Response to comments

Paper #: essd-2023-320
Title: GLC_FCS30D: The first global 30-m land-cover dynamic monitoring product with a fine classification system from 1985 to 2022 using dense time-series Landsat imagery and continuous change-detection method
Journal: Earth System Science Data

Reviewer #2

The authors present a very detailed manuscript on the generation of a global 30-m land cover product. It is original in the spatio-temporal density of Landsat satellite imagery used to generate annual maps over nearly 30 years. I recommend this paper for publication. There is one major issue that needs to be addressed and a number of minor or editorial issues to address:

Great thanks for the positive comments. The manuscript has been further improved based on your and another reviewers' comments and suggestions.

Major change:

Both the Continuous Change Detection and Classification (CCDC) Algorithm and the Random Forest Algorithm for subsequent land cover classification use several hyperparameters. Both models will be sensitive to the hyperparameters selected. As a minimum, the hyperparameters selected for both models need to be clearly defined and justified (this is already partially done for the CCDC algorithm). However, to fully justify the use of hyperparameters, sensitivity analysis should be provided of the values used, and validation that the optimum or a favorable set of hyperparameters values have been selected.

Great thanks for the comment. Based on your suggestion and another reviewer's comment, the analysis of how to determine the parameters of CCDC algorithm have been added in the Section 3.2 as:

Next, the CCD was also a multi-parameter change detection model and demonstrated to be sensitive to the parameter settings (Xiao et al., 2023; Zhu and Woodcock, 2014b). The CCDC algorithm on the Google Earth Engine platform (ee.Algorithms.TemporalSegmentation.Ccdc) contained three key adjustable parameters: minObservations, chiSquareProbability and minNumOfYearScaler. Zhu et al. (2019) analyzed the relationships between the omission error and commission error of land-cover changes with the variability of three parameters in the United States, and found their values affected the change detection accuracy. In this study, we also investigated the sensitivity between parameter settings with the change detection accuracies in Figure S1 (seen the Supplement material) using the time-series points from LCMAP_Val and LUCAS datasets after partly sampling. Notably, the sensitivity analysis was implemented in two large-areas for ensure the feasibility of optimal parameters, that is, which will be suitable for other areas in land-cover change detection. The results also showed the CCD is a parameter-sensitive algorithm and the optimal parameter values were 5, 0.95 and 2-year for minObservations, chiSquareProbability and minNumOfYearScaler.

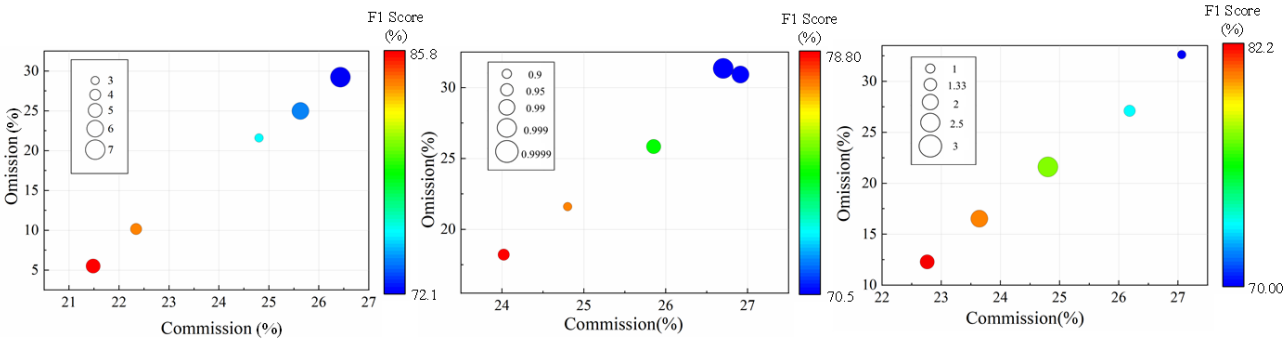


Figure S1. The sensitive analysis between the omission error and commission error with the minObservations, chiSquareProbability and minNumOfYearScaler using the time-series points from LCMAP_Val and LUCAS datasets after partly sampling.

Xiao, Y., Wang, Q., Tong, X., and Atkinson, P. M.: Thirty-meter map of young forest age in China, Earth Syst. Sci. Data, 15, 3365-3386, https://doi.org/10.5194/essd-15-3365-2023, 2023.

Zhu, Z., Zhang, J., Yang, Z., Aljaddani, A. H., Cohen, W. B., Qiu, S., and Zhou, C.: Continuous monitoring of land disturbance based on Landsat time series, Remote Sensing of Environment, https://doi.org/10.1016/j.rse.2019.03.009, 2019.

As for the parameters of random forest classifier, it only contains two adjustable parameters (the number of decision tree (Ntree) and predicted variables (Mtry)), and many previous studies have quantitatively or theoretically analyzed the influence of the parameters on the classification accuracy, and found that the classification accuracy was less sensitive to the selection of Ntree and Mtry. Thus, the default recommended setting of 500 and the square of the total number of input features were used. Correspondingly, the descriptions of how to determine these two parameters have been added in the manuscript as:

Thus, the RF algorithm was used to combine the training samples and multisourced features for updating the changed pixels. The RF algorithm allows for adjusting two key parameters **(the number of decision tree (Ntree) and predicted variables (Mtry)), and previous studies have quantitatively investigated the relationships between classification accuracy with the settings of these two parameters. Both theoretical and experimental results indicated that the selection of Mtry and Ntree had little influence on the classification accuracy (Belgiu and Drăguţ, 2016; Du et al., 2015). Thus, the default recommended values of 500 for Ntree and the square of the total number of input features for Mtry were used based on previous studies (Belgiu and Drăguţ, 2016; Zhang et al., 2019).**

Minor or editorial changes:

1. Introduction/ methods – various mentions of model ‘accuracy’ is used. This includes, but is not limited to lines 31-32, 69 and 155. Please be specific on the accuracy metric(s) used.

Great thanks for the comment. The ‘overall accuracy’ in the whole manuscript (line 31-32, 69 and 155) is an accuracy metric in the confusion matrix. In this manuscript, three accuracy metrics including: overall accuracy (O.A.), producer’s accuracy (P.A.) and user’s accuracy (U.A.) have been used and the corresponding formulas are also added in the Section 3.4 (accuracy assessment) as:

The validation process for the GLC_FCS30D dataset follows the recommended guidelines proposed by Pontus Olofsson (2014). These guidelines encompass two key components: area estimation (nonsite-specific accuracy) and accuracy assessment (site-specific accuracy). The site-specific accuracy assessment mainly focuses on estimating the confusion matrix and calculating some accuracy metrics including overall accuracy (O.A.),

producer's accuracy (P.A.), user's accuracy (U.A.) and the corresponding standard errors using a poststratified estimator (Pontus Olofsson, 2014).

$$P.A._k = \frac{p_{kk}}{\sum p_{k.}}, U.A._k = \frac{p_{kk}}{\sum p_{.k}}, O.A. = \sum_{k=1}^{m} p_{kk} \quad (4)$$

Where p_{kk} was the proportion of the area mapped as class k that had reference class k, $\sum p_{k.}$ and $\sum p_{.k}$ were the proportion of the area mapped as class k and the proportion of the reference area as class k, and the m denoted the number of land-cover types. Afterwards, because there is currently no global long-time series validation dataset, we used 84526 global validation points to assess the accuracy metrics of the GLC_FCS30D dataset in 2020 and used two third-party datasets to analyze the time-series accuracy variations. The GLC_FCS30D adopts a fine classification system containing 35 subcategories, for which we applied an analysis protocol into two validation systems (the level-0 classification system containing 10 major land-cover types and the LCCS level-1 validation system containing 17 land-cover types) to comprehensively understand the GLC_FCS30D dataset quality. The relationship between Level-0 and LCCS level-1 validation systems is explained in Table 1. Lastly, to quality the performance of land-cover changed pixels, we followed the proposal of Stehman et al. (2021) in assessing the LCMAP annual land-cover products 1985-2017, that is, the validation pixels were grouped into "changed" and "unchanged" categories and the corresponding confusion matrix were calculated. Meanwhile, to minimize the imbalance in the sample size of "change" and "no-change" samples, the metrics of F1 score was supplemented as:

$$F1 = \frac{P.A. \times U.A.}{P.A. + U.A.} \times 2 \times 100\% \quad (5)$$

2. Line 24- The use of the phrase 'In specific' is awkward and I suggest changing e.g. 'Specifically' or 'In particular…'

Great thanks for the suggestion. The 'In specific' has been changed as 'Specifically' and 'In particular' through the whole manuscript.

3. Line 201- 'The first time series validation set was assessed the performance…' remove 'was'.

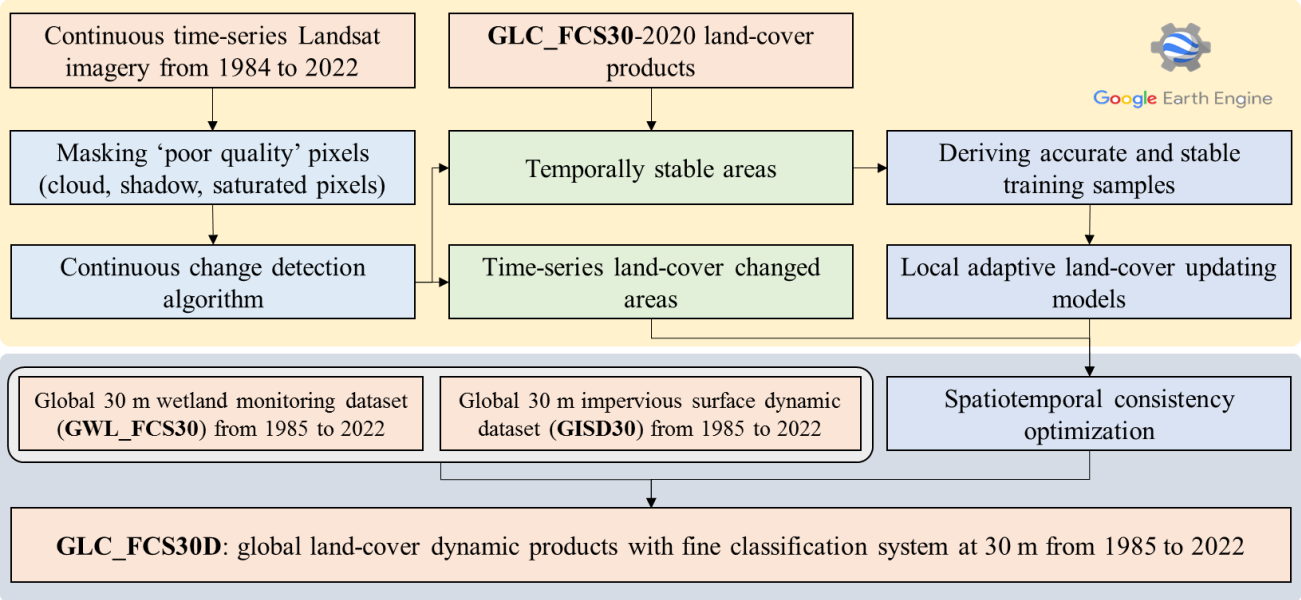
Great thanks for pointing out the mistake. It has been corrected.

4. Line 205- Change 'It developed by combining…' to 'It was developed by combining…'

Great thanks for pointing out the mistake. It has been corrected.

5. Figure 2: This Figure is very useful for help the reader understand the main processes carried out in this project. Please add the shortened names of each dataset to the flow chart to make it even easier for the reader to follow the text.

Great thanks for the comment. Based on your suggestion, the shortened names of each dataset have been added and also bolded into the flowchart as:



6. Figure 2: You refer to masking 'poor quality' pixels. Please be more specific on this.

Great thanks for the comment. The 'poor quality' refers to these cloud, shadow and saturated pixels, as well as the Scan Line Corrector Off pixels in Landsat 7, which was added in the revision version.

Does this just include applying a cloud mask, or does it also consider issues with the Scan Line Corrector on Landsat 7, for example. What cloud mask was used.

Yes, the Scan Line Corrector Off pixels are also masked. Specifically, the 'poor-quality' pixels were masked using the CFmask algorithm, which was demonstrated to achieve high accuracy and great robustness for masking these 'poor-quality' pixels.

How did you account for pixels that may be under light cloud/ haze which may not be picked up by a cloud mask (e.g. does the CCDC intend to overcome this?)

In terms of these light cloud/haze pixels, actually, the Tmask algorithm, which was integrated into the CCDC algorithm in the GEE platform, was used to further minimize their effects.

The explanations have been added in Section 3 as:

Before detecting the land-cover changed pixels, all 'poor quality' pixels (cloud, shadow and saturated pixels, as well as the Scan Line Corrector Off pixels in Landsat 7) in the continuous time-series Landsat imagery were firstly masked using the CFmask algorithm, which was demonstrated to achieve the overall accuracy of 96.4% and was adopted by the USGS as official cloud- and shadow detection algorithm (Zhu et al., 2015; Zhu and Woodcock, 2012). Then, in terms of these residual cloud pixels (light cloud and haze contaminated pixels), the Tmask (multiTemporal mask) algorithm, which used the temporal information from these clear-sky pixels to improve the cloud-detection capability (Zhu and Woodcock, 2014a), was used to mask the residual cloud pixels. It should be noted that the Tmask has been integrated into the CCD algorithm on the GEE platform as ee.Algorithms.TemporalSegmentation.Ccdc(), that is, the effect of 'poor-quality' pixels were minimized.

7. Table 1: Please add the abbreviations for each land cover type to this table (at later points you refer to Table 1 as containing these).

Great thanks for the suggestion. The abbreviations of each land-cover type have been into the Table 1 as:

Basic classification system		**Level-1 validation system**		**Fine classification system**	**Id**
Cropland	CRP	Rainfed cropland	RCP	Rainfed cropland	10

				Herbaceous cover cropland	11
				Tree or shrub cover cropland	12
		Irrigated cropland	ICP	Irrigated cropland	20
Forest	FST	Evergreen broadleaved forest	EBF	Closed evergreen broadleaved forest	51
				Open evergreen broadleaved forest	52
		Deciduous broadleaved forest	BDF	Closed deciduous broadleaved forest	61
				Open deciduous broadleaved forest	62
		Evergreen needleleaved forest	ENF	Closed evergreen needleleaved forest	71
				Open evergreen needleleaved forest	72
		Deciduous needleleaved forest	DNF	Closed deciduous needleleaved forest	81
				Open deciduous needleleaved forest	82
		Mixed-leaf forest	MFT	Closed mixed-leaf forest	91
				Open mixed-leaf forest	92
Shrubland	SHR	Shrubland	SHR	Shrubland	120
				Evergreen shrubland	121
				Deciduous shrubland	122
Grassland	GRS	Grassland	GRS	Grassland	130
Tundra	TUD	Lichens and mosses	LMS	Lichens and mosses	140
Wetland	WET	Inland wetland	IWL	Swamp	181
				Marsh	182
				Flooded flat	183
				Saline	184
		Coastal wetland	CWL	Mangrove	185
				Salt marsh	186
				Tidal flat	187
Impervious surface	IMP	Impervious surface	IMP	Impervious surface	190
Bare areas	BAL	Sparse vegetation	SVG	Sparse vegetation	150
				Sparse shrubland	152
				Sparse herbaceous cover	153
		Bare areas	BAL	Bare areas	200
				Consolidated bare areas	201
				Unconsolidated bare areas	202
Water body	WTR	Water body	WTR	Water body	210
Permanent snow and ice	PSI	Permanent snow and ice	PSI	Permanent snow and ice	220

8. Line 357- please provide more information on the indicator function, or at least a reference.

Great thanks for the comment. The description of the indicator function has been strength and the reference is also added as:

"and the I() denotes the indicator function for the equation of the status between two pixels. Namely, if $L_{x\prime,y\prime,t\prime}$ was equal to the $L_{x,y,t}$, then the value of indicator function was 1, otherwise it was equal to 0 (Kenny, 2003)"

Kenny, Q. Y.: Indicator function and its application in two-level factorial designs, The Annals of Statistics, 31, 984-994, https://doi.org/10.1214/aos/1056562470, 2003.

9. Figure 7- the very thick lines corresponding to pixels with a stable land cover overwhelm this image and make it difficult for the reader to decipher the most dominant types of land cover change. Please either rescale the image or consider removing the lines corresponding to no land cover change to make it easier for the reader to assess the dominant types of land cover change.

Great thanks for the comment. To make the Sankey diagram clear, the layout has been changed. The updated Figure 7 are following:

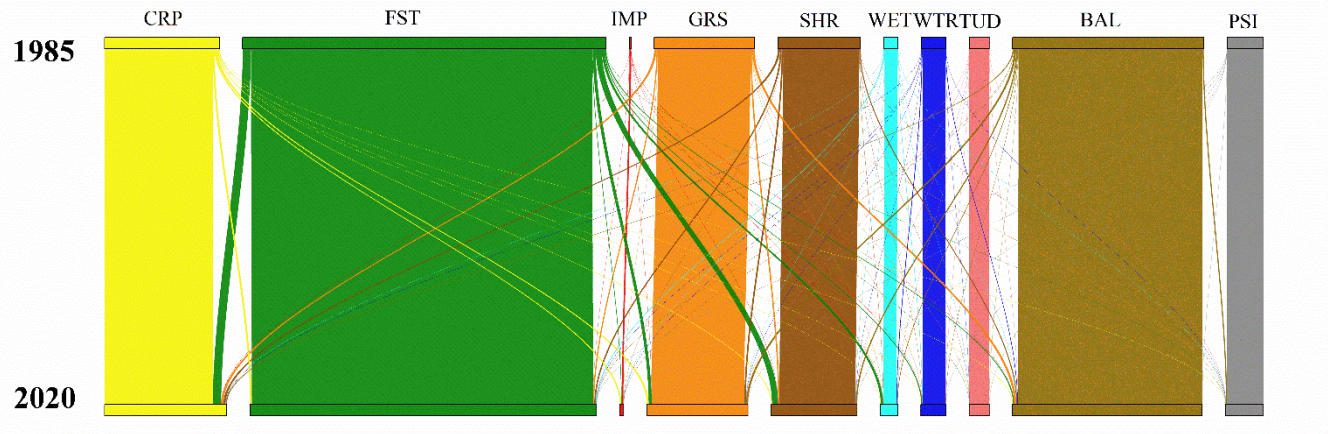


Figure 7. Sankey diagrams of the global land-cover changes during 1985-2022 in the GLC_FCS30D dataset.

10. Table 2 and 3- please use a method to highlight the relative performance of your algorithms. For example use a colour ramp or make particular values bold.

Great thanks for the comment. Since the ESSD journal does not allow colormaps to be added to the Table, some particular accuracy values (mentioned in the manuscript) have been bolded in Table 2 and 3 based on your suggestion as:

Table 2. Error matrix of the GLC_FCS30D dataset in 2020 based on the level-0 basic classification system. The reported Producer's Accuracy (P.A.) and User's Accuracy (U.A.) come with their corresponding standard errors (SE) shown in parentheses.

	Map											
	O.A. = 80.88% (±0.27%)											
Reference	CRP	FST	GRS	SHR	WET	WTR	TUD	IMP	BAL	PSI	Total	P.A.(SE)
CRP	15.442	0.792	0.679	0.388	0.086	0.027	0	0.174	0.117	0	17.704	87.22(0.54)
FST	0.513	28.712	0.315	0.811	0.371	0.021	0.008	0.063	0.113	0.002	30.93	**92.83(0.31)**
GRS	1.035	1.166	5.906	1.181	0.231	0.011	0.084	0.051	1.181	0.01	10.855	**54.41(1.02)**
SHR	0.555	1.798	0.863	5.392	0.161	0.013	0.019	0.05	0.502	0.002	9.356	**57.63(1.09)**
WET	0.068	0.465	0.156	0.157	4.047	0.347	0.031	0.021	0.222	0.001	5.516	73.37(1.27)
WTR	0.04	0.086	0.019	0.017	0.302	3.305	0.008	0.012	0.039	0.002	3.831	86.28(1.12)
TUD	0.01	0.123	0.168	0.167	0.018	0.03	2.444	0.002	0.473	0.02	3.454	70.76(1.65)
IMP	0.084	0.058	0.024	0.04	0.001	0.006	0.002	5.043	0.024	0	5.283	95.45(0.61)
BAL	0.13	0.049	0.783	0.585	0.043	0.045	0.577	0.048	9.239	0.131	11.628	79.45(0.8)
PSI	0	0.004	0.03	0.005	0	0.023	0.001	0	0.03	1.351	1.443	93.63(1.38)
Total	17.877	33.251	8.943	8.743	5.259	3.828	3.176	5.464	11.94	1.52		
U.A.(SE)	86.38 (0.55)	86.35 (0.4)	66.05 (1.07)	61.68 (1.11)	76.96 (1.2)	86.33 (1.35)	76.97 (1.6)	92.29 (0.77)	77.38 (0.82)	88.89 (1.72)		

Table 3. Error matrix of the GLC_FCS30D dataset in 2020 based on the LCCS level-1 validation system. The reported Producer's Accuracy (P.A.) and User's Accuracy (U.A.) come with their corresponding standard errors (SE) shown in parentheses.

Reference	RCP	ICP	EBF	DBF	ENF	DNF	MFT	SHR	GRS	LMS	SVG	IWL	CWL	IMP	BAL	WTR	PSI	Total	P.A. (SE)
RCP	12.225	**1.023**	0.239	0.358	0.102	0.016	0.009	0.382	0.66	0	0.078	0.056	0.005	0.124	0.028	0.001	0	15.332	79.7(0.7)
ICP	0.397	1.932	0.026	0.016	0.005	0	0	0.01	0.025	0	0.012	0.029	0.005	0.052	0	0.018	0	2.527	76.45(1.81)
EBF	0.2	0.048	9.091	1.098	0.262	0.103	**0.151**	0.371	0.084	0	0.012	0.136	0.028	0.029	0.001	0.004	0	11.514	78.96(0.82)
DBF	0.187	0.016	0.632	6.838	0.537	0.294	**0.396**	0.235	0.144	0.002	0.019	0.077	0.002	0.025	0.005	0.004	0.002	9.054	75.53(0.97)
ENF	0.046	0.004	0.174	0.316	5.681	0.328	**0.439**	0.128	0.034	0.006	0.043	0.094	0	0.008	0.01	0.01	0	6.895	**82.39(0.98)**
DNF	0.008	0	0.002	0.13	0.245	1.854	**0.073**	0.071	0.053	0	0.011	0.025	0	0.001	0.007	0.002	0	2.414	76.79(1.85)
MFT	0.004	0	**0.019**	**0.176**	**0.234**	**0.013**	**0.828**	0.014	0.004	0	0	0.010	0.05	0	0.001	0	0	1.308	**58.29(1.53)**
SHR	0.518	0.042	0.299	0.9	0.328	0.131	0.034	5.44	0.871	0.019	0.441	0.157	0.005	0.05	0.065	0.013	0.002	9.438	57.63(1.09)
GRS	0.947	0.097	0.167	0.582	0.209	0.154	0.024	1.191	5.958	0.085	0.974	0.229	0.006	0.052	0.217	0.008	0.01	10.95	54.41(1.02)
LMS	0.006	0.004	0.001	0.022	0.044	0.053	0.001	0.168	0.169	2.465	0.379	0.02	0.001	0.002	0.098	0.026	0.02	3.484	70.76(1.65)
SVG	0.064	0.01	0.008	0.006	0.007	0.01	0.001	0.397	0.462	0.025	2.71	0.012	0	0.013	0.643	0.002	0.024	4.399	61.6(1.57)
IWL	0.01	0.002	0.044	0.029	0.103	0.022	0.002	0.048	0.017	0.008	0.042	2.673	0.024	0.001	0.012	0.224	0	3.263	**81.91(1.45)**
CWL	0.004	0.002	0.008	0.002	0.004	0.002	0.004	0.008	0.006	0	0.008	0.188	1.476	0.007	0.007	0.059	0	1.783	**82.77(1.92)**
IMP	0.074	0.011	0.008	0.008	0.037	0.002	0	0.041	0.024	0.002	0.014	0.004	0	5.087	0.01	0.004	0	5.329	95.45(0.61)
BAL	0.048	0.01	0.002	0.004	0.002	0.001	0	0.193	0.328	0.557	0.582	0.043	0.002	0.035	5.384	0.029	0.108	7.33	73.45(1.11)
WTR	0.014	0.024	0.014	0.014	0.019	0.008	0.006	0.011	0.016	0.007	0.011	0.168	0.114	0.011	0.019	3.054	0.002	3.509	87.04(1.22)
PSI	0	0	0	0.001	0.002	0	0	0.005	0.03	0.001	0.011	0	0	0	0.019	0.023	1.363	1.455	93.65(1.37)
Total	14.757	3.224	10.753	10.56	7.833	3.724	1.97	8.711	8.883	3.179	5.353	3.927	1.668	5.497	6.526	3.482	1.532		
U.A. (SE)	82.85 (0.67)	**59.92 (1.85)**	84.55 (1.75)	64.76 (1)	72.52 (1.08)	49.77 (1.76)	**39.34 (1.38)**	62.44 (1.11)	67.07 (1.07)	77.55 (1.59)	**50.63 (1.47)**	**68.07 (1.6)**	**88.49 (1.68)**	92.54 (0.76)	82.5 (1.01)	87.73 (1.19)	88.96 (1.72)		
O.A.	**73.04% (±0.30%)**																		

11. Results and discussion are very thorough although there is no mention to coastal regions which will be areas of major change detectable at 30 m resolution.

Great thanks for the comment. Yes, we completely agree that the coastal regions experienced obvious land-cover changes. To intuitively understand these coastal changes, an example in Yellow River Estuary Delta was also added in the Figure 8 and the corresponding descriptions as:

Lastly, the Yellow River Delta, as one of the typical coastal regions, was selected to understand the GLC_FCS30D for capturing these coastal land-cover changes. Obviously, the land-cover changes in the GLC_FCS30D can be concluded into three aspects: 1) a large amount of flooded flats and flat flats were reclaimed as the aquaculture ponds, especially after 2000; 2) the mouth of the Yellow River turned from south to north (black rectangle), that is, there were large land-cover changes between tidal/flooded flats, water bodies and salt marshes; 3) a lot of impervious surfaces encroached the coastal water-bodies and flats. In short, if we combine real time-series remote-sensing observation data, the GLC_FCS30D effectively captures the spatiotemporal changes of the land surface.

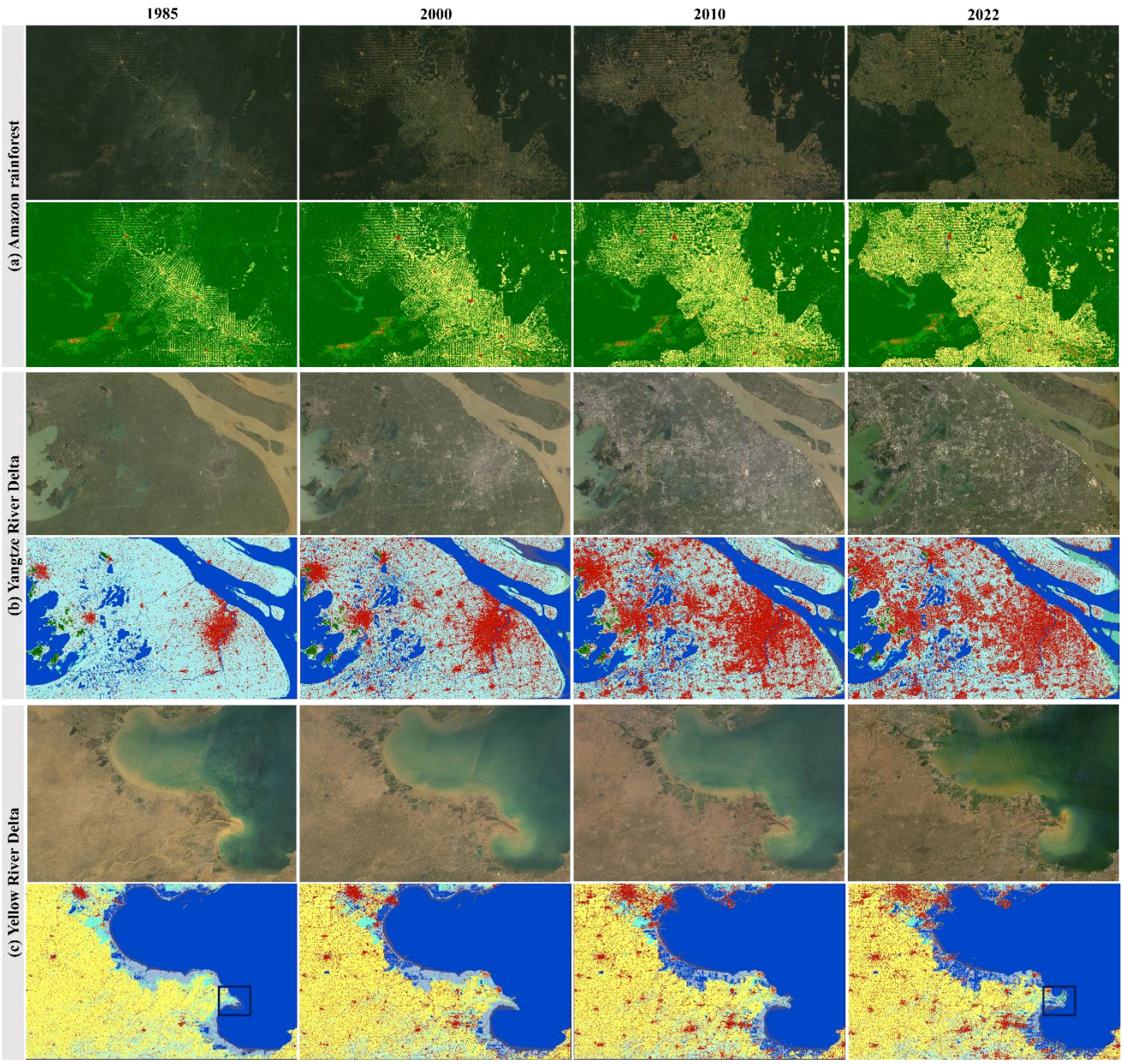


Figure 8. Three typical enlargements of land-cover changes in the GLC_FCS30D from 1985 to 2022 in (a) the Amazon rainforest, (b) the Yangtze River Delta in China, **and (c) the Yellow River Delta in China**. The color-coded legend is like the global map in Figure 4. In each case, the natural-color imagery from 1985 to 2022 is a composite taken from Landsat imagery.

12. Overall, I enjoyed reading this paper. The analysis was very thorough and easy to follow.

Great thanks for the positive comments. The analysis has been further improved based on your and another reviewer's comments.