https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

GEM-Forest: A Global satellite EMbedding–based map of forests and tree crops for 2020

Daniel Paluba[1], Valerio Marsocci[2], Katarína Onačillová[3], Yarin T. Puerta Quintana[4], Adam Hastie[4,5]

[1]Department of Applied Geoinformatics and Cartography, Faculty of Science, Charles University, Prague, 128 00, Czechia
[2]Φ-lab, European Space Agency (ESA/ESRIN), Frascati, 000 44, Italy
[3]Institute of Geography, Faculty of Science, Pavol Jozef Šafárik University in Košice, Košice, 040 01, Slovakia
[4]Department of Physical Geography and Geoecology, Faculty of Science, Charles University, Prague, 128 00, Czechia
[5]Department of Botany, Faculty of Science, Charles University, Prague, 128 00, Czechia

Correspondence to: Daniel Paluba (daniel.paluba@natur.cuni.cz)

Abstract. The advent of big data in Earth observation, coupled with recent advances in Artificial Intelligence, has led to the development of geospatial embeddings. These compact, information-rich feature vectors are designed for direct use in machine learning (ML) applications across a wide range of downstream tasks, such as forest monitoring. Motivated by the limitations of existing global forest products and policy requirements like the EU Deforestation Regulation (EUDR), we (1) evaluate whether lightweight classifiers applied to satellite embeddings from the Google DeepMind Alpha Earth Foundation (AEF) can accurately map global forest and tree crop extents, and (2) we introduce the resulting GEM-Forest dataset. GEM-Forest is a global satellite embedding–based dataset at 10 m spatial resolution for 2020 that provides a consistent classification across three classes: forest, non-forest, and tree crops. Our comparison of multiple ML approaches ranging from linear models to neural networks showed similar performance across classifiers, while linear models often outperformed more complex models. This consistency indicates that the embeddings encode highly informative and linearly separable structures for global forest discrimination, which includes tree crop separation. Based on these findings, a linear Support Vector Machine was used to generate the final GEM-Forest dataset, which outperformed eight existing global forest, tree cover, or land cover maps on two global validation datasets, while it placed second on the JRC's global forest validation dataset. Across all three datasets, the forest class achieved omission errors of 12–18% and commission errors of 16–21%, with overall accuracies from 88% to 92%. Misclassifications of tree crops as forests varied between 0.5% and 14.8%, with a producer's accuracy above 85% for most tree crop datasets, whereas the classification of European tree crops remains the most challenging. Globally, GEM-Forest maps 3,919 million hectares (Mha) of forest for 2020, representing a 5.9% underestimation relative to FAO reports. This variance is partly attributed to the exclusion of unstocked forest areas from our forest definition, discrepancies in country-based forest definitions, and misclassification errors that occurred primarily within open forests and forest–shrubland transition zones. Overall, these results demonstrate that satellite embeddings combined with simple ML approaches support highly accurate, computationally efficient global forest and tree crop mapping. The open-access release of the GEM-Forest dataset and its ML model weights (both available in Paluba et al., 2026; DOI: https://doi.org/10.5281/zenodo.18921586) can support international policy decisions and allows direct and straightforward temporal transferability for other years.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

1 Introduction

Forests, which cover around one-third of the Earth's land surface, are one of the world's most important ecosystems, playing a crucial role in climate regulation, the global carbon cycle and biodiversity conservation (Cook-Patton et al., 2020; Hughes et al., 2021; Lewis et al., 2019; Raven et al., 2020; Smith et al., 2023). However, forests are subject to continuous change, influenced by both human activities and natural processes. These influences lead to both forest loss, resulting from factors such as deforestation, climate change and natural disturbances, and forest gain, achieved through effective forest management and restoration efforts (Bartels et al., 2016; Forzieri et al., 2022; Ma et al., 2023). Notably, from 2011 to 2020, forests acted as a net carbon sink, absorbing more carbon than they released, which is largely attributed to recent afforestation activities and improved forest management practices (FAO, 2022). The advent of big data in Earth observation (EO) and cloud-based computational technologies enables near-real-time monitoring of status and dynamics of forests. With new policies, such as the European Union Deforestation Regulation (EUDR), the EU's new 90% greenhouse gas (GHG) emission reduction target (which includes the contribution of high-quality international carbon credits), the EU forest strategy for 2030 of the European Green Deal, and other national and regional policies, new challenges and opportunities arise in the era of big EO data.

Numerous national to global EO-based forest maps often serve as baseline layers in real-world applications, such as early-warning systems for deforestation monitoring and in policy- and decision-making on a national to global level. However, their accuracy and temporal availability vary across regions and climatic zones. Accurate and up-to-date forest cover information is essential for real-world applications. Recent years have been marked by attempts to monitor the state of global forests, especially aiming to support the EUDR. The EUDR aims to prevent the import of forest-related goods into the EU linked to deforestation or forest degradation after 31 December 2020 (Regulation (EU) 2023/1115). Therefore, recent and ongoing efforts are aimed at mapping the extent of global forests in 2020. Based on a recent 2025 review of 21 global forest/tree cover products, only two met the EUDR forest definition criteria. However, each evaluated dataset, including these two, failed to accurately distinguish between forests and other tree-based systems, e.g. agroforestry systems, generally resulting in overestimation of forest cover (Freitas Beyer et al., 2025). One of the two is the European Commission's Joint Research Centre's (JRC) Global Forest Cover map for 2020 at 10m spatial resolution (GFC2020) (Bourgoin et al., 2026a), which is specifically aimed at supporting the EUDR. GFC2020 was produced by combining multiple global forest-related, land cover, and ancillary datasets to map the maximum global forest extent for 2020. Its accuracy highly depends on the quality of the input datasets, and its approach is limited in transferability to other years, as this would require up-to-date data for those years.

Many other global forest maps/products have been developed over the recent years. One of the most frequently used datasets is the Global Forest Change (GFC) dataset (Hansen et al., 2013) that informs about yearly forest losses and tree cover gain from 2000 onward and is based on classification of multi-temporal Landsat data at a 30 m spatial resolution. However, this and other similar global forest products could be described as tree cover datasets, as they are often criticized for classifying tree crops as forest cover (Freitas Beyer et al., 2025; Tropek et al., 2014). For differentiating between forest and non-forest classes, general global land use /

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

land cover datasets can also be utilized, such as the ESA World Cover product that shows superior performance compared to other 10m global land cover products with high accuracy for the forest class (Xu et al., 2024).

Some of the datasets aimed to map the extent of natural forests, nevertheless, the final forest class does not include planted forests, such as the Natural Forests of the World 2020 by Neumann et al. (2025), or the planted forest category is merged with plantations used for agroforestry, e.g. in Global Forest Types dataset by Bourgoin et al. (2024 and 2025), or the layer is available only for 2015 at a limited 100 m spatial resolution in the Global forest management map by Lesiv et al. (2022). However, planted forests used for forestry purposes should be defined as forests based on the FAO and EUDR definitions. Other datasets have successfully mapped the extent of both natural and planted forests, such as the Global Natural and Planted Forests dataset by Xiao et al. (2024).

Recent advances in Artificial Intelligence (AI), combined with the growing availability of large-scale EO data, have led to the emergence of EO foundation models and their derived embedding products. These models are trained on massive, heterogeneous EO archives to learn general-purpose representations of the Earth's surface, capturing spatial, spectral, and contextual patterns across biomes, sensors, and acquisition conditions. The introduction of EO embeddings represents a significant step in this direction, providing globally consistent, high-dimensional feature representations that are designed to be transferable across tasks, regions, and time periods. Rather than optimizing for a single downstream application, embeddings aim to encode semantically meaningful information that can be reused across a wide range of EO analyses.

EO embeddings can be understood as ready-to-use feature representations that abstract raw satellite observations into a compact and information-rich space, enabling efficient application to downstream tasks such as clustering, classification, similarity search, and change detection. They can be provided both at pixel-level, such as in the case of Google DeepMind's Alpha Earth Foundation (AEF) (Brown et al., 2025) and TESSERA (Feng et al., 2025), and at patch-level, as intermediate features of geospatial foundation models (e.g., TerraMind (Jakubik et al., 2025) and THOR (Forgaard et al., 2026). This new paradigm substantially lowers the barrier for large-scale EO analytics, as it allows simple and computationally efficient machine learning (ML) models to achieve competitive performance without extensive task-specific feature engineering or model tuning.

In the context of forest monitoring, such embeddings are particularly valuable, as they offer the potential to consistently distinguish forests from structurally similar land-cover types, including tree crops, across diverse ecological and geographic settings. Recent studies have already demonstrated the practical utility of such embeddings for regional tree species monitoring in the United States (Gao et al., 2025) and in Italy (Ball et al., 2026), where EO embeddings outperformed composite-based multitemporal and multimodal EO baselines. As a result, EO foundation model embeddings like AEF provide a promising starting point for scalable, multi-temporal, and globally transferable forest and land-cover mapping workflows.

In this study, we introduce GEM-Forest, a global satellite embedding–based forest and tree crop dataset in 10 m spatial resolution for 2020. We introduce an effective automatic training data generation by intersecting multiple forest-related, land cover and tree crop datasets to reduce uncertainty and limitations of individual input datasets. We investigate whether lightweight ML methods can achieve competitive or superior performance in global forest/non-forest (F/nF) classification compared to more complex approaches, highlighting both the strengths and limitations of the embeddings and the classifiers themselves. Moreover, we compare

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

country- and continent-level forest cover extent based on GEM-Forest to FAO Global Forest Resources Assessments (FRA) statistics for 2020.

2 Data and methods

We used the AEF to develop a global F/nF classification at 10 m spatial resolution, with an additional tree crop class as a subclass of the non-forest class. Using AEF, we evaluated a set of lightweight supervised ML classifiers, including linear approaches (Logistic and Ridge Regression, linear Support Vector Machines - SVM), an instance-based approach (k-nearest neighbors - kNN) and compared their performance with more advanced non-linear classifiers, such as Random Forests (RF), Extreme Gradient Boosting (XGBoost) and neural networks (Multi-layer Perceptron - MLP). Training data are generated through an automated training process across 230 training areas that proportionally represent all global biomes. We assessed the accuracy of the final 2020 F/nF maps through an independent set of ~21,000 global F/nF reference samples for the year 2020, supplemented by eight open-access regional and global validation datasets for tropical tree crop plantations and a bespoke validation dataset for European tree crops. The final GEM-Forest dataset was produced by implementing model weights in GEE. By calculating the dot product between these weights and the AEF feature vectors, we identified the class with the highest probability for every global pixel. To meet the EUDR forest definition requirements, we mask out forests in urban areas and apply a majority filter to the forest class to mask out forest areas smaller than 0.5 ha. Each methodological step is detailed in the following paragraphs and visualised in Figure 1.

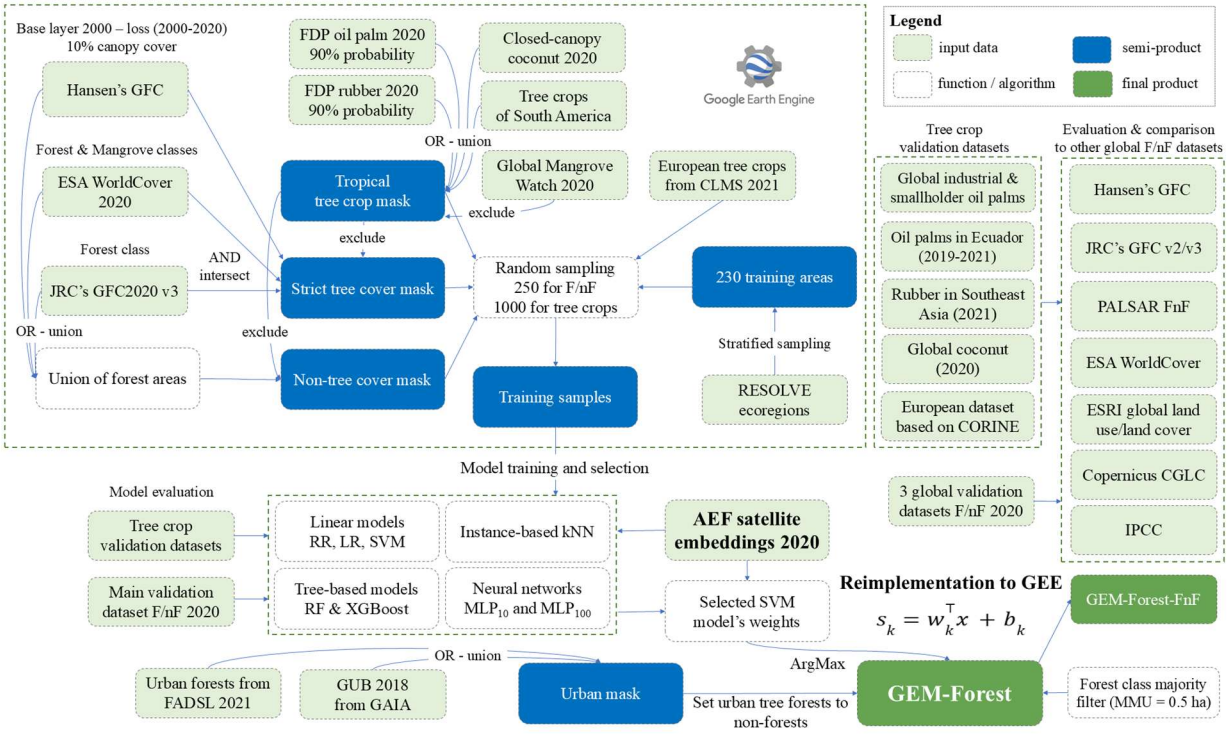

Figure 1. The full methodological pipeline used in this study.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

2.1 Forest definition

According to the FAO definition, forest is “Land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover of more than 10 percent, or trees able to reach these thresholds in situ. It does not include land that is predominantly under agricultural or urban land use.” (FAO, 2023, p.7). The EUDR adopts this forest definition, while the "other wooded land" is excluded from its forest definition. To align with the FAO/EUDR, we adopted these definitions, but with one main exception. We define forests as an area covered by trees with at least 10% canopy cover, with at least 5 m in height and representing an area of at least 0.5 ha, that is not under agricultural or urban land use in 2020. Notably, we chose to omit areas covered by young trees that have not yet reached the height and canopy coverage thresholds, and temporarily unstocked areas or clear-cuts, where forests are expected to meet the physical parameters in the future.

2.2 Input dataset - the Google DeepMind Alpha Earth Foundation’s Satellite Embedding dataset

The main data source is the AEF’s Satellite Embedding dataset, consisting of 64-dimensional feature vectors representing each 10 m pixel’s multi-modal, multi-temporal EO signal over a calendar year, with the current coverage from 2017 to 2025. These embeddings are produced by a deep geospatial representation learning model that fuses inputs from Landsat, Sentinel-2, and Sentinel-1 data, enhanced with temporal and spatial features (Brown et al., 2025). During training, the model learns to reconstruct target variables from other radar (from ALOS PALSAR-2), LiDAR (GEDI mission), environmental (topographic data from Copernicus DEM, meteorological data from ERA5-Land, gravity field information from GRACE), and annotated geolocated datasets (from Wikipedia and the Global Biodiversity Information Facility).

2.3 Preparing ‘strict’ tree cover, non-tree cover, tree crop and urban masks

In the first step, we prepare a strict tree cover mask for which we intersect three globally available forest- or tree cover-related and land cover datasets: 1) the ESA WorldCover’s (WorldCover) Forest and Mangrove classes in 10 m spatial resolution 2) JRC’s Global Forest Cover for 2020 (GFC2020) version 3 in 10 m spatial resolution and 3) the Hansen et al.’s GFC dataset in 30 m spatial resolution, with a minimum canopy cover threshold of 10%, where the forest loss layers from 2000 to 2020 were used to exclude non-forest areas from the base-layer from 2000. The intersection of these three datasets represents our ‘strict’ tree cover mask for 2020 and serves as a baseline for the subsequent training data generation.

To generate a tree crop mask, we used a union of global and semi-global maps of tree crops, specifically, 1) the 10m commodity probability layers for palm and rubber plantations, version 2025a, from the Forest Data Partnership (FDP) (Forest Data Partnership, 2025), specifically pixels that had higher than 90% probabilities of occurrence for the year 2020; 2) the 20 m global map of closed-canopy coconut palm (Descals, 2023; Descals et al., 2023), 3) a 10 m dataset for tree crops in the EU, i.e. trees predominantly used for agricultural practices (olive trees, fruit trees, nut trees) from the 2021 Forest Additional Support Layer (FADSL) from the High Resolution Layer Tree Cover and Forests product of the Copernicus Land Monitoring Service (CLMS) (European Environment Agency, 2024) and 4) the 10 m tree crop map of South America, version 2 from Jiang et al. (2026), specifically pixels that had

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

higher than 90% probabilities of occurrence for the year 2020. We have moreover used the Global Mangrove Watch data (Bunting et al., 2022b, a) to mask out (exclude) mangrove areas from the tree crop mask. Pixels corresponding to the tree crop mask were subsequently excluded from our strict tree cover mask.

The non-tree cover mask is defined as the complement to the union of all used forest/tree cover layers, that means everything except the union of three used forest/tree cover layers (GFC2020, GFC and WorldCover). Pixels corresponding to the tree crop mask were masked out (excluded) from the non-tree cover mask.

To align with the FAO forest definition regarding the exclusion of urban forests, we created an urban mask by combining the globally available Global Urban Boundaries for 2018 (GUB) from the global artificial impervious area (GAIA) data (Li et al., 2020) with urban forest layers from the FADSL (European Environment Agency, 2024) for the year 2021, which is available only for the EU. This urban mask was used in the post-classification step to remove urban areas and urban forests from the final classification, and these areas were reassigned to the non-forest class.

2.4 Training data preparation and ML model training in all biomes

To train a ML model, we first randomly generated 100x100 km areas across the globe, whose distributions were proportional to the extents of 14 biomes of the world based on the RESOLVE ecoregions (Dinerstein et al., 2017). This step included a generation of random points using a stratified sampling scheme within each biome with a subsequent application of a 100x100 km square buffer, while the overlapping areas were subsequently excluded from the final selection. Due to the random sampling approach, some areas became under-represented, especially areas with tree crops. Therefore, we added 21 representative 100x100 km areas representing large-scale palm, rubber and coconut plantations, and 13 areas representing European tree crops. This step reinforced the inclusion of tree crop areas in training data generation. In the end, a total of 230 training locations with a square 100x100 km area were selected.

Within these selected training locations, we generated random points based on the forest, non-forest and tree crop masks. The tree crop mask was divided into two masks: the tropical and the European tree crop masks. Based on the initial tests, further subcategorization of the crop mask did not yield better results qualitatively nor quantitatively. In each training location, 250 random points were generated for forest and non-forest classes, and 1000 points for the tree crops classes, each with a 10x10 m square buffer. Only points whose 10x10 m buffered area was entirely (100%) within a single land-cover class (forest, non-forest or tree crops) were retained for training to ensure the selection of highly representative training samples. In total, we obtained 47,308 training points globally (~14,500 forest, ~28,300 non-forest training and ~3,300 tropical and ~1,100 European tree crop points) that we used to train our ML models, specifically Logistic and Ridge Regression (LR and RR), linear Support Vector Machines (SVM), kNN, RF, XGBoost and neural networks, represented by MLPs with 10 and 100 neurons (MLP_{10} and MLP_{100}). We used the default hyperparameters for each ML model, based on the scikit-learn version 1.6.1 (scikit-learn.org, (Pedregosa et al., 2011) and the XGBoost version 3.2.0 (xgboost.readthedocs.io/en/release_3.2.0, (Chen and Guestrin, 2016) Python libraries. Building upon the workflow for automated training sample generation based on a strict tree cover mask previously implemented by Ončillová et al. (2023) for Europe, this study adapts and scales the logic for global application using globally available datasets.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

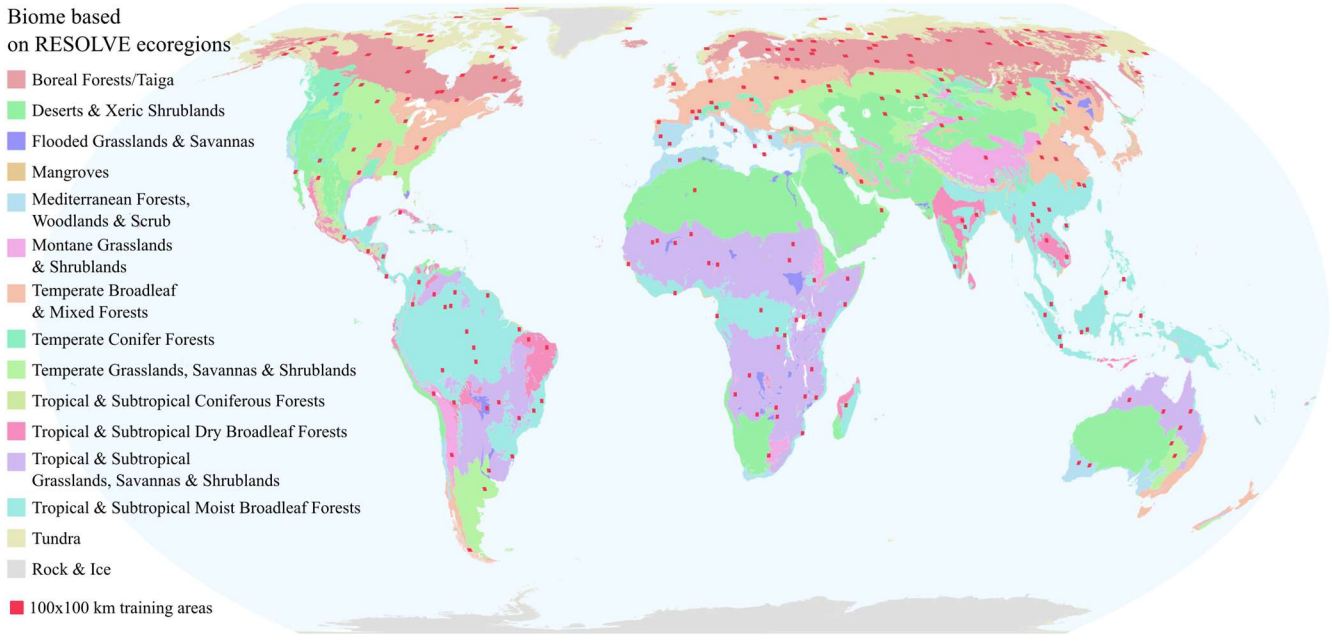

Figure 2. Distribution of 100x100 km training areas (red squares) over the world's biomes based on the RESOLVE ecoregions layer. Map projection: Equal Earth (EPSG: 8857)

In the initial stage, we classified forest, non-forest, and tropical and European tree crops, while the final datasets include three classes: forest, non-forest and tree-crops for 2020, where the tropical and European tree crops are merged into one tree crop class. Urban forests were masked out from both products based on the urban mask, and a 0.5 ha majority filter was applied to the forest class to align with the FAO and EUDR forest definitions. We compared our results to a model trained strictly on forest and non-forest samples (42,843 in total). This allowed us to test a two-class approach and determine if adding tree crop classes provides any extra value.

2.5 Accuracy assessment

2.5.1 Validation datasets for the evaluation of the binary F/nF version of the GEM-Forest

To evaluate the performance of the GEM-Forest product in discriminating between forest and non-forest areas, we reclassified the original three-class map into a binary F/nF scheme by merging the non-forest and tree crop classes. This binary F/nF map is further referred to as *GEM-Forest-FnF*. For this assessment, we used three global validation products.

The first one is the *Validation dataset for the global map of forest cover 2020, v.2* from JRC (Colditz et al., 2025) that was used to validate the GFC2020 product (Bourgoin et al., 2026a), further referred to as the *main validation dataset*. The main validation dataset included more than 21,000 samples covering the entire land area of the globe, except Antarctica. The categorical distribution of forest and non-forest area was 35% and 65% (~6,000 and ~15,000 samples), respectively, which represents the actual forest cover share over land globally. To replicate the validation process used in Bourgoin et al. (2026a), we excluded points from the original

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

dataset that had no assigned ground truth values strata information or were not overlapping with the FAO Global Administrative Unit Layer dataset (FAO, 2015), resulting in a total of 21,612 validation points. Based on an initial quality assessment using our urban mask and time series evaluation of high-resolution (HR) imagery from Google Earth, we identified 21 points in the main validation dataset that corresponded to urban forest areas but were labeled as forest. The ground truth labels of these points were subsequently changed to non-forest (class 0).

To enhance the robustness and generalization of our validation approach, we incorporated two additional globally available datasets. The second product is the *Stratified random sampling global land cover validation dataset* (SRS_Val) for 2020 from Zhao et al. (2023) containing 79,112 samples distributed across 16 classes, which include five distinct forest classes (Liu et al., 2023). The third is the *Land cover validation dataset over global mountains* (Mountain_Val) from Naboureh et al. (2025), a specialized dataset containing 55,758 samples across 7 classes restricted entirely to mountainous areas, which was derived and enhanced from the SRS_Val dataset (Naboureh et al., 2025b). These supplementary validation datasets were remapped to the binary tree cover and non-tree cover classes. This reclassification resulted in 26,711 tree cover and 51,780 non-tree cover samples for the SRS_Val dataset, and 18,657 tree cover and 37,101 non-tree cover samples for the Mountain_Val dataset. Because the Mountain_Val dataset was derived and enhanced directly from SRS_Val, a subset of approximately 20,000 sampling points remains identical across both datasets.

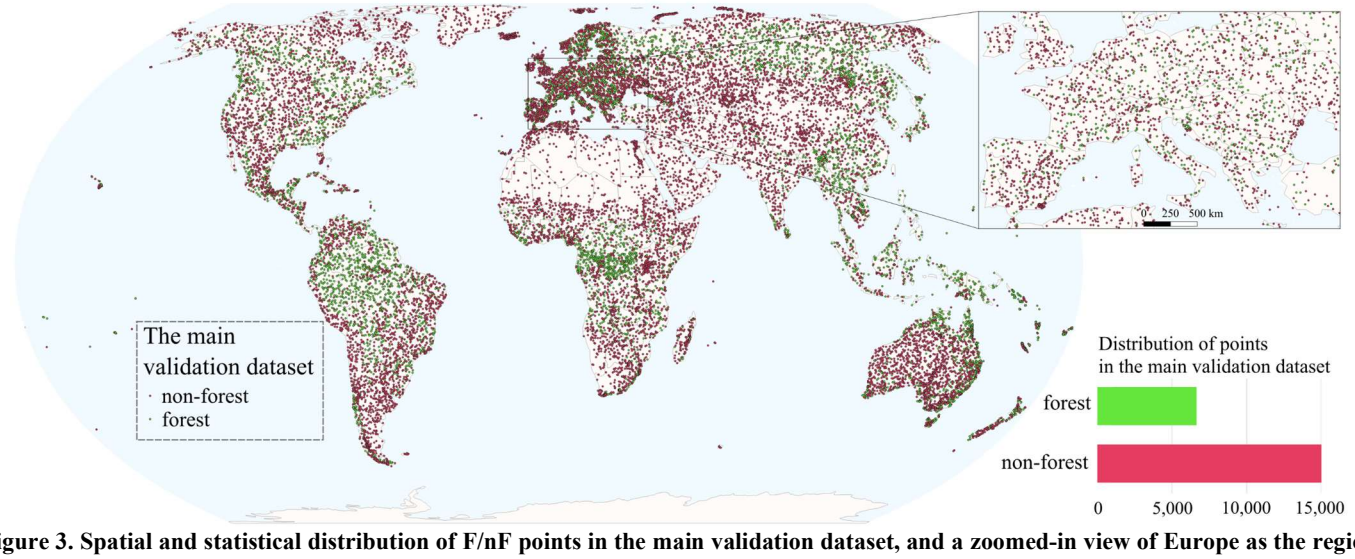

Figure 3. Spatial and statistical distribution of F/nF points in the main validation dataset, and a zoomed-in view of Europe as the region with the highest point density. Map projection: Equal Earth (EPSG: 8857)

2.5.2 Evaluation of misclassified areas in the GEM-Forest-FnF

To evaluate misclassified areas in the GEM-Forest-FnF, we derived a random subset of 433 samples from the main validation dataset that represented omission and commission errors (OEs and CEs) for the forest class. These OEs and CEs were independently visually interpreted for the presence or absence of forest cover for 2020 using high-resolution time series imagery in Google Earth Pro

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

(GEP). For ambiguous cases, we performed a visual interpretation using synthetic aperture radar (SAR) and optical time series using Sentinel-1 and Sentinel-2 data. This approach enabled the assessment of vegetation seasonality and temporal stability over a longer timeframe (since 2017 with both satellites) with high temporal resolution. This dataset thus allowed a two-way accuracy assessment and a quality check between the results of our binary F/nF classification and the main validation dataset.

In addition to standard accuracy metrics, all points were subjected to qualitative error attribution based on visual inspection and time-series analysis, with mismatches grouped into thematic categories such as sparse or scattered woody vegetation, forest edge or transition zones, shrub-tree confusion, land-use versus land-cover inconsistencies, riparian or wetland complexity, urban vegetation, and fragmentation effects. This error evaluation approach was used to explore and understand 1) true classification errors, 2) definition-driven mismatches related to forest structure and spatial continuity across different ecological and land-use contexts 3) and incorrectly assigned labels in the main validation dataset. This validation dataset is referred to as the error validation dataset.

2.5.3 Validation datasets for the evaluation of the tropical tree crop class

For the validation of the tree crop class of the GEM-Forest dataset in the tropics, we used global and regional open-source validation datasets for tropical plantations of oil palm, rubber, coconut and other palm species.

- For validation of oil palm plantations, we used three datasets: 1) 1,374 of globally distributed industrial oil palms for 2021 and 2) 531 globally distributed smallholder oil palms, both based on Descals (2024a) and Descals et al. (2024) *v1.2*, and 3) 164 points (dense and organized plantations, suggesting industrial oil palms) that were unchanged between 2019 and 2021 from Ecuador (Fundación EcoCiencia, 2025) *v1.0.*
- Rubber was validated using 217 monoculture rubber points in South-East Asia for 2021 (Descals, 2024b; Sheil et al., 2025) v1.1.
- Coconut areas were validated using 327 globally distributed points of dense open-canopy and closed-canopy coconut palm plantations for 2020 (Descals, 2023; Descals et al., 2023). These datasets were extended by palm species that are not coconut palm for 2020 from Descals (2023) and Descals et al. (2023).

2.5.4 Validation datasets for the evaluation of the European tree crops

For the validation of the tree crop class in Europe, we created a validation dataset based on the CORINE Land Cover 2018 (European Environment Agency, 2020) to identify stable tree crop areas across Europe. CORINE areas classified as olive plantations (class 223) and fruit tree and berry plantations (class 222) above a minimum area threshold of 25 ha were selected, with representative points generated either as centroids or as size-proportional random samples for very large polygons (for more details see the Appendix A). These selected points were further validated using a time series of high resolution imagery in GEP to ensure the presence of tree crops around the reference year 2020. The final dataset included 294 points for validation.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

2.5.5 Accuracy and error metrics used

The accuracy assessment in this study serves two purposes: first, selecting the optimal ML classifier using the main validation dataset, and second, benchmarking GEM-Forest against global datasets (see the section 2.6). Because the design, geographic scale, and thematic focus of the used validation datasets vary, we do not apply a single uniform metric. Instead, the evaluation is divided into area-weighted (specifically strata-based and area-proportional weighting) and unweighted (sample-based) approaches.

To achieve this, we calculated the unweighted macro F1-score (overall baseline average), area-weighted overall accuracy (OA), and the global area-weighted F1-score (w-F1). We also computed class-specific area-weighted OEs and CEs, which are the direct complements to producer's accuracy (recall or sensitivity in ML) and user's accuracy (precision in ML), respectively. Specifically, OE assesses the extent of true forest areas omitted from our classification, while CE quantifies the areas erroneously classified as forest that were not covered by forest in reality. Confidence intervals (95%) for OA, OEs, and CEs under the area-weighted frameworks are also provided, as well as full confusion matrices to allow users to calculate any alternative accuracy or error metrics.

Three distinct weighting frameworks were applied to evaluate map performance, determined by the design, additional data availability and spatial properties of each validation dataset:

1. Strata-based weighting: For the main validation dataset, accuracy and error metrics, specifically OA and w-F1, and class-specific area-weighted OEs and CEs were calculated following the stratification design described by Stehman (2014). Because the 149 sampling strata do not directly correspond to the final map classes, the explicit spatial areas of these 149 individual strata based on FAO The Global Administrative Unit Layers (GAUL), as provided in Colditz et al. (2025), were used as weights to ensure unbiased estimators. This approach was used both for evaluating the initial ML models and for comparing GEM-Forest with other global datasets. Accuracy assessments during the ML model selection phase were performed on the raw outputs, without applying a 0.5 ha majority filter to the forest class. Conversely, the final metrics of the GEM-Forest dataset used for comparisons against other global products were obtained after applying a 0.5 ha majority filter to the forest class to ensure structural and thematic consistency.
2. Area-proportional weighting: For the two additional global validation datasets (SRS_Val and Mountain_Val), the area-proportional estimation framework of Olofsson et al. (2013, 2014) was implemented. Map class weights (W_j) were derived from the FAO FRA 2025 report, where total forest area for 2020 was designated as 4,165 Mha (W_{forest} ~ 31.9%) and non-forest land area, calculated by subtracting global forest from total land area, was designated as 8,866 Mha ($W_{non\text{-}forest}$ ~ 68.1%). This approach was used only for comparing the GEM-Forest dataset, with applied 0.5 ha majority filter on the forest class, with other global datasets.
3. Sample-based weighting: For the validation of tree crop classes, standard unweighted, count-based sample metrics were used. Because the global total area proportions of these specific crop classes were unknown a priori, area-proportional weighting could not be mathematically applied. This is the reason why we used a pure sample-based evaluation approach. This approach was used both for evaluating the ML models and for comparing GEM-Forest with other global datasets. The tree crop class, treated as a subclass of the broader non-forest category, was validated using independent datasets containing

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

purely "true" tree crop validation points. Therefore, we evaluated the producer's accuracy (PA) of the tree crop class as the ratio of correctly identified tree crops to the total number of validation points in each tree crop validation dataset. Furthermore, we analysed inter-class confusion by quantifying instances where true tree crop points were misattributed to the 'forest' class. From the perspective of the forest product's reliability, these instances represent a commission error, where tree crops are erroneously included in the forest extent.

2.6 Comparison GEM-Forest-FnF dataset to other global F/nF and tree cover datasets

To compare GEM-Forest-FnF with other globally available F/nF and tree cover datasets, we compared our dataset to the following seven global datasets, detailed in Table 1: 1) GFC with a base layer from 2020, but with excluded areas deforested between 2000 and 2020 (Hansen et al., 2013), 2) the Global 4-class PALSAR-2/PALSAR Forest/Non-Forest Map (PALSAR FnF) for 2020 (Shimada et al., 2014), 3) Copernicus Global Land Cover Layers (CGLCL) for 2019 (Copernicus Land Monitoring Service, 2015), 4) Global 2020 Forest Classification for Intergovernmental Panel on Climate Change Aboveground Biomass Tier 1 Estimates (hereafter referred to as IPCC) (Hunka et al., 2024a, b), 5) ESA WorldCover for 2020 (Zanaga et al., 2022), 6) ESRI Global Land Use/Land Cover (LU/LC) (Karra et al., 2021) 7) GFC2020 v2 and 8) GFC2020 v3 (Bourgoin et al., 2026a). All datasets were accessed through the official GEE data catalog, except for the ESRI LU/LC dataset that was accessed through the Awesome GEE Community Catalog (Roy et al., 2025).

Table 1. Characteristics of global datasets used in the comparison

Dataset name, version	Year	Spatial resolution	FAO forest definition	Citation
GFC, v1.12	2000	30 m	no	Hansen et al. (2013)
PALSAR FnF, v2.0.0	2020	25 m	yes	Shimada et al. (2014)
CGLCL, v3	2019	100 m	no	Copernicus Land Monitoring Service (2015)
IPCC, v1	2020	30 m	yes	Hunka et al. (2024a, 2024b)
WorldCover, v100	2020	10 m	no	Zanaga et al. (2022)
ESRI global LU/LC	2020	10 m	no	Karra et al. (2021)
GFC2020, v2 & v3	2020	10 m	yes	Bourgoin et al. (2026a)

2.7 Computational effectiveness of data preparation, validation and data export process

To ensure computational efficiency and feasibility of the data preparation workflow, including the generation of forest, non-forest, and plantation masks; the preparation of training locations; the creation of training datasets; and the extraction of features from the

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

AEF for both training and validation points, we used the Google Earth Engine (GEE) platform (Gorelick et al., 2017), which hosts most required datasets described in the previous sections as well as the AEF. Some datasets were needed to be ingested to GEE. Subsequently, ML model training and validation were conducted in a local Python environment. As a final step, the learned weights and biases of the linear models were re-implemented within a GEE processing pipeline applied to the AEF to generate the final GEM-Forest dataset. These products were exported from GEE and archived in a publicly accessible repository in Zenodo. For data export, the study area was tiled into 5° × 5° spatial grids that enabled an efficient export of results at 10 m spatial resolution.

2.8 Model inference using linear models

An important advantage of linear models, particularly RR and SVM, is their straightforward integration into the proposed processing pipeline for large-scale F/nF and tree crop mapping, e.g. in GEE. In these cases, only the classifier weight vectors and intercepts for each class are required that allows the models to be implemented as simple linear equations applied to the embedding feature vectors. For each pixel, the class score s is computed as the dot product of the embedding feature vector x with the classifier weight vector w_k, plus the intercept b_k:

$$s_k = w_k^{\top} x + b_k \,, \tag{1}$$

Class decision scores are then computed for each pixel, and the final class is assigned by selecting the class with the highest value, consistent with the One-vs-Rest (OvR) formulation used during training. This approach allows efficient large-scale deployment, since the models require minimal computation at inference.

2.9 Country- and continent-level forest area comparisons with FAO FRA 2020

Country- and continent-level forest areas were derived from the binary F/nF version of the GEM-Forest dataset, using the FAO Global Administrative Unit Layers 2015 (FAO, 2015) as country boundaries. All calculations were performed in the native WGS84 geographic coordinate system. To correct for latitude-caused area distortion of WGS84, a row-wise cosine correction was applied, where each pixel row was assigned an area proportional to the cosine of its central latitude.

Country-level forest area estimates were subsequently compared to the FAO FRA 2020 (FAO, 2025) official statistics. Agreement between GEM-Forest-FnF and FAO FRA 2020 was assessed at both country and continent levels using the coefficient of determination (R^2) and the percentage difference (PD), calculated as:

$$PD = \frac{\text{GEM-Forest-FnF} - \text{FAO FRA}}{\text{FAO FRA}} \; x \; 100\% \,, \tag{2}$$

At the continental level, PD was computed from the aggregated total forest areas of all countries within each continent.

In addition to the raw mapped forest cover extent, we report the unbiased, sample-based forest cover estimate alongside its 95% confidence interval. For this area estimation, we followed the stratified framework of Stehman (2014) using the reference samples from the main validation dataset. By weighting the sample data against the total land area of the 149 GAUL administrative strata provided by Colditz et al. (2025), we derived the sample-based forest area for our final SVM-based GEM-Forest-FnF dataset.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

3. Results and discussion: technical validation

3.1 Summary of validation results for the GEM-Forest-FnF

To assess the generalization capacity and discriminative power of the AEF satellite embeddings on previously unseen data, we evaluated a range of traditional ML classifiers on an independent global validation dataset, on the main validation dataset. The embeddings were used directly as input features, without task-specific fine-tuning, allowing the evaluation to isolate the representational quality of the AEF feature space.

Our final GEM-Forest-FnF map is produced by merging the tropical and European tree crop classes with the non-forest class, which achieved high accuracies both on the main F/nF validation dataset and on the selected tree crop validation datasets. Across all evaluated models, OAs consistently achieved 91–92% ± 0.4%, with macro F1-scores between 89% and 91%, with a wider variation in forest OEs and CEs with 14–19% ± 0.8% and 11–14% ± 0.8–0.9%, respectively (Table 2). This stable performance across fundamentally different classifiers indicates that the AEF embeddings encode a highly informative and task-relevant structure for global F/nF discrimination. At the same time, the dataset remains challenging for certain region- and tree crop-specific classes and provides a meaningful benchmark for evaluating the limits of current EO embeddings, as summarized in the following section 3.2.

Table 2. Overall and class-based accuracy and error metrics for the tested algorithms. Bolded results represent the best achieved accuracies, while SVM (underscored) was selected for further evaluation.

Model	OA ± CI (%)	macro-F1 (%)	w-F1 (%)	OE F ± CI (%)	CE F ± CI (%)	OE nF ± CI (%)	CE nF ± CI (%)
kNN	90.89 ± 0.46	89.07	89.07	16.49 ± 1.06	14.28 ± 1.03	5.95 ± 0.44	6.98 ± 0.48
LR	91.82 ± 0.43	90.18	90.18	14.86 ± 1.02	12.76 ± 0.97	5.33 ± 0.42	6.29 ± 0.46
RR	91.48 ± 0.44	89.65	89.65	17.42 ± 1.07	**11.78 ± 0.97**	**4.72 ± 0.40**	7.25 ± 0.48
SVM	91.59 ± 0.44	89.89	89.89	15.60 ± 1.04	12.88 ± 0.98	5.34 ± 0.41	6.58 ± 0.47
RF	91.32 ± 0.44	89.46	89.46	17.76 ± 1.08	11.99 ± 0.98	4.79 ± 0.40	7.39 ± 0.48
XGBoost	91.79 ± 0.43	90.13	90.13	15.33 ± 1.02	12.47 ± 0.97	5.16 ± 0.41	6.47 ± 0.46
MLP_{10}	**92.06 ± 0.43**	**90.51**	**90.51**	**13.73 ± 0.99**	12.90 ± 0.97	5.46 ± 0.42	**5.85 ± 0.45**
MLP_{100}	91.81 ± 0.43	90.20	90.20	14.51 ± 1.01	13.04 ± 0.97	5.48 ± 0.42	6.16 ± 0.46

Note: w-F1 = weighted F1-score, F = forest, nF = non-forest, CI = 95% confidence interval range.

The comparison with the F/nF classification trained without the inclusion of tree crop training data showed comparable results on the main validation dataset. Although the overall and class-specific accuracies were promising, and in most cases better (OA of 91–92%, macro F1-score of 90–91%, with slightly lower forest OEs and higher forest CEs compared to GEM-Forest-FnF; see Appendix B), subsequent qualitative assessment in regions dominated by tropical plantations, together with statistical evaluation using tree crop validation datasets, revealed clear limitations, with forest CEs higher than 70% in most cases (Appendix C). Specifically,

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

training the models exclusively on forest and non-forest samples did not provide sufficient information to reliably classify tree crops as non-forest areas. These results highlight the limitations of using a global F/nF validation dataset alone, particularly for assessing product performance in identifying tree crop areas as non-forest. Even with ~21,000 validation points, the distribution of non-forest areas representing tree crops may be sparse, causing their misclassification to have minimal impact on overall metrics. For this reason, we employed additional tree crop validation datasets and included tree crop training data, classifying the tree crop subclass separately in the initial GEM-Forest-FnF classification, which was subsequently merged with the non-forest class.

3.2 Summary of validation results for the tree crop class

The evaluation of the tree crop class in the GEM-Forest dataset using regionally or globally available tree crop validation datasets shows variability across crop types and geographic contexts. The highest accuracies were achieved for the global industrial and smallholder oil palm validation dataset (PA: 87–96%; forest CEs: below 7%), with generally slightly higher PAs and lower forest CEs for industrial oil palms in most models (Figure 4 and Figure 5). This reflects that the large, homogeneous, closed-canopy structure of industrial plantations is easier to differentiate from forests using the AEF embedding space. Oil palms in Ecuador exhibited slightly lower OAs and higher forest CEs, which we attribute to reduced separability of semi-wild plantations, greater landscape heterogeneity, smaller plantation sizes and mixed management systems. These factors reduce the distinguishability of oil palms and tree crops in general from forests in the AEF feature space. Notably, smallholder oil palms also achieved very high accuracies in this evaluation, indicating that the method effectively captures both plantation types. These results are consistent with findings using EO data, specifically Sentinel-1 and Landsat time series, in Descals et al. (2024), who reported 91% class-specific PAs for industrial oil palms and 71% for smallholders, although in our study smallholder oil palms were classified comparably good compared to industrial plantations. Using the same oil palm validation dataset, Clinton et al. (2024) reported a PA of 82% for their global palm community model for 2020 at a palm probability threshold of 0.82 and higher. It should be noted that the comparison of our achieved PA is based only on positive samples (points classified as palm) and is not directly comparable to their overall PAs reported on both palm and non-palm samples.

Similarly, high PAs and low forest CEs were also achieved by rubber points in Southeast Asia (PA of 84–94% and forest CEs 5–11%, except for RF and RR models). Based on the rubber validation dataset used in Southeast Asia, the reported PAs for Wang et al. 's map (Wang et al., 2023) of rubber monoculture distribution showed substantially lower accuracy of 68.1 ± 3.9%, based on the validation in Sheil et al. (2025).

Coconut plantations show relatively high PAs (73–93%), but higher forest CEs (6–19%), depending on canopy closure, with closed-canopy coconuts outperforming open-canopy coconuts, consistent with results from the global coconut plantation layer (Descals et al., 2023). This highlights the importance of structural density in distinguishing forests from tree crops. Moderate accuracies with the highest forest CEs were achieved by the other palms dataset.

Even for the European tree crop validation dataset, where moderate PAs were achieved (66–45%, depending on the ML model), the proportion of forest CEs remains low, i.e., 10–16%. Consequently, a larger share of misclassifications occurred as non-forest (non-forest CEs). This is not problematic when the primary objective is F/nF separation, as the GEM-Forest-FnF map correctly classifies

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

these areas as non-forest. Lower PAs for European tree crops are likely due to limited training data, as only 1,133 European tree crop points were used to train the models. Lower PAs may also reflect the fragmented parcel structure and mixed agricultural mosaics typical for European landscapes.

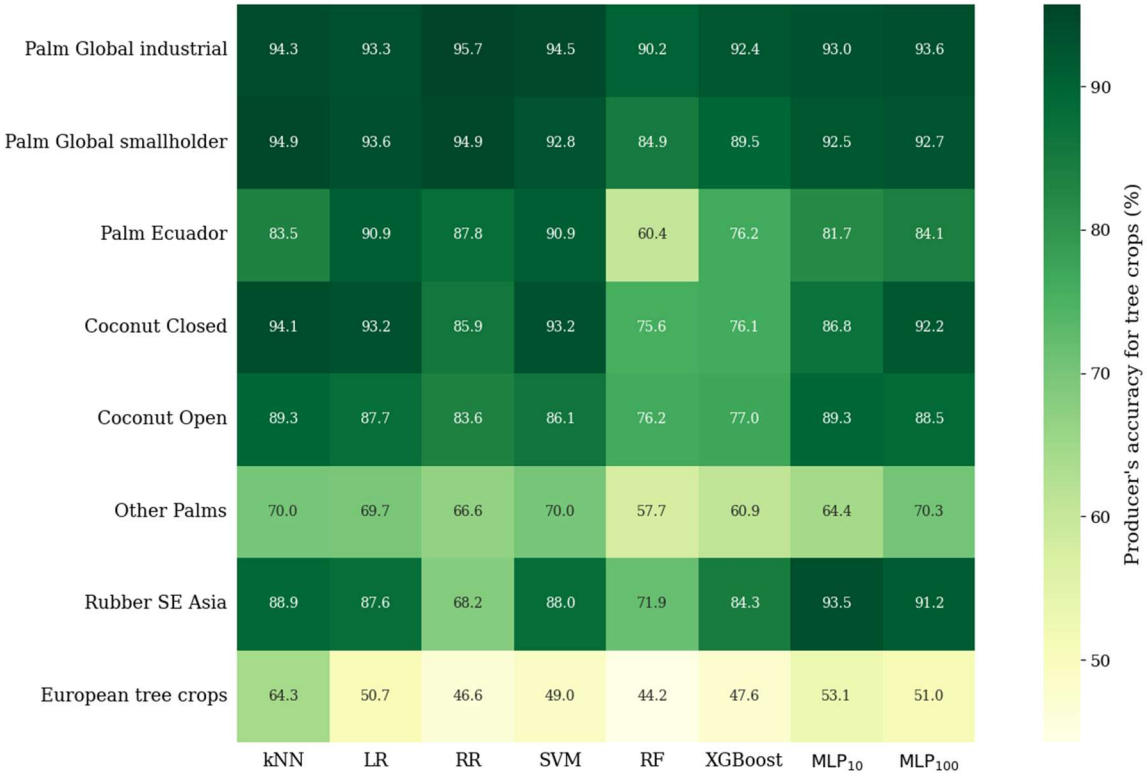

Figure 4. Overall accuracy of used tree crop validation datasets.

Overall, these results show that the GEM-Forest dataset can effectively separate tree crops from forests across diverse contexts, with performance determined by plantation size, canopy structure, landscape heterogeneity and training data availability, while it highlights both the strengths and limitations of the AEF-based classification approach for global tree crop mapping.

Across models, the highest average performance (considering PA and CEs, i.e., tree crops incorrectly classified as forest) was achieved by kNN, followed by the linear models (RR, LR and SVM) and neural network approaches (MLP), whereas tree-based ensembles (RF and XGBoost) performed worst (Figure 4, Figure 5 and Appendix D). These results support the assumption that the use of more complex nonlinear models does not necessarily improve classification performance in this case and may even underperform compared to linear or instance-based approaches. It further suggests that the AEF embedding space already provides a well-structured representation in which class separability can be effectively exploited by relatively simple models.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

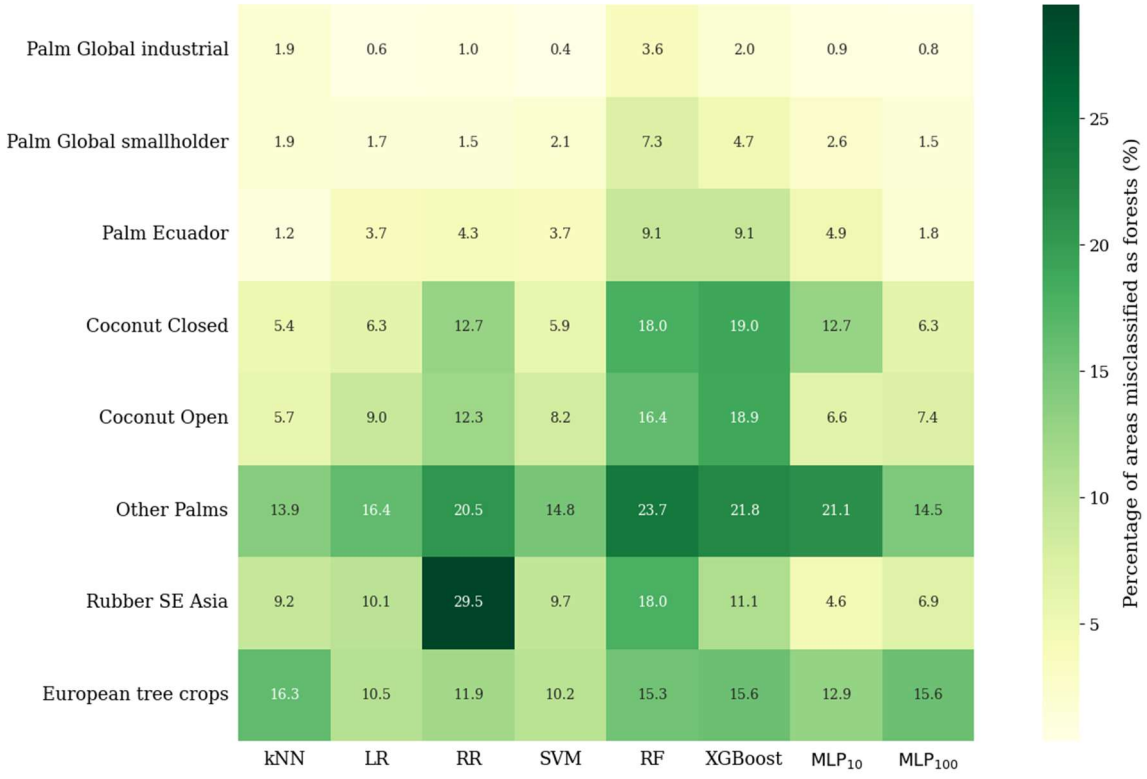

Figure 5. percentage of areas misclassified as forest class based on the tree crop validation datasets.

3.3 Detailed performance evaluation of ML models for the binary GEM-Forest-FnF map and the tree crop class

3.3.1 Instance-based interpretation: k-nearest neighbors

The kNN classifier achieved an OA of 90.89 ± 0.46% and a macro F1-score of 89.07% on the primary validation dataset. It showed the best performance on 4 out of 9 tree crop validation datasets, specifically for both coconut, smallholder oil palms and European tree crops, while for coconut plantations showed the lowest misclassifications as forests. It outperformed all others ML models in F/nF classification when tree crops were reclassified as non-forest (Appendix D). Moreover, as an instance-based learner that relies solely on local neighborhood similarity, its strong performance provides evidence that the AEF embeddings cluster semantically similar land-cover types in the 64-dimensional feature space. In contrast, European tree crops exhibited the highest misclassification rate for kNN, with 16.3% of samples incorrectly identified as forest.

The combination of strong overall performance and confusion mainly between forests and specific tree crops indicates that the representations separate most land-cover types well, while classes with similar vegetation structure remain less distinct. This likely reflects overlap between natural forests and certain perennial crops in the embedding feature space. Because kNN does not learn feature transformations or class-specific decision boundaries, these results provide a conservative estimate of the separability achieved by the AEF embeddings.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

3.3.2 Linear models: logistic regression, ridge regression and linear SVM

LR, RR, and Linear SVM achieved very similar performance, with overall accuracies of 91.48–91.82 ± 0.43–0.44% and macro-averaged F1-scores of 89.65–90.18%. Among them, LR generally ranked among the top performers in overall and class-specific results for F/nF separability (Table 2). RR achieved the highest overall accuracy for the industrial palm oil validation dataset but performed worst for rubber in SE Asia. The strong performance of linear classifiers indicates that forest-relevant information is already well aligned with simple linear decision boundaries that require only minimal model complexity to extract. Remaining errors appear to result from similarities in the embedding representations rather than limitations of the models themselves.

We selected the linear SVM to produce the final GEM-Forest product. This model provided high accuracy for the forest class, with balanced OE (15.60 ± 1.04%) and CE (12.88 ± 0.98%). Furthermore, it maintained consistently strong performance across all tested tree crops, with an average PA of 78% and a low average misclassification rate as forest (7%, see Appendix D). Beyond its predictive performance, the linear SVM offered a significant technical advantage due to its linear formulation. This structure permitted a straightforward implementation of the model weights via Eq. 1. Consequently, we calculated class decision boundary scores directly within GEE, which ensured computational efficiency for global applications.

3.3.3 Non-linear models: RF, MLP and XGBoost

Non-linear models such as MLP_{100}, MLP_{10}, RF and XGBoost achieved competitive OAs (91.32–92.06 ± 0.43–0.44%) and macro F1-scores (89.46–90.51%), comparable to linear models. The overall best results on F/nF separation were achieved by single-layered MLP with 10 neurons (Table 2). Despite their greater expressive capacity, none of the non-linear models provided substantial improvements over linear approaches or kNN. MLP_{10} only marginally outperformed individual linear models, compared to SVM for OA (0.45% difference) and for macro F1-score (0.31% difference).

A detailed analysis of tree crop class-based performance showed instability between different tree crops. RF and XGBoost showed the lowest average PA for tree crop classifications with the highest forest CE rates (Appendix D). In contrast to the Linear SVM, which provided consistent results across all tree crops, the non-linear models exhibited lower accuracies and higher forest misclassification rates for specific tree crop classes. Notably, despite having the highest aggregate OA for F/nF classification, the MLP_{10} underperformed compared to the SVM in five out of eight tree crop validation datasets. Similarly, tree-based models like RF and XGBoost showed a distinct OA drop-off for oil palms in Ecuador and for other global palms, with PAs between 58% and 76%. non-linear models also often achieved the highest forest CE rates, with up to 24% for other palms using RF. This strengthens that the dominant discriminative structure in the AEF embedding space is effectively captured by relatively simple / linear decision surfaces. Increasing model complexity fails to resolve underlying class overlaps; instead, it tends to degrade generalization through overfitting, as evidenced by the inconsistent performance of non-linear models across different tree crop classes.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

3.4 Feature space structure analysis of the validation datasets

The t-SNE, directly performed on the AEF embedding of each data point, visualization of the validation feature space provides complementary qualitative insight (Figure 6). The dominant non-forest class forms a dense and well-separated cluster, explaining the consistently high OAs. In contrast, forests and tree crops exhibit substantial overlap, consistent with the observed misclassification patterns across all classifiers. These errors therefore reflect genuine semantic similarity in the embedding space rather than overfitting or model failure.

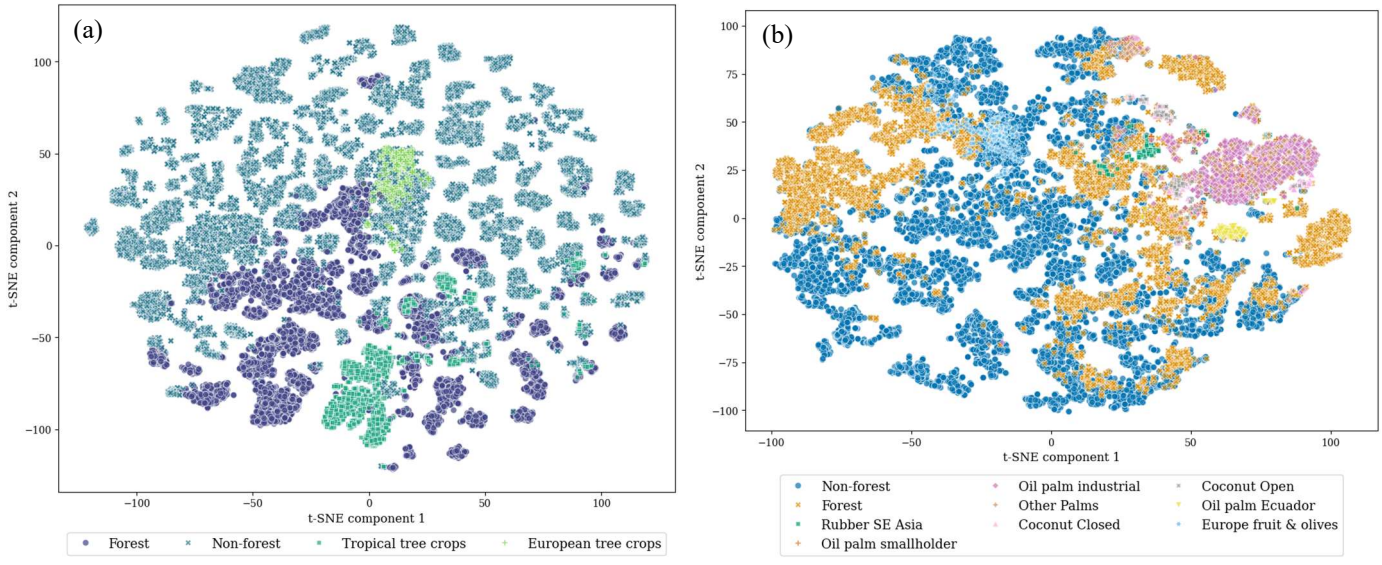

Figure 6. Separability of the training dataset (a) and validation datasets (b), including both the main validation dataset and class-wise tree crop validation datasets, through the two-dimensional feature space of the t-SNE plot.

3.5 Exploration of identified omission and commission errors

The error validation dataset, that helped to explore the primary sources of disagreement, consisted of 241 forest and 192 non-forest points according to the main validation dataset, representing 230 OEs and 203 CEs. Based on 433 randomly selected error samples, the dominant source of disagreement (62%) in the final GEM-Forest-FnF map was related to areas with sparse or scattered woody vegetation and shrub-tree confusion. These cases likely reflect the inherent difficulty of differentiating transitional vegetation types within the AEF embedding representation.

Approximately 27% of the disagreements were associated with forest edges and transition zones, which may represent problematic assignments due to the 10 m spatial resolution of the input AEF embedding dataset in combination with differences in interpretation scale. While the GEM-Forest-FnF map applied a minimum mapping unit (MMU) of 0.5 ha, the reference labeling procedure in the main validation dataset in some cases also considered the surrounding 100x100 m context that could potentially contribute to mismatches between map predictions and reference labels.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

The remaining disagreement cases (11%) were linked to land use versus land cover inconsistencies, particularly trees located within agricultural land use that were classified as forest (6%), indicating a limited but expected confusion between tree crops and forest areas. Additional sources of land use/land cover confusion included forested patches within urban environments (3%), that reflect potential limitations of the applied urban mask in excluding tree-covered urban areas. Locations affected by riparian or wetland complexity represented 2% and likely reflect the structural heterogeneity and transitional nature of riparian and wetland ecosystems, which complicate F/nF discrimination at the spatial resolution and MMU of the AEF input data.

A potential source of disagreement arises from the forest definition used in the main validation dataset, which includes areas where trees are temporarily below the minimum height threshold of 5 m (e.g., in a regeneration or regrowth phase) or where the land is currently unstocked but intended for forest use in the future. This aspect of the definition was not applied in our GEM-Forest-FnF product, as it cannot be directly observed from EO data, which form the basis for the AEF embeddings. In such areas, trees may not yet reach the physical definitional thresholds in situ in 2020 (5 m height, ≥10% canopy cover, and a MMU of 0.5 ha), potentially resulting in OEs when comparing mapped forest in our binary GEM-Forest-FnF product to reference labels in the main validation dataset.

Distinct regional clusters of disagreement further illustrate the interaction between forest definition and ecological context. In boreal regions of northern Europe and North America, discrepancies were concentrated in sparse and open forest formations and forest-tundra transition zones, where low canopy density challenges the delineation of continuous forest cover. In tropical and subtropical regions, particularly in savanna-dominated landscapes of Africa and northern Australia, disagreements were frequently associated with open woodlands and shrub-tree mosaics, where scattered trees generate forest-like spectral signatures without meeting canopy continuity requirements. In intensively managed agricultural regions, especially in Europe, misclassifications were dominated by riparian vegetation, shelterbelts, and small fragmented patches, emphasizing the influence of minimum width and area thresholds embedded in the FAO/EUDR forest definition.

It should be noted that when comparing the labels from the main validation dataset with our interpretation, only 101 points showed clear agreement between the two (71 forest and 30 non-forest). A substantial proportion of points (209) were located in areas with sparse or scattered woody vegetation, where shrub-tree confusion and transitional land-cover characteristics made assignment to forest or non-forest categories uncertain, even when using high-resolution imagery in GEP and radar and optical time series analysis, without in situ information on vegetation height or canopy density. The remaining 123 points showed discrepancies between the two interpretations, corresponding to potential commission- and omission-type differences. This indicates that these 123 points, that represents 28% of the identified OEs and CEs, may in fact reflect correctly classified F/nF areas in our binary GEM-Forest-FnF map, suggesting that the actual accuracy of the product could be higher than indicated by comparisons with the main validation dataset alone. Therefore, the use of local or regional F/nF validation datasets is recommended to further assess accuracy and identify the underlying causes of errors in the GEM-Forest-FnF map and in the tree crop class of GEM-Forest.

Of the 101 points showing clear agreement, 30 (30%) were commission-type mismatches and 71 (70%) were omission-type mismatches. The main sources of these mismatches were sparse or scattered woody vegetation and shrub-tree confusion (58%), forest edge or transitional zones (32%), with smaller contributions from land-use vs. land-cover inconsistencies. The identified OEs

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

were spatially concentrated along forest boundaries and in areas with gradual canopy transitions, rather than within homogeneous core forest or clearly non-forested regions.

Overall, most disagreements in the GEM-Forest-FnF map arise from transitional or ambiguous land-cover types, which pose inherent challenges for the AEF vector space in distinguishing forest and non-forest, consistent with previous EO-based global forest mapping approaches (Bourgoin et al., 2026a; Neumann et al., 2025; Shimada et al., 2014; Xiao et al., 2024). Sparse or scattered woody vegetation, shrub-tree mosaics, and forest edges account for the majority of errors, while land use conflicts and differences in forest definitions could explain the remaining errors. Disagreements are concentrated in open boreal forests, tropical savanna mosaics and fragmented agricultural landscapes. Up to 28% of identified OEs and CEs may reflect correct classifications, indicating that the actual accuracy of GEM-Forest-FnF is likely higher than suggested by comparison with the main validation dataset.

3.6 Comparison of the GEM-Forest-FnF and the tree crop class quality with other global forest maps

In comparison to other global F/nF datasets, based on the main validation dataset, our SVM-based GEM-Forest-FnF map ranks second in OA, with differences of 1.1% in OA, and 1.4% in macro F1-score, compared to the JRC's GFC2020 v3 dataset (Figure 7). The GFC2020 v2 achieved a very similar, but slightly lower OA than our GEM-Forest-FnF map, with macro F1-score of 90.27%. While the GFC2020 is a composite of existing data layers, GEM-Forest-FnF outperforms all other EO-based products, making it the best-performing map generated through direct EO input analysis. The other F/nF datasets achieved relatively high and comparable overall accuracies on the main validation dataset, with OA ranged from 86% to 91% and macro F1-scores from 83.87% to 89.96%. Regarding error rates, our dataset achieves balanced OE and CE rates for the forest class, at 15.68% and 12.36%, respectively, which are comparable to the GFC2020 v3 error rates of 10.78% and 12.94%. In contrast, the other F/nF datasets show higher bias in CE compared to OE for the forest class, indicating a tendency to classify non-forest areas as forests. Higher CE rates could be related to the definition of the forest/tree cover classes in the compared datasets, as also demonstrated in the following analysis. The only exception is the IPCC dataset, which exhibits a very high OE for the forest class (28.35%) and the lowest CE of 11.50%.

When evaluated against the two additional global validation datasets, GEM-Forest-FnF achieved higher OAs and macro-F1 scores than JRC's GFC2020 and most other datasets (Table 3). The only exception was ESA WorldCover, which achieved a slightly higher OA in the Mountain_Val dataset (88.09 ± 0.25%) compared to GEM-Forest-FnF (87.88 ± 0.26%). Across the Mountain_Val dataset, that include areas in world's mountain ranges, forest CEs were generally higher than OEs. This trend was expected due to the inherent difficulty of differentiating forests at tree lines, such as forest-shrubland boundaries. Individually, CGLC yielded the lowest forest OE across both validation datasets but suffered from the second-highest forest CE. Conversely, the lowest forest CEs were achieved by IPCC for SRS_Val (which also had the highest OE) and by GEM-Forest-FnF for Mountain_Val. Notably, GEM-Forest-FnF demonstrated the most balanced error profile among all tested datasets, with an OE of 17.50 ± 0.52% and a CE of 21.19 ± 0.56%. Finally, it should be noted that the SRS_Val and Mountain_Val datasets may include tree-covered areas under agricultural or urban land use, while potentially assigning unstocked forest areas to non-forest categories. Consequently, the high accuracies

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

achieved by products representing tree cover rather than forest land land use, such as ESA WorldCover and ESRI LU/LC, may be attributed to this class definition.

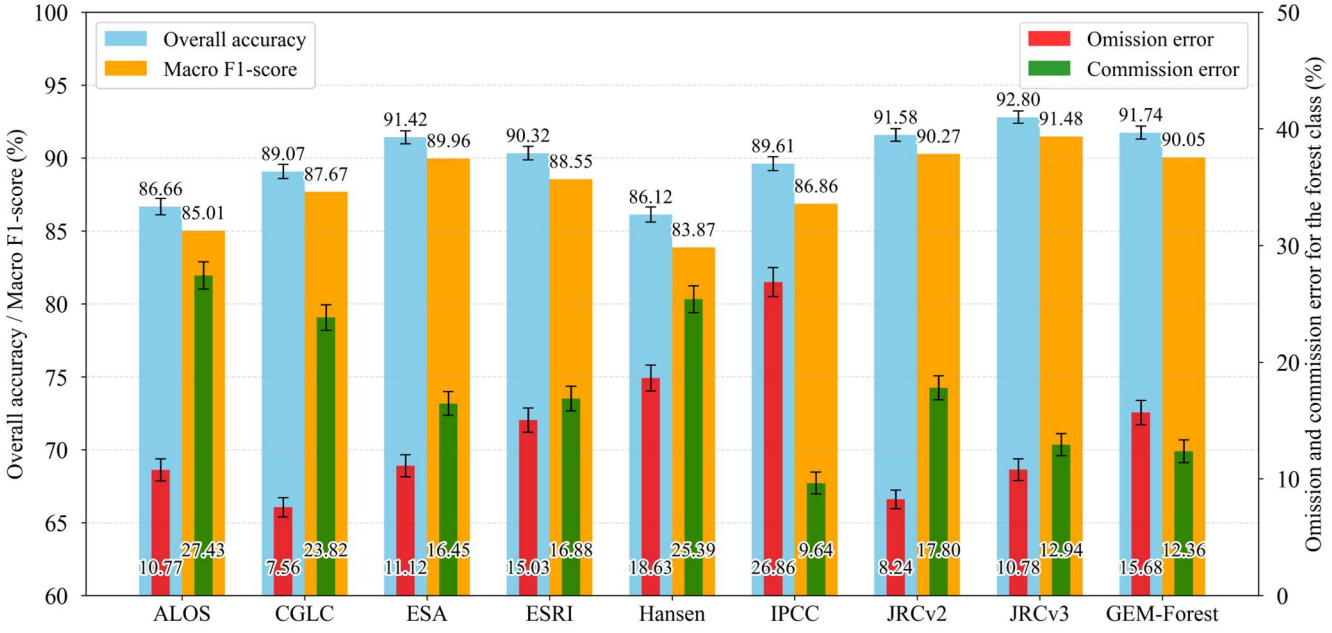

Figure 7. Comparison of macro F1-score, strata-based weighted OA, OE and CE for the forest class with associated 95% CIs among global F/nF products and our SVM-based GEM-Forest-FnF map based on the main validation dataset.

Moderate to very high accuracies across the tree crop validation datasets were achieved by products aligned with the FAO forest definition, which excludes tree-covered areas under agricultural use. These include the GFC2020 v2/v3, IPCC, PALSAR FnF, and our GEM-Forest-FnF map. Our GEM-Forest-FnF product consistently ranks among the top performers, placing between first and fourth across the evaluated datasets among JRC's GFC2020 v2/v3 and IPCC (Figure 8 and Figure 9), while only our GEM-Forest-FnF map and the GFC2020 v3 achieves tree crop PAs higher than 80% across all validated tree crops. Although the GFC2020 v3 product performs best in most cases and the IPCC dataset achieves the highest accuracy for the European tree crop validation dataset, the absolute differences relative to our GEM-Forest-FnF map are small. The PALSAR FnF product achieved moderate PAs exceeding 60% for most tree crop datasets. In contrast, as expected, the lowest PAs and highest forest CEs, were observed for datasets in which the forest or tree cover class effectively represents general tree cover, including perennial woody crops. The CGLCL and ESA WorldCover datasets, where tree crops are classified as forest/tree cover (see Buchhorn et al. (2020) and Van De Kerchove et al. (2020), respectively) achieved the lowest accuracies (3-30%) and the highest shares of forest CEs (above 80%) for all datasets except European tree crops. The Hansen et al.'s GFC product also represents tree cover rather than forest according to the FAO definition. Although this distinction was not explicitly stated in the original publication or its supplementary materials (Hansen et al., 2013), it has been highlighted by the research community directly after its publication (Tropek et al., 2014) or recently in Freitas Beyer et al. (2025). Consequently, GFC also performed poorly in terms of tree crop PA and forest CE across most tree crop validation datasets, with CE values exceeding 40%.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

Table 3. Comparison of macro F1-score, area-weighted OA, OE and CE for the forest class with associated 95% CIs among global F/nF products and our SVM-based GEM-Forest-FnF map based on SRS_Val and Mountain_Val global validation datasets. Bolded results represent the best achieved accuracies.

	SRS_Val dataset (Zhao et al., 2023)				Mountain_Val dataset (Naboureh et al., 2025)			
Dataset	OA ± CI (%)	macro-F1	OE F ± CI (%)	CE F ± CI (%)	OA ± CI (%)	macro-F1	OE F ± CI (%)	CE F ± CI (%)
ALOS	86.10 ± 0.22	83.42	17.55 ± 0.49	28.20 ± 0.49	83.99 ± 0.27	80.44	18.58 ± 0.64	35.32 ± 0.59
CGLC	88.83 ± 0.20	86.53	**11.04 ± 0.43**	25.72 ± 0.47	87.70 ± 0.24	84.96	**11.22 ± 0.54**	29.58 ± 0.57
ESA	89.85 ± 0.20	88.15	13.83 ± 0.42	18.70 ± 0.45	**88.09 ± 0.25**	85.75	13.65 ± 0.53	25.50 ± 0.57
ESRI	89.01 ± 0.21	87.33	16.75 ± 0.42	17.84 ± 0.45	87.62 ± 0.26	85.49	17.07 ± 0.53	22.87 ± 0.57
Hansen	85.41 ± 0.24	83.1	22.04 ± 0.46	24.21 ± 0.50	84.01 ± 0.29	81.25	23.10 ± 0.58	28.58 ± 0.61
IPCC	86.94 ± 0.24	85.67	25.31 ± 0.39	**10.56 ± 0.42**	82.88 ± 0.31	80.82	28.93 ± 0.51	21.65 ± 0.63
GFC2020 v2	89.45 ± 0.20	87.6	13.60 ± 0.42	20.49 ± 0.46	87.43 ± 0.25	84.96	14.82 ± 0.55	26.56 ± 0.57
GFC2020 v3	89.72 ± 0.21	88.17	15.95 ± 0.42	16.28 ± 0.44	87.27 ± 0.27	85.1	17.80 ± 0.54	23.20 ± 0.57
GEM-Forest-FnF	**90.39 ± 0.20**	**89.05**	16.25 ± 0.40	13.24 ± 0.41	87.88 ± 0.26	**85.9**	17.50 ± 0.52	**21.19 ± 0.56**

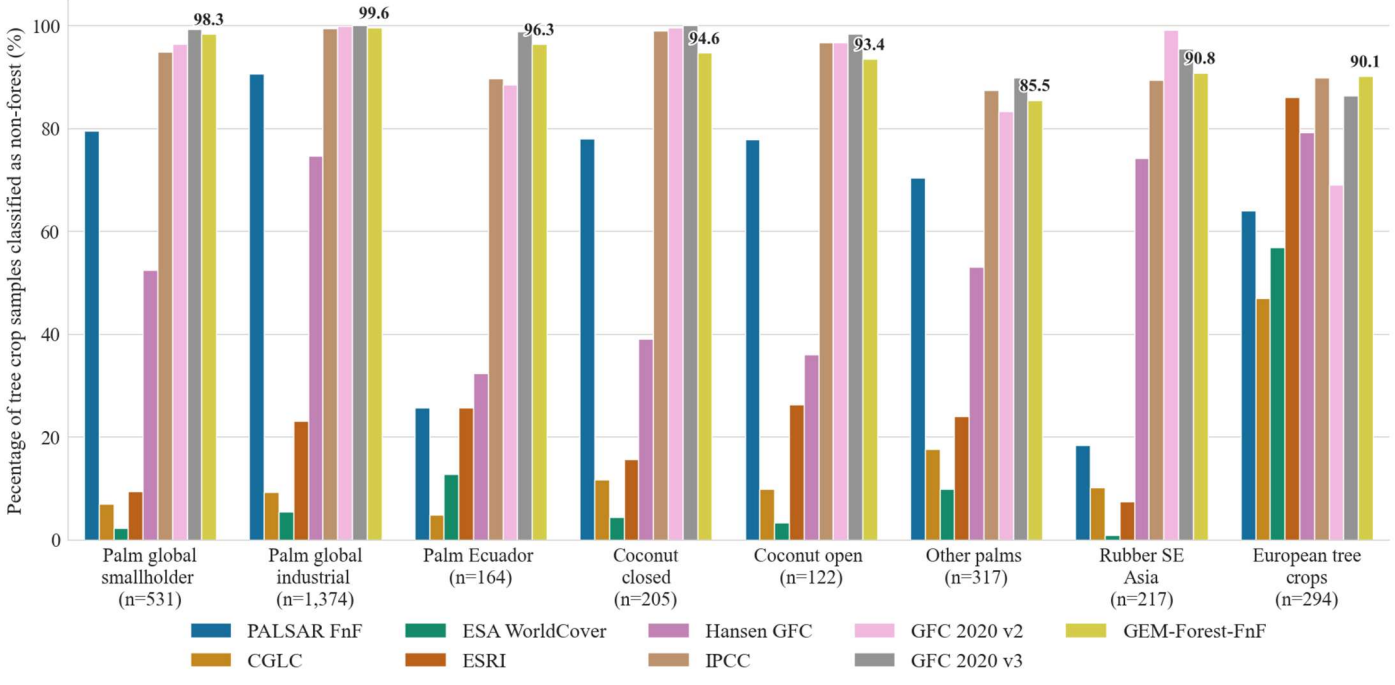

Figure 8. Producer's accuracies of the tree crop class as non-forest in global FnF products and SVM-based GEM-Forest-FnF, based on tree crop validation datasets.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

It should be noted that both the GFC2020 and IPCC products, which outperform our GEM-Forest-FnF map for selected tree crop types, were created by compositing multiple input datasets. The IPCC dataset combines various land use, land cover, and forest parameter products through Boolean analysis (Hunka et al., 2024b), while the GFC2020 products integrate numerous global and regional land cover, forest, and plantation datasets together with additional thematic layers, with plantation areas explicitly masked and reclassified as non-forest (Bourgoin et al., 2026a). As a result, they are highly optimized for a specific reference period and can achieve superior performance, but their application for different years requires updated ancillary inputs. Moreover, their quality and spatial resolution are limited by the characteristics of the underlying input datasets.

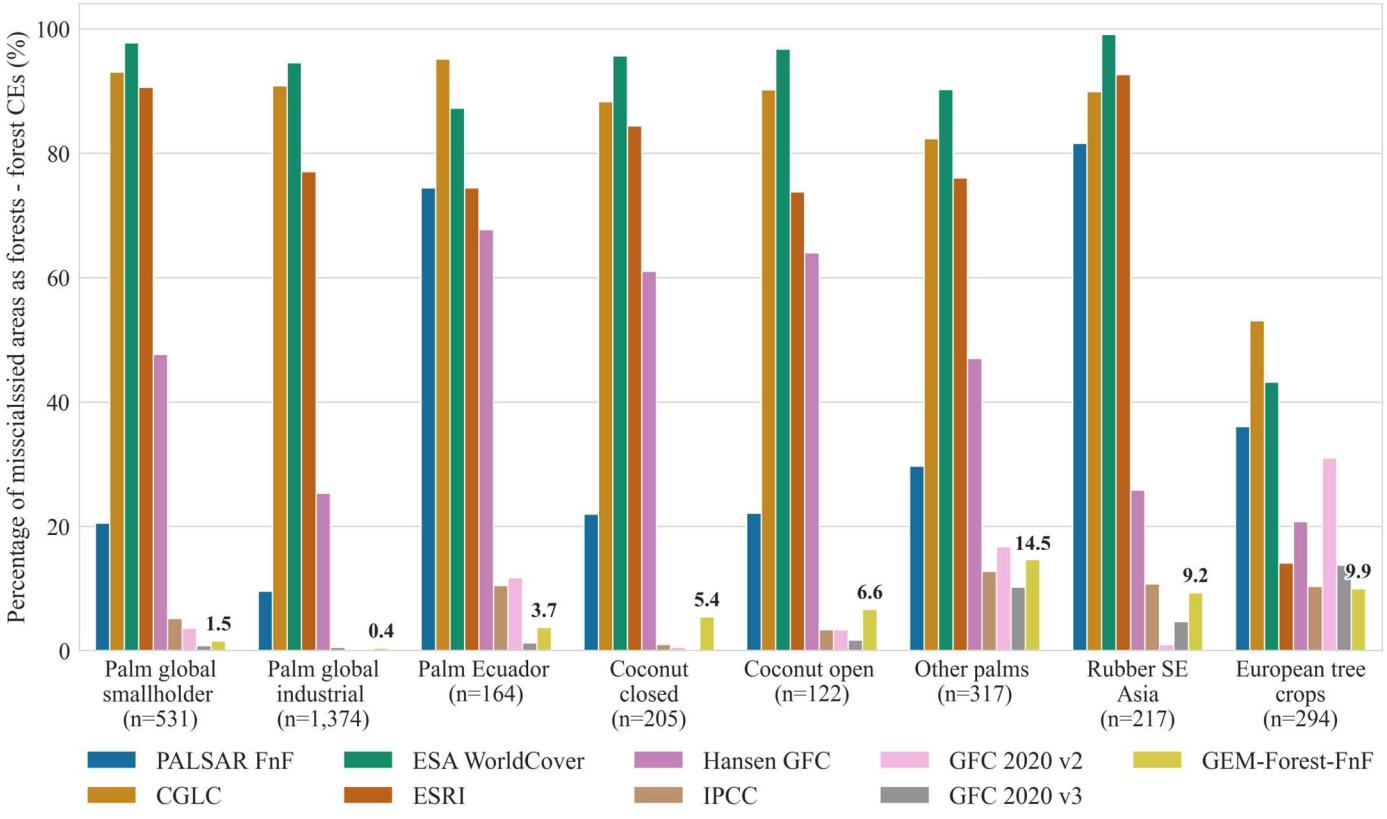

Figure 9. Rates of misclassification of plantation areas as forests (forest CEs) for the selected global FnF datasets and SVM-based GEM-Forest-FnF, based on tree crop validation datasets.

In contrast, our embedding-based approach enables consistent multi-year mapping using efficient and transparent machine learning methods, which is critical for monitoring applications such as deforestation assessment, carbon accounting and land use change analysis. Notably, the automatic training procedure in our approach, that is based on strict tree cover, non-tree cover and tree crop masks, helps to mitigate the limitations of individual input datasets, although some outlier forest, non-forest, or tree crop samples may still be included in the training data. As an additional contribution, we provide the trained model weights that allow users to generate F/nF and F/nF/tree crop maps for the period 2017–2025, when AEF embeddings are available. Importantly, our product

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

performs best among datasets derived solely from EO data and highlights the effectiveness of the proposed automatic training data generation and classification framework without reliance on external thematic inputs.

Overall, the results demonstrate that while composite products currently achieve the highest absolute accuracy with the lowest forest CEs, EO-driven approaches, such as the novel GEM-Forest-FnF map, present an attractive alternative due to their competitive performance combined with methodological transparency and temporal scalability.

3.7 GEM-Forest forest area comparison with FAO FRA 2020

GEM-Forest detects 3,919 Mha of forest globally that represents a 5.9% overall underestimation compared to the FAO FRA 2025 statistics for 2020 (4,165 Mha). The sample-weighted estimate yields to 4,017 ± 83 Mha of forest cover by GEM-Forest that represents 3.6 ± 2% underestimation compared to FAO FRA statistics. These results are comparable to recent studies; for instance, the GFC2020 v2 product accounts for 4,562 Mha of forest (overestimating FAO FRA by 9.5%), while the GFC2020 v3 product reports 4,164 Mha that represents a minor global underestimation of just 0.02% (Bourgoin et al., 2026b). At the continental level, Asia, Europe, North America, and South America showed high overall agreement, with R^2 values between 0.969 and 0.999 and absolute PD of 0.5–2.4%. Conversely, Africa and Oceania exhibited higher underestimations of 26% and 43%, respectively. The largest absolute area underestimation occurred in Australia (72 Mha, PD = –54%), followed by the USA (26 Mha, PD ≈ –8.6%) and Sudan (22 Mha, where GEM-Forest mapped only 0.1 Mha). Notably, extreme underestimations were detected in arid countries where GEM-Forest mapped near-zero or minimal forest cover compared to the substantial forest areas reported by the FAO FRA. This discrepancy is evident in Western Sahara (0 ha vs. 0.5 Mha in GEM-Forest vs. FAO FRA, respectively), Mauritania (103 ha vs. 1 Mha), Niger (667 ha vs. 1 Mha), Saudi Arabia (7,392 ha vs. 2.8 Mha), Turkmenistan (8,948 ha vs. 2.3 Mha), Burkina Faso (14,663 ha vs. 3.8 Mha) or Mali (0.1 Mha vs. 10.9 Mha). These underestimations by GEM-Forest relative to FAO FRA statistics can be attributed to several factors. First, satellite EO faces difficulties in mapping sparse, low-density woody vegetation against bright, arid backgrounds. Second, specific national forest definitions thresholds may vary widely, such as Australia's requirement of a 3 m minimum height and 20% canopy cover, or Sudan's 2 m height and MMU > 0.4 ha. Third, nationally reported forest areas vary substantially in their accuracy and indeed in some of these arid regions the FAO FRA values are outliers compared to both GEM-Forest and GFC2020. In Western Sahara for example GEM-Forest and GFC2020 both report zero forest cover whilst the FAO report >500,000 ha. Additionally, GEM-Forest does not account for temporarily unstocked forest areas. For example, severe wildfires cause large-scale disturbances across Australia, that resulted in approximately 3.9 Mha of tree cover loss in areas with >10% canopy cover during 2019 and 2020 (based on the Global forest loss due to fire dataset in Global Forest Watch, from Tyukavina et al., 2022). Finally, the spatial distribution of training samples that are generated from the intersection of three global forest layers may exclude and thus underrepresent these sparsely forested and arid landscapes, such as southern Australia or the Sahelian transition zones.

Conversely, substantial overestimations exceeding 20 Mha were observed in dense or high-latitude forest expanses, specifically in the Democratic Republic of the Congo (by 31 Mha, PD ≈ +22%), Indonesia (by 27 Mha, PD ≈ +28%), and Canada (by 22 Mha, PD ≈ +6%), some other countries exhibited very highe relative overestimation, such as Côte d'Ivoire (by 4.3 Mha, PD ≈ +101%) or

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

Myanmar (by 8 Mha, PD ≈ +28%). These discrepancies are likely driven by the misclassification of agricultural tree crops as natural forests in tropical regions, as well as confusion within the transitional zones between shrublands and forests in high-latitude boreal landscapes like Canada.

Comparisons of all country-level areas mapped by GEM-Forest with FAO FRA statistics are provided in the Supplement.

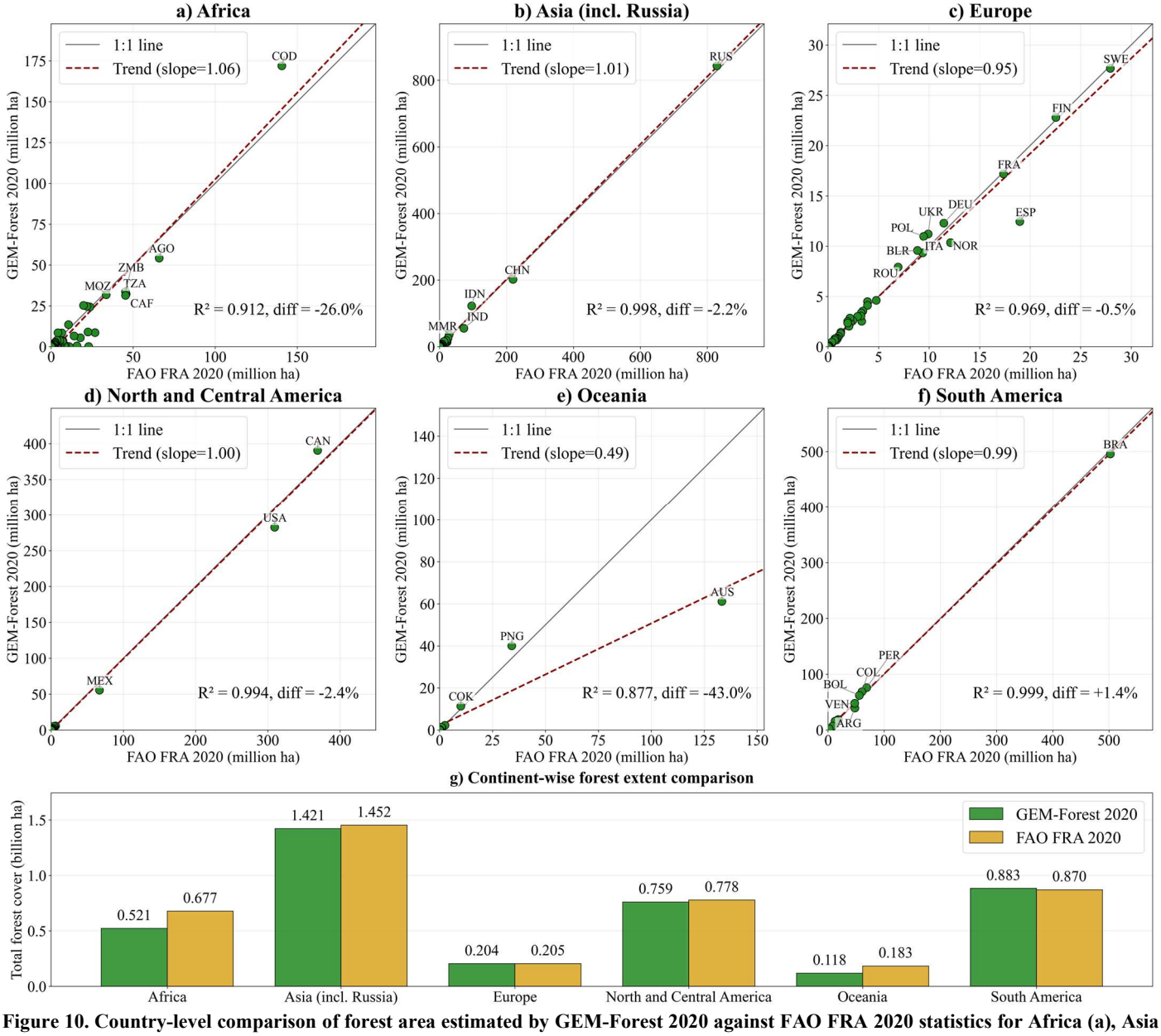

Figure 10. Country-level comparison of forest area estimated by GEM-Forest 2020 against FAO FRA 2020 statistics for Africa (a), Asia (including Russia) (b), Europe (c), North and Central America (d), Oceania (e), and South America (f), with a continent-level summary (g). The dashed red line shows the ordinary least-squares regression fit; the solid grey line is the 1:1 reference. The coefficient of determination (R^2) and percentage difference (diff.) are reported for each panel. Country ISO3 codes are shown for countries exceeding 5 million ha forest cover in Europe and 30 million ha in all other continents.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

3.8 Implications for global forest mapping

Overall, these results demonstrate that AEF embeddings provide a highly informative and transferable representation for global F/nF and tree crop mapping. Their effectiveness across instance-based, linear and non-linear models confirms that much of the relevant information for forest discrimination is already encoded in the embedding space. This supports the use of simple, transparent classifiers for large-scale and multi-temporal forest monitoring.

At the same time, confusion between forests and certain tree crops highlights a fundamental challenge for global forest mapping. While AEF embeddings substantially improve separability compared to raw spectral inputs, further gains will likely require additional and more specific training data for specific tree crops, more contextual information, such as temporal dynamics, spatial context or region-specific calibration, rather than increased classifier complexity alone. In this direction, our dataset is a first effort in this direction, on the one hand showing the capabilities of AEF embeddings, on the other, showing the need of better data to get better results.

Our GEM-Forest-FnF map (Figure 11) achieves comparable overall accuracy on the main F/nF validation dataset and on tree crop validation datasets relative to the best-performing GFC2020 v3 product, while outperforming it on two other global validation datasets. GEM-Forest's primary advantage lies in its straightforward transferability to other years for which AEF embeddings are available or will become available (currently 2017–2025). Recent work by Rahman (2026) supports this transferability by demonstrating that the relationship between embedding vectors and environmental variables remains highly stable across 2017–2023, with an average Pearson's correlation coefficient of $r = 0.963$. In contrast, composite products such as GFC2020 v3 would require updated ancillary input layers to be reproduced for additional time periods.

In light of policy regulations such as the EUDR and other national strategies, the presented GEM-Forest dataset and its potential future multi-temporal extensions can support deforestation assessment and other decision-making processes by users and policy-makers, alongside other global or more detailed regional products (not limited to EO-based data), ground-based observations or references, as highlighted in van Noordwijk et al. (2025).

3.9 Limitations

Despite the high potential for reproducibility of the GEM-Forest dataset for years other than 2020, and thus its potential use for change detection analyses, an important limitation remains: the GEM-Forest dataset and its derived products can only be generated for years in which AEF satellite embeddings are available, currently 2017–2025. While AEF is a powerful dataset, it has a few key limitations. First, interpretation of its embedding vectors in downstream analyses remains challenging. Recent work demonstrated that AEF encodes physically interpretable land surface properties across the Continental United States, showing a spatially robust and temporally stable relationship (Rahman, 2026). However, this global interpretability still needs to be validated across other continents and diverse climatic conditions. Second, AEF's annual temporal resolution prevents its use in near-real-time or intra-annual applications.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

Commission errors in the forest class of the GEM-Forest-FnF product occur in some urban areas where the combined GUB GAIA and FADSL datasets do not fully capture the extent of urban land cover.

It should be noted that while GEM-Forest incorporates specific land-use distinctions for forest, such as excluding agricultural tree crops and filtering out urban forests, it evaluates forest presence based on vegetative characteristics within a single observation year. Consequently, it does not explicitly account for forest lands that are temporarily unstocked due to management activities or natural disturbances, nor does it capture very young regrowth. While the FAO and EUDR frameworks classify these unstocked areas as forest, GEM-Forest classifies them as non-forest due to their lack of tree cover or not meeting the physical criteria (>5m tall, 10% canopy cover or 0.5 ha in extent) in 2020. Identifying these temporarily unstocked forest lands requires long-term time-series data or auxiliary layers, while our approach is optimized around a single year of AEF embeddings. By comparison, Bourgoin et al. (2026) addressed this temporal dimension in the GFC2020 v2 product by combining forest loss drivers and land-use datasets, while they allocated approximately 25 Mha of unstocked areas globally to their forest class.

The tree crop classes used in the training process include oil palm, rubber, coconut, and European tree crops, specifically fruit trees and olive orchards; therefore, the tree crop class of the GEM-Forest dataset primarily reflects these classes. It should be noted that cocoa and coffee plantations were not used in the training process due to the lack of high quality open-access global cocoa or coffee datasets and the absence of reliable validation datasets. Moreover, cocoa cultivation often occurs beneath the canopy of other trees, which makes the distinction between cocoa plantations and natural tree cover difficult using EO data. For example, full-sun (non-shaded) cocoa plantations, those that could be reliably monitored by EO-based data, represent only about 9% and 25% of cocoa systems in Nicaragua and Peru, respectively (Orozco-Aguilar et al., 2021). Even in Ghana and Côte d'Ivoire, where over 80% share of full-sun and close to full-sun cocoa plantations were reported (Becker et al., 2025), existing EO-based studies (Abu et al., 2021) report limited performance in cocoa mapping, with low OEs but moderate CEs (17% and 38%, respectively). Similarly, the coffee probability dataset from the FDP version 2025a shows a trade-off between OE and CE only at around 54% (Forest Data Partnership, 2025). Initial tests incorporating coffee and cocoa probability layers from this dataset did not yield satisfactory results and increased OE for the forest class. Therefore, cocoa and coffee were excluded from the training process. Nevertheless, upcoming iterations of GEM-Forest could overcome these limitations by including recently released AEF-based global datasets. This includes the FDP version 2025b (announced May 13, 2026; Sullivan and Tarrio, 2026), which features more extensive training than its predecessor, version 2025a, and the Forest Typology (ForTy) v1 dataset (announced May 23, 2026; Neumann et al., 2026). By incorporating these new products, we could improve the sampling of problematic classes like coffee and cocoa, while enhancing the classification accuracy across all target classes.

Low error rates in forest classification over tree crop areas were observed based on the eight tree crop validation datasets. However, additional validation datasets would allow a more comprehensive assessment of the tree crop class and could help identify the main problematic regions, while it could support further improvement of the GEM-Forest dataset. Based on qualitative evaluation, tree crop areas seem underrepresented in some regions and overrepresented in others. The most challenging areas occur in small-scale, fragmented tree crop landscapes, particularly in Europe, where the lowest OAs were recorded. In these regions, most tree crop OEs

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

correspond to areas classified as non-forest and indicates that the trained ML models demonstrate limited ability to reliably distinguish between non-forest and European tree crop systems.

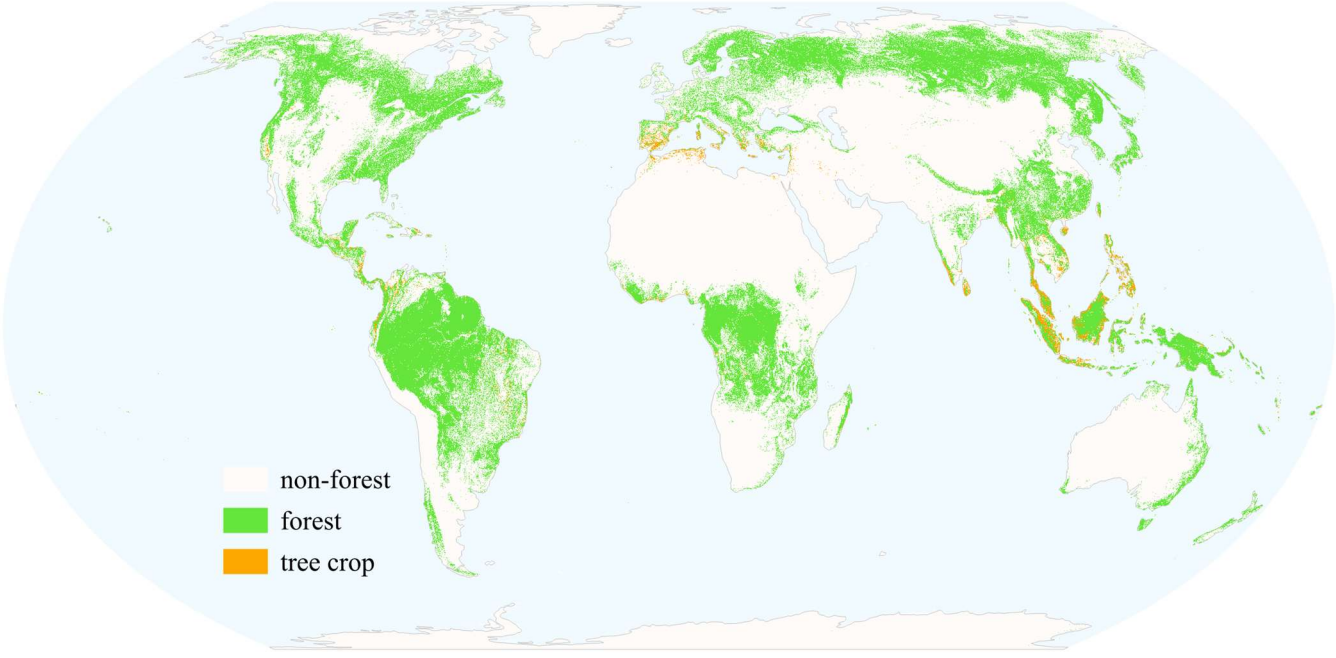

Figure 11. Global forest and tree crop extent based on GEM-Forest for 2020 in a 500 m/pixel aggregation, visualised in the Equal Earth projection (EPSG:8857).

Data availability

Data availability and format

The GEM-Forest (2020) dataset is provided as a global suite of Cloud-optimized GeoTIFFs (COG), utilizing Lempel-Ziv-Welch (LZW) lossless compression through the Zenodo repository under the following link https://doi.org/10.5281/zenodo.18921586 (Paluba et al., 2026). The data is set in the WGS84 geographic coordinate system with a spatial resolution of approximately 10 m (0.00008983°).

- Dataset structure: The dataset consists of 945 tiles, each covering 5° x 5° tiles (55,660 x 55,660 pixels). Individual tiles are grouped into 30° x 30° ZIP archives.
- Naming convention: ZIP archives of 30° x 30° are named by their Northwest (NW) and Southeast (SE) corners (e.g., GEM_Forest_N060W090_to_N030E000.zip), while individual GeoTIFFs are identified by their NW corner (e.g., GEM_Forest_N45E010.tif).
- Classification: Pixels are encoded as: 0 (non-forest), 1 (forest), and 2 (tree crops). This schema allows for the straightforward derivation of the GEM-Forest-FnF dataset by merging tree crops with the non-forest class.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

Supporting data

Training and validation datasets (containing AEF embedding vectors), and pre-trained machine learning weights (for RR, LR, SVM, RF, XGBoost and MLP models), are also archived on Zenodo, alongside the full GEM-Forest dataset: https://doi.org/10.5281/zenodo.18921586 (Paluba et al., 2026). Confusion matrices for both ML model evaluation and global dataset comparisons, and country-level mapped forest areas using GEM-Forest and their comparison to FAO FRA statistics are available in the Supplement.

Cloud integration and data exploration

- For further analysis or visualisation, the GEM-Forest is also accessible as an Image Collection Asset via Google Earth Engine through the following link: https://code.earthengine.google.com/?asset=projects/danielp-cuni/assets/GEM-Forest/GEM-Forest_2020 (last access: 12 May 2026). The dataset can be loaded into GEE as ee.ImageCollection("projects/danielp-cuni/assets/GEM-Forest/GEM-Forest_2020") and is part of the Awesome GEE Community Catalog (Roy et al., 2025).
- An interactive web application that enables the exploration of GEM-Forest, alongside the probability layers for forest, non-forest, tropical and European tree crop layers, available at danielp-cuni.projects.earthengine.app/view/gem-forest (last access: 12 June 2026).
- The GEM-Forest dataset can be loaded directly into desktop GIS software, such as QGIS, ArcGIS Pro and others, via the ArcGIS REST Web Service: https://cuni.maps.arcgis.com/home/item.html?id=2da0bbb4fcd14024aa3053cc01cf102d, accessible via https://tiles.arcgis.com/tiles/LPm07959azIAvFRD/arcgis/rest/services/GEM_Forest_v1_0/MapServer (last access: 12 June 2026).

Code availability

All the supporting codes to generate the training areas, training data, ML models training and testing, and for the reimplementation of the SVM weights to GEE will be made publicly available after the review process on GitHub. The supporting codes during the pre-review process are available from the corresponding author on reasonable request.

4. Conclusions

Overall, these results demonstrate that AEF embeddings provide a highly informative representation for global F/nF and tree crop mapping at 10 m spatial resolution. We show that simple, resource-efficient linear ML approaches combined with AEF can produce highly accurate global F/nF maps, often outperforming more complex algorithms. Our dataset, GEM-Forest, outperformed eight existing global forest, tree cover, and land cover datasets for F/nF differentiation across two global validation datasets, while ranking

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

second on the JRC's global forest validation dataset. Moreover, the tree crop class shows good performance in delineating tree crops and further confirms the low misclassification rates of agricultural tree crop plantations as forests.

The proposed approach has a strong potential for transfer across the 2017–2025 period, which enables multi-year applications based on models trained for a single year and represents a key next step in our research. In addition, reliance on a single global F/nF validation dataset has clear limitations; we therefore recommend the inclusion of complementary thematic validation datasets to better capture common classification challenges, particularly the confusion between forests and agricultural tree crops.

These results and the presented approach can support policy and regulatory applications, including the EUDR, while the open-access release of datasets and trained models facilitates global use, further evaluation, and methodological development by the EO and forest monitoring communities.

Appendices

Appendix A. Description of the methodology used for the generation of the validation dataset of the European tree crops (fruit and olive orchards)

The CORINE Land Cover 2018 (CLC) dataset was used to create an independent validation dataset for assessing the accuracy of the tree crop class in the GEM-Forest dataset over Europe. Two CLC classes representing permanent tree crops were selected: fruit trees and berry plantations (class 222) and olive groves (class 223). To ensure spatial relevance and reduce the influence of very small or potentially mixed land cover patches, only polygons with a minimum area of 25 ha were retained for further analysis.

Following this filtering step, a stratified point sampling strategy was applied to generate validation points while accounting for the wide range of polygon sizes. Polygons were grouped into size classes, and the number of validation points assigned to each group was proportional to polygon area in order to better represent large and spatially heterogeneous plantations, while avoiding over-representation of small polygons. The sampling strategy was defined as follows:

- Polygons with an area between 25 ha and 10,000 ha (40,843 polygons): a total of 100 validation points were randomly selected.
- Polygons with an area between 10,000 ha and 20,000 ha (50 polygons): two validation points were generated per polygon, resulting in 100 points.
- Polygons with an area between 20,000 ha and 50,000 ha (25 polygons): three validation points were generated per polygon, resulting in 75 points.
- Polygons with an area between 50,000 ha and 100,000 ha (7 polygons): four validation points were generated per polygon, resulting in 28 points.

In total, 303 validation points were generated across Europe. All points were subsequently visually interpreted using high-resolution satellite imagery available in Google Earth Pro, with particular attention to the temporal consistency of tree crop presence around the reference year 2020. Points located in areas showing land cover change, mixed land use, or ambiguous tree crop patterns were excluded to ensure a high-confidence validation dataset. During this final validation step, 9 points were excluded because it was not

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

possible to unambiguously determine whether the land cover represented a tree crop plantation or another type of woody or mixed vegetation. The final CORINE-based validation dataset therefore consisted of 294 high-confidence validation points.

Appendix B. Accuracy metrics for the F/nF classification where only forest and non-forest training samples, excluding tree crop training data, were used to train the ML models.

Model	OA ± CI (%)	macro-F1	w-F1	OE F ± CI (%)	CE F ± CI (%)	OE nF ± CI (%)	CE nF ± CI (%)
KNN	91.14 ± 0.45	89.49	89.49	13.89 ± 1.00	15.41 ± 1.03	6.71 ± 0.46	5.99 ± 0.46
LR	91.99 ± 0.43	90.46	90.46	13.24 ± 0.97	13.47 ± 0.98	5.78 ± 0.43	5.67 ± 0.44
RR	91.86 ± 0.43	90.24	90.24	14.61 ± 1.00	12.83 ± 0.98	5.38 ± 0.42	6.19 ± 0.45
SVM	92.00 ± 0.43	90.5	90.5	12.78 ± 0.97	13.76 ± 0.98	5.95 ± 0.43	5.49 ± 0.44
RF	91.52 ± 0.44	89.78	89.78	16.12 ± 1.04	12.68 ± 0.99	5.21 ± 0.42	6.78 ± 0.47
XGBoost	91.84 ± 0.43	90.24	90.24	14.12 ± 0.99	13.26 ± 0.98	5.62 ± 0.43	6.01 ± 0.45
MLP_{10}	91.93 ± 0.43	90.41	90.41	12.98 ± 0.98	13.82 ± 0.98	5.97 ± 0.43	5.58 ± 0.45
MLP_{100}	92.14 ± 0.43	90.66	90.66	12.68 ± 0.96	13.43 ± 0.97	5.80 ± 0.43	5.45 ± 0.44

Appendix C. Percentage of areas misclassified as forest class (forest CEs) based on the tree crop validation datasets for the F/nF classification where only forest and non-forest training samples were used to train the ML models.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

Appendix D. Mean producer's accuracy, mean commission errors for the tree crop validation datasets and mean producer's accuracy when tree crops (TC) were merged to the non-forest (nF) class. Bolded results represent the best achieved accuracies, while SVM (underscored) was selected for further evaluation.

Model	PA	CE as F	CE as nF	PA (TC as nF)
KNN	**80.80**	**6.66**	12.54	**93.34**
LR	78.84	7.51	13.65	92.49
RR	78.45	9.58	**11.96**	90.42
SVM	77.80	7.14	15.06	92.86
RF	66.11	14.24	19.65	85.76
XGBoost	71.43	12.64	15.93	87.36
MLP_{10}	77.07	8.59	14.34	91.41
MLP_{100}	75.45	8.29	16.25	91.71

Author contribution

D.P. and A.H. conceived and designed the study and its methodology, A.H. supervised the study, D.P. and V.M. conducted the experiments, performed the analysis and prepared the supporting codes. D.P. lead the data curation, manuscript preparation and validation of the results. K.O. and Y.T.P.Q. contributed to data curation and validation of the results. All authors contributed to manuscript review and editing and approved the final manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Financial support

This research was supported by the Johannes Amos Comenius Programme (P JAC), project No. CZ.02.01.01/00/22_008/0004605, Natural and anthropogenic georisks. AH & YTPQ also acknowledge funding from Charles University PRIMUS programme (PRIMUS/23/SCI/013), the Grant Agency of Charles University (GAUK), project no. [448325], and the Charles University Research Centre programme (UNCE/24/SCI/006). This research was also supported by the Slovak Research and Development Agency (APVV) under contract No. APVV-23-0210: Urban Overheating: Consequences, Mitigation and Perception, and by the Cultural and Educational Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic (KEGA) under contract No. KEGA 023UPJŠ-4/2025: Remote Sensing Education with a Focus on Satellite and Unmanned Aerial Platforms – Development of a Textbook and Online Educational Resources.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

Acknowledgements

This publication has been prepared using European Union's Copernicus Land Monitoring Service information; specifically the Forest Additional Support Layer (FADSL) from the High Resolution Layer Tree Cover and Forests product, DOI: https://doi.org/10.2909/4605463b-7150-49c6-9b45-01bf542891a9. We would like to thank Maxim Neumann and Yuchang Jiang for providing early access to their 10 m-resolution tree crop map of South America from (Jiang et al., (2026). We would also like to thank Jakub Lysák, Hugo Majer, Zoltan Szantoi and Eric Jensen for their useful suggestions.

References

Abu, I.-O., Szantoi, Z., Brink, A., Robuchon, M., and Thiel, M.: Detecting cocoa plantations in Côte d'Ivoire and Ghana and their implications on protected areas, Ecol. Indic., 129, 107863, https://doi.org/10.1016/j.ecolind.2021.107863, 2021.

Ball, J. G., Wicklein, J. A., Feng, Z., Knezevic, J., Jaffer, S., Atzberger, C., Dalponte, M., and Coomes, D.: Geospatial foundation models enable data-efficient tree species mapping in temperate montane forests, https://doi.org/10.64898/2026.02.23.707022, 2026.

Bartels, S. F., Chen, H. Y. H., Wulder, M. A., and White, J. C.: Trends in post-disturbance recovery rates of Canada's forests following wildfire and harvest, For. Ecol. Manag., 361, 194–207, https://doi.org/10.1016/j.foreco.2015.11.015, 2016.

Becker, A., Wegner, J. D., Dawoe, E., Schindler, K., Thompson, W. J., Bunn, C., Garrett, R. D., Castro-Llanos, F., Hart, S. P., and Blaser-Hart, W. J.: The unrealized potential of agroforestry for an emissions-intensive agricultural commodity, Nat. Sustain., 8, 994–1003, https://doi.org/10.1038/s41893-025-01608-7, 2025.

Bourgoin, C., Verhegghen, A., Carboni, S., Ameztoy, I., Ceccherini, G., Colditz, R., and Achard, F.: Global map of forest types 2020 - version 0 (version 0), https://doi.org/http://data.europa.eu/89h/037ca376-ba92-49db-a8f7-0c277c1e5436, 2024.

Bourgoin, C., Verhegghen, A., Carboni, S., Degreve, L., Ameztoy, I., Ceccherini, G., Colditz, R., and Achard, F.: Global forest maps for the year 2020 to support the EU regulation on deforestation-free supply chains: improved map of global forest cover(GFC2020) and preliminary map of global forest types (GFT2020)., Publications Office, LU, https://doi.org/10.2760/1975879, 2025.

Bourgoin, C., Verhegghen, A., Carboni, S., Ameztoy, I., Degreve, L., Fritz, S., Herold, M., Tsendbazar, N., Lesiv, M., Achard, F., and Colditz, R.: GFC2020: a global map of forest land use for year 2020 to support the EU Deforestation Regulation, Earth Syst. Sci. Data, 18, 1331–1365, https://doi.org/10.5194/essd-18-1331-2026, 2026a.

Bourgoin, C., Verhegghen, A., Ameztoy, I., Beuchle, R., Carboni, S., Carreiras, J., Lesiv, M., Lupi, A., Marinelli, D., Richter, J., Simonetti, D., Turubanova, S., Achard, F., and Colditz, R.: Maps of global forest cover 2020 version 3 and global forest type 2020 version 1 supporting the EU deforestation regulation, Publications Office of the European Union, LU, https://doi.org/10.2760/9982436, 2026b.

Brown, C. F., Kazmierski, M. R., Pasquarella, V. J., Rucklidge, W. J., Samsikova, M., Zhang, C., Shelhamer, E., Lahera, E., Wiles, O., Ilyushchenko, S., Gorelick, N., Zhang, L. L., Alj, S., Schechter, E., Askay, S., Guinan, O., Moore, R., Boukouvalas, A., and Kohli, P.: AlphaEarth Foundations: An embedding field model for accurate and efficient global mapping from sparse label data, https://doi.org/10.48550/arXiv.2507.22291, 2025.

Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tsendbazar, N.-E., Li, L., and Tarko, A.: Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Product User Manual, Zenodo, https://doi.org/10.5281/ZENODO.3938963, 2020.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

Bunting, P., Rosenqvist, A., Hilarides, L., Lucas, R. M., Thomas, N., Tadono, T., Worthington, T. A., Spalding, M., Murray, N. J., and Rebelo, L.-M.: Global Mangrove Extent Change 1996–2020: Global Mangrove Watch Version 3.0, Remote Sens., 14, 3657, https://doi.org/10.3390/rs14153657, 2022a.

Bunting, P., Rosenqvist, A., Hilarides, L., Lucas, R., Thomas, N., Tadono, T., Worthington, T., Spalding, M., Murray, N., and Rebelo, L.-M.: Global Mangrove Watch (1996 - 2020) Version 3.0 Dataset (3.0), https://doi.org/10.5281/ZENODO.6894273, 2022b.

Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794, https://doi.org/10.1145/2939672.2939785, 2016.

Clinton, N., Vollrath, A., D'annunzio, R., Liu, D., Glick, H. B., Descals, A., Sullivan, A., Guinan, O., Abramowitz, J., Stolle, F., Goodman, C., Birch, T., Quinn, D., Danylo, O., Lips, T., Coelho, D., Bihari, E., Cronkite-Ratcliff, B., Poortinga, A., Haghighattalab, A., Notman, E., DeWitt, M., Yonas, A., Donchyts, G., Shah, D., Saah, D., Tenneson, K., Quyen, N. H., Verma, M., and Wilcox, A.: A community palm model, https://doi.org/10.48550/arXiv.2405.09530, 20 November 2024.

Colditz, R., Verhegghen, A., Carboni, S., Bourgoin, C., and Achard, F.: Validation dataset for the global map of forest cover 2020 - version 2 (2), 2025.

Cook-Patton, S. C., Leavitt, S. M., Gibbs, D., Harris, N. L., Lister, K., Anderson-Teixeira, K. J., Briggs, R. D., Chazdon, R. L., Crowther, T. W., Ellis, P. W., Griscom, H. P., Herrmann, V., Holl, K. D., Houghton, R. A., Larrosa, C., Lomax, G., Lucas, R., Madsen, P., Malhi, Y., Paquette, A., Parker, J. D., Paul, K., Routh, D., Roxburgh, S., Saatchi, S., van den Hoogen, J., Walker, W. S., Wheeler, C. E., Wood, S. A., Xu, L., and Griscom, B. W.: Mapping carbon accumulation potential from global natural forest regrowth, Nature, 585, 545–550, https://doi.org/10.1038/s41586-020-2686-x, 2020.

Copernicus Land Monitoring Service: Land Cover 2015-2019 (raster 100 m), global, annual - version 3 (3.0), https://doi.org/10.2909/C6377C6E-76CC-4D03-8330-628A03693042, 2015.

Descals, A.: High-resolution global map of closed-canopy coconut palm (v1-2), https://doi.org/10.5281/ZENODO.8128183, 2023.

Descals, A.: Global oil palm extent and planting year from 1990 to 2021, https://doi.org/10.5281/zenodo.13379129, 2024a.

Descals, A.: Validation dataset for the article "Rubber planting and deforestation," https://doi.org/10.5281/zenodo.10646349, 2024b.

Descals, A., Wich, S., Szantoi, Z., Struebig, M. J., Dennis, R., Hatton, Z., Ariffin, T., Unus, N., Gaveau, D. L. A., and Meijaard, E.: High-resolution global map of closed-canopy coconut palm, Earth Syst. Sci. Data, 15, 3991–4010, https://doi.org/10.5194/essd-15-3991-2023, 2023.

Descals, A., Gaveau, D. L. A., Wich, S., Szantoi, Z., and Meijaard, E.: Global mapping of oil palm planting year from 1990 to 2021, Earth Syst. Sci. Data, 16, 5111–5129, https://doi.org/10.5194/essd-16-5111-2024, 2024.

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., Hansen, M., Locke, H., Ellis, E. C., Jones, B., Barber, C. V., Hayes, R., Kormos, C., Martin, V., Crist, E., Sechrest, W., Price, L., Baillie, J. E. M., Weeden, D., Suckling, K., Davis, C., Sizer, N., Moore, R., Thau, D., Birch, T., Potapov, P., Turubanova, S., Tyukavina, A., de Souza, N., Pintea, L., Brito, J. C., Llewellyn, O. A., Miller, A. G., Patzelt, A., Ghazanfar, S. A., Timberlake, J., Klöser, H., Shennan-Farpón, Y., Kindt, R., Lillesø, J.-P. B., van Breugel, P., Graudal, L., Voge, M., Al-Shammari, K. F., and Saleem, M.: An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm, BioScience, 67, 534–545, https://doi.org/10.1093/biosci/bix014, 2017.

European Environment Agency: CORINE Land Cover 2018 (vector), Europe, 6-yearly (20.01), https://doi.org/10.2909/71C95A07-E296-44FC-B22B-415F42ACFDF0, 2020.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

European Environment Agency: Forest Additional Support Layer 2018 - Present (raster 10m), Europe, 3-yearly, Nov. 2024 (01.00), https://doi.org/10.2909/4605463B-7150-49C6-9B45-01BF542891A9, 2024.

FAO: FAO GAUL: Global Administrative Unit Layers 2015, Country Boundaries, 2015.

FAO: The State of the World's Forests 2022, FAO, https://doi.org/10.4060/cb9360en, 2022.

FAO: Global Forest Resources Assessment - FRA 2025 - Terms and Definitions, 2023.

FAO: Global Forest Resources Assessment 2025, FAO, https://doi.org/10.4060/cd6709en, 2025.

Feng, Z., Atzberger, C., Jaffer, S., Knezevic, J., Sormunen, S., Young, R., Lisaius, M. C., Immitzer, M., Jackson, T., Ball, J., Coomes, D. A., Madhavapeddy, A., Blake, A., and Keshav, S.: TESSERA: Temporal Embeddings of Surface Spectra for Earth Representation and Analysis, https://doi.org/10.48550/ARXIV.2506.20380, 2025.

Community models 2025a: https://github.com/google/forest-data-partnership/edit/main/models/README.md.

Forgaard, T., Reksten, J. H., Waldeland, A. U., Marsocci, V., Longépé, N., Kampffmeyer, M., and Salberg, A.-B.: THOR: A Versatile Foundation Model for Earth Observation Climate and Society Applications, https://doi.org/10.48550/ARXIV.2601.16011, 2026.

Forzieri, G., Dakos, V., McDowell, N. G., Ramdane, A., and Cescatti, A.: Emerging signals of declining forest resilience under climate change, Nature, 608, 534–539, https://doi.org/10.1038/s41586-022-04959-9, 2022.

Freitas Beyer, J., Köthke, M., and Lippe, M.: Assessing the Suitability of Available Global Forest Maps as Reference Tools for EUDR-Compliant Deforestation Monitoring, Remote Sens., 17, 3012, https://doi.org/10.3390/rs17173012, 2025.

Fundación EcoCiencia: Documento de validación:"Machine Learning Training Data for Continental Ecuador and Galapagos from 1985 to 2023" (v1.0), 2025.

Gao, X., Pasquarella, V., Laflower, D., and Thompson, J. R.: Mapping Forest Communities, Including Species Composition, Structure, and Carbon, at 10-m Resolution Using Geospatial Embeddings, https://doi.org/10.2139/ssrn.5936862, 18 December 2025.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote Sens. Environ., 202, 18–27, https://doi.org/10.1016/j.rse.2017.06.031, 2017.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., and Townshend, J. R. G.: High-Resolution Global Maps of 21st-Century Forest Cover Change, Science, 342, 850–853, https://doi.org/10.1126/science.1244693, 2013.

Hughes, A. C., Orr, M. C., Yang, Q., and Qiao, H.: Effectively and accurately mapping global biodiversity patterns for different regions and taxa, Glob. Ecol. Biogeogr., 30, 1375–1388, https://doi.org/10.1111/geb.13304, 2021.

Hunka, N., Duncanson, L., Armston, J., Dubayah, R. O., Healey, S. P., Santoro, M., May, P. B., Araza, A., Bourgain, C., Montesano, P. M., Neigh, C. S., Grantham, H., Potapov, V., Turubanova, S., Tyukavina, A., Richter, J., Harris, N., Urbazaev, M., Pascual, A., Requena Suarez, D., Herold, M., Poulter, B., Wilson, S. N., Grassi, G., Federici, S., Sanz Sanchez, M. J., and Melo, J.: Carbon Monitoring System (CMS)Classification of Global Forests for IPCC Aboveground Biomass Tier 1 Estimates, 2020, https://doi.org/10.3334/ORNLDAAC/2345, 1 January 2024a.

Hunka, N., Duncanson, L., Armston, J., Dubayah, R., Healey, S. P., Santoro, M., May, P., Araza, A., Bourgoin, C., Montesano, P. M., Neigh, C. S. R., Grantham, H., Potapov, P., Turubanova, S., Tyukavina, A., Richter, J., Harris, N., Urbazaev, M., Pascual, A., Suarez, D. R., Herold, M., Poulter, B., Wilson, S. N., Grassi, G., Federici, S., Sanz, M. J., and Melo, J.: Intergovernmental Panel on Climate Change (IPCC) Tier 1 forest biomass estimates from Earth Observation, Sci. Data, 11, 1127, https://doi.org/10.1038/s41597-024-03930-9, 2024b.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

Jakubik, J., Yang, F., Blumenstiel, B., Scheurer, E., Sedona, R., Maurogiovanni, S., Bosmans, J., Dionelis, N., Marsocci, V., Kopp, N., Ramachandran, R., Fraccaro, P., Brunschwiler, T., Cavallaro, G., Bernabe-Moreno, J., and Longépé, N.: TerraMind: Large-Scale Generative Multimodality for Earth Observation, https://doi.org/10.48550/ARXIV.2504.11171, 2025.

Jiang, Y., Raichuk, A., Tong, X., Garnot, V. S. F., Ortiz-Gonzalo, D., Morris, D., Schindler, K., Wegner, J. D., and Neumann, M.: Tree crop mapping of South America reveals links to deforestation and conservation, https://doi.org/10.48550/arXiv.2602.17372, 2026.

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M., and Brumby, S. P.: Global land use / land cover with Sentinel 2 and deep learning, in: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, ISSN: 2153-7003, 4704–4707, https://doi.org/10.1109/IGARSS47720.2021.9553499, 2021.

Lesiv, M., Schepaschenko, D., Buchhorn, M., See, L., Dürauer, M., Georgieva, I., Jung, M., Hofhansl, F., Schulze, K., Bilous, A., Blyshchyk, V., Mukhortova, L., Brenes, C. L. M., Krivobokov, L., Ntie, S., Tsogt, K., Pietsch, S. A., Tikhonova, E., Kim, M., Di Fulvio, F., Su, Y.-F., Zadorozhniuk, R., Sirbu, F. S., Panging, K., Bilous, S., Kovalevskii, S. B., Kraxner, F., Rabia, A. H., Vasylyshyn, R., Ahmed, R., Diachuk, P., Kovalevskyi, S. S., Bungnamei, K., Bordoloi, K., Churilov, A., Vasylyshyn, O., Sahariah, D., Tertyshnyi, A. P., Saikia, A., Malek, Ž., Singha, K., Feshchenko, R., Prestele, R., Akhtar, I. ul H., Sharma, K., Domashovets, G., Spawn-Lee, S. A., Blyshchyk, O., Slyva, O., Ilkiv, M., Melnyk, O., Sliusarchuk, V., Karpuk, A., Terentiev, A., Bilous, V., Blyshchyk, K., Bilous, M., Bogovyk, N., Blyshchyk, I., Bartalev, S., Yatskov, M., Smets, B., Visconti, P., Mccallum, I., Obersteiner, M., and Fritz, S.: Global forest management data for 2015 at a 100 m resolution, Sci. Data, 9, 199, https://doi.org/10.1038/s41597-022-01332-3, 2022.

Lewis, S. L., Wheeler, C. E., Mitchard, E. T. A., and Koch, A.: Restoring natural forests is the best way to remove atmospheric carbon, Nature, 568, 25–28, https://doi.org/10.1038/d41586-019-01026-8, 2019.

Li, X., Gong, P., Zhou, Y., Wang, J., Bai, Y., Chen, B., Hu, T., Xiao, Y., Xu, B., Yang, J., Liu, X., Cai, W., Huang, H., Wu, T., Wang, X., Lin, P., Li, X., Chen, J., He, C., Li, X., Yu, L., Clinton, N., and Zhu, Z.: Mapping global urban boundaries from the global artificial impervious area (GAIA) data, Environ. Res. Lett., 15, 094044, https://doi.org/10.1088/1748-9326/ab9be3, 2020.

Liu, L., Zhao, T., and Zhang, X.: A novel stratified random sampling global validation dataset in 2020—SRS_Val dataset (3), https://doi.org/10.5281/zenodo.7846090, 2023.

Ma, J., Li, J., Wu, W., and Liu, J.: Global forest fragmentation change from 2000 to 2020, Nat. Commun., 14, 3752, https://doi.org/10.1038/s41467-023-39221-x, 2023.

Naboureh, A., Li, A., Bian, J., Moharrami, M., Ebrahimy, H., Lei, G., Nan, X., Zhang, Z., Feizizadeh, B., Dabove, P., Makki, M., Sauter, T., Attarchi, S., and Amani, M.: Accuracies, discrepancies, and challenges of the 10 m global land cover products in mountains, GIScience Remote Sens., 62, 2556064, https://doi.org/10.1080/15481603.2025.2556064, 2025a.

Naboureh, A., Li, A., Bian, J., and Meysam, M.: Land cover validation dataset over global mountains, https://doi.org/10.5281/zenodo.15354081, 2025b.

Neumann, M., Raichuk, A., Jiang, Y., Rey, M., Stanimirova, R., Sims, M. J., Carter, S., Goldman, E., Anderson, K., Poklukar, P., Tarrio, K., Lesiv, M., Fritz, S., Clinton, N., Stanton, C., Morris, D., and Purves, D.: Natural forests of the world – a 2020 baseline for deforestation and degradation monitoring, Sci. Data, 12, 1715, https://doi.org/10.1038/s41597-025-06097-z, 2025.

Neumann, M., Raichuk, A., Potapov, P., Lesiv, M., Overlan, M., Rey, M., Rajakumar, R., Conserva, M., Stanimirova, R., Sims, M., Carter, S., Goldman, E., Jiang, Y., Scheibenreif, L., Georgieva, I., Shchepashchenko, M., Fritz, S., Clinton, N., Stanton, C., Morris, D., and Purves,

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

D.: Global forest typology at 10-meter resolution for forest and land-use monitoring, https://eartharxiv.org/repository/view/13130/, 22 May 2026.

van Noordwijk, M., Dewi, S., Minang, P. A., Harrison, R. D., Leimona, B., Ekadinata, A., Burgers, P., Slingerland, M., Sassen, M., Watson, C., and Sayer, J.: Beyond imperfect maps: Evidence for EUDR-compliant agroforestry, People Nat., 7, 1713–1723, https://doi.org/10.1002/pan3.70088, 2025.

Olofsson, P., Foody, G. M., Stehman, S. V., and Woodcock, C. E.: Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation, Remote Sens. Environ., 129, 122–131, https://doi.org/10.1016/j.rse.2012.10.031, 2013.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., and Wulder, M. A.: Good practices for estimating area and assessing accuracy of land change, Remote Sens. Environ., 148, 42–57, https://doi.org/10.1016/j.rse.2014.02.015, 2014.

Onačillová, K., Krištofová, V., and Paluba, D.: Automatic forest cover classification using Sentinel-2 multispectral satellite data and machine learning algorithms in Google Earth Engine, Acta Geogr. Univ. Comen., 67, 163–185, 2023.

Orozco-Aguilar, L., López-Sampson, A., Leandro-Muñoz, M. E., Robiglio, V., Reyes, M., Bordeaux, M., Sepúlveda, N., and Somarriba, E.: Elucidating Pathways and Discourses Linking Cocoa Cultivation to Deforestation, Reforestation, and Tree Cover Change in Nicaragua and Peru, Front. Sustain. Food Syst., 5, https://doi.org/10.3389/fsufs.2021.635779, 2021.

Paluba, D., Marsocci, V., Onačillová, K., Puerta Quintana, Y. T., and Hastie, A.: GEM-Forest: A Global satellite EMbedding–based map of forests and tree crops for 2020 (GEM-Forest products, training & validation data, and model weights), https://doi.org/10.5281/zenodo.18921586, 2026.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É.: Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res., 12, 2825–2830, 2011.

Rahman, M.: Physically interpretable AlphaEarth foundation model embeddings enable LLM-based land surface intelligence, Remote Sens. Appl. Soc. Environ., 42, 102045, https://doi.org/10.1016/j.rsase.2026.102045, 2026.

Raven, P. H., Gereau, R. E., Phillipson, P. B., Chatelain, C., Jenkins, C. N., and Ulloa Ulloa, C.: The distribution of biodiversity richness in the tropics, Sci. Adv., 6, eabc6228, https://doi.org/10.1126/sciadv.abc6228, 2020.

Roy, S., Jensen, E., Saah, A., and Swetnam, T.: samapriya/awesome-gee-community-datasets: Community Catalog (3.5.0) (3.5.0), https://doi.org/https://doi.org/10.5281/zenodo.15587604, 2025.

Sheil, D., Descals, A., Meijaard, E., and Gaveau, D.: Rubber planting and deforestation, Nature, 644, E20–E22, https://doi.org/10.1038/s41586-025-08848-9, 2025.

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., and Lucas, R.: New global forest/non-forest maps from ALOS PALSAR data (2007–2010), Remote Sens. Environ., 155, 13–31, https://doi.org/10.1016/j.rse.2014.04.014, 2014.

Smith, C., Baker, J. C. A., and Spracklen, D. V.: Tropical deforestation causes large reductions in observed precipitation, Nature, 615, 270–275, https://doi.org/10.1038/s41586-022-05690-1, 2023.

Stehman, S. V.: Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes, Int. J. Remote Sens., 35, 4923–4939, https://doi.org/10.1080/01431161.2014.930207, 2014.

https://doi.org/10.5194/egusphere-2026-1401
Preprint. Discussion started: 19 March 2026

Sullivan, A. and Tarrio, K.: Scaling Transparency: Annual, Pan-Tropical Commodity Maps Powered by AlphaEarth Foundations, Google Earth Earth Engine, 2026.

Tropek, R., Sedláček, O., Beck, J., Keil, P., Musilová, Z., Šímová, I., and Storch, D.: Comment on "High-resolution global maps of 21st-century forest cover change," Science, 344, 981–981, https://doi.org/10.1126/science.1248753, 2014.

Tyukavina, A., Potapov, P., Hansen, M. C., Pickens, A. H., Stehman, S. V., Turubanova, S., Parker, D., Zalles, V., Lima, A., Kommareddy, I., Song, X.-P., Wang, L., and Harris, N.: Global Trends of Forest Loss Due to Fire From 2001 to 2019, Front. Remote Sens., 3, 825190, https://doi.org/10.3389/frsen.2022.825190, 2022.

Van De Kerchove, R., Zanaga, D., Li, L., Tsendbazar, N., and Lesiv, M.: WorldCover, v1.0. Product User Manual, VITO, 2020.

Wang, Y., Hollingsworth, P. M., Zhai, D., West, C. D., Green, J. M. H., Chen, H., Hurni, K., Su, Y., Warren-Thomas, E., Xu, J., and Ahrends, A.: High-resolution maps show that rubber causes substantial deforestation, Nature, 623, 340–346, https://doi.org/10.1038/s41586-023-06642-z, 2023.

Xiao, Y., Wang, Q., and Zhang, H. K.: Global Natural and Planted Forests Mapping at Fine Spatial Resolution of 30 m, J. Remote Sens., 4, 0204, https://doi.org/10.34133/remotesensing.0204, 2024.

Xu, P., Tsendbazar, N.-E., Herold, M., De Bruin, S., Koopmans, M., Birch, T., Carter, S., Fritz, S., Lesiv, M., Mazur, E., Pickens, A., Potapov, P., Stolle, F., Tyukavina, A., Van De Kerchove, R., and Zanaga, D.: Comparative validation of recent 10 m-resolution global land cover maps, Remote Sens. Environ., 311, 114316, https://doi.org/10.1016/j.rse.2024.114316, 2024.

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsendbazar, N.-E., Xu, P., Ramoino, F., and Arino, O.: ESA WorldCover 10 m 2021 v200 (v200), https://doi.org/10.5281/ZENODO.7254221, 2022.

Zhao, T., Zhang, X., Gao, Y., Mi, J., Liu, W., Wang, J., Jiang, M., and Liu, L.: Assessing the Accuracy and Consistency of Six Fine-Resolution Global Land Cover Products Using a Novel Stratified Random Sampling Validation Dataset, Remote Sens., 15, 2285, https://doi.org/10.3390/rs15092285, 2023.