC 2006 accuracy is similar to the CLC 2000 accuracy.

For CLC 2012, validation is done by evaluation of more than 25,000 sampling locations which were evaluated by experts (Jaffrain et al., 2017). Overall accuracy was found to be 85 %.

CORINE CLC is considered as one of the most consistent and most carefully prepared land cover product for Europe. Nevertheless, the product is only available for a few countries in Europe and few time steps. Therefore, CLC might be rather unsuitable for the use in high resolution RCMs investigating LULC change induced feedback mechanisms over continental scale domains. Yet, CLC provides a valuable source of high resolution LULC information. This information can be used to compare coarser, global LULC products over these countries to investigate their quality in a comparative assessment.

Its availability for three timesteps of the ESA CCI LC time series (i.e. 2000, 2006 and 2012) makes CLC a most valuable product for validating ESA CCI LC. CLC is classified into 3 levels with 44 land cover and land use classes on level 3 (Heymann, 1994).

2.3. Dataset harmonization

In land cover comparison, agreement measures depend on the semantic resolution of the chosen land cover typology where a lower number of classes is resulting in higher agreement/ accuracy (Bechtel et al., 2019). In order to avoid this issue, we decided to use an established harmonization method. Harmonization of classifications is done following Vilar et al. (2019) who provided a robust categorization method for both, ESA CCI LC and CLC to eight LULC categories in total (Table 3).

Overall, modification of utilized datasets was kept to a minimum leaving the resolution of both datasets unchanged.

ESA CCI LC and CLC are available in different projections. Therefore, the projection of the ESA CCI LC maps was transformed to fit the CLC ETRS89 Lambert Azimuthal Equal Area (LAEA). The projection is suitable for all approaches where true area representation is required and is further suited to the research area. Since the map product classification consists of discrete categories on a nominal scale, a nearest neighbor resampling strategy was applied. Both datasets were clipped to the extent of the research area in Eastern Europe. In order to account for geoprecision of the used data, the offset between datasets was looked into by comparing certain landmarks like coastlines and rectangular features. The offset was found not to exceed 20 m, therefore the geoprecision was found to be sufficient for the present analysis.

Table 3
Classification harmonization of ESA CCI LC and CLC[a].

	Category	ESA CCI LC	CLC
1	Agriculture	10, 11, 12, 20, 30, 40	2
2	Forest	50, 60, 61, 62, 70, 71, 72, 80, 81, 82, 90, 100, 160, 170	3.1
3	Grassland	110, 130	3.2.1
4	Wetland	180	4
5	Settlement	190	1
6	Shrubland	120, 121, 122	3.2.2, 3.2.3, 3.2.4
7	Sparse vegetation, bare areas, permanent snow and ice	140, 150, 152, 153, 200, 201, 202, 220	3.3
8	Water bodies	210	5

[a] Nomenclature can be found at http://maps.elie.ucl.ac.be/CCI/viewer/download/CCI-LC_Maps_Legend.pdf (ESA CCI LC) and https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines/html (CLC).

2.4. Methodology

In the present study two LC data comparison approaches are carried out, a PAC and a comparison method using only the majority CLC class per ESA CCI LC pixel. For both approaches, overall accuracy (OA), user's accuracy (UA) and producer's accuracy (PA) are evaluated for all assessed timesteps and countries. In the following assessment the measures OA, UA and PA are referred to as overall agreement (OA = OA), precision (PR = UA), recall (RE = PA), since they are used as comparative measures and not as accuracy measures. Spatial comparison of ESA CCI LC and CLC is carried out using the PAC maps. All calculations were performed using SAGA GIS (Conrad et al., 2015).

In the PAC, the proportion of each CLC category per ESA CCI LC pixel is counted and added. Table 4 shows the calculated area of three CLC classes for the proportional area method (Ref_{prop}) and for the majority method (Ref_{maj}) for four example ESA CCI LC pixels (Fig. 1).

The area proportions are added per assessed class or category. Agreement is measured by utilizing an area based confusion matrix (Story and Congalton, 1986; Stehman, 1997) where the rows of the matrix correspond to the assessed dataset (ESA CCI LC) and the columns correspond to the reference dataset. In the following comparison assessment, the confusion matrix is referred to as contingency table, since the confusion matrix term is rather to be used in the context of an accuracy assessment. The cell values are calculated as follows in Eq. (1):

$$c_{ij} = \sum_{s=1}^{r} pp_{ij}(s) \tag{1}$$

Table 2
CORINE (CLC) Product information.

	Satellite data	Time consistency	Geometric accuracy satellite data	Geometric accuracy	Thematic accuracy
CLC 1990	Landsat-5 MSS/TM single date	1986–1998	≤50 m	100 m	≥85 % (probably not achieved)
CLC 2000	Landsat-7 ETM single date	2000 +/- 1 year	≤25 m	>100 m	≥85 %
CLC 2006	SPOT-4/5 and IRS LISS III dual date	2006 +/- 1 year	≤25 m	>100 m	≥85 %
CLC 2012	IRS LISS III and RapidEye dual date	2011–2012	≤25 m	>100 m	≥85 %
CLC 2018	Sentinel 2A/ 2B	2017 mid-spring to mid-autumn	≤10 m	>100 m	≥85 %

Table 4
Proportional area (Ref_{prop}) of three LULC categories (see Table 3 for category descriptions) for four example ESA CCI LC pixels (Fig. 1) in comparison to a majority comparison method (Ref_{maj}).

Pixel ID *(k)*	Map Class *i* (ESA CCI LC)	Ref_{prop}		Ref_{maj}	
		Reference Class *j* (CLC)	Proportion in pixel pp_{ij} *(k)*	Reference Class *j* (CLC)	Proportion in pixel pp_{ij} *(k)*
1	2	1	pp_{11} (1) = 0.20	1	pp_{11} (1) = 0
		2	pp_{12} (1) = 0.45	2	pp_{12} (1) = 1
		5	pp_{15} (1) = 0.35	5	pp_{15} (1) = 0
2	1	1	pp_{11} (2) = 0.60	1	pp_{11} (2) = 1
		2	pp_{12} (2) = 0.25	2	pp_{12} (2) = 0
		5	pp_{15} (2) = 0.15	5	pp_{15} (2) = 0
3	5	1	pp_{11} (3) = 0.30	1	pp_{11} (3) = 0
		2	pp_{12} (3) = 0.10	2	pp_{12} (3) = 0
		5	pp_{15} (3) = 0.60	5	pp_{15} (3) = 1
4	2	1	pp_{11} (4) = 0.25	1	pp_{11} (4) = 0
		2	pp_{12} (4) = 0.6	2	pp_{12} (4) = 1
		5	pp_{15} (4) = 0.15	5	pp_{15} (4) = 0

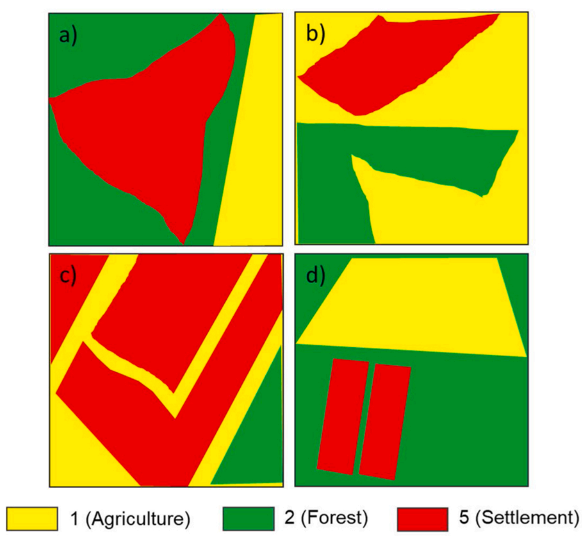

Fig. 1. PAC with three LULC classes for four example pixels (a-d). Proportional area covered by CLC categories agriculture, forest and settlement can be found in Table 4. Pixel IDs refer to example pixels as follows: a) Pixel ID = 1; b) Pixel ID = 2; c) Pixel ID = 3; d) Pixel ID = 4.

where c_{ij} is the value of cell in row i and column j, r is the number of spatial units in the reference dataset and $pp_{ij}(s)$ is the proportion of class j in the spatial unit s assigned to class i in the assessment.

The contingency table for the four example cells is shown in Table 5. In comparison, Table 6 shows the contingency table for the majority method. The main differences between the two matrices are the column sums appearing in decimals that correspond to the added proportions of area per class.

Indices of comparison are then derived from the resulting contingency tables. OA, PR and RE are calculated using Eqs. (2),(3) and (4).

$$OA_i = \frac{\sum_{k=1}^{n} c_{kk}}{n} \tag{2}$$

$$PR_i = \frac{c_{ii}}{\sum_{k=1}^{n} c_{ik}} \tag{3}$$

Table 5
Example contingency table using the PAC. Values derived from Table 4 (Ref_{prop}).

	1 (Agriculture)	2 (Forest)	5 (Settlements)	Sum
1 (Agriculture)	0.6	0.25	0.15	1
2 (Forest)	0.45	1.05	0.5	2
3 (Settlements)	0.3	0.1	0.6	1
Sum	1.35	1.4	1.25	4

Table 6
Example contingency table using the majority method. Values derived from Table 4 (Ref_{maj}).

	1 (Agriculture)	2 (Forest)	5 (Settlements)	Sum
1 (Agriculture)	1	0	0	1
2 (Forest)	0	2	0	2
3 (Settlements)	0	0	1	1
Sum	1	2	1	4

$$RE_j = \frac{c_{jj}}{\sum_{k=1}^{n} c_{kj}} \tag{4}$$

where c_{ij} is the value of the cell in row i in column j of the contingency table and n is the number of classes or categories in the map.

In addition, visual analysis of proportional overlay maps (according to Fig. 1) can indicate spatial patterns of agreement between ESA CCI LC and CLC and therefore reveal inconsistencies for certain categories.

2.5. Research area

The total size of the Research Area (RA) in Eastern Europe is ~867.000 km^2. Countries included in the RA investigated in this paper are Estonia, Latvia, Lithuania, Poland, Hungary, Romania and Slovakia (Fig. 2) for the available CLC time steps 2000, 2006 and 2012. Cells that are not covered by one of the used datasets are left out of the analysis.

The RA in Eastern Europe and the Baltic States was chosen following preliminary investigations that show extensive land use changes suggested by ESA CCI LC (Figs. 2 & 3). The bars in Fig. 3 shows the area gains and losses of the eight LULC categories according to the dataset. The colors show the respective categories that the area was gained from or lost to. Most dynamics are found in the categories agriculture and forest. Most agricultural area loss is due to forest gain and most of agricultural area gain is due to forest area loss. The dynamics might indicate a spatial shift of LULC categories, where net area is not lost but moved. Most of settlement area gain is due to agricultural area loss which is a common dynamic all over the European continent, where urban areas are known to expand.

3. Results

3.1. Contingency tables and agreement measures

Table 7 shows the proportion of each harmonized LULC category for each assessed time step according to ESA CCI LC and CLC. Dominant categories are agriculture and forest with total share of more than 80 % in both datasets. Sparse vegetation is almost non-existent in the RA in both datasets after classification harmonization while shrubland is completely vanished in ESA CCI LC. For agriculture, the difference between the two datasets is ~2 % and for forest ~4% respectively. For grassland and shrubland the proportional areas differ widely.

The proportional areas of categories do not change substantially between the assessed time steps for most of the categories in the individual data sets. The largest relative changes are found for settlement areas, which double the percentage share over the assessed time period in the ESA CCI LC dataset. However, CLC settlement area proportions are ~5 % while the area in ESA CCI LC is much lower with 1.5–2.5 %.

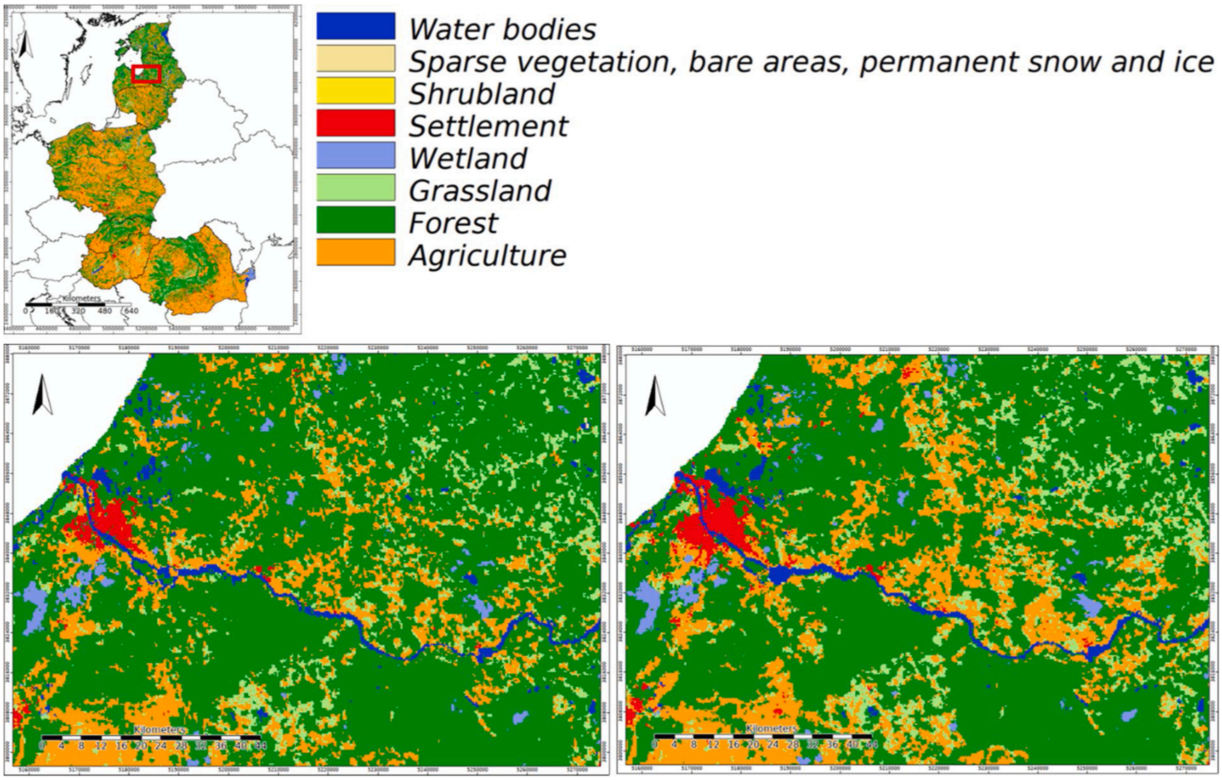

Fig. 2. Research area (upper left) with ESA CCI LC representation of the city Riga (Latvia) in 1992 (lower left) and 2015 (lower right) as an example for extensive LULC changes in the whole RA.

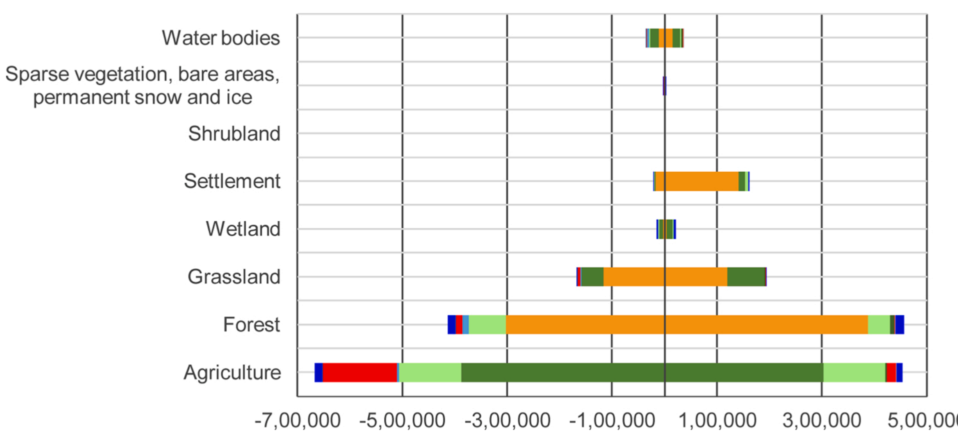

Fig. 3. Category to category changes from 1992-2015 in the RA according to ESA-CCI LC (in 300 × 300 m cell units).

A contingency table gives quantitative information on agreement between the datasets and on issues between LULC categories. Table 8 shows the contingency table for the PAC between ESA CCI LC and CLC for the year 2012. Table 9 shows the same matrix but for the MA. Since there are no significant changes in agreement measures RE and PR among the assessed time steps, only one pair of matrices is exemplarily shown.

Agreement (PA) of ESA CCI LC agriculture and forest with CLC is over 80 %. ~58.000 pixels of CLC agriculture are classified as settlement by ESA CCI LC which is not a major share for agriculture areas but makes a considerable difference for settlements. ESA CCI LC forest areas show the highest agreement with CLC. Nevertheless, the classification of CLC forest areas as agriculture and grassland areas might be not negligible for the performance of RCMs and needs to be discussed. The low accuracies of categories 6 and 7 (shrubland and sparse vegetation) might occur because they are not (or almost not) present in the RA in one or both datasets used after classification harmonization. With ~36 %, agreement for settlements is very low. Most of the CLC settlement areas are classified as agriculture by ESA CCI LC. Grassland and wetlands accuracies are also not on a high level. For both categories, ~50 % of the respective area is classified as forest or agriculture. Possible reasons for this very high disagreement in three categories between the two datasets are investigated in the visual map analysis. Table 9 shows a very similar picture in the contingency table for the majority method. The differences between the methods are not apparent using only the confusion matrices. Nevertheless, the modification step of using only the dominant LULC class of the reference dataset in the MA can give a biased picture of the agreement between two datasets, depending on resolution and

Table 7
Proportional area of every LULC category for ESA CCI LC and CLC [%]. Assessed time steps are 2000, 2006 & 2012.

		Area per category in the RA [%]					
		2000		2006		2012	
	Category	ESA CCI LC	CLC	ESA CCI LC	CLC	ESA CCI LC	CLC
1	Agriculture	54.99	57.74	53.81	56.48	53.63	55.54
2	Forest	34.82	30.87	35.02	31.24	35.05	31.38
3	Grassland	6.08	0.79	6.11	1.05	6.15	1.04
4	Wetland	0.90	1.13	0.90	1.09	0.93	1.07
5	Settlement	1.54	4.57	2.47	4.86	2.58	5.19
6	Shrubland	0	3.02	0	3.38	0	3.86
7	Sparse vegetation, bare areas, permanent snow and ice	0.04	0.10	0.04	0.07	0.04	0.07
8	Water bodies	1.60	1.74	1.61	1.79	1.60	1.81

complexity of classification.

A summary of agreement measures OA, RE and PR is given in Tables 10–12. Like for the confusion matrices, only minor differences are found between the two methods for the agreement measures. OA (Table 10) gives a quantification of the overall agreement of the two datasets. It is steady over all assessed timesteps and for both methods. A more distinguished picture is given by PA and PR (Tables 11 and 12).

Table 11 shows the PR, the proportional share of reference pixels of CLC correctly classified by ESA CCI LC for each category. PR for agricultural and forest is highest. Also, the PR for settlements is high in 2000 but is decreasing considerably per timestep for both approaches. PR for water bodies appears to increase slightly but remains below 75 %. A relatively low proportion of wetlands and an extremely low proportion of grassland is classified correctly by ESA CCI LC.

The RE in Table 12 shows the proportion of area classified by ESA CCI LC that agrees with the respective category in CLC. While PR gives information on the reliability of the assessed dataset, RE gives

Table 10
Overall agreement for all assessed categories and time steps.

Overall agreement (%)					
2000		2006		2012	
Ref_{prop}	Ref_{maj}	Ref_{prop}	Ref_{maj}	Ref_{prop}	Ref_{maj}
76.38	76.59	76.64	76.68	76.21	76.26

Table 11
Precision for all assessed categories and time steps.

	Precision (%)					
	2000		2006		2012	
	Ref_{prop}	Ref_{maj}	Ref_{prop}	Ref_{maj}	Ref_{prop}	Ref_{maj}
1 Agriculture	85.30	85.48	85.51	85.47	84.72	84.69
2 Forest	75.12	75.91	75.90	76.64	75.99	76.68
3 Grassland	5.83	5.93	7.75	7.94	7.64	7.74
4 Wetland	61.66	62.88	65.76	65.89	64.04	64.37
5 Settlement	79.24	78.61	71.46	71.00	72.60	72.06
6 Shrubland	0.00	–	0.00	–	0.00	–
7 Sparse vegetation	36.08	37.35	33.90	33.92	35.01	36.55
8 Water bodies	72.47	72.48	73.25	73.38	73.84	73.88

Table 12
Recall for all assessed categories and time steps.

	Recall (%)					
	2000		2006		2012	
	Ref_{prop}	Ref_{maj}	Ref_{prop}	Ref_{maj}	Ref_{prop}	Ref_{maj}
1 Agriculture	81.24	80.86	81.48	81.04	81.81	79.46
2 Forest	84.73	85.51	85.09	85.77	84.88	84.75
3 Grassland	44.81	46.08	44.87	46.05	45.01	45.78
4 Wetland	48.93	50.69	54.25	55.46	55.36	56.68
5 Settlement	26.79	27.68	36.35	37.13	36.05	36.87
6 Shrubland	–	0	–	0	–	0
7 Sparse vegetation	15.33	18.13	20.55	23.09	20.55	22.20
8 Water bodies	66.29	68.30	65.34	67.48	65.34	66.76

Table 8
contingency table for the year 2012 - Ref_{prop}. Proportional area and SUM are given in thousands (e.g. 8,6 = 8600 spatial units). Recall (RE) and precision (PR) are given in percentage.

	1	2	3	4	5	6	7	8	SUM	PR [%]
1 Agriculture	4373.9	379.0	20.3	14.6	264.0	79.5	1.8	29.9	5163.0	84.72
2 Forest	437.9	2564.0	32.6	21.6	34.0	262.3	1.7	19.8	3374.0	75.99
3 Grassland	450.8	51.0	45.2	5.5	15.7	18.7	1.7	3.3	592.0	7.64
4 Wetland	8.5	8.1	1.6	57.4	0.8	8.6	0.2	4.4	89.6	64.04
5 Settlement	58.7	5.2	0.0	0.3	180.4	1.0	0.1	2.7	248.4	72.60
6 Shrubland	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00
7 Sparse vegetation	0.2	0.2	0.2	0.1	1.5	0.5	1.5	0.3	4.4	35.01
8 Water bodies	16.3	13.2	0.5	4.1	4.1	1.8	0.3	113.9	154.2	73.84
SUM	5346.4	3020.7	100.5	103.6	500.4	372.3	7.5	174.3		
RE [%]	81.81	84.88	45.01	55.36	36.05	0.00	20.55	65.34		

Table 9
contingency table for the year 2012 - Ref_{maj}. Area of categories and SUM are given in thousands (e.g. 8,6 = 8600 spatial units). Recall (RE) and precision (PR) are given in percentage.

	1	2	3	4	5	6	7	8	SUM	PR [%]
1 Agriculture	4173.1	358.7	20.6	13.7	255.7	77.6	1.8	26.7	4927.8	84.68
2 Forest	415.6	2551.4	30.7	20.8	32.5	256.8	1.7	19.2	3328.9	76.64
3 Grassland	452.6	47.5	45.5	5.5	15.1	17.9	1.5	3.1	588.6	7.74
4 Wetland	8.6	7.3	1.7	58.2	0.7	8.6	0.3	5.0	90.5	64.37
5 Settlement	60.4	5.6	0.1	0.4	180.5	0.9	0.1	2.6	250.5	72.08
6 Shrubland	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	50.00
7 Sparse vegetation	0.3	0.2	0.2	0.1	1.4	0.4	1.6	0.3	4.4	36.56
8 Water bodies	16.4	13.6	0.5	4.1	3.8	1.6	0.4	114.3	154.7	73.90
SUM	5127.0	2984.2	99.4	102.7	489.9	363.7	7.3	171.2		
RE [%]	81.39	85.50	45.84	56.70	36.86	0.00	22.20	66.77		

information on the probability of a pixel in the reference map being classified correctly by the assessed dataset. Previously identified issues with grassland, wetlands, settlements and water bodies appear for each assessed timestep. A considerable increase of RE for wetlands and settlements can be noted between 2000 and 2006 but the RE is still below a reasonable level.

It needs to be considered, that slight changes in proportional area influence the agreement measures especially for categories that are rather rarely present in the RA (categories 3, 4,5 and 8). Therefore, in addition to the agreement measures, we provide the visual map analysis, which can give insight on the spatial occurrence of disagreement, also for categories that are rarely occurring in the respective RA.

3.2. Visual map analysis

A spatial analysis of the data, in addition to the raw quantification shown in the confusion matrices, reveals in which areas the disagreements between ESA CCI LC and CLC occur. Since agricultural areas and forest areas are the most well represented categories in the RA with a RE of 81.8 % and 84.88 % respectively, special focus of the maps is on grassland, wetlands, settlements and water bodies. The categories shrubland and sparse vegetation, bare areas, permanent snow and ice are not relevant for the RA and the present analysis and will therefore be neglected in the visual map analysis. The effects of disagreement occur throughout the whole RA. Since the mostly small LULC features are not displayable in a sufficient way for the whole RA, Figs. 4–7 show map sections which highlight typical effects of the disagreeing categories grassland, wetlands, settlements and water bodies for the most recent assessed timestep 2012.

3.2.1. Grassland

RE of grassland for ESA CCI LC reaches around 45 % in the RA. For 2012, ~20 % of the CLC grassland areas are classified as agricultural areas and ~32 % are classified as forest while only ~45 % are classified correctly. According to CLC most of the grassland areas in the RA are in Hungary and around the Carpathian Arc. Fig. 4 shows that the grassland areas classified as forest and agriculture by ESA CCI LC are located around correctly classified areas. Due to the location of the disagreeing areas mostly in mountainous regions, the orography of the area might have major influence on classification results of grassland in ESA CCI LC. Another reason for the low agreement is the classification harmonization. The ESA CCI LC classification includes many mixed classes where for example grassland and low tree density occur as one class. In the harmonization, many of these classes are assigned to the forest class, although the classes might not incorporate actual forest surface properties. The same applies to agriculture in ESA CCI LC, where mixed agriculture-shrubland or agriculture-forest classes are present.

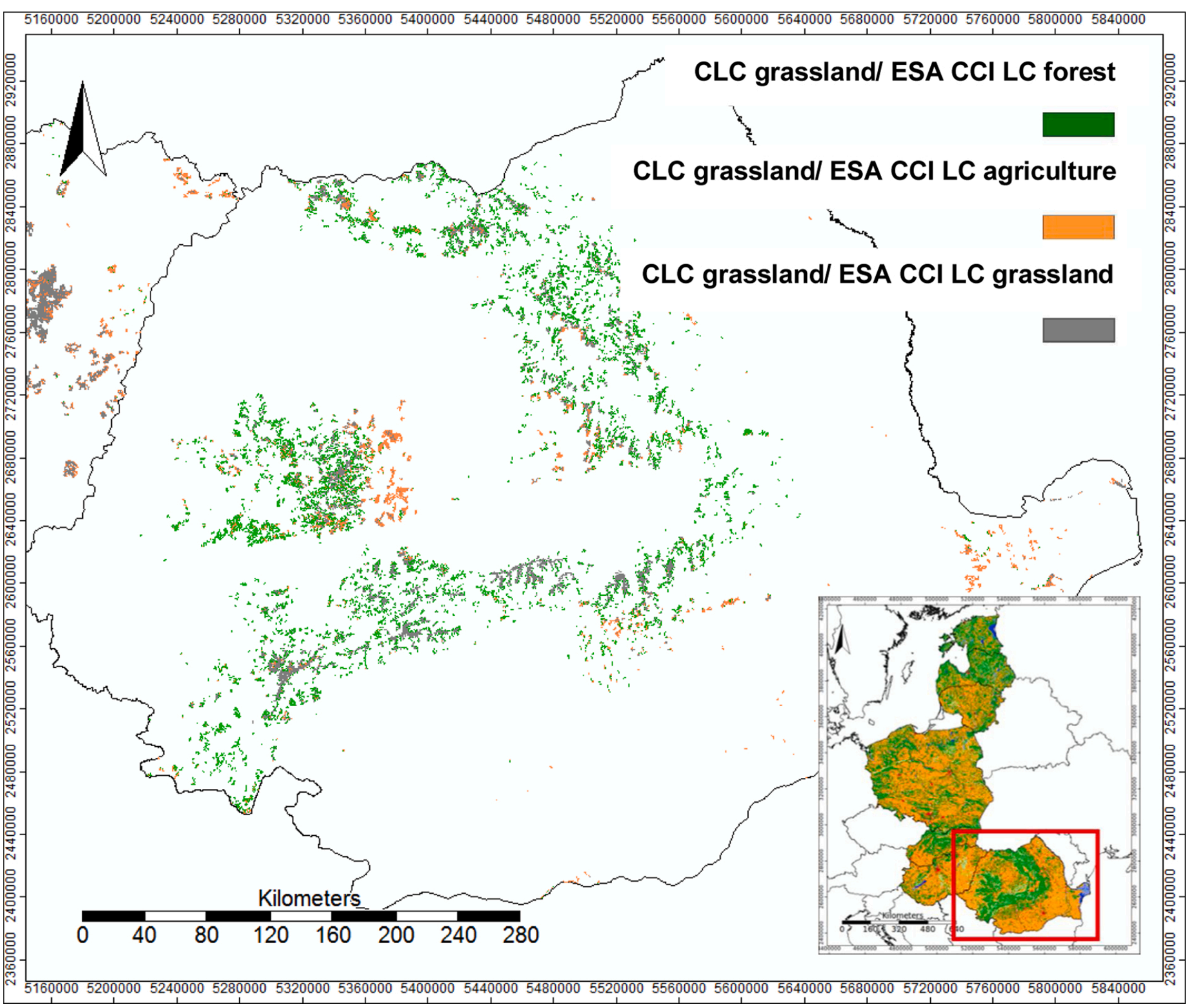

Fig. 4. Romania and parts of Hungary including the grassland proportional area overlay. Grey areas show the agreement of ESA CCI LC with CLC in the grassland category, colors green and orange show the disagreement and the respective other ESA CCI LC category (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

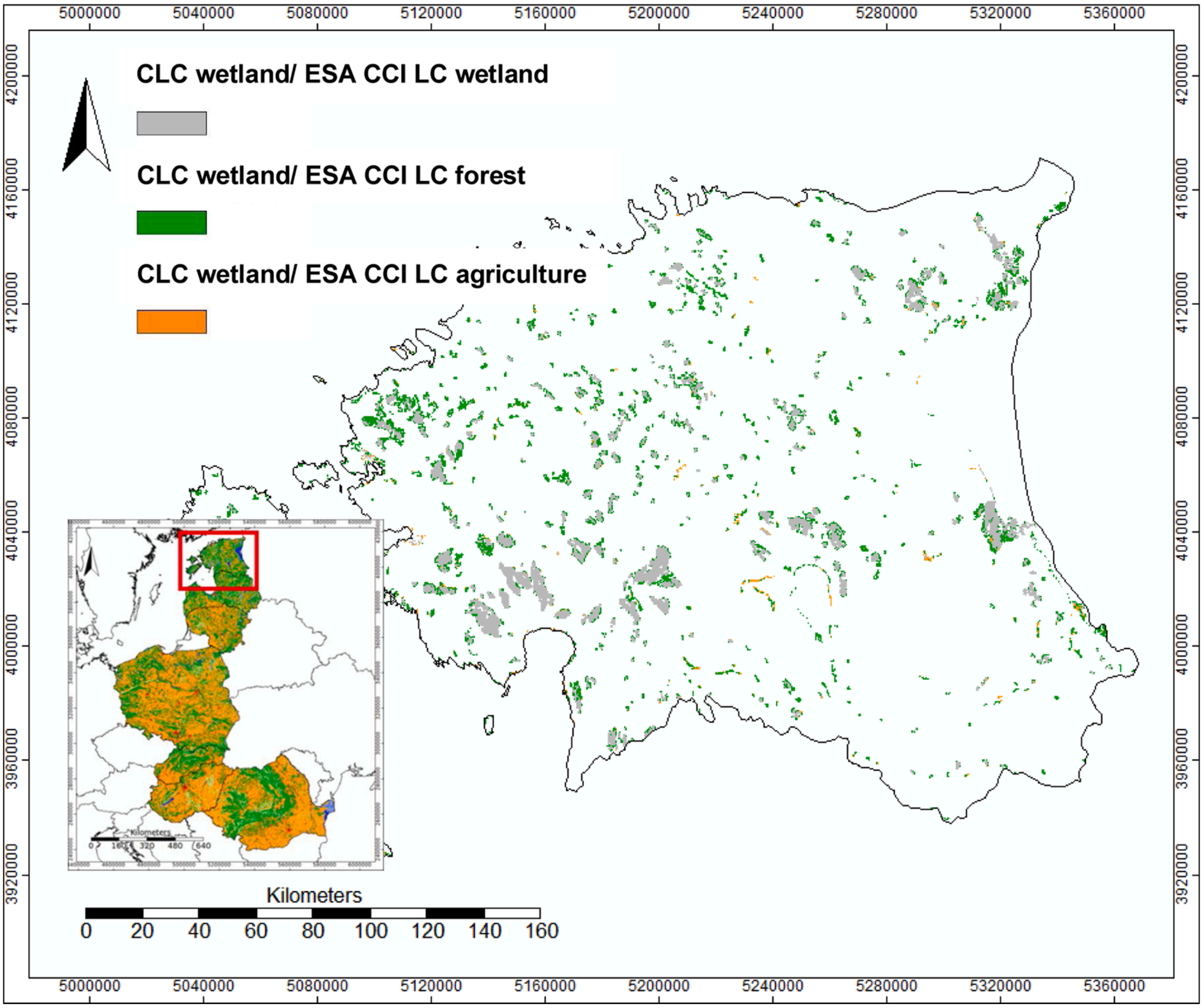

Fig. 5. Proportional area overly map for wetlands in Continental Estonia. Grey areas show the agreement of both data sets.

3.2.2. *Wetlands*

Wetlands in the RA are found in the Danube river delta in Romania and in the Baltic States Estonia, Latvia and Lithuania where they appear as fens and bogs (Fig. 4). Most of the falsely classified areas are assigned to the categories forest and agriculture by ESA CCI LC (reference year 2012: ~14 % to agriculture and ~21 % to forest, respectively. ~57 % of wetland areas were correctly classified). The map shows the proportional overlay map of ESA CCI LC and CLC for wetlands in Estonia for 2012. Here, the wetlands themselves are reasonably well recognized but the surrounding transition areas are often confused with forest or agricultural areas. Although the overall proportion of wetlands in the RA is relatively small (~1 %) an underestimation of wetlands might have an influence on RCM performance.

3.2.3. *Settlements*

In the reference year 2012, ~52 % of the total settlement area was classified as agriculture by ESA CCI LC. Fig. 6 shows the Romanian cities Bucharest, Ploiești and the surrounding area as an example. ESA CCI LC works quite well for metropolitan areas but cannot capture smaller, rural settlements which are rather classified as agriculture. Further, linearly distributed features and infrastructure like streets are not identified as settlements or urban areas. There is no clear pattern indicated regarding an influencing spatial factor. The 300 m pixel size of ESA CCI LC has most definitely an influence on the classification results. However, there are also settlement structures seen in the map that are not recognized as such but that are clearly bigger than the minimum mapping unit (MMU) of ESA CCI LC (9 ha). ESA CCI LC includes only one urban LULC class while CLC includes a set of more distinguished artificial areas, which might also be partly responsible for the low agreement of the datasets in the RA.

3.2.4. *Water bodies*

Around 66 % of water bodies in the RA are correctly classified by ESA CCI LC. As for all other categories, most of the disagreeing areas (~28 % in total) are assigned to agriculture and forest. Fig. 6 shows that most of the disagreeing areas are in or around narrow streams that could not be identified as such by ESA CCI LC, presumably due to surface reflection or surrounding vegetation. Nevertheless, the biggest issue here might be the resolution. The ESA CCI LC resolution of 300 m is in fact not able to capture micro scale landscape features, especially when the features are line shaped.

4. Discussion

The present work investigated the ESA CCI LC dataset regarding agreement with CLC and spatial disagreement patterns of LULC in Eastern Europe and the Baltic States for 2000, 2006 and 2012. The aim of the study was to test if ESA CCI LC agrees with a regional, high-resolution LULC product to investigate the datasets suitability for implementation in regional climate models. Classification harmonization was achieved through transformation of both dataset classifications into eight LULC categories. A PAC was tested against a majority

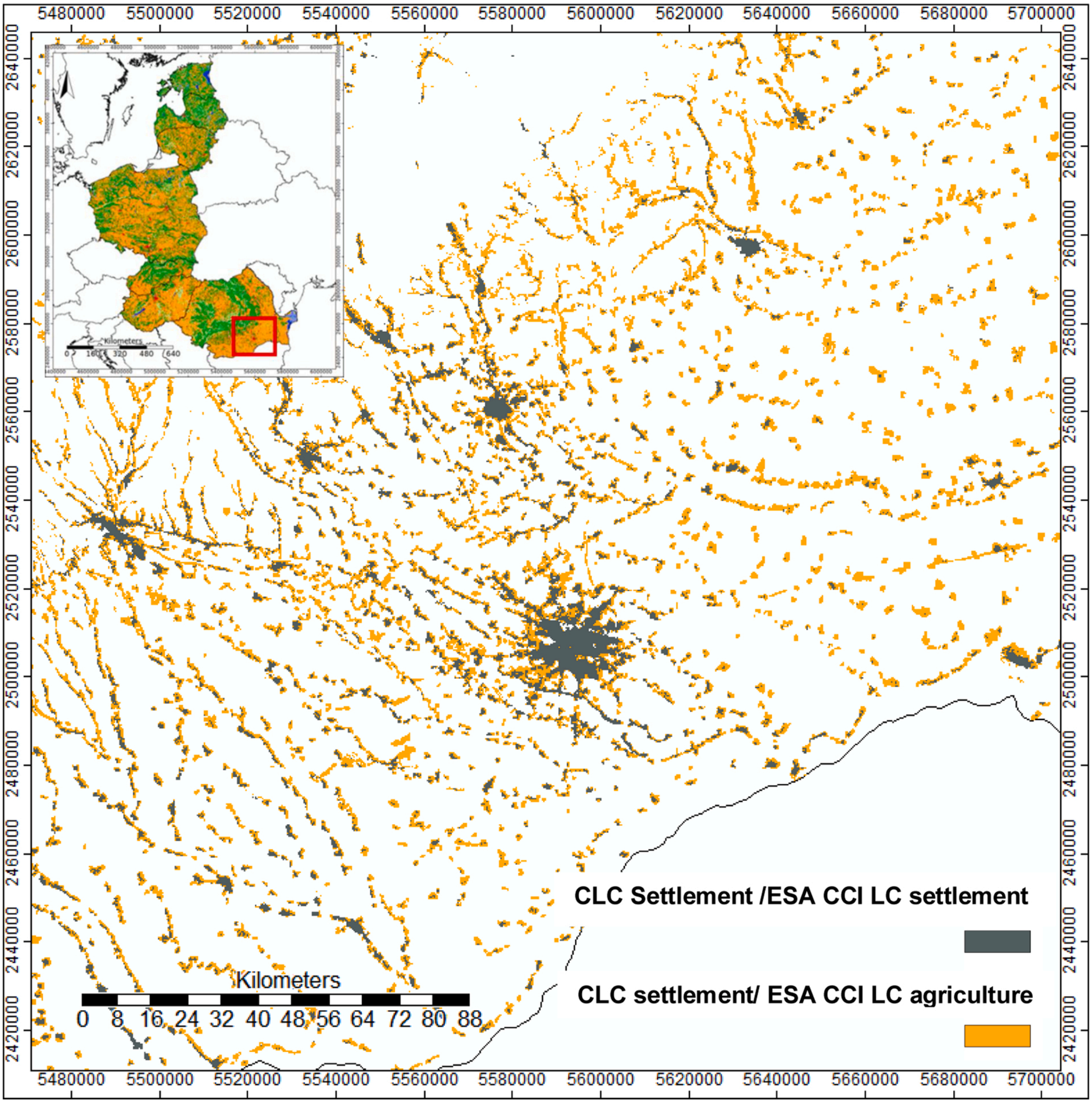

Fig. 6. Proportional overlay map for settlements. ~50 % of the CLC settlement areas are classified as agriculture by ESA CCI LC. Biggest urban agglomerations are the Romanian cities Bucharest (middle of map) and Ploiești (North of Bucharest).

comparison. Main benefit of the PAC for a LC comparative assessment of two or more LC products is that it makes datasets with different resolution and structure comparable without performing preliminary spatial modification. Following the quantitative assessment, spatial disagreement issues for the all categories were investigated through visual analyses. Several issues could be identified for all categories except for shrublands and sparse vegetation, which are not relevant for the RA.

Previous work found the ESA CCI LC epoch time steps (2000, 2005, 2010 and 2015) to be relatively accurate on a global scale (75.4 % for 2015 according to Achard et al., 2017). The present analysis shows consistent OA results for the assessed region (~76 %) for both approaches and all timesteps respectively. However, the comparability between the agreement measures is limited because in all comparative assessment approaches, different data harmonization techniques are used. Nevertheless, the results of the global studies can give a good measure of reliability of the present results. A further restriction is the absence of the categories shrublands and sparse vegetation in the RA. To achieve full comparability with other approaches regarding overall agreement, all existing categories need to be present in the investigated region. Since the aim of the present analysis was to investigate agreement of ESA CCI LC with a regional LULC product in a certain region and from a climate modelling perspective, this criterion can be neglected.

Based on the results for the RA, there are no significant differences in agreement measures between the majority and the PAC for the investigated time steps and categories. The comparison was carried out on the basis of recent findings that showed a considerable bias in agreement measures for the MA due to the loss of small landscape features (Fonte et al., 2020). The findings could not be confirmed in the present study. A reason could be the MMU of CLC which is in fact larger than the MMU of ESA CCI LC (9 ha and 25 ha respectively). Due to the manufacturing process of CLC, small landscape features still appear which makes the product have greater detail than ESA CCI LC. It is expected that when a high resolution product in vector format is used as reference that the differences in agreement measures between the two methods become more apparent. The PAC is therefore still recommended because it can be applied to combine gridded or vector data sets regardless their

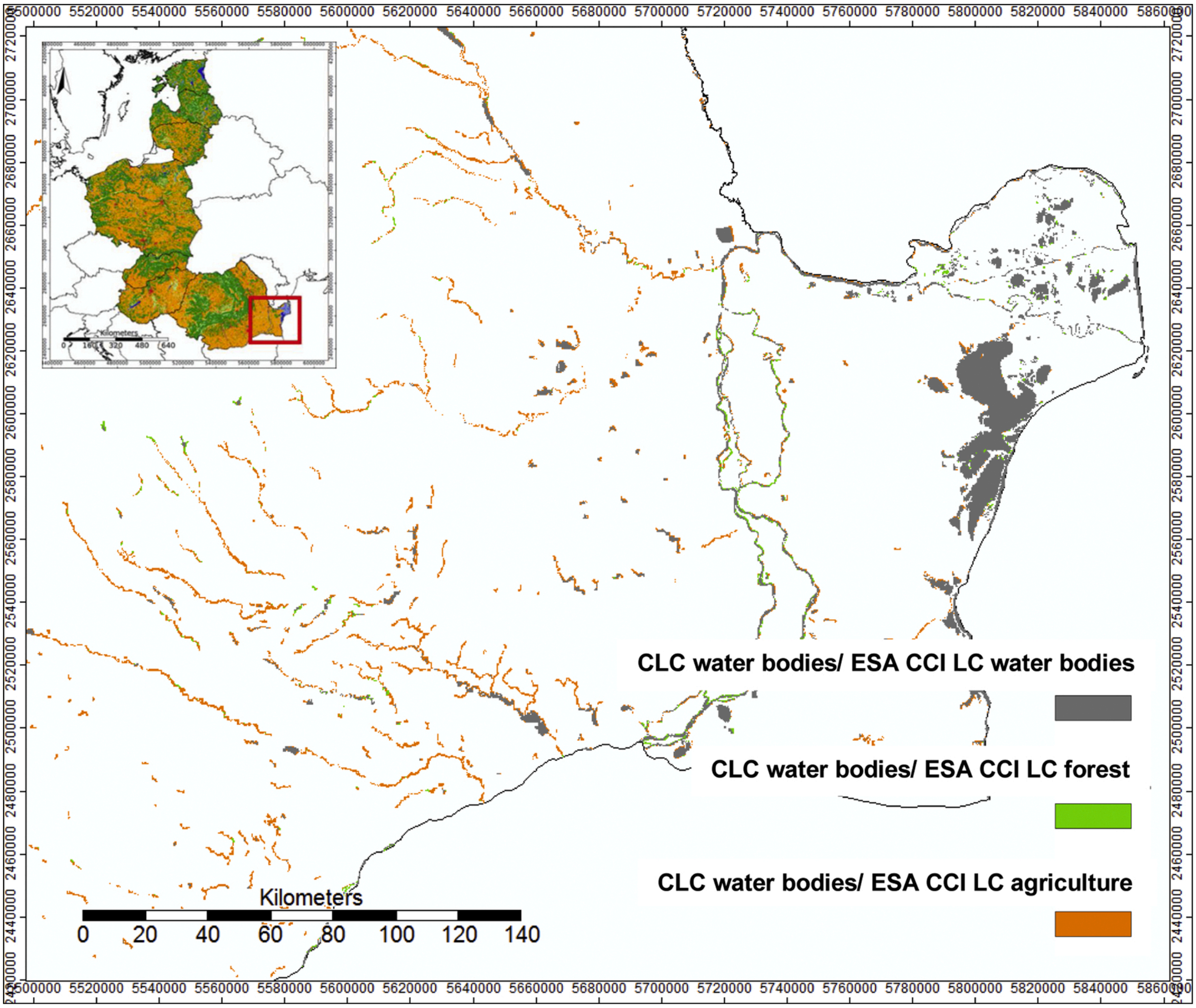

Fig. 7. Proportional overlay map for water bodies. ~28 % of the CLC water bodies are classified as agriculture or forest by ESA CCI LC. The map shows the southeastern part of Romania including the water body proportional agreement results.

structure and resolution. The possibility to compare LULC data products with even less modifications makes the approach advantageous against common LULC dataset comparison techniques. It is therefore highly recommended to be used in LULC dataset comparisons including high-resolution gridded or vector format land cover data.

No substantial changes in overall agreement of ESA CCI LC and CLC over time were found. ESA CCI LC shows a consistent agreement of ~76 % with CLC for every assessed time step. Since the accuracy of CLC was not consistently assessed for all time steps, the reference data set can be biasing the present analysis. Although the CLC database is considered to be one of the most accurate and consistent land cover products for Europe, errors can occur during the production process. Since CLC is only available for certain countries in Europe, independent regional land cover data sets and other available ground truth data need to be used when transferring the PAC to other regions of the globe accordingly. The fact that the data structure of the reference data set does not matter for the PAC solves the issue of finding suitable reference data partly. All types of reference data, which might come in different formats and structures should be used simultaneously for comparison without spatial transformation biasing the agreement measures.

Agriculture and forest are the predominant categories in the RA. Compared to CLC, Recall of ESA CCI LC is over 80 % for both which can be considered as relatively high. However, consistency between other global state of the art LC products seems problematic for both categories (Fritz et al., 2011) Especially for land-climate interaction scenarios that depend highly on LU and LC, overestimation of forest and agriculture has, due to specific LULC related surface properties (e.g. high surface roughness and albedo), most certainly an impact on RCM performance. Therefore, it can be beneficial to use a modification or a combination of the ESA CCI LC data set with other reliable, independent LULC data (Chu et al., 2011). Crowdsourcing approaches like the Geo-Wiki (https://www.geo-wiki.org/) can further be a valuable source of high resolution input data to test reliability and to refine LULC maps in the desired region of interest.

For grassland, mostly present in Hungary and Romania, inaccuracies tend to be found in areas around mountains and in valleys along the Carpathian Arc. Over 50 % of the grassland is differently classified by ESA CCI LC, ~32 % of it as forest. These inaccuracies might occur due to shadowing in the valleys caused by steep and narrow slopes. The constraints of satellite image classifications in mountainous regions are widely known in the remote sensing community and also permanent improvement was achieved during the last decades (Giles, 2001; Mostafa, 2017; Shahtahmassebi et al., 2013). It needs to be investigated if the issue for grassland occurs only in the investigated region or also in other mountainous regions to see whether this is a global issue of ESA CCI LC. Further, the mixed classes of ESA CCI LC are a biasing factor. It needs to be checked, if these disagreements also occur with different reference datasets or with a different dataset harmonization method.

Issues with wetlands are mostly found in the Baltic Countries where a considerable area of wetlands are classified as forest by ESA CCI LC.

These findings are consistent with (Törmä et al., 2015) who discovered misclassification of bogs, mires and marshes as forest in Finland. Against the background of widely differing surface properties of forest and wetland respectively, it should be investigated how this issue can be handled for implementation in RCMs.

Almost half of the settlement areas in the RA (~47 %) are classified as agricultural areas by ESA CCI LC. Mostly missing are rural settlements and infrastructure like roads without an agglomeration center and of linear shape. This might be because ESA CCI LC only incorporates one explicit urban LULC class. Since RCMs are moving to finer resolution, explicit representation of settlement structures is becoming critical. In particular, when RCMs are used on high spatial resolutions down to convection permitting scales below 4 km and are increasingly employing an explicit urban parameterization, relevance of high quality urban input data is extremely increased (Trusilova et al., 2013; Daniel et al., 2019; Langendijk et al., 2019). Modifications of the ESA CCI LC maps, which take the linear shaped infrastructure and rural settlements into account, might be necessary. For instance, LULC datasets that are specialized on urban representation like the urban atlas for Europe (Montero et al., 2014) or equivalent regional and global data products for other regions could be combined with the ESA CCI LC dataset.

The water bodies show a similar picture like the settlements regarding spatial disagreement issues. Coherent features like lakes or larger basins are captured very well but when it comes to rivers and streams, ESA CCI LC classifies water bodies as agriculture. Considering the surface properties of water bodies as well as the influence of rivers and streams on the surrounding landscape features, the missing features in ESA CCI LC will be relevant when implementing the data set into an RCM. Since the disagreement seems to be limited to the streams and rivers that are represented very well by CLC, it might be beneficial to improve the ESA CCI LC dataset through the integration of CLC or other suitable reference data that represents the river network in a more sufficient way. ESA CCI LC addressed the issue themselves with publishing a global water bodies map on 150 m horizontal resolution (Lamarche et al., 2017). That map could be integrated into ESA CCI LC before aggregation into coarser resolution, to preserve the small water body proportions for further use of the data.

5. Conclusion

The present work investigated the agreement of the ESA CCI LC dataset with CLC over Eastern Europe and the Baltic states to explore ESA CCI LCs suitability for implementation into RCMs over Europe. Three timesteps of the annual ESA CCI LC dataset were compared to CORINE LC, applying a PAC and a majority method, respectively. Classification harmonization of the assessed and the reference dataset was achieved through transformation into eight LULC categories.

Taking all results regarding overall agreement of ~76 % and temporal consistency into account the ESA CCI LC is considered to be suitable for implementation into RCMs by taking the following issues found in the present study under consideration. Disagreement with CORINE LC, which is considered a reliable reference for Europe, is for ESA CCI LC ~55 % for grassland, ~43 % for wetland, ~64 % for settlements and ~34 % for water bodies in the investigated RA.

Regional quality of the dataset must be confirmed for each region of interest separately with comparison to independent reference data. Although ESA CCI LC was found to be overall suitable for implementation in RCMs, spatial disagreement patterns were found that might influence RCM performance on certain scales which must also be investigated in each region of interest.

To get a deeper understanding of spatial disagreement not only for the RA but for the whole European Continent a consistent reference database for Europe should be developed. In addition to the continuous LULC product CLC, regional, independent datasets or also datasets specified on one LULC aspect should be included in the analysis which then can be compared using the PAC, regardless spatial structure and resolution.

In order to investigate the effects of detected LULC issues on the performance of RCMs and on regional climate in a region of interest, different LULC distributions and maps could be implemented into an RCM and tested. The testing would include different intensities of LULC over- and underestimation as well as varying spatial resolutions to quantify the impact of inaccuracies of LULC on different scales and to specify how to treat inaccuracies in LULCC products in regional climate modelling.

CRediT authorship contribution statement

V. Reinhart: Writing - original draft, Conceptualization, Validation, Formal analysis, Investigation. **C.C. Fonte:** Methodology, Writing - review & editing. **P. Hoffmann:** Writing - review & editing, Investigation. **B. Bechtel:** Methodology, Writing - review & editing. **D. Rechid:** Writing - review & editing, Supervision, Project administration. **J. Boehner:** Writing - review & editing, Supervision, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was conducted and financed within the framework of the Helmholtz Institute for Climate Service Science (HICSS), a cooperation between Climate Service Center Germany (GERICS) and Universität Hamburg as part of the project LANDMATE. Geospatial analyses were done with overwhelming support of the SAGA GIS User Group (http://www.saga-gis.org/en/index.html) at the Universität Hamburg (Institute of Geography, Section of Physical Geography). We thank Dr. Laurens Bower (GERICS) for the support in the internal review process. Finally, we thank the anonymous reviewers for their highly valuable feedback to our work.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jag.2020.102221.

References

Achard, F., Bontemps, S., Lamarche, C., Da Maet, T., Mayaux, P., Van Bogaert, E., Defourny, P., 2017. Quality Assessment of the CCI Land Cover Maps. https://www.esa-landcover-cci.org/?q=webfm_send/159.

Bechtel, B., Demuzere, M., Stewart, I.D., 2019. A weighted accuracy measure for land cover mapping: comment on Johnson et al. Local climate zone (LCZ) map accuracy assessments should account for land cover physical characteristics that affect the local thermal environment. Remote Sens. (Basel) 12 (11), 1769. https://doi.org/10.3390/rs12111769.

Bontemps, S., Defourny, P., Van Bogaert, E., Arino, O., Kalogirou, V., Perez, J.R., 2011. GLOBCOVER 2009-Products Description and Validation Report. URL: Http://Ionia1.Esrin. Esa.Int/Docs/GLOBCOVER2009_Validation_Report_2,*2*.

Bontemps, S., Defourny, P., Radoux, J., Van Bogaert, E., Lamarche, C., Achard, F., Mayaux, P., Boettcher, M., Brockmann, C., Kirches, G., Zülkhe, M., Kalogirou, V., Arino, O., 2013. Consistent global land cover maps for climate modeling communities: current achievements of the ESA's land cover CCI. ESA Living Planet Symposium 2013 (September), 9–13.

Brovkin, V., Boysen, L., Arora, V.K., Boisier, J.P., Cadule, P., Chini, L., Claussen, M., Friedlingstein, P., Gayler, V., Van den hurk, B.J.J.M., Hurtt, G.C., Jones, C.D., Kato, E., De noblet-ducoudre, N., Pacifico, F., Pongratz, J., Weiss, M., 2013. Effect of anthropogenic land-use and land-cover changes on climate and land carbon storage in CMIP5 projections for the twenty-first century. J. Clim. 26 (18), 6859–6881. https://doi.org/10.1175/JCLI-D-12-00623.1.

Büttner, G., Maucha, G., 2006. The thematic accuracy of Corine land cover 2000. Assessment using LUCAS (land use/cover area frame statistical survey). EEA Tech. Rep. 7/2006 (7), 85. www.eea.europa.eu.

Büttner, G., Maucha, G., Kosztra, B., 2011.). European Validation of Land Cover Changes in CLC2006 Project. EARSeL Symposium, Prague.
Castilla, G., Hay, G.J., 2007. Uncertainties in land use data To cite this version: HAL Id: Hal-00305113 Uncertainties in land use data. Hydrol. Earth Syst. Sci. Discuss. 11 (6), 1857–1868.
Chu, J., Syktus, J., McAlpine, C., Thatcher, M., Scarth, P., Jeffrey, S., Katzfey, J., Zhang, H., McGregor, J., Adams-Hosking, C., 2011. Validation of land surface products for modelling the climate impacts of large-scale revegetation in Queensland. In: MODSIM 2011 - 19th International Congress on Modelling and Simulation - Sustaining Our Future: Understanding and Living With Uncertainty. December 2011, pp. 2676–2682.
Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1. 4. Geosci. Model. Dev. Discuss. 8 (2).
Copernicus Land Monitoring Service, 2020. Evolution of Corine Land Cover. https://land.copernicus.eu/pan-european/corine-land-cover.
Daniel, M., Lemonsu, A., Déqué, M., Somot, S., Alias, A., Masson, V., 2019. Benefits of explicit urban parameterization in regional climate modeling to study climate and city interactions. Clim. Dyn. 52 (5–6), 2745–2764. https://doi.org/10.1007/s00382-018-4289-x.
Davin, E.L., Rechid, D., Breil, M., Cardoso, R.M., Coppola, E., Hoffmann, P., Jach, L.L., Katragkou, E., de Noblet-Ducoudré, N., Radtke, K., Raffa, M., Soares, P.M.M., Sofiadis, G., Strada, S., Strandberg, G., Tölle, M.H., Warrach-Sagi, K., Wulfmeyer, V., 2019. Biogeophysical impacts of forestation in Europe: First results from the LUCAS Regional Climate Model intercomparison. Earth Syst. Dyn. Discuss. (February), 1–31. https://doi.org/10.5194/esd-2019-4.
ESA, 2017. Land Cover CCI Product User Guide Version 2.0. http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf.
ESA, 2019. ESA/CCI Viewer. http://maps.elie.ucl.ac.be/CCI/viewer/download.php.
Fonte, C.C., See, L., Laso-Bayas, J.C., Lesiv, M., Fritz, S., 2020. Assessing the accuracy of land use land cover (LULC) maps using class proportions in the reference data. Isprs Ann. Photogramm. Remote. Sens. Spat. Inf. Sci. V-3–2020, 669–674. https://doi.org/10.5194/isprs-annals-V-3-2020-669-2020.
Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80 (1), 185–201. https://doi.org/10.1016/S0034-4257(01)00295-4.
Fritz, S., See, L., Rembold, F., 2010. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. Int. J. Remote Sens. 31 (9), 2237–2256. https://doi.org/10.1080/01431160902946598.
Fritz, S., See, L., McCallum, I., Schill, C., Obersteiner, M., Velde, Mvander, Boettcher, H., Havlík, P., Achard, F., 2011. Highlighting continued uncertainty in global land cover maps for the user community. Environ. Res. Lett. 6 (4), 044005. https://doi.org/10.1088/1748-9326/6/4/044005.
Fuchs, R., Herold, M., Verburg, P.H., Clevers, J.G.P.W., 2013. A high-resolution and harmonized model approach for reconstructing and analysing historic land changes in Europe. Biogeosciences 10 (3), 1543–1559. https://doi.org/10.5194/bg-10-1543-2013.
Gao, Y., Weiher, S., Markkanen, T., Pietikäinen, J.-P., Gregow, H., Henttonen, H.M., Jacob, D., Laaksonen, A., 2015. Implementation of the CORINE Land Use Classification in the Regional Climate Model REMO.
Georgievski, G., Hagemann, S., 2019. Characterizing uncertainties in the ESA-CCI land cover map of the epoch 2010 and their impacts on MPI-ESM climate simulations. Theor. Appl. Climatol. 137 (1–2), 1587–1603. https://doi.org/10.1007/s00704-018-2675-2.
Giles, P.T., 2001. Remote sensing and cast shadows in mountainous terrain. Photogramm. Eng. Remote Sensing 67 (7), 833–840.
GISAT, 2019. CLC Seamless Data Coverage. https://land.copernicus.eu/user-corner/technical-library/clc-country-coverage-v18.1.
Heymann, Y., 1994. CORINE Land Cover: Technical Guide. Office for Official Publ. of the Europ. Communities.
Houghton, R.A., House, J.I., Pongratz, J., Van Der Werf, G.R., Defries, R.S., Hansen, M.C., Le Quéré, C., Ramankutty, N., 2012. Carbon emissions from land use and land-cover change. Biogeosciences 9 (12), 5125–5142. https://doi.org/10.5194/bg-9-5125-2012.
Hua, T., Zhao, W., Liu, Y., Wang, S., Yang, S., 2018. Spatial consistency assessments for global land-cover datasets: a comparison among GLC2000, CCI LC, MCD12, GLOBCOVER and GLCNMO. Remote Sens. 10 (11), 1846.
Hurtt, G.C., Chini, L.P., Frolking, S., Betts, R.A., Feddema, J., Fischer, G., Fisk, J.P., Hibbard, K., Houghton, R.A., Janetos, A., Jones, C.D., Kindermann, G., Kinoshita, T., Klein Goldewijk, K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., et al., 2011. Harmonization of land-use scenarios for the period 1500-2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. Clim. Change 109 (1), 117–161. https://doi.org/10.1007/s10584-011-0153-2.
Jaffrain, G., Sannier, C., Pennec, A., Dufourmont, H., 2017. CORINE land cover 2012 final validation report. Copernicus Land Monitoring 214.
Koubodana, D.H., Diekkrüger, B., Näschen, K., Adounkpe, J., Atchonouglo, K., 2019. Impact of the accuracy of land cover data sets on the accuracy of land cover change scenarios in the Mono River Basin, Togo, West Africa. Int. J. Adv. Remote. Sens. Gis 8 (1), 3073–3095. https://doi.org/10.23953/cloud.ijarsg.422.
Lamarche, C., Santoro, M., Bontemps, S., d'Andrimont, R., Radoux, J., Giustarini, L., Brockmann, C., Wevers, J., Defourny, P., Arino, O., 2017. Compilation and validation of SAR and optical data products for a complete and global map of inland/ocean water tailored to the climate modeling community. Remote Sens. (Basel) 9 (1), 36.
Langendijk, G.S., Rechid, D., Jacob, D., 2019. Urban areas and urban–rural contrasts under climate change: what does the EURO-CORDEX ensemble tell us?—investigating near surface humidity in Berlin and its surroundings. Atmosphere 10 (12), 730.
Li, W., MacBean, N., Ciais, P., Defourny, P., Lamarche, C., Bontemps, S., Houghton, R.A., Peng, S., 2017. Gross and net land cover changes based on plant functional types derived from the annual ESA CCI land cover maps. Earth Syst. Sci. Data Discuss. 1–23. https://doi.org/10.5194/essd-2017-74.
Loveland, T.R., Belward, A.S., 1997. The IGBP-DIS global 1km land cover data set, DISCover: First results. Int. J. Remote Sens. 18 (15), 3289–3295.
Montero, E., Van Wolvelaer, J., Garzón, A., 2014. The European urban atlas. Land Use and Land Cover Mapping in Europe. Springer, pp. 115–124.
Mostafa, Y., 2017. A review on various shadow detection and compensation techniques in remote sensing images. Can. J. Remote. Sens. 43 (6), 545–562.
Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57. https://doi.org/10.1016/j.rse.2014.02.015.
Pérez-Hoyos, A., Rembold, F., Kerdiles, H., Gallego, J., 2017. Comparison of global land cover datasets for cropland monitoring. Remote Sens. (Basel) 9 (11). https://doi.org/10.3390/rs9111118.
Petit, C.C., Lambin, E.F., 2002. Long-term land-cover changes in the Belgian Ardennes (1775–1929): model-based reconstruction vs. Historical maps. Glob. Chang. Biol. 8 (7), 616–630. https://doi.org/10.1046/j.1365-2486.2002.00500.x.
Pongratz, J., Reick, C., Raddatz, T., Claussen, M., 2008. A reconstruction of global agricultural areas and land cover for the last millennium. Global Biogeochem. Cycles 22 (3). https://doi.org/10.1029/2007GB003153.
Ramankutty, N., Foley, J.A., 1999. Estimating historical changes in land cover:North American croplands from 1850 to 1992. Glob. Ecol. Biogeogr. 8 (5), 381–396. https://doi.org/10.1046/j.1365-2699.1999.00141.x.
Samasse, K., Hanan, N.P., Tappan, G., Diallo, Y., 2018. Assessing cropland area in West Africa for agricultural yield analysis. Remote Sens. (Basel) 10 (11). https://doi.org/10.3390/rs10111785.
Santos-Alamillos, F., Pozo-Vazquez, D., Ruiz-Arias, J., Tovar-Pescador, J., 2015. Influence of land-use misrepresentation on the accuracy of WRF wind estimates: evaluation of GLCC and CORINE land-use maps in southern Spain. Atmos. Res. 157 https://doi.org/10.1016/j.atmosres.2015.01.006.
Sarmento, P., Fonte, C.C., Dinis, J., Stehman, S.V., Caetano, M., 2015. Assessing the impacts of human uncertainty in the accuracy assessment of land-cover maps using linguistic scales and fuzzy intervals. Int. J. Remote Sens. 36 (10), 2524–2547. https://doi.org/10.1080/01431161.2015.1043034.
Sertel, E., Robock, A., Ormeci, C., 2010. Impacts of land cover data quality on regional climate simulations. Int. J. Climatol. 30 (13), 1942–1953. https://doi.org/10.1002/joc.2036.
Shahtahmassebi, A., Yang, N., Wang, K., Moore, N., Shen, Z., 2013. Review of shadow detection and de-shadowing methods in remote sensing. Chin. Geogr. Sci. 23 (4), 403–420.
Stehman, S.V., 1997. Selecting and interpreting measures of thematic classification accuracy. Remote Sens. Environ. 62 (1), 77–89. https://doi.org/10.1016/S0034-4257(97)00083-7.
Story, M., Congalton, R.G., 1986. Remote sensing brief accuracy assessment: a user's perspective. Photogramm. Eng. Remote Sensing 52 (3), 397–399.
Tchuenté, A.T.K., Roujean, J.-L., De Jong, S.M., 2011. Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale. Int. J. Appl. Earth Obs. Geoinf. 13 (2), 207–219.
Törmä, Markus, Markkanen, Tiina, Hatunen, Suvi, Härmä, Pekka, Mattila, Olli-Pekka, Arslan, Ali Nadir, 2015. Assessment of land-cover data for land-surface modelling in regional climate studies. Boreal Environment Research 20, 243–260.
Trusilova, K., Früh, B., Brienen, S., Walter, A., Masson, V., Pigeon, G., Becker, P., 2013. Implementation of an urban parameterization scheme into the regional climate model COSMO-CLM. J. Appl. Meteorol. Climatol. 52 (10), 2296–2311. https://doi.org/10.1175/JAMC-D-12-0209.1.
Verburg, P.H., Neumann, K., Nol, L., 2011. Challenges in using land use and land cover data for global change studies. Glob. Chang. Biol. 17 (2), 974–989. https://doi.org/10.1111/j.1365-2486.2010.02307.x.
Vilar, L., Garrido, J., Echavarría, P., Martínez-Vega, J., Martín, M.P., 2019. Comparative analysis of CORINE and climate change initiative land cover maps in Europe: implications for wildfire occurrence estimation at regional and local scales. Int. J. Appl. Earth Obs. Geoinf. 78 (November 2018), 102–117. https://doi.org/10.1016/j.jag.2019.01.019.
Yang, Y., Xiao, P., Feng, X., Li, H., 2017. Accuracy assessment of seven global land cover datasets over China. Isprs J. Photogramm. Remote. Sens. 125, 156–173. https://doi.org/10.1016/j.isprsjprs.2017.01.016.