

Reviewer #2

R2C1: This manuscript provides a detailed analysis of how land use changes in the Southeast Asia region affect vegetation greenness. Utilizing multi-source land cover datasets, it reveals how the transformation of land use since the 21st century has impacted vegetation greenness, based on machine learning algorithms and the SHAP interpreter. The topic of this manuscript is interesting, explaining why China and India, despite both being countries with rapidly developing agriculture, make significant contributions to greening trends, while the greening trend in Southeast Asia remains stagnant. However, certain aspects may need addressing before publication.

Response: We appreciate the accurate summary and positive comments from the reviewer, and thank them for recognizing the importance of our work on studying regional greening trends.

Major comments

R2C2: The literature review concerning the driving mechanisms behind vegetation greenness changes in Southeast Asia appears to be incomplete and insufficiently detailed. It is essential to provide a more comprehensive overview of existing research to adequately situate the study within the current body of knowledge.

Response: According to the suggestion, we further enhanced the literature review on the drivers of vegetation greenness in Southeast Asia. It is unfortunate that few studies specifically focused on this region (e.g., Satriawan et al., 2024), and we gained most of our knowledge on Southeast Asia from global scale studies (e.g., Zhu et al., 2016; Piao et al., 2019; Chen et al., 2019; Chen et al., 2022). Specifically, these global studies reveal that CO₂ fertilization is a primary driver of the greening trend globally, including in Southeast Asia (Zhu et al., 2016, Chen et al., 2022). Climate change, especially temperature rise, could reduce vegetation growth in the tropics (Piao et al., 2019) or drive green-up in maritime Southeast Asia during El Niño (Satriawan et al., 2024). However, land-use change, especially deforestation, is the predominant factor driving the greenness decline in tropical countries like Indonesia (Piao et al., 2019; Chen et al., 2019).

We have included the corresponding references in our manuscript and the context is added in Line 35 to Line 45, "Southeast Asia harbours diverse biodiversity and ecosystems. Yet, the trends and drivers of regional greenness remain largely underexplored. Previous studies reveal that CO₂ fertilization is a primary driver of the greening trend in Southeast Asia within a global context (Zhu et al., 2016, Chen et al., 2022). The impact of climate change on vegetation growth, however, remains uncertain (Piao et al., 2019, Satriawan et al., 2024), although some studies have reported that tropical temperature approaching critical thresholds may lead to leaf browning (Doughty et al., 2023). Land-use change, particularly deforestation, has

been found to be a predominant factor causing the decline of greenness in some tropical regions (Piao et al., 2019; Chen et al., 2019). However, these studies primarily focused on a global scale while the regional mechanisms (i.e., complexity in land use change) for greenness change were not fully examined”

Satriawan T W, Luo X, Tian J, et al (2024). Strong green-up of tropical Asia during the 2015/16 El Niño. *Geophysical Research Letters*, 51(8): e2023GL106955

Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., &Canadell, J. G., et al. (2016). Greening of the Earth and its drivers. *Nature Climate Change*, 6(8), 791-795. <http://doi.org/10.1038/nclimate004>

Piao, S., Wang, X., Park, T., Chen, C., Lian, X., &He, Y., et al. (2019). Characteristics, drivers and feedbacks of global greening. *Nature reviews. Earth & environment*, 1(1), 14-27. <http://doi.org/10.1038/s43017-019-0001-x>

Chen, C., Park, T., Wang, X., Piao, S., Xu, B., &Chaturvedi, R. K., et al. (2019). China and India lead in greening of the world through land-use management. *Nature Sustainability*, 2(2), 122-129. <http://doi.org/10.1038/s41893-019-0220-7>

Chen, C., Riley, W. J., Prentice, I. C., &Keenan, T. F. (2022). CO2 fertilization of terrestrial photosynthesis inferred from site to global scales. *Proceedings of the National Academy of Sciences*, 119(10). <http://doi.org/10.1073/pnas.2115627119>

Chen, J. M., &Black, T. A. (1992). Defining leaf area index for non-flat leaves. *Plant, cell and environment*, 15(4), 421-429. <http://doi.org/10.1111/j.1365-3040.1992.tb00992.x>

R2C3: The methodology section requires significant revision due to several critical issues: · The use of citations is improper, with several missing references that need to be included to support the study's claims and methodology.

Response: We double-checked the citations in the methodology section and added the necessary references. Specifically, we added references (i.e., Euler et al., 2016; Chen et al., 2024) to support our statements about land use change in Southeast Asia, and references (i.e., Fable, 2020; Sulla-Menashe and Friedl, 2018; Hurtt et al., 2020; Lundberg et al., 2018; Sitch et al., 2015) about the methodology regarding harmonization of different land use datasets, and the XGBoost-SHAP framework and scenario simulations (i.e., Lundberg et al., 2018; Sitch et al., 2015).

Euler M, Schwarze S, Siregar H, et al. Oil palm expansion among smallholder farmers in Sumatra, Indonesia[J]. *Journal of Agricultural Economics*, 2016, 67(3): 658-676.

Chen S, Woodcock C, Dong L, et al. Review of drivers of forest degradation and deforestation in Southeast Asia[J]. Remote Sensing Applications: Society and Environment, 2023: 101129.

Fable (2020). Pathways to Sustainable Land-Use and Food Systems. 2020 Report of the FABLE Consortium. International Institute for Applied Systems Analysis (IIASA) and Sustainable Development Solutions Network (SDSN), Laxenburg and Paris. 10.22022/ESM/12-2020.16896. Indonesia chapter.

Sulla-Menashe D, Friedl M A. User guide to collection 6 MODIS land cover (MCD12Q1 and MCD12C1) product[J]. Usgs: Reston, Va, Usa, 2018, 1: 18.

Hurtt G C, Chini L, Sahajpal R, et al. Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6[J]. Geoscientific Model Development, 2020, 13(11): 5425-5464.

Lundberg S M, Erion G G, Lee S I. Consistent individualized feature attribution for tree ensembles[J]. arXiv preprint arXiv:1802.03888, 2018.

Sitch S, Friedlingstein P, Gruber N, et al. Recent trends and drivers of regional sources and sinks of carbon dioxide[J]. Biogeosciences, 2015, 12(3): 653-679.

R2C4: · Details regarding the specific version of the dataset used and the preprocessing steps undertaken are absent, which is crucial for the reproducibility and integrity of the research.

Response: We apologize for the missing information about the version of the dataset we used. In the updated manuscript, we have included details about GLOBMAP LAI (v3), Global forest change maps (v1.11), and MODIS land cover product (v6.1). We have also made sure that version numbers for other land use datasets and climate datasets used in our study were provided.

We hope to clarify that the datasets we collected are all established products. The main processing we carried out is to harmonize them into a common grids. To better describe the process, we have included a figure and statements in SI (see below).

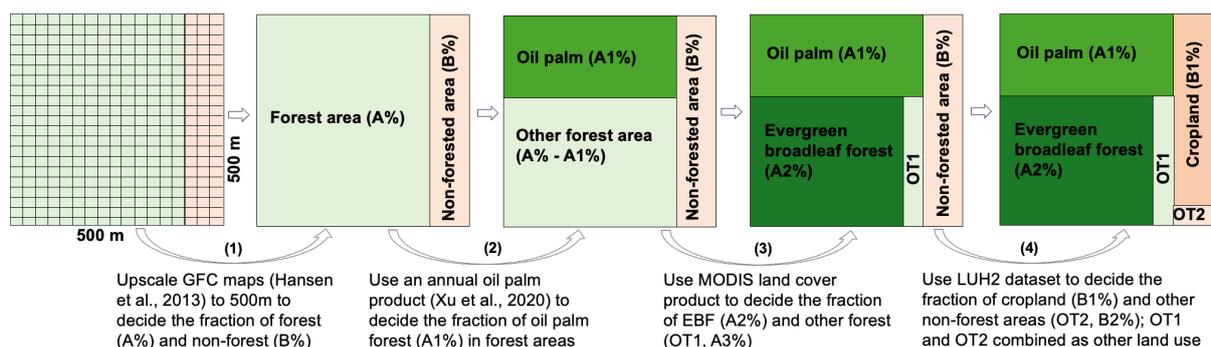


Figure. S1: A conceptual figure illustrating the processes of harmonizing land cover datasets with in a grid with a spatial resolution of 500 meter in this study. Step (1): upscale fine-resolution global forest change maps (GFC) to 500 meter to determine forest (A%) and non-forest (B%). Step (2) and step (3): calculate percentages of oil palm, evergreen broadleaf forest (EBF), and other forest types (OT1) within forest areas. Oil palm percentages are derived by upscaling a 100-meter resolution oil palm product to 500 meters. EBF and OT1 percentages are sourced from the MODIS dataset. Step (4): Determine cropland and other land uses (OT2) percentages using the LUH2 dataset, assuming LUH2 data at a 0.25° grid applies to 500-meter grid cells within each 0.25° grid cell. Note, the conceptual figure illustrates only the percentage of each land use, not their specific locations.

R2C5: · The explanation of how multiple land cover datasets were harmonized lacks clarity, making it difficult to understand the approach taken.

Response: To enhance clarity, we moved the flowchart that illustrates how multiple land cover datasets were harmonized from the supplementary file to the Methods section. Additionally, we added order numbers and corresponding statements from Line 100 to Line 115 in this section for further clarification: “As shown in Fig. 1, the workflow for harmonizing multiple land cover datasets involved the following steps:

(1) We first determined the annual percentage of forested (A%) and non-forested areas (B%) within each 500m grid cell by aggregating the mean of the annual 30 m resolution Global Forest Change v1.11 (GFC) maps (Hansen et al., 2013), using the ‘reduceResolution()’ function in Google Earth Engine (<https://developers.google.com/earth-engine/guides/resample>).

(2) Within the forested fraction of each grid cell, we estimated the proportion of oil palm (OP) plantations (A1%) using an openly available dataset that covers OP distribution from 2001 to 2016 across Malaysia and Indonesia (Xu et al., 2020). To estimate the proportion of OP, we calculated the frequency of oil palm pixels in each 500 m × 500 m window.

(3) After accounting for the area of OP, the remaining forested area in each grid was further categorized into the evergreen broadleaf forest (EBF) (A2%) and other forest types (A3%) (i.e., deciduous broadleaf forest, coniferous forest, mixed forest, etc.), based on the ratio of EBF to the total forested area provided by MODIS Land Cover Type Product (MCD12Q1 v6.1) (Sulla-Menashe and Friedl, 2018).

(4) For the non-forested fraction of each grid cell, we used the latest version of the Land-use harmonization datasets (LUH2) dataset (Hurt et al., 2020) to estimate the percentage of cropland (CRO) (B1%) and other non-forest land uses (B2%) (i.e., pasture, grass, etc.). ”

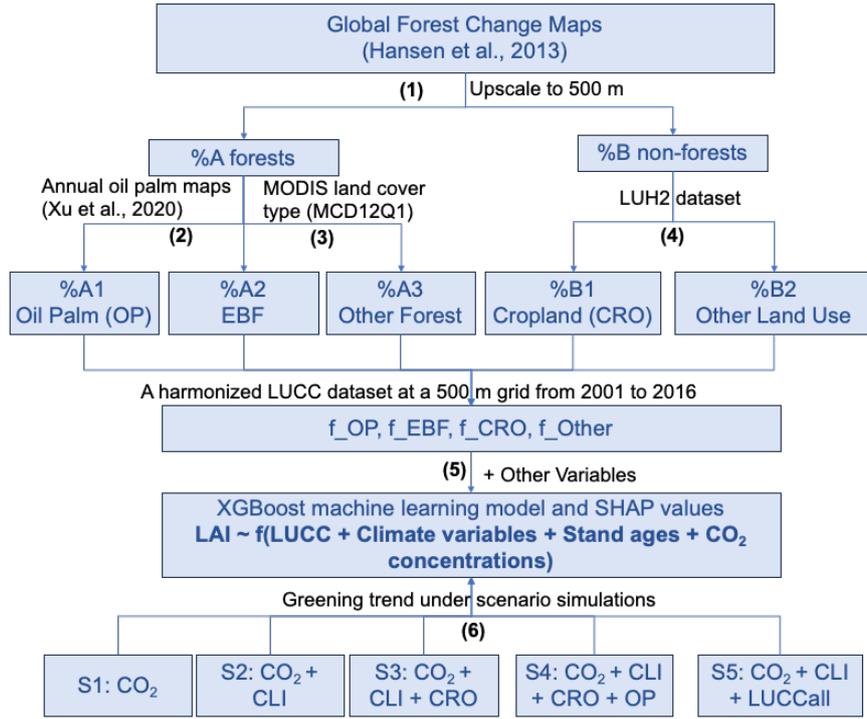


Figure 1: Workflow of the study. Steps (1) to (4) outline the processes for harmonizing multiple land cover datasets. Steps (5) to (6) show the establishment and interpretation of the LAI prediction machine learning model and the process of scenario simulations.

R2C6: · Descriptions of scenario simulations are unclear. When introducing scenario simulation schemes, it is imperative to explicitly detail the calculation methods for assessing the impact of each factor, which would greatly enhance the manuscript's credibility and reliability.

Response: We apologize for the confusion. In the updated version, we explicitly included equations detailing the calculation process and improved the clarity of our statements for this section as follows:

“To quantify and compare the impacts of specific LUCC processes, climate change, and elevated CO₂ concentrations on vegetation greenness changes, we adopted the scenario simulation framework from several factorial attribution analyses (Sitch et al., 2015). Specifically, we first estimated the LAI trend under five hypothetical scenarios (S1 to S5) using the established XGBoost model. The equations are as below,

$$LAI_{i,t,S_n} = XGBoost(CO_{2,i,t,S_n}, CLI_{i,t,S_n}, f_EBF_{i,t,S_n}, f_CRO_{i,t,S_n}, f_OP_{i,t,S_n}, f_Other_{i,t,S_n}) \quad (3)$$

$$\beta LAI_{i,t,S_n} = slope(LAI_{i,t,S_n}) \quad (4)$$

Where, LAI_{i,t,S_n} represents the simulated LAI for the i^{th} grid at year of t under scenario S_n and $\beta LAI_{i,t,S_n}$ indicates the LAI trend. The XGBoost stands for the established model for LAI prediction using CO₂ concentration, climate variable (CLI),

and land cover types such as fraction of evergreen broadleaf forest (f_EBF), cropland (f_CRO), oil palm (f_OP), other land uses (f_Other) (see Method 2.4).

For different scenarios, we adjusted the input variables according to specific assumptions to progressively incorporated different factors. For S1, we assumed only CO₂ concentration varies from 2001 to 2016, while climate and land uses variables (i.e., CLI, f_EBF, f_CRO, f_OP and f_Other) remained constant at their values in 2001. For S2, CO₂ and climate change over time, with land uses remaining unchanged since 2001. S3 to S5 sequentially considered different land use processes. S3 involved changes from EBF to CRO using time-varying CO₂, climate, and CRO area, while keeping OP and other land use types constant post-2001; S4 included conversions from EBF to both CRO and OP using time-varying CO₂, climate, CRO and OP areas, while other land uses unchanged since 2001; S5 encompassed all LUCC changes, with all variables including CO₂, climate, and all types of LUCC varying over time.

We then quantified the impacts of each factor on vegetation greening based on differences in LAI trends between scenarios,

$$Driver_n = \delta LAI trend = \beta LAI_{i,t,S_n} - \beta LAI_{i,t,S_{n-1}} \quad (5)$$

Here, $Driver_n$ measures the impact of the nth driver (ranging from CO₂, climate change, CRO expansion, OP expansion, to Other LUCC) on LAI trends. Notably, $Driver_1$ quantifies the impact of CO₂, equal to $\beta LAI_{i,t,S_1}$.

Minor comments

R2C7: 1. The discussion mentions, "It is also important to note that our estimation of CRO or OP expansion was based on the assumption that the increased areas of CRO or OP since 2001 came from EBF." Such a crucial assumption should be stated in the methodology section.

Response: Thank you for the suggestion. We have moved this statement to the methodology section Line 220 to Line 222.

R2C8: 2. On the basis of Figure 5, it would be beneficial to add the spatial distribution of dominant factors for each pixel. This enhancement would more clearly reveal whether the LAI trend for each pixel is positive or negative and which factors primarily drive these changes.

Response: Following the suggestion, we included the spatial distribution of dominant factors for each pixel by comparing the impacts of factors on the LAI trend (Fig. S6). Consistent with Figure 5, we found that the effect of CO₂ fertilization

dominated the increase in LAI in most areas, accounting for 62.10% of the study area. Conversely, CRO expansion was a dominant driver for greenness decline in many regions, accounting for 26.33% of the study area.

We have added the figure below in the supplementary file and included the following statement in the Result section, from Line xx to Line xx: “From a spatial perspective, we found that elevated CO₂ dominated the increase in LAI in most areas, accounting for 62.10% of the study area, while CRO expansion was the primary driver in LAI decrease in other regions (26.33%), especially coastal areas (Fig. S6)”.

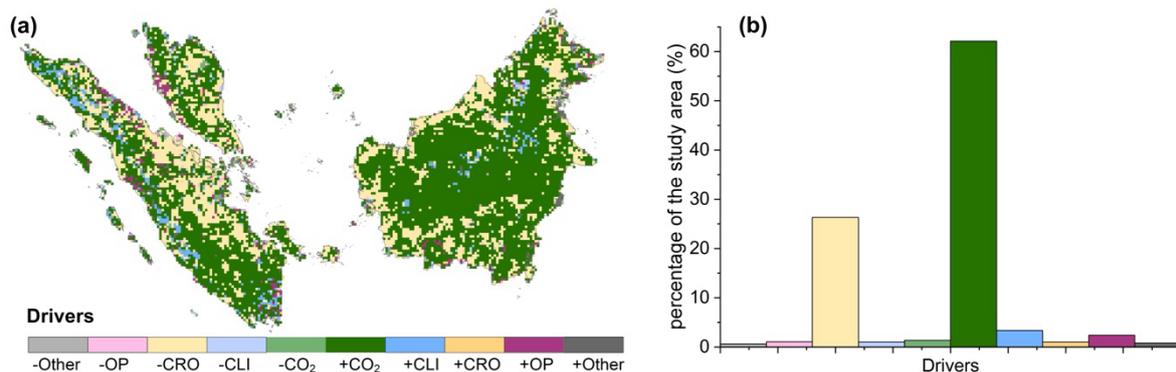


Figure S6: Spatial pattern of dominant drivers of trend in LAI (a), and the percentage of the study area dominated by each driver (b). The drivers include elevated CO₂ (CO₂), climate change (CLI), crop expansion (CRO), oil palm expansion (OP) and other land use changes (Other). A prefix ‘+’ of the drivers indicates a positive impact on LAI trends, whereas ‘-’ indicates a negative impact.

R2C9: 3. It would be preferable to represent Figure S2 as a scatter density plot (like Figure 4c,d) to facilitate the observation of changes in SHAP values with features, and to prevent potential misinterpretation arising from the clustering of scatter points.

Response: We agree that representing Figure S2 as a density plot will help avoid misinterpretation. Accordingly, we have revised this figure in the updated manuscript, as shown below.

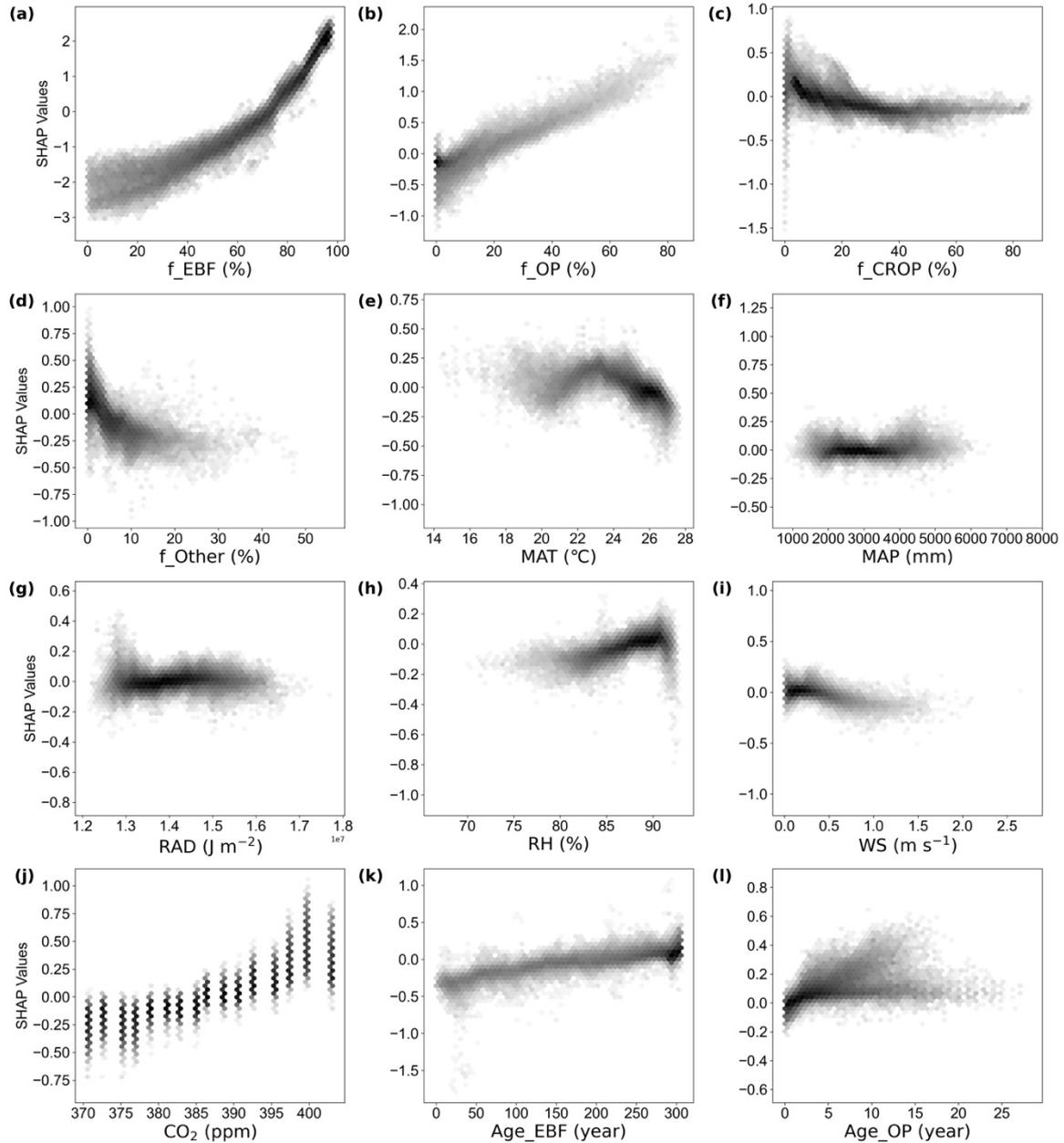


Figure S2: The density plots show the changes in SHAP values of each factor on LAI with corresponding factor variations. The abbreviations for each factor are available in Table S3.

R2C10: 4. There is an error in Equation (2) that needs to be corrected.

Response: Thanks for pointing this out. We have corrected the Equation (2).

$$\phi_{i,j}(x) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|! (|N| - |S| - 2)!}{|N|!} [f(S \cup \{i,j\}) - f(S \cup \{i\}) - f(S \cup \{j\}) + f(S)] \quad (2)$$

R2C11: 5. Figure 4c,d depicts the coupling effects of f_EBF with f_OP and f_CRO rather than the interaction effects mentioned in the caption, making it seem

indistinguishable from Figure S2a. It is recommended to add SHAP dependence plots illustrating the interaction effects for a more in-depth analysis.

Response: Following the suggestion from the reviewer, we added SHAP interaction plots (Fig. 4e-f) on top of the SHAP dependence (coupling) plots (Fig. 4c-d) in the main text, to distinguish the figure from Fig. S2. We have also revised Fig. S2 to highlight the difference. We ensure that the caption provided for figure 4 is correct to avoid confusions. However, we refrained from overinterpreting the interaction plots (such as the interactions between f_EBF , f_OP and f_CRO), as by nature, we suspect that these three factors are likely dependent on each other (e.g., less f_EBF , more f_OP), not necessarily interact with each other (i.e., independent variables) in meaningful way.

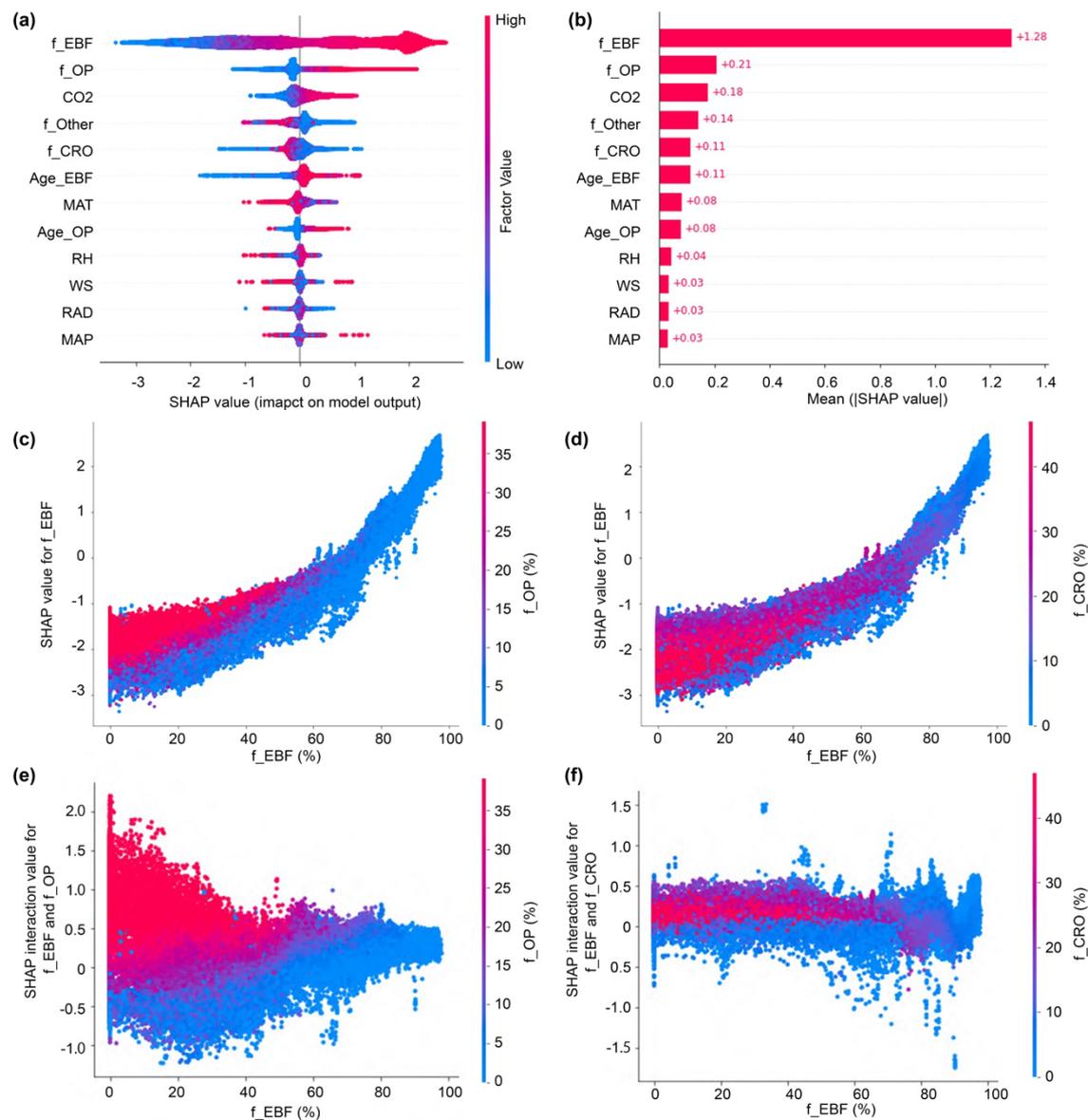


Figure 5: The impact of factors on LAI. (a) Bee swarm plots show the SHAP values of each factor on LAI for each sample. The SHAP value indicates the magnitude and direction of the impact on LAI (see Methods). Each dot represents an individual sample, with the color indicating the relative values of the specific factor. (b) The bar plot of the mean absolute SHAP

values of each factor for LAI. (c) The interaction of f_OP and f_EBF, and (d) the interaction of f_CRO and f_EBF on LAI. The abbreviations for each factor are available in Table S3.

R2C12: 6. Previous studies have highlighted discrepancies between the cropland area changes provided by LUH2 and actual conditions in China and the United States. It is worth investigating whether a similar discrepancy exists in Southeast Asia. Meanwhile, the spatial resolution of the LUH2 dataset is too coarse for the purposes of this study.

Response: Regarding the accuracy of LUH2 in Southeast Asia, Mao et al. (2023) conducted a comparative analysis of LUH2 and eight other land-use products against a constructed land-use product for Southeast Asia. They compared several datasets, including remote sensing datasets like the MODIS Land Cover dataset (MCD), ESA CCI land cover maps (CCI), GLC_FCS30 (GLC), Copernicus Global Land Service Land Cover product (CGLS), and GlobeLand30 (GL), along with datasets from FAO, HYDE, and SAGE (Mao et al., 2023). Their analysis found that the cropland area estimates for this region were most closely aligned with those from LUH2, with a correlation coefficient (r) of 0.98 (Mao et al., 2023; Figure R1). This consistency indicates that LUH2 provides reliable cropland data for Southeast Asia.

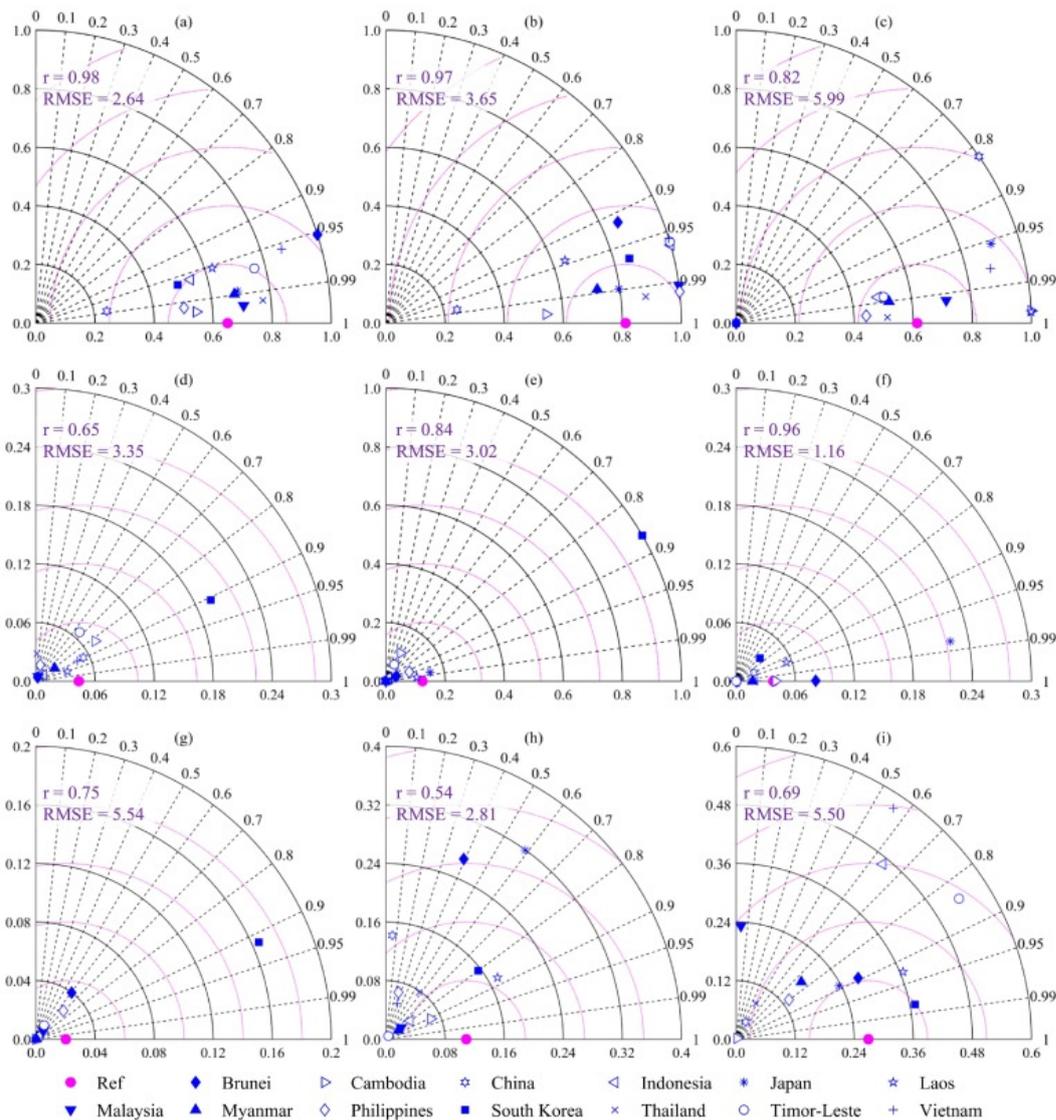


Figure. R1. Taylor diagrams comparing cropland area estimates with (a) LUH2, (b) HYDE, (c) SAGE (d) MCD, (e) GL, (f) CCI, (g) CGLS, (h) GLC, and (i) FAO data for subtropical East Asia and Southeast Asia. (Source: Mao et al., 2023)

Mao F, Li X, Zhou G, et al. Land use and cover in subtropical East Asia and Southeast Asia from 1700 to 2018[J]. *Global and Planetary Change*, 2023, 226: 104157.

To examine the impact of spatial resolution on our results, we conducted a parallel analysis using spatial resolution of both 500m and 0.25 degree. We found our findings were robust to variations in spatial resolution, though the resolution of LUH2 is coarse. Specifically, we found the land use change and their impacts on the greening trends in our study area remain consistent across both 0.25-degree (Fig. S2, Fig. S6) and 500-meter (Fig. 2, Fig. 6) grid resolutions.

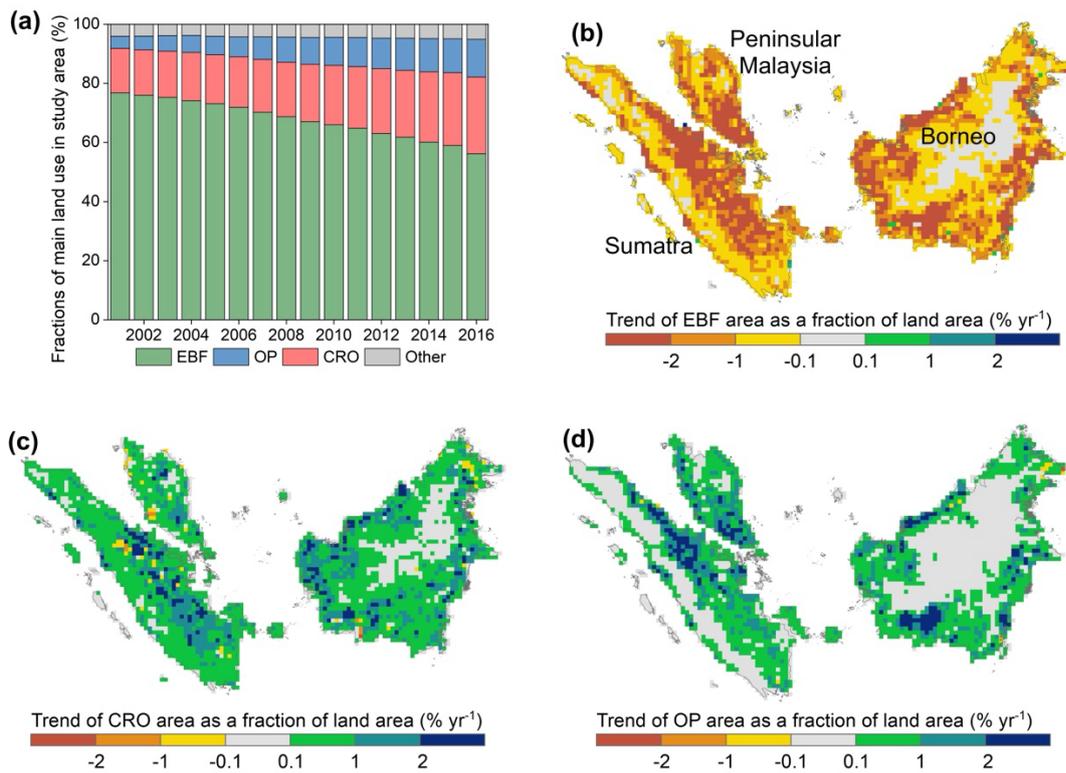


Figure S2: Land use composition and its changes from 2001 to 2016 in the study area, analyzed at a 0.25-degree resolution.

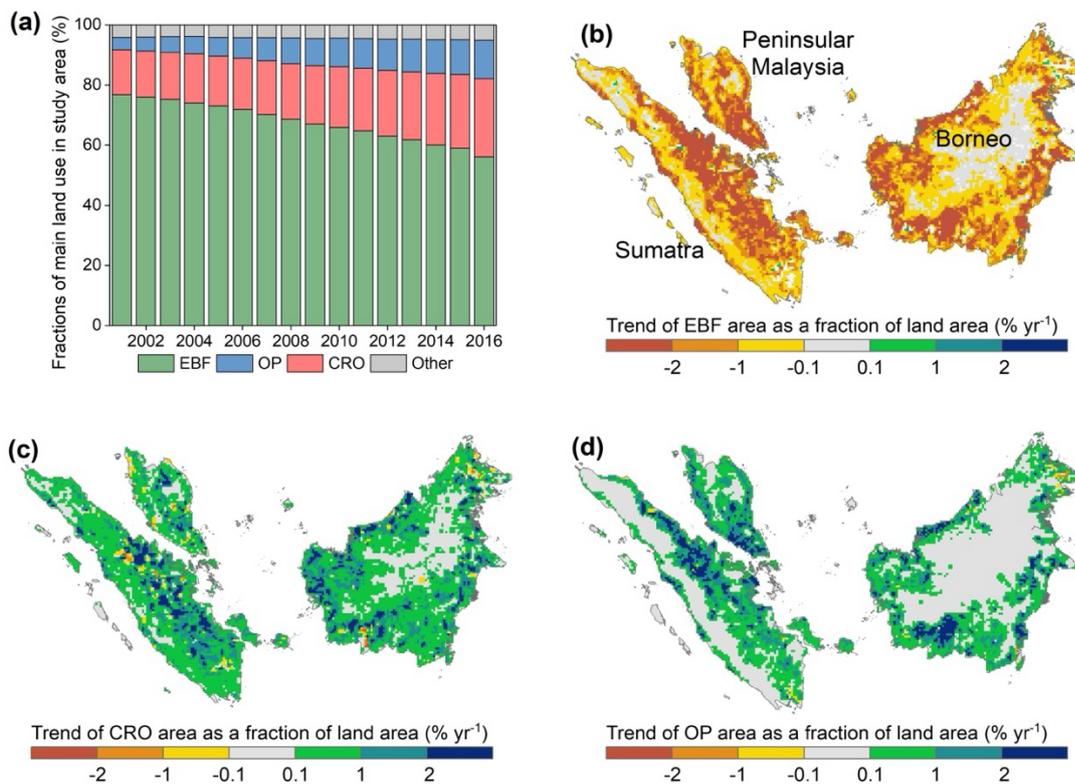


Figure 2: Land use composition and its changes from 2001 to 2016 in the study area, analyzed at a 500-m resolution.

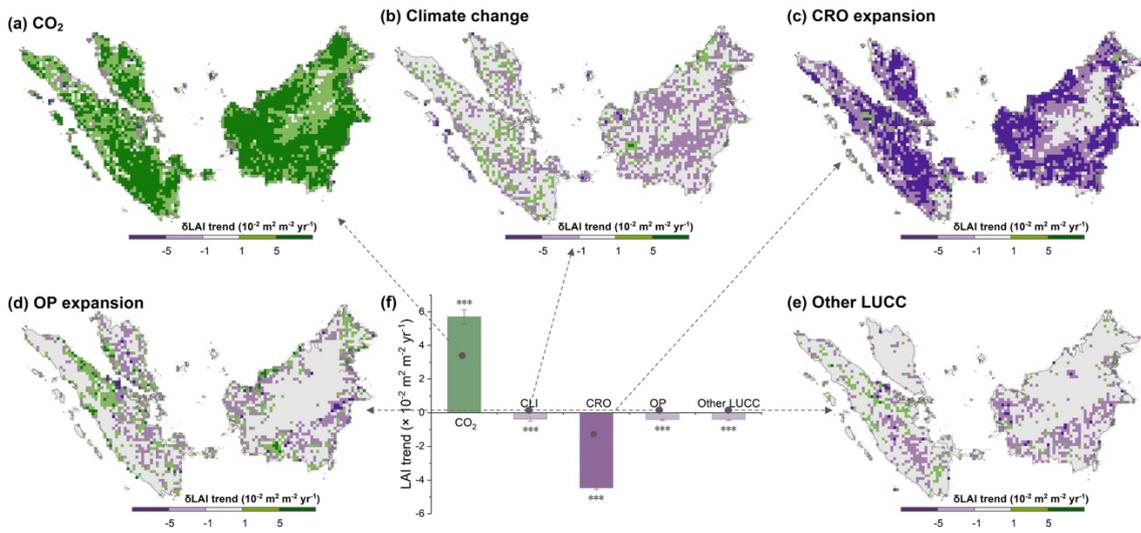


Figure S6: The spatial distribution of the pixel-wise impacts of each process on the greening trends, analyzed at a 0.25-degree resolution.

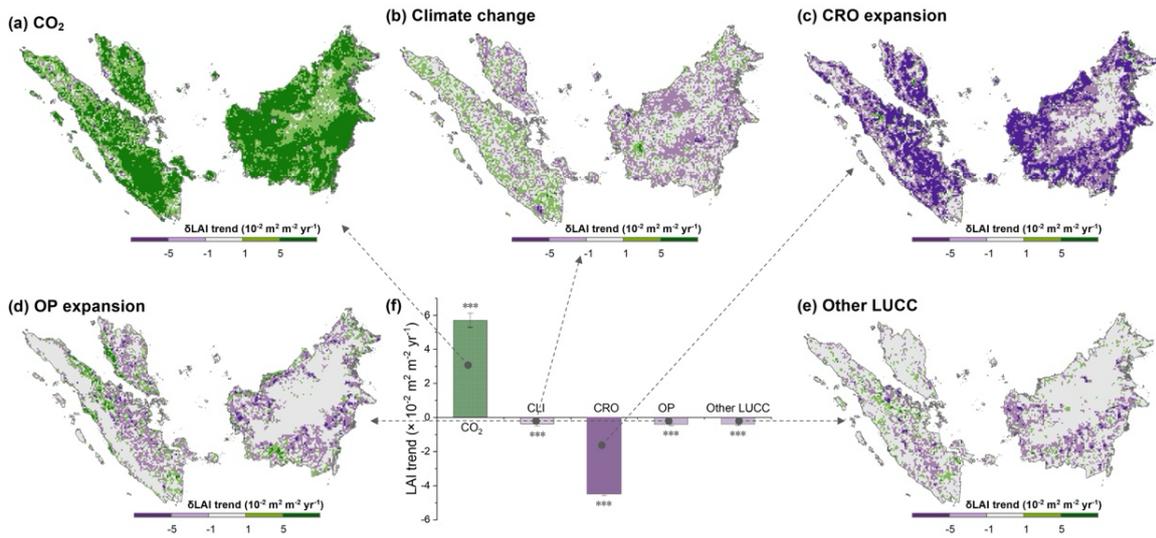


Figure 6: The spatial distribution of the pixel-wise impacts of each process on the greening trends, analyzed at 500-m resolution.