Response letter of manuscript ESSD-2025-807

25-year, quarterly land change maps of China's Loess Plateau reveal long-term and substantial water-induced soil erosion mitigation

Authors: **Mofan Cheng, Zhuohong Li, Linxin Li, Wei He[†], Liangpei Zhang, and Hongyan Zhang[†]**

Dear Editors and Reviewers:

The authors would like to thank the editors and reviewers for their constructive comments and suggestions that have helped improve the quality of this manuscript. The manuscript has undergone a thorough revision according to the editors and reviewers' comments. Please see our responses below. For the reviewers' convenience, we have highlighted significant changes in the revised manuscript in blue.

Reviewer 1

This manuscript presents a substantial advancement in both methodology and data availability for investigating soil erosion dynamics in fragile ecosystems, with a particular focus on the Loess Plateau. By integrating long-term time-series Landsat and Sentinel imagery with relevant environmental datasets, the authors develop a high-resolution land-cover and soil erosion dataset spanning 25 years, while further shortening the update frequency to a quarterly scale. Such temporal coverage and resolution represent a notable improvement over existing regional product. The reported mapping performance is robust for a large-scale application, with an overall accuracy of 81.44% for land-cover classification and a mean absolute error of 4.5% for soil erosion estimates. Based on an examination of the released dataset, this work is expected to provide a valuable data foundation for future studies on land-surface processes, ecological restoration, and environmental change across the Loess Plateau. Leveraging this self-produced dataset, the authors further analyze the spatial and temporal evolution of land cover and soil erosion, revealing a pronounced overall reduction in erosion intensity over the study period. The long-term and high-frequency observations enable novel insights into the seasonal heterogeneity of erosion driven by precipitation, the role of vegetation dynamics in erosion mitigation, and the influence of topographic factors. In addition, the attempt to assess potential erosion under an optimized vegetation configuration provides practical implications for soil conservation and land management strategies in the region.

The manuscript is clearly written and accompanied by high-quality data visualizations. Overall, this study makes a valuable contribution in terms of both data products and analytical perspectives, and I believe it is suitable for publication after revisions.

Reply: Thank you so much for your positive comments. We believed the manuscript is much improved after addressing your comments. Below, please find our point-to-point response.

Reviewer Comment 1.1 — The land-cover dataset developed in this study is directly applied to soil erosion estimation, while the land-cover products referenced in the Introduction and used for accuracy comparison are mostly general-purpose datasets with coarse thematic categories and broad spatial coverage. Given the strong dependence of soil erosion modeling on land-cover characteristics, it would be helpful to clarify whether any land-cover datasets specifically designed for soil erosion assessment have been developed or applied in the Loess Plateau region. If such erosion-oriented land-cover datasets exist, how do their spatial resolution and mapping accuracy compare with the proposed LP-QLC10 dataset?

Table R1: Quantitative results of the land-cover products on the time-series valuation points.

Datasets	Metrics	2001	2006	2011	2016	2021	Average
CLCDs		0.737	**0.791**	0.759	0.790	0.781	0.7716
GLC_FCS30D		0.753	0.743	0.757	0.749	0.742	0.7488
YRCC_LPLC	OA	0.707	0.752	0.757	0.791	0.784	0.7582
Esri_GLC10		–	–	–	–	0.744	–
LP-QLC10 (ours)		**0.785**	0.787	**0.800**	**0.855**	**0.845**	**0.8144**
CLCDs		0.610	0.668	0.651	0.685	0.669	0.6566
GLC_FCS30D		0.648	0.621	0.663	0.645	0.633	0.6420
YRCC_LPLC	Kappa	0.583	0.625	0.654	0.694	0.682	0.6476
Esri_GLC10		–	–	–	–	0.611	–
LP-QLC10 (ours)		**0.686**	**0.673**	**0.715**	**0.783**	**0.772**	**0.7258**
CLCDs		0.523	0.609	0.566	0.589	0.594	0.5762
GLC_FCS30D		0.556	0.548	0.555	0.556	0.570	0.5570
YRCC_LPLC	FWIoU	0.512	0.570	0.564	0.600	0.604	0.5700
Esri_GLC10		–	–	–	–	**0.739**	–
LP-QLC10 (ours)		**0.663**	**0.643**	**0.582**	**0.647**	0.692	**0.6454**

Table R2: Quantitative results of the land-cover products on the publicly available validation points.

Datasets	SinoLC-1 (7382 points in 2020)			Geo-wiki (1183 points in 2012)		
	OA	Kappa	FWIoU	OA	Kappa	FWIoU
CLCDs	0.696	0.559	0.520	0.480	0.311	0.317
GLC_FCS30D	0.662	0.529	0.501	0.475	0.307	0.326
YRCC_LPLC	0.706	0.581	0.535	0.496	0.330	0.334
Esri_LC10	0.699	0.555	0.623	–	–	–
LP-QLC10 (ours)	**0.798**	**0.709**	**0.660**	**0.505**	**0.346**	**0.357**

Reply: Thank you for this insightful comment. To further demonstrate the advantages of our proposed dataset, we included an additional published dataset specifically developed for land-cover change analysis on the Loess Plateau for both quantitative and qualitative comparisons. In the Loess Plateau region, most existing land-cover datasets are general-purpose products developed at national or global scales[R1,R2] and are not specifically designed for soil erosion assessment. Although several studies[R3,R4] have produced land-use or erosion-related land-cover information at local or catchment scales, these datasets typically have limited spatial coverage or insufficient temporal continuity, which constrains their consistent application in large-scale, long-term analyses such as the entire Loess Plateau.

In addition, YRCC_LPLC[R5] was developed specifically for the Loess Plateau to characterize land-cover changes induced by ecological restoration and provides annual land-cover maps at 30 m resolution from 1990 to 2022. Following the suggestion, we included YRCC_LPLC in our comparative evaluation and updated the quantitative and qualitative results. As shown in Table R1, the multi-year validation indicates that YRCC_LPLC performs comparably to national-scale 30 m products such as CLCDs, achieving an average OA of 0.7582. In contrast, our LP-QLC10 achieves superior quantitative performance, with an average OA of 0.8144. Moreover, the evaluation based on third-party sample points further supports this conclusion,

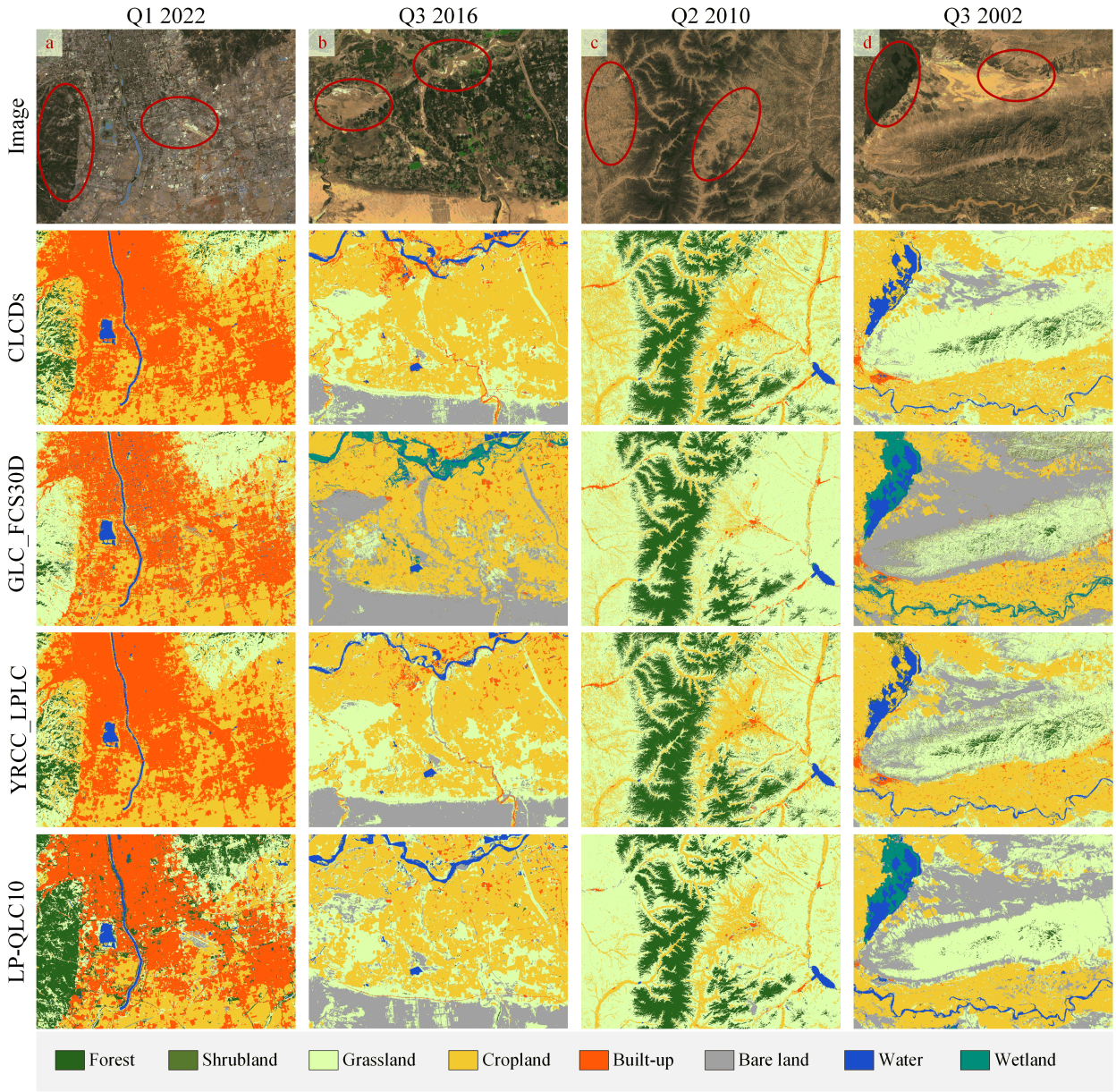

Figure R1: Visual comparison of land-cover classification results over four representative regions: (a) urban areas, (b) agricultural zones, (c) mountainous regions, and (d) wetlands. The satellite images in the figure are from © Google Earth.

as shown in Table R2, where LP-QLC10 outperforms the other datasets across all three metrics. In addition to improving overall classification accuracy, LP-QLC10 is designed with soil-erosion applications in mind, emphasizing the reliable representation of erosion-relevant surface characteristics and their seasonal dynamics. In particular, the quarterly temporal resolution enables the dataset to capture intra-annual variability in vegetation cover types such as cropland and grassland, which directly influences erosion-related processes and associated modeling factors. These design features make LP-QLC10 particularly suitable for long-term, large-scale soil erosion studies in the Loess Plateau, where consistent classification, temporal stability, and

the ability to resolve seasonal land-surface dynamics are essential.

Meanwhile, we have also added the visualization results of YRCC_LPLC to the qualitative comparison, as shown in Fig. R1. YRCC_LPLC exhibits patterns similar to those of CLCDs. Specifically, it shows noticeable confusion among forest, grassland, and cropland in transitional landscapes that characterized by fragmented land parcels. In addition, YRCC_LPLC does not include wetlands in its classification system, which limits its ability to represent wetland-related land-surface characteristics.

In the revised manuscript, comparison results have been modified and presented in Tables 3, 4, and Figure 7 in Section 4.1 and 4.2.

Reference of this reply:

[R1] Zhang, X., Zhao, T., Xu, H., Liu, W., Wang, J., Chen, X. and Liu, L., 2024. GLC_FCS30D: the first global 30 m land-cover dynamics monitoring product with a fine classification system for the period from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method. Earth System Science Data, 16(3), pp.1353-1381.

[R2] Yang, J. and Huang, X., 2021. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. Earth System Science Data, 13(8), pp.3907-3925.

[R3] Tian, P., Tian, X., Geng, R., Zhao, G., Yang, L., Mu, X., Gao, P., Sun, W. and Liu, Y., 2023. Response of soil erosion to vegetation restoration and terracing on the Loess Plateau. Catena, 227, p.107103.

[R4] Klik, A. and Rosner, J., 2020. Long-term experience with conservation tillage practices in Austria: Impacts on soil erosion processes. Soil and Tillage Research, 203, p.104669.

[R5] Wang, Z., Shi, X., Dou, S., Cheng, M. and Miao, L., 2025. The 30 m land cover dataset for capturing land cover changes induced by ecological restoration from 1990 to 2022 on the Chinese Loess Plateau. Scientific Data, 12(1), p.252.

Reviewer Comment 1.2 — The land-cover mapping in this study is based on eight categories. Please clarify the rationale for selecting these specific categories. In particular, it would be helpful to explain how these categories are relevant to soil erosion research and whether they are designed to effectively represent erosion-related surface characteristics.

Reply: Thank you for this question. The eight land-cover categories were selected based on their functional relevance to soil erosion processes rather than purely thematic or land-use considerations. In soil-erosion research, the key concern is how different surface types regulate rainfall interception, surface runoff generation, soil exposure, and human disturbance. Accordingly, these categories represent major erosion-related surface characteristics across the Loess Plateau, including vegetation-covered surfaces with different protective capacities (e.g., forest, shrubland, and grassland), cultivated land with active human disturbance (e.g., cropland), exposed or weakly protected surfaces (e.g., bare land), and surfaces where water erosion is negligible or not applicable (e.g., water and wetlands). Built-up areas are treated separately due to their distinct hydrological and erosion responses associated with surface hardening. This classification scheme strikes a balance between physical interpretability and temporal consistency over long time series.

Moreover, the selected categories are directly compatible with soil-erosion modeling frameworks such as RUSLE, in which land-cover classes are mapped to erosion-related parameters, particularly the cover management factor (C) and the support practice factor (P). Specifically, according to ecological specification of China[R1], C factor is determined by land-cover category and Fractional Vegetation Cover (FVC), as summarized in Table R3 and Eq. 1, and all land-cover types required by this standard are included in the LP-QLC10 classification system. Meanwhile, the P factor is assigned based on land cover category and slope class, as detailed in Table R4. Therefore, the eight-category classification is not only physically

Table R3: *C* factor values under different levels of Fractional Vegetation Cover (FVC).

Category	FVC					
	<0.1	0.1–0.3	0.3–0.5	0.5–0.7	0.7–0.9	>0.9
Forest	0.100	0.080	0.060	0.020	0.004	0.001
Shrubland	0.400	0.220	0.140	0.085	0.040	0.011
Grassland	0.450	0.240	0.150	0.090	0.043	0.011
Bare land	0.700					
Built-up	0.010					

Table R4: *P* factor values assigned to different land-cover types and slopes.

Category	Slope (°)	*P*
Forest	-	1
Shrubland	-	0.29
Grassland	-	0.41
	0–5	0.49
	5–7	0.59
	7–9	0.65
Cropland	9–12	0.70
	12–20	0.81
	20–24	0.95
	>24	1
Built-up	-	0
Bare land	-	1
Water	-	0
Wetland	-	0

meaningful for erosion mechanisms, but also ensures straightforward for long-term soil erosion modeling on the Loess Plateau.

$$C_{\text{crop}} = 0.221 - 0.595 \log c. \tag{1}$$

Reference of this reply:

[R1] Ministry of Ecology and Environment of the People's Republic of China (MEE). Technical Specification for Investigation and Assessment of National Ecological Status: Ecosystem Services Assessment. Tech. Rep. HJ 1173–2021, Beijing, China, 2021 (in Chinese).

Reviewer Comment 1.3 — The imagery used for land-cover mapping is derived from multiple satellite sources (e.g., Landsat and Sentinel) and forms dense temporal sequences. To improve the transparency and usability of the dataset, it is recommended to provide a concise visualization of the image acquisition timeline, such as stacked bar charts or timeline histograms showing the temporal distribution of input images. Displaying acquisition dates in this way would help users better understand data availability, temporal coverage, and seasonal sampling characteristics, thereby facilitating more effective use of the dataset.

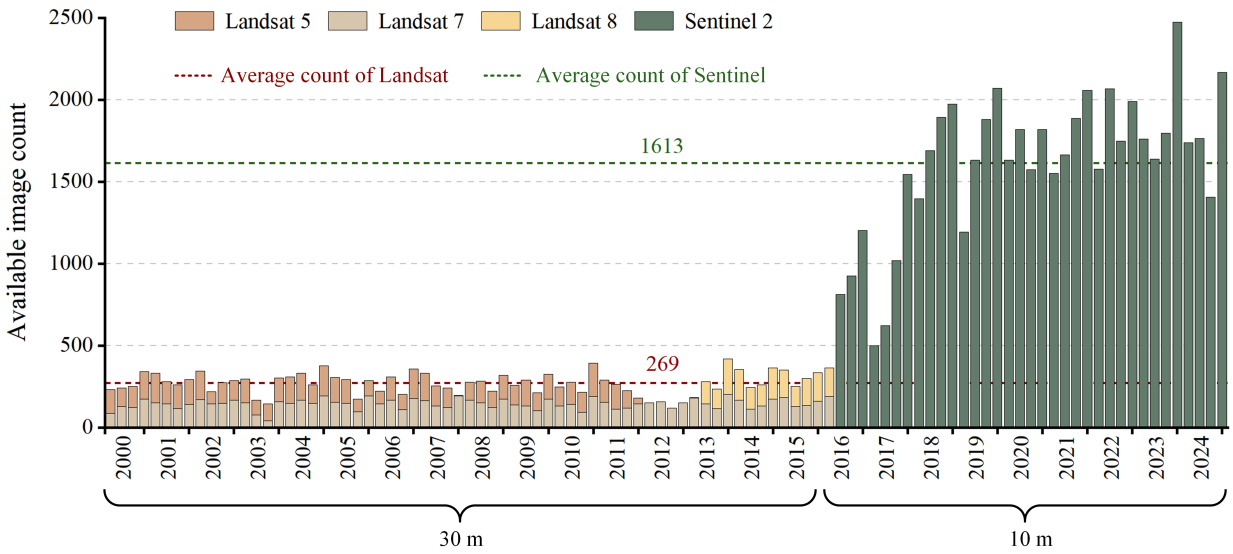

Figure R2: The obtained image count from multiple platforms within each temporal period. The statistic denotes the number of quality-screened satellite image tiles in Google Earth Engine that intersect the Loess Plateau. A Landsat scene covers approximately 3.3×10^4 km^2, whereas a Sentinel tile covers approximately 1.0×10^4 km^2.

Reply: Thank you for your comments. We compiled statistics on the acquired data counts and dates from multiple platforms and illustrated them in figure to clarify the data distribution. Specifically, during the data collection phase, we divided the time span into three consecutive months, corresponding to the four quarters of each year. For each period, we retrieved all available imagery from the Google Earth Engine (GEE) platform and applied a median compositing strategy to generate a single representative image for that quarter. Because median composites rely on the number of valid observations, the quantity of available images directly affects the robustness and quality of the synthesized quarterly image. Therefore, for each period, we reported the number of available images from each satellite as a reference for data availability.

The acquisition period and the number of available images are illustrated in Figure R2. From 2000 to 2016, the quarterly composites were generated using the Landsat series, combining observations from different sensors (e.g., Landsat 5, 7, and 8, depending on the year). After 2016, we used Sentinel-2 imagery. The statistics indicate that, over the Loess Plateau, the Landsat series provides about 300 usable images per quarter, whereas Sentinel-2 provides approximately 1600 tiles per quarter, substantially increasing the sampling density. Consequently, the Sentinel-2 observations not only provide finer spatial resolution (10 m) but also a higher observation frequency, together enabling more robust median compositing and supporting higher-quality, more spatially complete land-cover mapping after 2016.

In the revised manuscript, the statistical details are presented in Section 2.2 and Figure A2 as follows: *"The acquisition periods and the number of images used from multiple platforms are detailed in Figure A2."*

Reviewer Comment 1.4 — Given the long-term and high-frequency nature of the proposed land-cover and soil erosion dataset, the manuscript would benefit from a more explicit discussion on how these results could be translated into policy-relevant applications. In particular, it is recommended that the authors clarify how the dataset and the identified erosion patterns could support practical decision-making at the policy level, such as evaluating the effectiveness of existing soil conservation measures, and guiding ecological

restoration planning in the Loess Plateau.

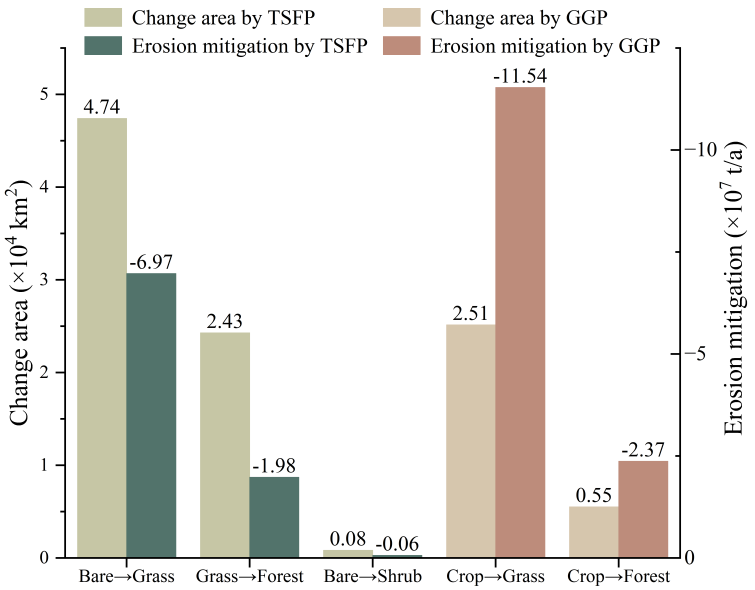

Figure R3: The land-cover changes and associated soil-erosion mitigation from 2000 to 2024, driven by major national programs including the Three-North Shelter Forest Program (TSFP) and the Grain for Green Program (GGP).

Reply: Thanks for the constructive comments. To further analyze the impact of policies on the Loess Plateau region, we explore the progress achieved in policy-guided soil erosion mitigation and identify areas where governance requires further improvement. First, we quantified the land-cover change area and the associated erosion reduction of major vegetation transitions from 2000 to 2024. In this region, ecological restoration has been primarily guided by two national programs: the Three-North Shelter Forest Program (TSFP) [R1] and the Grain for Green Program (GGP) [R2]. Based on the objectives and implementation focus of these policies, we attribute major land-cover transitions to the two programs: transitions from bare land to grassland, grassland to forest, and bare land to shrubland are mainly associated with TSFP, whereas transitions from cropland to grassland and cropland to forest are mainly associated with GGP. We note that these attributions may not capture all land-cover changes across the entire Loess Plateau, as additional environmental drivers and local human activities also contribute to observed transitions. Nevertheless, this rule-based attribution provides a practical way for policy evaluation using the proposed LP-QLC10 dataset, and offers a quantitative reference for assessing restoration effectiveness.

As illustrated in Figure R3, the TSFP primarily promotes the conversion of bare land to grassland (4.74×10^4 km^2) and grassland to forest (2.43×10^4 km^2) across the Loess Plateau, with a smaller extent of bare land converting to shrubland. Among these transitions, the conversion from bare land to grassland yields substantial erosion mitigation, reducing soil loss by -6.97×10^7 t/a. In contrast, although the land-cover changes associated with the GGP occur over a smaller total area, their erosion-mitigation benefits are more pronounced. The cropland-to-grassland transition covers 2.51×10^4 km^2 and reduces erosion by -11.54×10^7 t/a. The area of cropland-to-forest transition is 0.55×10^4 km^2, yet it decreases erosion by -2.37×10^7 t/a. Overall, the estimated erosion reduction attributed to the GGP is about 1.5 times that attributed to the TSFP over 2000–2024.

Second, based on the land-cover pattern in Q4 2024, we generated an optimized land-cover map un-

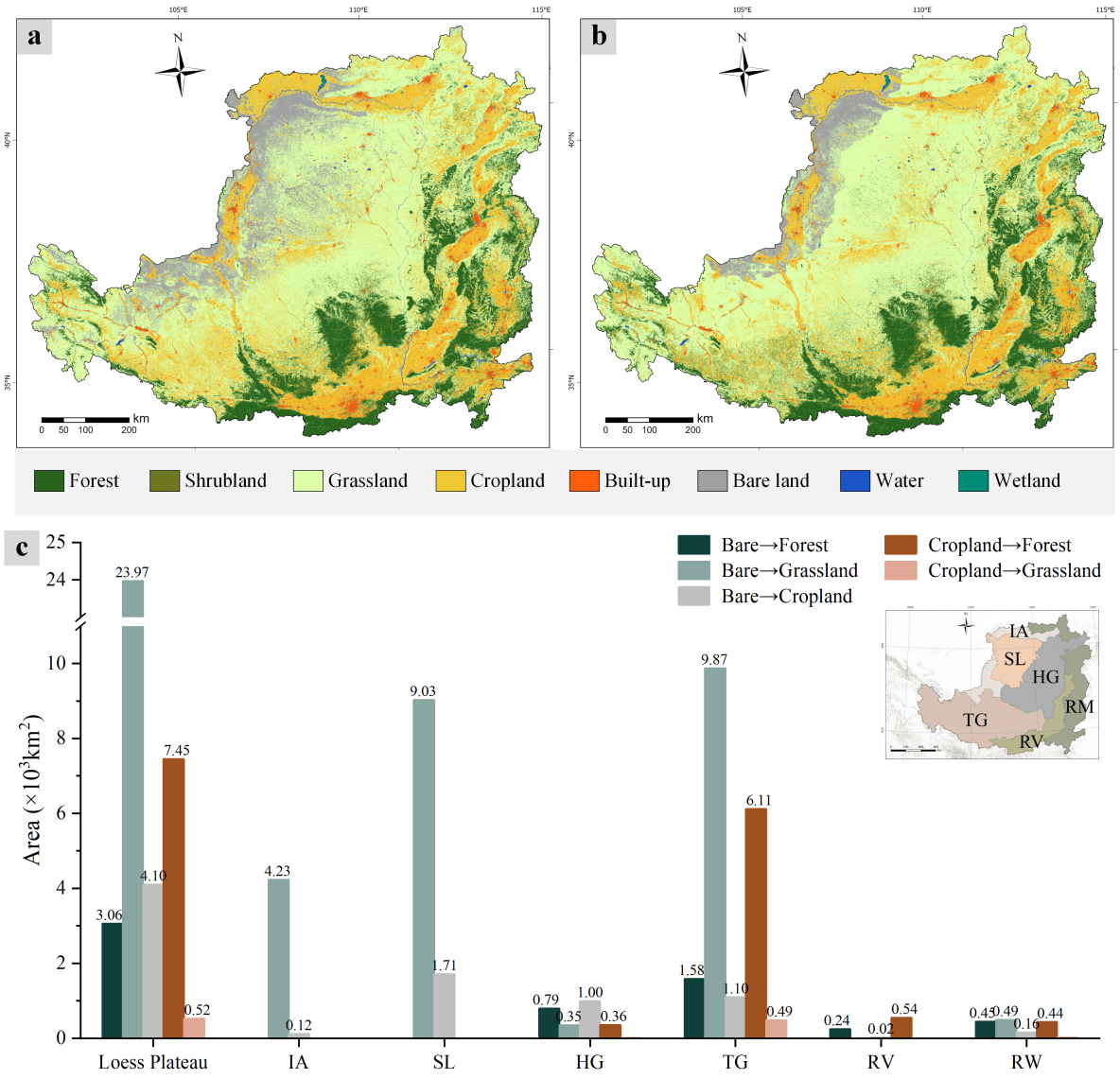

Figure R4: Potential minimum soil erosion in the Loess Plateau. (a) Current soil erosion in 2024, (b) estimated minimum soil erosion under optimized land-cover conditions, and (c) statistical distribution of erosion intensity levels.

der the environmental suitability constraints defined in Section 5.2 of the manuscript. Compared with the current land-cover pattern in Fig. R4(a), the optimized map in Fig. R4(b) indicates substantial landscape adjustments, and a detailed statistical summary of the changes is provided in Fig. R4(c). The transitions from the current state to the minimum-erosion scenario can also be grouped according to the two national programs. Specifically, bare-to-forest and bare-to-grassland transitions are consistent with the primary objectives of the TSFP, whereas cropland-to-forest and cropland-to-grassland transitions align with the GGP. In addition, bare-to-cropland serves as a compensatory change to offset cropland losses in the optimized scenario. Among these transitions, bare-to-grassland under the TSFP accounts for the largest changing area, requiring an additional 2.4×10^4 km^2 conversion. Meanwhile, under the GGP, cropland-to-forest represents

a particularly challenging transition: not only it involves an area of 7×10^3 km^2, but the establishment of forest cover typically requires sustained management over years to decades, requiring long-term planning and continuous implementation.

In the revised manuscript, we added a paragraph in Section 5.1 to analyze the effectiveness of existing soil conservation measures: *"As illustrated in Figure 14(b), the TSFP primarily promotes the conversion of bare land to grassland (4.74×10^4 km^2) and grassland to forest (2.43×10^4 km^2) across the Loess Plateau, with a smaller extent of bare land converting to shrubland. Among these transitions, conversion from bare land to grassland yields substantial erosion mitigation (-6.97×10^7 t/a, ranging -7.66– -6.50$\times 10^7$ t/a considering the land-cover uncertainty). In contrast, although the land cover changes associated with the GGP occur over a smaller total area, their erosion mitigation benefits are more pronounced. The cropland to grassland transition covers 2.51×10^4 km^2 and reduces erosion by -11.54×10^7 t/a. The cropland to forest transition covers only 0.55×10^4 km^2, yet it decreases erosion by -2.37×10^7 t/a. Overall, the estimated erosion reduction attributed to the GGP is approximately 1.5 times that attributed to the TSFP over 2000–2024."*

The suggestion for future ecological restoration is modified in Section 5.2 as follows: *"Bare to forest and bare to grassland transitions are consistent with the primary objectives of the TSFP, whereas cropland to forest and cropland to grassland transitions align with the GGP. In addition, bare to cropland serves as a compensatory adjustment to offset cropland losses in the optimized scenario. Among these transitions, bare to grassland under the TSFP accounts for the largest converted area (2.4×10^4 km^2), concentrated in the central Plateau, including the SL and TG regions. Meanwhile, under the GGP, cropland to forest represents a particularly challenging transition: it not only involves 7×10^3 km^2, but also typically requires sustained management over years to decades to establish stable forest cover, underscoring the need for long-term planning and continued implementation."*

Reference of this reply:

[R1] Zhai, J., Wang, L., Liu, Y., Wang, C. and Mao, X., 2023. Assessing the effects of China's three-north shelter forest program over 40 years. Science of the Total Environment, 857, p.159354.

[R2] Lei, D.E.N.G., Shangguan, Z.P. and Rui, L.I., 2012. Effects of the grain-for-green program on soil erosion in China. International Journal of Sediment Research, 27(1), pp.120-127.

Reviewer 2

This manuscript synthesized Sentinal-2 and Landsat series high resolution data and applied an advanced framework, cross-temporal consistency-constraint learning, to map land cover (LC) time series 2000-2024 over China's Loess Plateau region. In this framework, encoder is applied on the time series to obtain consistency-constrained time series (CCTS), which enhanced the temporal coherence and semantic stability of land-cover representations, and first build segmentation model accounting for inter-temporal consistency in feature levels, then trained segmented features under the supervision of labeled datasets, and finally perturbed CCTS accounting for different noise in the original dataset to perform intra-temporal refinement over the classification and produce the final LC time series. After analysis of the land use change over the past two decades, authors applied the LC product to estimate the average annual soil loss per unit area and assess the water-induced soil erosion over the whole Loess Plateau area. Overall, this is solid work introducing a new fine-resolution dataset with comprehensive and in-depth application case study, which matches the scope of ESSD. There are several concerns from my side, which suggest further improvement before the manuscript can be accepted for publication.

Reply: We greatly appreciate the time and effort you dedicated to reviewing our manuscript. We have carefully revised the paper to address all the points raised. Please find our point-to-point response in the following.

Reviewer Comment 2.1 — As the major scope of ESSD, work shall focus on your data product first. The strategy using CCTS is claimed to bring more consistency, which is the biggest highlight in your data product, but lacks quantitative evaluation in the whole manuscript. I suggest authors quantitatively evaluate the improvement of spatial and temporal consistency in LC mapping with/without building CCTS.

Table R5: Quantitative comparison with baseline model.

Method	Metrics	2001	2006	2011	2016	2021	Average
HRFormer (baseline, without CCTS)	OA	0.733	0.748	0.767	0.838	0.842	0.7856
LP-QLC10 (ours, with CCTS)		**0.785**	**0.787**	**0.800**	**0.855**	**0.845**	**0.8144**
HRFormer	Kappa	0.602	0.617	0.669	0.760	0.768	0.6832
LP-QLC10		**0.686**	**0.673**	**0.715**	**0.783**	**0.772**	**0.7258**
HRFormer	FWIoU	0.492	0.578	0.542	0.638	0.683	0.5866
LP-QLC10		**0.663**	**0.643**	**0.582**	**0.647**	**0.692**	**0.6454**

Reply: We really appreciate the reviewer's suggestion on the scope of ESSD, which could further improve the reliability of CCTS in producing the products. To further compare the improvement of spatial and temporal consistency brought by CCTS, we conducted both quantitative and qualitative evaluations to assess the framework with/without building CCTS for long-term land-cover mapping. To be more specific, as the proposed CCTS used in this manuscript adopted **HRFormer**[R4] as its backbone segmentation model, we trained a pure, original HRFormer (without CCTS) on imagery of 2021 by using the ESA WorldCover v200 product[R5] as labels, and then used the well-trained model to generate baseline maps from 2000 to 2024. The quantitative evaluation of the baseline maps, evaluated using the multi-temporal reference set

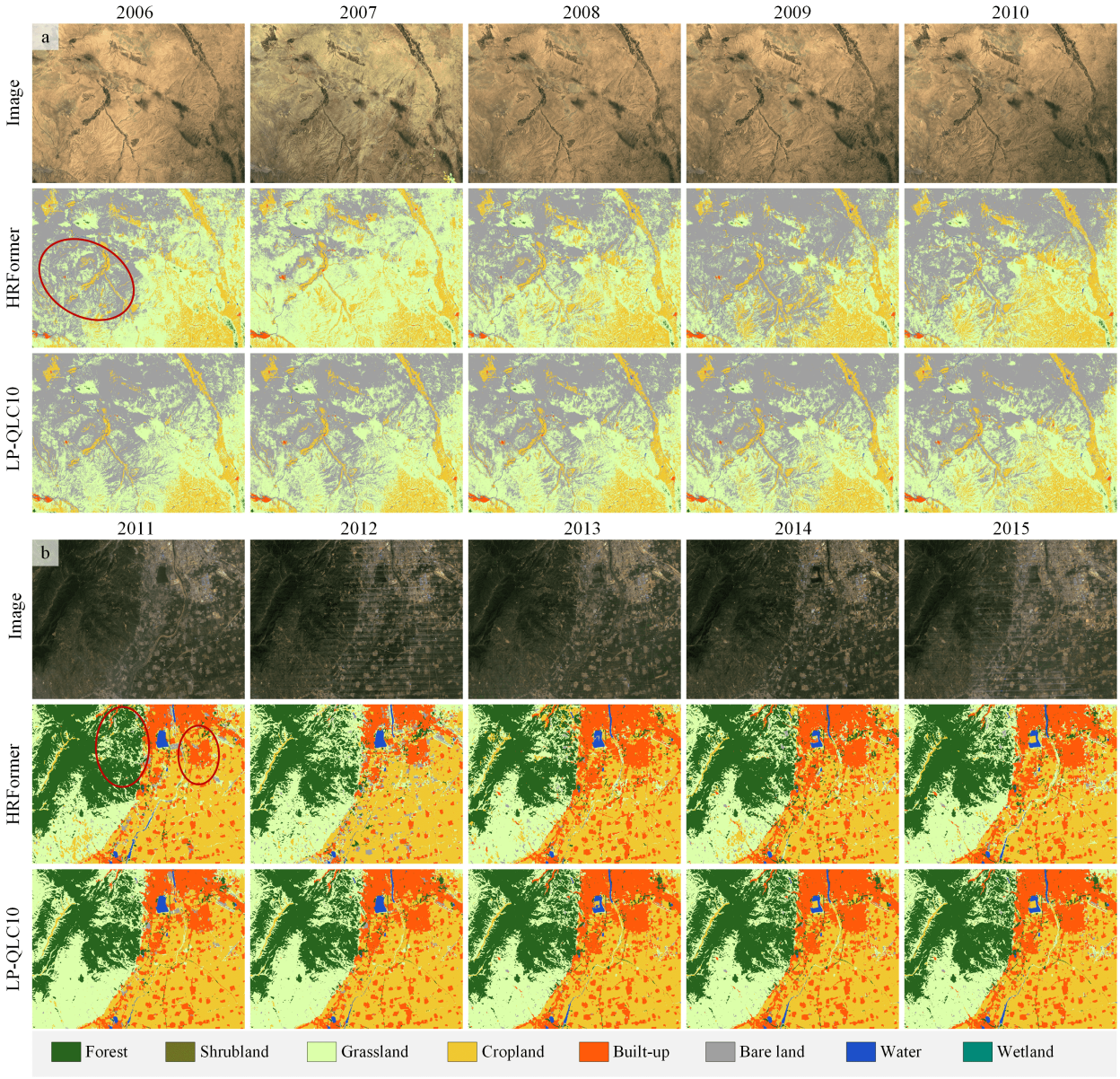

Figure R5: Qualitative comparison with/without CCTS. HRFormer serves as the baseline model without using CCTS, while LP-QLC10 denotes the dataset generated using CCTS. To facilitate the identification of actual land-cover categories in remote sensing imagery, all visualizations are based on composite images from the third quarter.

constructed in this study, is presented in Table R5. From the comparison, the LP-QLC10 dataset (the final data product of this manuscript) generated by the CCTS framework consistently outperformed the baseline model across all evaluated years. Notably, the baseline achieved performance comparable to LP-QLC10 in 2021 because the training and evaluation data came from the same year and therefore had highly consistent feature distributions. However, this advantage did not generalize to other years. When applied to earlier periods, particularly those dominated by Landsat imagery, HRFormer was more sensitive to temporal differences in imaging conditions and sensor characteristics, resulting in lower performance in the

2001, 2006, and 2011 validations. In contrast, CCTS improved both accuracy and temporal stability by learning more consistent cross-year representations, achieving approximately 4% higher OA than the baseline in these years. This quantitative comparison further demonstrates the reliability and improvement of spatial/temporal consistency by building CCTS during the multi-temporal mapping.

To further confirm the improvement of building the CCTS in the qualitative comparison, we illustrated the visual results of in Figure R5. Based on the baseline maps shown in the second row of the two areas, the variations in imaging conditions and sensor characteristics substantially compromised the stability of multi-temporal predictions generated by the baseline land-cover model. Figure R5(a) presents an area that experienced only limited land-cover change. Although the actual semantic changes were minor, the HRFormer (baseline without CCTS) results showed pronounced interannual fluctuations, particularly in vegetation-related categories such as grassland and cropland. In contrast, by incorporating consistency constraints during training, the CCTS framework produced temporally stable and spatially coherent predictions throughout the entire multi-year inference period. Furthermore, Figure R5(b) shows an area characterized by the gradual expansion of built-up area and forest. For categories such as forest, which generally require many years to establish, and built-up land, which typically changes incrementally, the baseline model produced markedly inconsistent predictions across years in some areas due to multi-temporal spectral variability, clearly indicating misclassification. This instability not only reduced spatial prediction accuracy but also introduced cumulative errors into area statistics, thereby limiting the ability to reliably characterize long-term land-cover change trajectories. Such limitations are especially problematic for analyses spanning multiple decades. By contrast, the CCTS framework enabled more temporally consistent inference and more faithfully captured the expansion of urban land and the increase in forest area from 2011 to 2015.

Besides, CCTS incorporates a semi-supervised learning strategy to extend single-temporal land-cover maps across the temporal dimension, thereby establishing a novel paradigm for long-term time-series mapping. Accordingly, in the revised manuscript, we compared our LP-QLC10 with published and widely used long-term land-cover products[R1−R3] and demonstrated its superiority. Meanwhile, we added a paragraph in Section 4.1 to highlight the consistency enhancements introduced by CCTS: *"In addition, the LP-QLC10 dataset developed under the CCTS framework significantly outperforms the baseline model in terms of both spatial accuracy and temporal continuity. Quantitatively, this advantage leads to an average overall accuracy (OA) improvement of 2.88%. Qualitatively, it demonstrates enhanced multi-temporal mapping stability for land-cover types, such as grassland, cropland, and forest, with details provided in Appendix A1."* The improvements in long-term stability achieved by the CCTS framework are described in detail in Appendix A1.

Reference of this reply: [R1] Zhang, X., Zhao, T., Xu, H., Liu, W., Wang, J., Chen, X. and Liu, L., 2024. GLC_FCS30D: the first global 30 m land-cover dynamics monitoring product with a fine classification system for the period from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method. Earth System Science Data, 16(3), pp.1353-1381.
[R2] Yang, J. and Huang, X., 2021. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. Earth System Science Data, 13(8), pp.3907-3925.
[R3] Wang, Z., Shi, X., Dou, S., Cheng, M. and Miao, L., 2025. The 30 m land cover dataset for capturing land cover changes induced by ecological restoration from 1990 to 2022 on the Chinese Loess Plateau. Scientific Data, 12(1), p.252.
[R4] Yuan, Y., Fu, R., Huang, L., Lin, W., Zhang, C., Chen, X. and Wang, J., 2021. Hrformer: High-resolution vision transformer for dense predict. Advances in neural information processing systems, 34, pp.7281-7293.
[R5] Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G.,

Table R6: Confusion matrix on SinoLC-1 validation set with per-class Producer Accuracy (PA), User Accuracy (UA), and Intersection over Union (IoU).

Class	Forest	Shrubland	Grassland	Cropland	Built-up	Bare land	Water	Wetland	PA
Tree cover	831	0	42	6	1	0	0	0	0.944
Shrubland	2	15	41	0	0	1	0	1	0.250
Grassland	177	2	2882	90	1	102	0	0	0.886
Cropland	7	0	255	1424	2	14	2	0	0.836
Building & route	16	0	95	79	161	20	0	0	0.434
Barren & sparse veg.	3	1	378	97	40	546	4	0	0.511
Water	0	0	4	4	0	3	31	2	0.705
Wetland	0	0	0	0	0	0	0	0	0
UA	0.802	0.833	0.780	0.838	0.785	0.796	0.838	0	OA 0.798
IoU	0.766	0.238	0.708	0.719	0.388	0.452	0.620	0	

Wevers, J., Cartus, O., Santoro, M., Fritz, S. and Lesiv, M., 2022. ESA WorldCover 10 m 2021 v200.

Reviewer Comment 2.2 — Grassland, cropland and barren land all have low PA, which will bring substantial uncertainty to your conclusions. Please consider the possible uncertainties that might change the conclusion of your analysis.

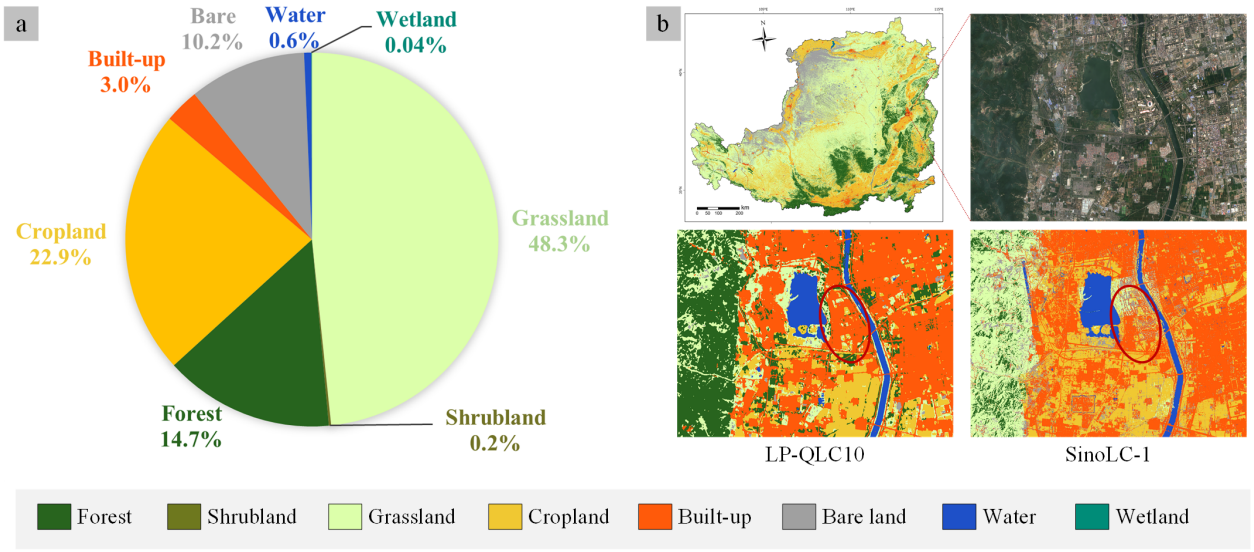

Figure R6: Statistics and visualizations of eight land-cover categories in 2021. (a) Class proportions of each class. (b) Visualization comparison of urban area with SinoLC-1.

Reply: We appreciate this constructive comment. To consider the possible uncertainties, we first explain the potential reasons that may affect the evaluation, and then we further offer a prospective assessment of their potential impact in this response. We observed that, as indicated in Table R6, lower PA is primarily associated with shrubland, built-up areas, and bare land, whereas grassland and cropland consistently

Table R7: Confusion matrix of the 2021 mapping results, evaluated using the time-series validation set, with per-class Producer Accuracy (PA) and User Accuracy (UA).

Class	Forest	Shrubland	Grassland	Cropland	Built-up	Bare land	Water	Wetland	PA
Forest	947	0	82	10	1	0	0	0	0.911
Shrubland	0	18	53	1	1	2	0	0	0.240
Grassland	168	3	3310	150	7	136	2	1	0.876
Cropland	14	0	259	1656	1	15	2	0	0.850
Built-up	6	0	65	29	191	20	0	0	0.614
Bare land	3	1	151	34	18	591	0	0	0.741
Water	1	0	5	1	0	0	44	0	0.863
Wetland	0	0	0	0	0	0	0	3	1.000
UA	0.831	0.818	0.843	0.880	0.872	0.774	0.917	0.750	OA 0.845

achieve PA values above 0.8. First, the low PA for shrubland is primarily associated with its extremely limited spatial extent in the Loess Plateau, which results in insufficient distribution. As shown in Figure R6(a), shrubland represented only 0.2% of the Loess Plateau in 2021. Although training strategies such as OHEM cross-entropy loss[R1] were introduced to mitigate class imbalance and emphasize difficult samples, the pronounced insufficient distribution remains a major constraint on the robust identification of minority classes. In addition, the 10–30m spatial resolution limits the separability of shrubland, thereby increasing spectral confusion. Shrubland (lower than 3 m height and with branch structures) is spectrally similar to grassland, leading to considerable confusion between the two classes. This assessment was also reflected in Table R6, most misclassified shrubland samples were assigned to the grassland category (41 out of 60 points). Second, the PA of built-up land is limited by image resolution. The 10–30 m imagery used in this study can't provide sufficient spatial detail to adequately characterize urban structures. Since small buildings and narrow roads occupy only a few pixels in the images, they are prone to omission or confusion with adjacent pixels during the encoding–decoding process. As shown in Figure R6(b), within urban areas, the LP-QLC10 dataset exhibits blurred boundaries for small building clusters and shows lower delineation accuracy than our previous work "SinoLC-1[R1]", which is derived from 1-m resolution satellite imagery. Third, the low PA of bare land is primarily attributable to systematic bias introduced by differences in class definitions. Following the classification system of the ESA Landcover v200 product[R2], bare land in the LP-QLC10 dataset is defined as "Lands with exposed soil, sand, or rocks and that never have more than 10% vegetated cover during any time of the year." However, the corresponding category, "barren and sparse vegetation" in SinoLC-1, is defined as "Areas covered by sparse vegetation or bare land covered by sand, gravel, or rocks." Thus, the definition of bare land in our classification system is more stringent. Sparsely vegetated areas are classified as grassland in our dataset, but are included in the "barren and sparse vegetation" category in SinoLC-1. Therefore, as shown in Table R6, a substantial proportion of the "barren and sparse vegetation samples" in the validation set were identified as grassland in our dataset (378 points). Besides, to illustrate the impact of inconsistencies in class definitions, the 2021 mapping results were evaluated using time-series validation points annotated by experts in this study, which follow consistent class definitions. As reported in Table R7, the PA for bare land reached approximately 0.741, which is substantially higher than the SinoLC-1 validition. In the revised manuscript, we modified the names of the land-cover types in Table 5 (Table R6 in the response letter) to ensure consistency with those of the sample points in SinoLC-1.

To further offer a prospective assessment of their potential impact, we found out that these classification

uncertainties can propagate into the subsequent soil-erosion analysis, but their effects differ in magnitude and spatial significance among categories. First, for shrubland, the associated uncertainty is expected to have only a slight influence on the regional-scale erosion conclusions, because shrubland occupies an extremely small proportion of the Loess Plateau. Moreover, most shrubland misclassification occurred as grassland, and both classes generally represent vegetation-covered surfaces with relatively similar ecological functions in erosion mitigation, as shown in Tables R3 and R4. Second, for built-up area, the main uncertainty arises in the delineation of small urban features and narrow roads. This may introduce local bias into the estimation of erosion intensity and land-cover transitions at urban boundaries, where mixed pixels are common. However, because built-up land occupies a small proportion (3%) of the Loess Plateau and is not the dominant land-cover type controlling regional erosion dynamics, its influence is expected to be mainly confined to the local spatial detail of erosion change rather than to the long-term regional evolution of soil erosion. Third, uncertainty in the bare land class may also influence the erosion assessment. Misclassification between bare land and grassland may affect the quantified contribution of bare land reduction to erosion mitigation. This effect can be approximately evaluated through uncertainty propagation of the C factor in the RUSLE model. Because soil erosion is linearly proportional to the C factor, misclassification between bare land and low-FVC grassland directly translates into proportional bias in the estimated erosion intensity when the other factors remain unchanged. According to the national specification[R4], if a true bare-land pixel is misclassified as grassland, the erosion intensity estimated from the C factor would be underestimated by $(0.45-0.7)/0.7 = -35.7\%$. Based on the confusion matrix in Table R7, $151/798 = 18.9\%$ of the true bare land samples were mapped as grassland. Under the simplifying assumption that these omitted bare-land pixels are mainly assigned to low-FVC grassland, an approximate first-order estimate suggests that the average erosion bias of the bare land category may reach about $18.9\% \times (-35.7\%) = -6.7\%$. Conversely, from the perspective of the mapped bare land category, the corresponding first-order overestimation caused by grassland pixels misclassified as bare land is about 9.9%. Thus, the erosion mitigation caused by bare-to-grassland transform in Figure 14 can be expressed with an approximate range of -7.66– -6.50$\times 10^7$ t/a. However, the impact on regional total erosion further depends on the spatial proportion of misclassified pixels and the corresponding *R, K, L, S,* and *P* values at those locations. This estimate is provided solely for uncertainty analysis.

In the revised manuscript, we supplemented the impact of bare land mapping uncertainties on erosion assessment, as follows: *"Among these transitions, conversion from bare land to grassland yields substantial erosion mitigation (-6.97$\times 10^7$ t/a, ranging -7.66– -6.50$\times 10^7$ t/a considering the land-cover uncertainty)."*

Reference of this reply:

[R1] Shrivastava, A., Gupta, A. and Girshick, R., 2016. Training region-based object detectors with online hard example mining. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 761-769).

[R2] Li, Z., He, W., Cheng, M., Hu, J., Yang, G. and Zhang, H., 2023. SinoLC-1: The first 1 m resolution national-scale land-cover map of China created with a deep learning framework and open-access data. Earth system science data, 15(11), pp.4749-4780.

[R3] Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S. and Lesiv, M., 2022. ESA WorldCover 10 m 2021 v200.

[R4] Ministry of Ecology and Environment of the People's Republic of China (MEE). Technical Specification for Investigation and Assessment of National Ecological Status: Ecosystem Services Assessment. Tech. Rep. HJ 1173–2021, Beijing, China, 2021 (in Chinese).

Reviewer Comment 2.3 — Please indicate in the title that the soil erosion discussed in this manuscript is

specifically water-induced erosion.

Reply: We thank the reviewer for the suggestion to describe the title more precise and specific. We have revised the title to: *"25-year, quarterly land change maps of China's Loess Plateau reveal long-term and substantial water-induced soil erosion mitigation"*.

Reviewer Comment 2.4 — Line 202: How exactly did you perturb the image to consider intra-temporal consistency?

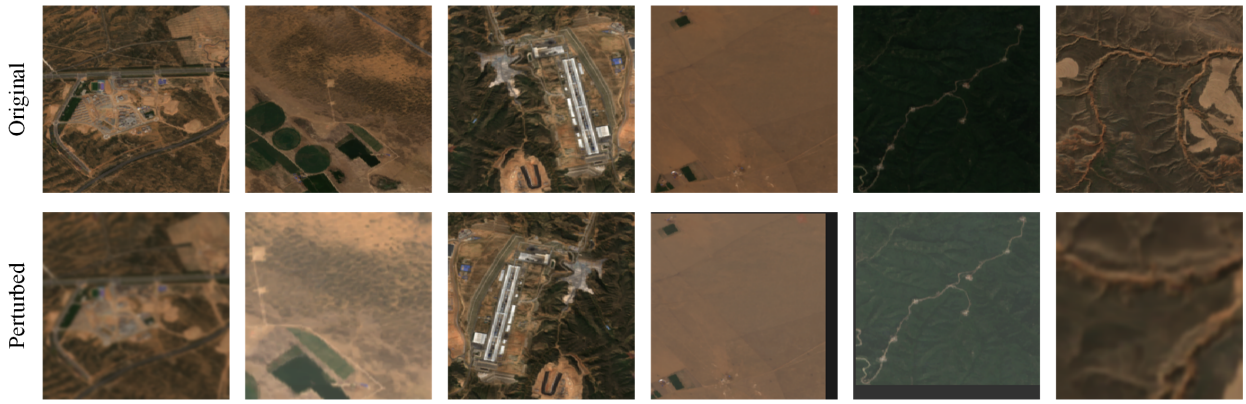

Figure R7: The visualization of perturbed imagery.

Reply: We thank the reviewer for this question. In this study, image perturbation is divided into two categories: spatial-domain perturbation and spectral-domain perturbation. Specifically, spatial-domain perturbation includes image flipping (horizontal and vertical), shifting (up to 5% of the image size), rotation (90°, 180°, or 270°), and cropping with a scale range of 20%–100%. Spectral-domain perturbation includes brightness variation (up to ±20%), contrast variation with a range of 20%–80%, and Gaussian blur (intensity range of 0–2.0). During preprocessing, one spatial-domain perturbation and one spectral-domain perturbation are randomly selected and then jointly applied to the original image to generate a perturbed sample, as shown in Figure R7. Meanwhile, the same spatial-domain perturbation is applied during pseudo-label generation, ensuring that the pseudo-label remains aligned with the perturbed image in spatial distribution. The purpose of these perturbations is to mimic the differences commonly observed in multi-temporal imagery due to changes in imaging conditions and satellite platforms, while preserving consistent land-cover semantics. By exposing the model to such variations and enforcing consistency constraints, the encoder is encouraged to map various image appearances into a shared semantic feature space, thereby improving the robustness and temporal stability of land-cover mapping.

Reviewer Comment 2.5 — Line 286: It seems like the meaningful spatial resolution is only 30m for 2000-2015, then your title claiming the data product to be a 10-m resolution is misleading?

Reply: Thank you for your question. We used a combination of Landsat and Sentinel satellite data to produce the time-series land-cover maps from 2000 to 2024. Owing to differences in the native spatial resolution of these satellite sources, the effective mapping resolution was 30 m for the period from 2000 to 2015 and 10 m from 2016 onward. Therefore, we agree that the original title may have caused some misunderstanding regarding the spatial resolution of the final product. Following your suggestion, we have

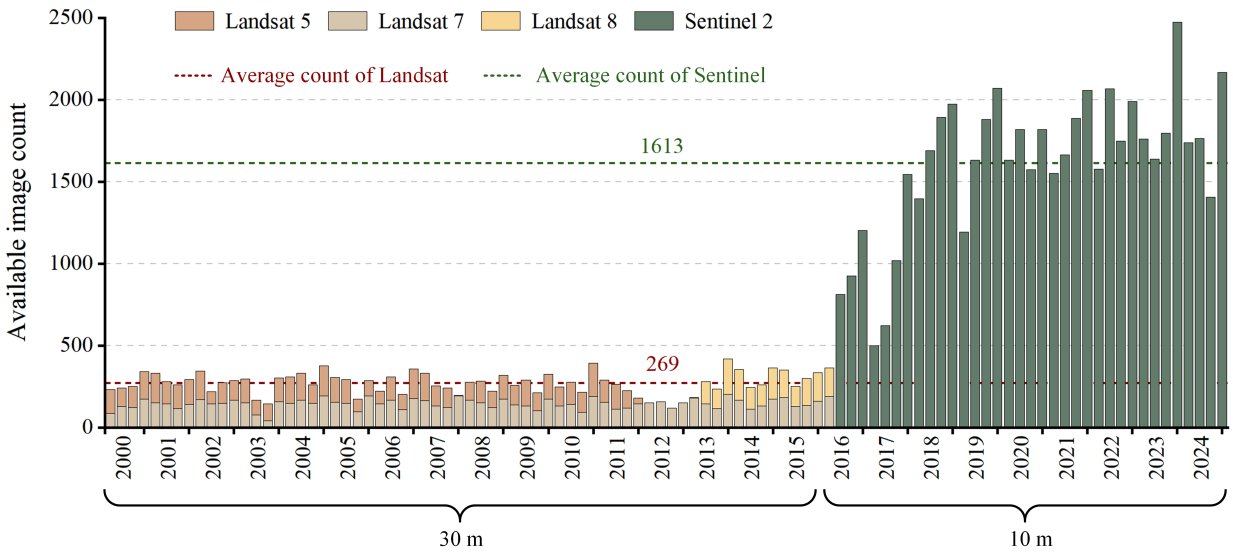

Figure R8: The obtained image count from multiple platforms within each temporal period. The statistic denotes the number of quality-screened satellite image tiles in Google Earth Engine that intersect the Loess Plateau. A Landsat scene covers approximately 3.3×10^4 km^2, whereas a Sentinel tile covers approximately 1.0×10^4 km^2.

revised the title to: *"25-year, quarterly land change maps of China's Loess Plateau reveal long-term and substantial water-induced soil erosion mitigation."*

We also placed a more detailed explanation about the image volume and distribution to further answer the constructive questions. In terms of data volume, as shown in Figure R8, the 10-m imagery acquired after 2016 constitutes the major part of the overall mapping effort. Specifically, an average of 1,613 images per quarter were processed for the period after 2016, whereas only 269 images per quarter were required on average before 2016. This marked difference indicates that the 10-m resolution component dominates the dataset in terms of image volume and mapping workload. In addition, our mapping framework is not restricted to the current spatial resolution and can be readily extended to higher-resolution long-term mapping when suitable image sources are available.

In the manuscript, we also clarified the temporal variation in spatial resolution in the product description within the main text to avoid potential misunderstanding, such as: *"Specifically, the dataset comprises 100 temporal phases and was produced at a spatial resolution of 30 m for 2000–2015 and 10 m for 2016–2024, reflecting the native characteristics of the underlying satellite sensors."*

Reviewer Comment 2.6 — Line 308: Why does alpha have two different values for warm and cold periods separately? Is it because of snow fall contributing almost no weathering or just the weathering rate reduced under low temperature?

Reply: Thank you for this important question. The rainfall erosivity factor *(R)* describes the capacity of rainfall to detach and transport soil and is primarily controlled by rainfall kinetic energy and intensity, rather than by snowfall. At the event scale, erosivity is strongly influenced by short-term variations in rainfall amount, duration, and intensity[R1]. However, in large-scale assessments, precipitation information is usually limited to daily rainfall totals[R2], whereas detailed observations of intra-day rainfall intensity are often unavailable. This limitation makes it difficult to directly characterize rainfall type and its associated erosive

force using daily data. To address this issue, some studies[R1,R3] have adopted daily rainfall erosivity models with season-specific α values to represent differences in rainfall regimes between warm and cold seasons. Specifically, in the warm season, rainfall is more frequently characterized by short-duration, high-intensity storms with stronger erosive effects, whereas in the cold season, precipitation is generally of lower intensity and longer duration, resulting in relatively weaker erosivity. Considering these clear seasonal differences in rainfall characteristics and erosivity, this study followed the national specification of China[R4], with α values of 0.3937 for the warm season and 0.3101 for the cold season.

Reference of this reply:

[R1] Weng, X., Ye, Y., Ye, Y., Zeng, B. and Long, X.J., 2022. Spatiotemporal characteristics of rainfall erosivity in Xiaoanxi Basin using multiple algorithms. Trans. Chin. Soc. Agric. Eng. Trans. CSAE, 38, pp.143-150.
[R2] Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. and Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Scientific data, 2(1), p.150066.
[R3] Gao, G., Yin, S., Chen, T., Huang, D. and Wang, W., 2022. Spatiotemporal variation and cause analysis of rainfall erosivity in the Yangtze River Basin of China. Trans. CSAE, 38, pp.84-92.
[R4] Ministry of Ecology and Environment of the People's Republic of China (MEE). Technical Specification for Investigation and Assessment of National Ecological Status: Ecosystem Services Assessment. Tech. Rep. HJ 1173–2021, Beijing, China, 2021 (in Chinese).

Reviewer Comment 2.7 — Line 309: Is snowfall included or excluded?

Reply: Thanks for your question. As noted in our response to Comment 2.6, the rainfall erosivity factor is intended to characterize the capacity of rainfall to detach and transport soil, and it is primarily governed by rainfall kinetic energy and intensity rather than by snowfall processes. Therefore, the erosivity equation used in this study is specifically designed to represent rainfall-driven erosion, snowfall is not considered in its calculation. In addition, the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS)[R1] employed in this study provides daily quasi-global rainfall estimates, which are consistent with both the spatial coverage of our study and the requirements of the adopted rainfall erosivity calculation. To avoid conceptual ambiguity and improve terminological precision, we have revised the wording in Section 3.2.1 of the manuscript by replacing *"precipitation"* with *"rainfall"*. This revision ensures that the description is more scientifically rigorous and consistent with the physical meaning of the erosivity factor.

Reference of this reply:

[R1] Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. and Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Scientific data, 2(1), p.150066.

Reviewer Comment 2.8 — Line 317: What's the reason for re-calibrating Chinese soil? It seems to be a linear bias correction formula that due to some systematic bias? I assume soil in China has no difference from other regions when represented by texture (sand, silt, clay, OC and bedrock)

Reply: Thank you for the question. To address this, we further clarify the rationale of this calculation and its range of applicability. Soil erodibility (K) factor characterizes the susceptibility of soil particles to

detachment and transport by water. In this study, it was computed with the following equations:

$$K_{\mathrm{EPIC}} = \{0.2 + 0.3\exp[-0.0256 m_{\mathrm{s}}(1 - m_{\mathrm{silt}}/100)]\} \times [m_{\mathrm{silt}}/(m_{\mathrm{c}} + m_{\mathrm{silt}})]^{0.3} \times \{1 - 0.25 m_{\mathrm{o}}/[m_{\mathrm{o}} + \exp(3.72 - 2.95 m_{\mathrm{o}})]\} \times \left\{1 - 0.7(1 - m_{\mathrm{s}}/100) / \left[(1 - m_{\mathrm{s}}/100) + \exp[-5.51 + 22.9(1 - m_{\mathrm{s}}/100)]\right]\right\}, \quad (2)$$

$$K = (-0.01383 + 0.51575 K_{\mathrm{EPIC}}) \times 0.1317. \quad (3)$$

The Environmental Policy Integrated Climate (EPIC) model[R1] in Eq. 2 estimates the soil erodibility *(K)* factor using soil organic carbon content and particle-size composition. This model is an empirical parameterization developed for large-scale mapping applications and was originally calibrated and tested across seventeen major land resource areas in the United States. However, previous studies[R2] in China have shown that direct application of the EPIC model tends to overestimate soil erodibility compared with values derived from field sampling and observational surveys. Such discrepancies are mainly attributable to differences in erosion assessment conditions and agricultural practices between China and the United States. First, China and the United States adopt slightly different standards for defining and calculating erosive rainfall, which may lead to inconsistent erosivity estimates for the same rainfall event. Second, runoff observation records in China are often temporally incomplete, with limited or absent winter observations in many monitoring sites. Third, contour farming is more commonly practiced in many parts of China, whereas downslope cultivation has historically been more common in the United States[R2]. These differences in rainfall erosivity definition, hydrological observation conditions, and cultivation practices contribute to systematic differences in estimated soil erodibility between the two countries. Therefore, the original EPIC model cannot be directly applied to estimate soil erodibility in China. Fortunately, a stable linear relationship has been reported between EPIC-derived values and field-measured values. Based on this relationship, Zhang et al.[R2] established a correction equation for the EPIC-derived *K* factor using observations from six erosion monitoring stations across China, as shown in Eq. 3. The corrected *K* factor shows better agreement with field-measured values and is therefore considered more suitable for estimating soil erodibility under Chinese environmental conditions.

Reference of this reply:

[R1] Williams, J.R., Renard, K.G. and Dyke, P.T., 1983. EPIC: A new method for assessing erosion's effect on soil productivity. Journal of Soil and water Conservation, 38(5), pp.381-383.
[R2] Zhang, K.L., Shu, A.P., Xu, X.L., Yang, Q.K. and Yu, B., 2008. Soil erodibility and its estimation for agricultural soils in China. Journal of arid environments, 72(6), pp.1002-1011.

Reviewer Comment 2.9 — Line 323: Cropland has a strong seasonal signal. But here in your equation you mentioned A is an annual value. Are you using the same NDVI and parameters for different seasons? Please clarify.

Reply: Thanks for this question. All time-varying factors in this study, including not only the *C* factor but also the *R* and *P* factors, were estimated at quarterly intervals. Accordingly, the resulting soil erosion *A* was also produced at a quarterly temporal resolution. Specifically, for the *C* factor, we first derived quarterly NDVI from remote sensing imagery, then calculated fractional vegetation cover (FVC) using Eq. 4, and finally generated the quarterly *C* factor by integrating the land-cover maps with the corresponding FVC maps for each quarter. This workflow captures not only the seasonal variability of cropland, but also the intra-annual dynamics of other vegetation types, including forest, shrubland, and grassland, thereby providing a more realistic representation of vegetation-related erosion regulation throughout the year. Benefiting from

this finer temporal resolution, our soil erosion maps are able to characterize seasonal variations in erosion processes and achieve substantially higher predictive accuracy than existing products.

$$\mathrm{FVC} = \frac{\mathrm{NDVI} - \mathrm{NDVI}_{\mathrm{bare}}}{\mathrm{NDVI}_{\mathrm{veg}} - \mathrm{NDVI}_{\mathrm{bare}}}. \tag{4}$$

In the revised manuscript, we have added a paragraph to clarify the temporal interval used for factor calculation: *"In this study, all time-varying factors were estimated at quarterly intervals so as to capture intra-annual variability. Since soil erosion is cumulative over time, the annual soil erosion used for subsequent assessment was obtained by summing the four quarterly estimates within each year. This strategy preserves seasonal information during the calculation process while ensuring consistency with annual-scale evaluation of long-term erosion dynamics."*

Reviewer Comment 2.10 — Line 391: Under fine resolution such as 10m, the regular 8 LC types might not be enough for mosaic landscape cases, e.g., the green space in built-up regions, such as parks, that can be classified as either forest or cropland. I would suggest authors to discuss this point.

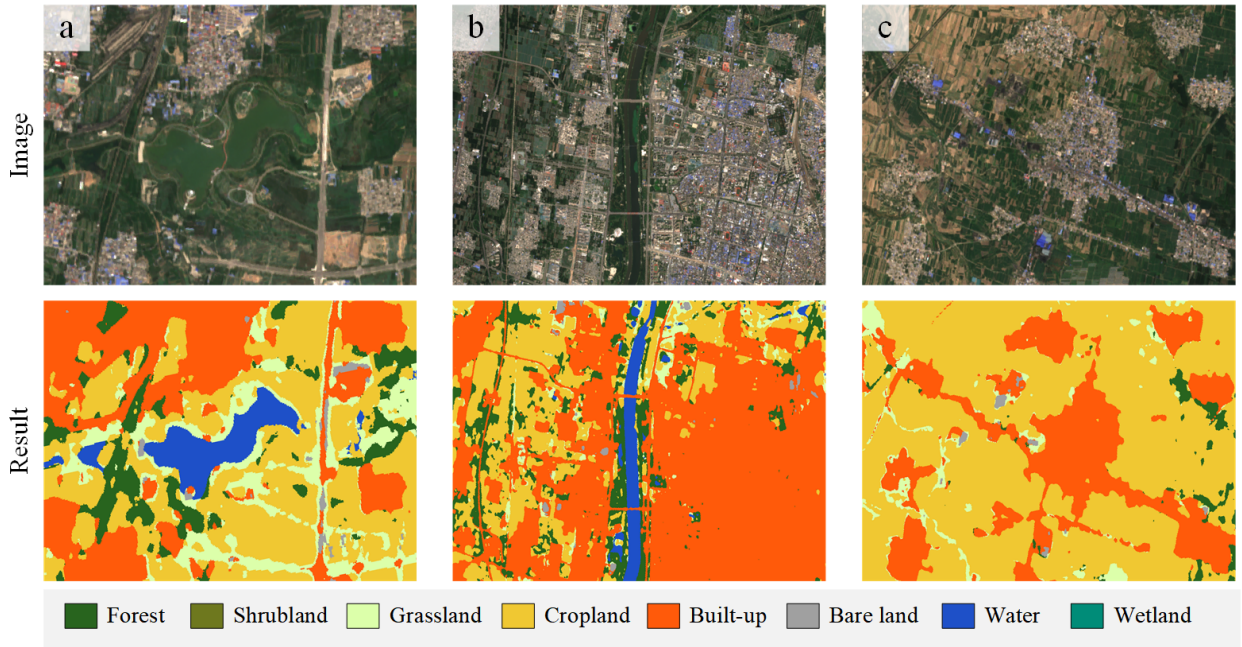

Figure R9: The land-cover mapping results for mosaic landscape cases.

Reply: Thank you for this insightful comment. We clarify this issue by discussing the selection criteria of land-cover categories and presenting examples of mixed scenarios. The classification system used in this study focuses on land cover, which describes the physical and biological characteristics of the land surface at the pixel level. In contrast, finer-grained classifications of mixed landscapes are more closely related to land use, reflecting the functional roles of land in economic and social activities. The eight land-cover categories selected in this study are based on the globally recognized Level-1 classification system[R1,R2], and have also been widely adopted in previous studies[R3–R5]. Moreover, in soil-erosion research, the primary concern is how different surface types influence rainfall interception, surface runoff generation, soil

exposure, and human disturbance. Therefore, these categories represent the major surface characteristics related to erosion across the Loess Plateau, including vegetation-covered surfaces with varying protective capacities (e.g., forest, shrubland, and grassland), cultivated land with active human disturbance (e.g., cropland), exposed or weakly protected surfaces (e.g., bare land), and surfaces where water erosion is minimal or not applicable (e.g., water and wetlands). Built-up areas are treated separately due to their distinct erosion responses resulting from surface hardening. This classification scheme strikes a balance between physical interpretability and temporal consistency across long-term time series.

Furthermore, in mixed landscape scenarios, such as urban areas, the classification results reflect the coexistence of multiple land-cover types. For example, Figure R9(a) illustrates a typical park scene in the city center, which includes water, built-up area, and grassland. Instead of classifying the entire area as a built-up zone, our mapping approach assigns one of eight labels to each pixel based on the land cover's physical properties. Similarly, as shown in Figures R9(b) and (c), urban areas often contain a mixture of vegetation types, such as cropland, grassland, and forest. The land-cover type for each pixel is determined by the specific physical characteristics of the surface at that location.

Reference of this reply:

[R1] Anderson, J.R., 1976. A land use and land cover classification system for use with remote sensor data (Vol. 964). US Government Printing Office.

[R2] Di Gregorio, A., 2005. Land cover classification system: classification concepts and user manual: LCCS (Vol. 8). Food & Agriculture Org..

[R3] Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J.C., Mathis, M. and Brumby, S.P., 2021, July. Global land use/land cover with Sentinel 2 and deep learning. In 2021 IEEE international geoscience and remote sensing symposium IGARSS (pp. 4704-4707). IEEE.

[R4] Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S. and Lesiv, M., 2022. ESA WorldCover 10 m 2021 v200.

[R5] Zhang, X., Liu, L., Chen, X., Gao, Y., Xie, S. and Mi, J., 2021. GLC_FCS30: Global land-cover product with fine classification system at 30 m using time-series Landsat imagery. Earth System Science Data, 13(6), pp.2753-2776.

Reviewer Comment 2.11 — Line 391: Are both validation LC products having the same LC type definition and follow the same criteria to map LC?

Reply: Thanks for your question. The land-cover categories selected in this study are fundamental classifications commonly used in land-cover mapping[R1−R4], and all the comparison methods include the eight categories that found in the Loess Plateau. In addition, we have introduced an additional mapping product named YRCC_LPLC[R3], which is specifically developed to assess land-cover change in the Loess Plateau. This dataset shares the same classification system with CLCDs[R2], and is included in the accuracy assessment and comparative analysis in the revised manuscript (Figure 7, Tables 3 and 4). Moreover, to standardize the definitions of land-cover types in this study and align with the classification systems of other published products, Table R8 presents both the category descriptions of our dataset and the corresponding class names from other products. In the revised manuscript, the classification system and its relationships with other products are provided in Table A2.

Reference of this reply:

[R1] Zhang, X., Zhao, T., Xu, H., Liu, W., Wang, J., Chen, X. and Liu, L., 2024. GLC_FCS30D: the first global 30 m land-cover dynamics monitoring product with a fine classification system for the period

Table R8: Land-cover types in LP-QLC10: value, definition, and correspondence with existing products. CLCDs and YRCC_LPLC share the same classification system.

Value	Category	Definition in LP-QLC10	Category relations with		
			CLCDs	GLC_FCS30D	SinoLC-1
1	Forest	Geographic areas dominated by trees ($\geq$ 10% cover). Other land-cover types possibly present beneath the canopy.	Forest	Forest	Tree cover
2	Shrubland	Areas dominated by natural shrubs ($\geq$ 10% cover). Shrubs are defined as woody perennial plants under 5 m tall with no distinct main stem.	Shrubland	Shrubland	Shrubland
3	Grassland	Geographic areas dominated by natural herbaceous vegetation ($\geq$ 10% cover), possibly including uncultivated cropland.	Grassland	Grassland	Grassland
4	Cropland	Land covered with annual cropland that is sowed/planted and harvestable at least once within the 12 months.	Cropland	Cropland	Cropland
5	Built-up	Land covered by buildings, roads and other man-made structures. Urban green (parks, sport facilities) is not included.	Impervious	Impervious	Building & traffic route
6	Bare land	Lands with exposed soil, sand, or rocks and never has more than 10 % vegetated cover during any time of the year.	Barren	Bare areas	Barren and sparse vegetation
7	Water	Geographic area covered for most of the year (more than 9 months) by water bodies: lakes, reservoirs, and rivers.	Water	Water body	Water
8	Wetland	Land dominated by natural herbaceous vegetation ($\geq$ 10% cover) that is permanently or regularly flooded by fresh, brackish or salt water.	Wetland	Wetland	Wetland

from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method. Earth System Science Data, 16(3), pp.1353-1381.
[R2] Yang, J. and Huang, X., 2021. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. Earth System Science Data, 13(8), pp.3907-3925.
[R3] Tian, P., Tian, X., Geng, R., Zhao, G., Yang, L., Mu, X., Gao, P., Sun, W. and Liu, Y., 2023. Response of soil erosion to vegetation restoration and terracing on the Loess Plateau. Catena, 227, p.107103.
[R4] Li, Z., He, W., Cheng, M., Hu, J., Yang, G. and Zhang, H., 2023. SinoLC-1: The first 1 m resolution national-scale land-cover map of China created with a deep learning framework and open-access data. Earth system science data, 15(11), pp.4749-4780.

Reviewer Comment 2.12 — Line 399: "By contrast, GLC_FCS30D often misinterprets water bodies as wetland due to surrounding vegetation, and further misclassifies substantial grassland areas as cropland or bare land" Do you have reference to prove this conclusion? Is nearby vegetation the reason for misclassification?

Reply: Thank you for your constructive comment. We conducted a comparative analysis of mapping results along riverbanks to investigate the causes of the erroneous predictions in GLC_FCS30D[R1]. One major reason is that GLC_FCS30D uses a fine classification system for wetlands[R2], which subdivides wetlands into categories such as swamp, marsh, flooded flat, saline, and other sub-types. Specifically, the mapping errors in river areas were concentrated in two main scenarios. First, cropland near the riverbank was misclassified as marsh, as highlighted in the blue box of Figure R10. The definition of marsh is "inland natural

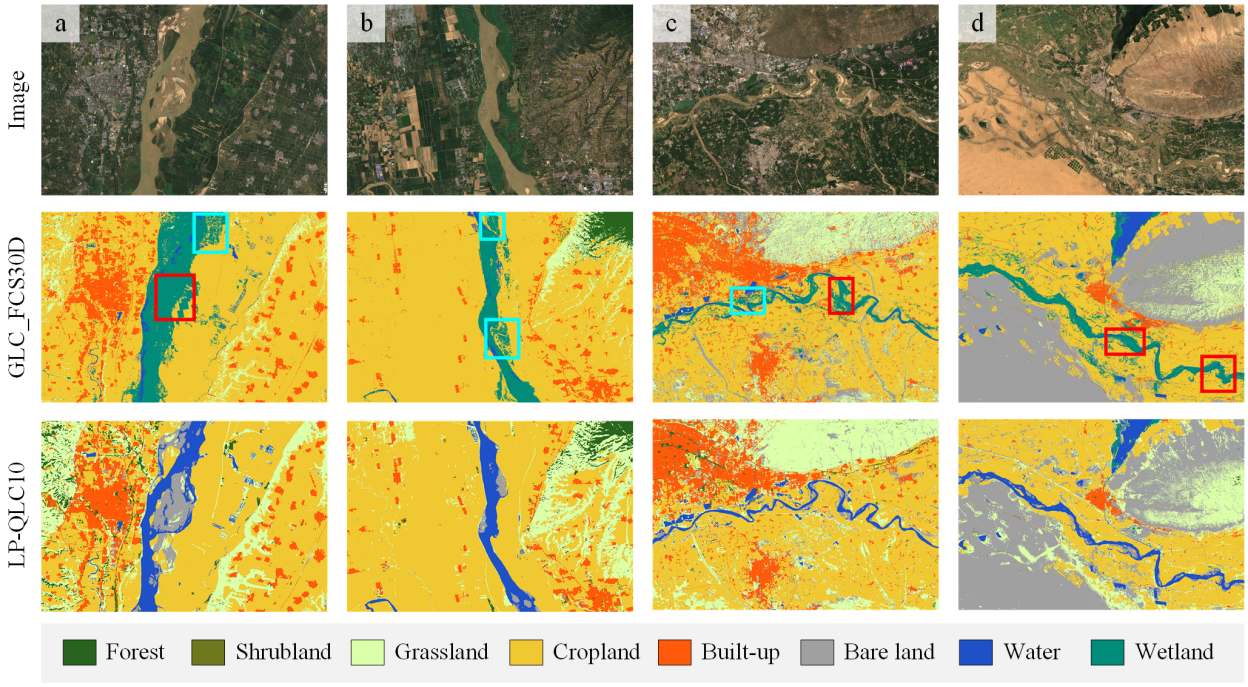

Figure R10: The land-cover mapping results for mosaic landscape cases.

wetlands dominated by herbaceous vegetation", which often overlaps with the spectral signature of paddy field, contributing to this misidentification. Second, exposed floodplains within the river channel were identified as flooded flats, as shown in the red box. A flooded flat is characterized as "non-vegetated inland areas regularly inundated with water", which is identified as bare land in our classification system, as it does not provide the same soil and water conservation benefits as wetlands.

Furthermore, the fine-grained wetland classification system places excessive emphasis on non-water areas in valleys, resulting in under estimation of areas that are actually covered by water. For broader river surfaces, such as the lower reaches of the Yellow River (Figures R10(a) and (b)), GLC_FCS30D tends to predict only a small portion of the river as water bodies. However, for narrower river channels, such as the middle reaches of the Yellow River (Figures R10(c) and (d)), the dataset incorrectly predicted the entire river channel as wetland, more precisely, flooded flat. This misclassification is likely caused by the high sediment content in the Yellow River, which is influenced by soil erosion, leading to confusion with flooded flat. In the Loess Plateau region, where most rivers carry large amounts of sediment, such systematic classification errors are not acceptable. Accordingly, our LP-QLC10 dataset offers significant advantages by accurately classifying river water bodies, surrounding cropland, and bare floodplains. This capability enhances the precision of change detection and provides a reliable dataset for soil erosion assessments.

In the revised manuscript, we modified and refined the corresponding results analysis as : *"GLC_FCS30D successfully identified wetland areas. However, errors were observed along rivers and their banks. Due to the fine-grained wetland classification system in GLC_FCS30D, confusion occurred among wetland sub-classes. Specifically, cropland was misclassified as marshes, while water-covered surfaces were misidentified as flooded flats (sub-classes of wetland), particularly in sediment-rich river channels."*

Reference of this reply:

[R1] Zhang, X., Zhao, T., Xu, H., Liu, W., Wang, J., Chen, X. and Liu, L., 2024. GLC_FCS30D: the

first global 30 m land-cover dynamics monitoring product with a fine classification system for the period from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method. Earth System Science Data, 16(3), pp.1353-1381.
[R2] Zhang, X., Liu, L., Zhao, T., Wang, J., Liu, W. and Chen, X., 2024. Global annual wetland dataset at 30 m with a fine classification system from 2000 to 2022. Scientific Data, 11(1), p.310.

Reviewer Comment 2.13 — Line 458: "In particular, 8.6% of the region was transformed from bare land to grassland". In table 5, Barren land PA is not high and mostly confounded by grassland. This also jeopardizes your further conclusion, for example, "the transformation from bare land to grassland produces the most erosion reduction" in line 527. Please discuss and provide uncertainty.

Reply: Thanks for this constructive question. To demonstrate the uncertainty discussion between grassland and bare land, we further analyze the reasons and measure their potential impact on the results. From the land-cover type definition, sparsely vegetated areas are classified as grassland in our dataset, but are included in the "barren and sparse vegetation" category in SinoLC-1. As a result, a substantial proportion of the "barren and sparse vegetation" samples in the SinoLC-1 validation set were identified as grassland, as shown in Table R6. When the 2021 mapping results were evaluated using the time-series validation set, which follows consistent class definitions, the UA and PA of bare land reached 0.774 and 0.741 (Table R7), respectively. Based on the confusion matrix, $136/764 = 17.8\%$ of the pixels mapped as bare land were in fact grassland, implying that the reported bare land-to-grassland transition could be overestimated by about $8.6\% \times 17.8\% = 1.53\%$. At the same time, $151/798 = 18.9\%$ of the true bare land samples were mapped as grassland, suggesting that part of the actual bare land-to-grassland transition may have been missed, with a potential underestimation of up to about 2%. Therefore, rather than treating the reported 8.6% value as exact, it is more appropriate to interpret it with an approximate uncertainty range of 7.1–10.6%, under the observed bare-grass classification confusion.

In the revised manuscript, we added the impact of bare land mapping uncertainties on land-cover transition, as follows: *" In particular, 8.6% (with estimates ranging 7.1–10.6% when accounting for land-cover uncertainty) of the region was transformed from bare land to grassland."*