WORLD RESOURCES INSTITUTE

TECHNICAL NOTE

Transfer learning to detect natural, monoculture, and agroforestry tree-based systems in Ghana using remote sensing

Jessica Ertel, John Brandt, Rhiannon Rognstad, and Erin Glen

CONTENTS

Technical notes document the research or analytical methodology underpinning a publication, interactive application, or tool.

Suggested Citation: Ertel, J., J. Brandt, R. Rognstad, and E. Glen 2025. "Transfer learning to detect natural, monoculture, and agroforestry tree-based systems in Ghana using remote sensing." Technical Note. Washington, DC: World Resources Institute. Available online at doi.org/10.46830/writn.24.00030.

Abstract

Differentiating between natural and agricultural trees using remote sensing is essential for assessing ecosystem services, commodity-driven deforestation, and restoration progress. Existing approaches focus on identifying a single tree commodity, rather than a system classification that is agnostic to species. This study presents a transfer learning approach to classify tree-based systems, leveraging extracted spatial embeddings from a high-performing neural network to improve classification accuracy in label-scarce environments. We applied a CatBoost classifier to a combination of Sentinel imagery, gray-level co-occurrence matrix texture features, and extracted spatial embeddings to classify four land use classes: natural, agroforestry, monoculture, and other (background). Through comparative modeling and feature selection exercises, we validate performance gains resulting from transfer learning and texture features. Building on previous efforts to model tree extent across the tropics (Brandt et al. 2023), we explore whether the spatial features extracted from Brandt et al.'s (2023) convolutional neural network can be repurposed to help classify tree-based systems. This method is demonstrated for 26 priority districts in Ghana, resulting in a 10-meter resolution land use map for 2020. Our findings suggest the spatial embeddings extracted from Brandt et al.'s (2023) tree cover model offer value beyond their original task and represent a scalable path forward for broader monitoring efforts.

Introduction

Earth observation data that can differentiate between natural and agricultural trees are crucial to effectively assess ecosystem services, commodity-driven deforestation, and restoration progress. While natural and agricultural tree systems appear visually similar in satellite imagery, they contribute in notably different ways to biodiversity and climate change mitigation (Naudts et al. 2016). Accurate and reliable information about the type and extent of these systems is critical to understand land use dynamics and ensure that national restoration targets are credibly monitored and reported. In Ghana, distinguishing between natural and agricultural tree systems remains uniquely challenging because of multiple factors. These factors include high spectral similarity between certain systems; the small minimum mapping unit required to capture heterogenous, smallholder agricultural landscapes; and challenges such as persistent cloud cover and haze.

We present a transfer learning approach to classify tree-based systems. This approach leverages extracted spatial embeddings from a high-performing neural network to improve classification accuracy in label-scarce environments. We refer to label scarcity as the limited availability of consistent, high-quality training labels, particularly due to ambiguity in class definitions for complex systems like agroforestry. Building on previous efforts to model tree extent across the tropics (Brandt et al. 2023), we explore whether the spatial features extracted from Brandt et al.'s (2023) convolutional neural network (CNN) can be repurposed to support a distinct but related downstream task. We piloted the approach in Ghana, as part of a multi-year restoration monitoring partnership between World Resources Institute (WRI) and Ghana's Environmental Protection Authority (EPA).

To test our hypothesis, we applied a gradient boosting classification algorithm (CatBoost) to a combination of Sentinel-2 images, spatial embeddings, and gray-level co-occurrence matrix (GLCM) texture features derived from Sentinel images. Referencing existing efforts from Ghana's remote sensing community (Abu et al. 2021; Ashiagbor et al. 2020; Benefoh et al. 2018; Numbisi et al. 2019), we determined a set of approaches and discrete criteria to capture our target land use classes, which reflect the structure of different tree-based systems. We assessed whether using extracted embeddings and GLCM texture features enhances model learning. We performed benchmarking against a standard deep learning technique and a CatBoost model trained without embeddings. Through a feature selection exercise, we validated the contribution of texture and embedding features to model performance.

The transfer learning method was demonstrated across 26 priority administrative districts that Ghana's EPA identified. The final product is a 10-meter (m) resolution land use map for 2020 that distinguishes between tree-based systems, with the goal of informing future efforts to distinguish natural and planted vegetation using remote sensing.

Challenges

In the Ghanaian context, three key challenges affect the ability to develop accurate land use maps that distinguish between natural and planted tree systems. The first challenge is the limited availability of high-quality optical satellite imagery. Persistent cloud cover and haze, which reduce the accuracy of optical imagery, can cause low-quality images. While the dry season (December to March) may offer more cloud-free imagery, this season brings Harmattan haze, which is caused by dust-laden winds originating from the Sahara Desert. Harmattan winds can create a haze effect in optical remote sensing imagery, due to the winds suspending dust particles in the atmosphere for extended periods of time. However, some studies suggest the spectral distinction between certain vegetation (such as cashew and forest) can become more pronounced during the dry season, making the vegetation easier to differentiate (Pereira et al. 2022). This seasonal tradeoff between imagery availability and class separability underlines how the timing of image acquisition can have a sizable impact on the detectability and classification of certain tree systems.

The second challenge is deciphering trends within Ghana's highly diverse agricultural landscape. Half of smallholder agricultural systems in Africa are smaller than 1 hectare (ha) (Estes et al. 2022). Small-scale oil palm plantations vary in size across the tropics, but they are noticeably smaller in Ghana, where they range from 0.5 to 5 ha (Chamberlin 2008). To classify small-scale agricultural activity, a small minimum mapping unit is necessary for remote sensing-based datasets to capture the high spatial variability in land use dynamics, irregular field boundaries, and heterogenous smallholder characteristics. These characteristics increase the potential for error when creating land use maps.

The third challenge in distinguishing between natural and planted tree systems in Ghana relates to the structural and spectral similarity between shaded cocoa plantations and open canopy forests. Certain agroforestry systems, such as silvopastoral systems, boundary plantings, home gardens, and woodlots (Daniel et al. 2018), can be easier to detect due to their clear and distinguishable spatial patterns. In contrast, agricultural crops sit under existing or intentionally planted shade trees in shaded cocoa plantations, resulting in multi-strata canopy structures that closely resemble open canopy forests in satellite imagery. This structural and spectral similarity makes these plantations difficult to isolate from neighboring forests and other tree crops using optical imagery alone (Abu et al. 2021; Ashiagbor et al. 2020; Benefoh et al.

2018). The presence of the same tree species in both shaded cocoa plantations and open canopy forests also accounts for these spectral similarities. Farmers consistently rank *Terminalia superba* and *Terminalia ivorensis* among the top preferred shade tree species across all cocoa production stages (Asigbaase et al. 2025). These are also two dominant species in Ghana's forests, especially in the moist semi-deciduous ecological zone (Hall and Swaine 1981). Ghana is uniquely known within the remote sensing community for this class separability problem, with studies finding negligible differences between the spectral signature of Ghana's natural forests and full sun or shaded cocoa systems compared to the same systems in Ecuador (Filella 2018). The spectral similarities between agroforestry and open canopy natural systems result in low classification accuracy among existing land use and land cover maps of Ghana (Ashiagbor et al. 2020). These three remote sensing challenges underscore the difficulty in using traditional classification approaches to map land use in this country.

Existing approaches

Due to its strong performance in label-scarce environments, transfer learning is gaining traction in the machine learning field. Transfer learning techniques involve applying the knowledge from a pre-trained algorithm to a separate but related task. In a remote sensing context, CNN-based approaches map input optical imagery to output map classes using spatially explicit feature maps at various spatial resolutions. These features transform the input into a high dimensional feature space with rich representations of the input imagery, which can be used as a starting point for other downstream tasks via transfer learning. Transfer learning approaches can reduce expensive and time-consuming training steps, while also minimizing the need for a large number of labeled samples to achieve strong performance.

Several types of transfer learning techniques have been applied in the field of remote sensing for land cover mapping (Ma et al. 2024). A typical approach to transfer learning involves adapting a pre-trained algorithm by fine-tuning its parameters and re-training layers to perform a new prediction task. However, transfer learning can also involve extracting learned features from new data and using the features as input to a new model. Alem and Kumar (2022) explore how a bottleneck feature extraction technique applied to three pre-trained models can improve classification performance for land use and land cover (LULC) classification in remote sensing images. Hamrouni et al. (2021) illustrate how adapting a local classifier with new relevant training samples can help classify poplar plantations at scale in Sentinel-2 imagery.

Several recent remote sensing-based efforts have made significant progress in mapping agroforestry and plantation systems across Ghana and the broader West Africa region. However, most research has focused on identifying a single tree commodity, rather than a system classification that is agnostic to species. To track the rapid expansion of oil palm production in Ghana, Abramowitz et al. (2023) applied a Random Forest classifier to Sentinel imagery to differentiate closed-canopy industrial from smallholder oil palm. Several studies have explored methods to distinguish natural from agroforestry land uses at a national scale using freely available Sentinel-1 and Sentinel-2 imagery. Abu et al. (2021) applied a Random Forest model to detect cocoa plantation encroachment into protected areas in Côte d'Ivoire and Ghana, using GLCM texture features to detect small-scale cocoa farms. Ashiagbor et al. (2020) combined a pixel and object-based approach to distinguish agroforestry cocoa from forest and other land use classes in Ghana. Numbisi et al. (2019) also explored using texture features to discriminate cocoa agroforests from transition forests in Cameroon. Benefoh et al. (2018) applied an image-fusion technique to isolate cocoa plantations from other vegetation and estimate cocoa-led deforestation. By attempting to detect and classify broader systems rather than a single commodity type, we build on the technical and definitional foundations these researchers established, particularly in using GLCM texture metrics and operational remote sensing-based definitions of agroforestry.

In 2023, following extensive model development and research, Brandt et al. (2023) released the Tropical Tree Cover (TTC) dataset, a Sentinel-based tree cover product that covers 2020. The dataset was generated using a multi-temporal CNN that achieved 94 percent overall accuracy across 4.35 billion ha in the tropics. The dataset shows consistently high performance across multiple land types, including areas with open canopy forest, dryland, and cropland. As a 10-m resolution product, TTC outperformed existing 30-m resolution datasets in detecting small patches of tree cover, particularly in complex and fragmented landscapes. Given this strong performance, we hypothesized that the spatial embeddings extracted from a model adept at tree detection across varied land cover types could be repurposed to inform a separate but related exercise. One significant shortcoming of the TTC dataset is the lack of differentiation between natural and agricultural tree cover. This distinction is critical for restoration monitoring applications, as gains in tree cover cannot be meaningfully assessed without understanding whether they result from a successful restoration intervention or agricultural expansion. The inability of most global- or tropical-scale tree cover products to differentiate between natural and planted trees (Fagan et al. 2022) limits their relevance for restoration monitoring. This study explores the potential to use embeddings from the TTC model to inform tree-based system classification in Ghana. By using a "light touch" transfer learning approach that integrates TTC spatial embeddings in a new classification pipeline, we extend the TTC dataset's usefulness (Brandt et al. 2023), while ensuring methodological compatibility to lay the foundation for TTC's future change detection applications.

Methods

Study area

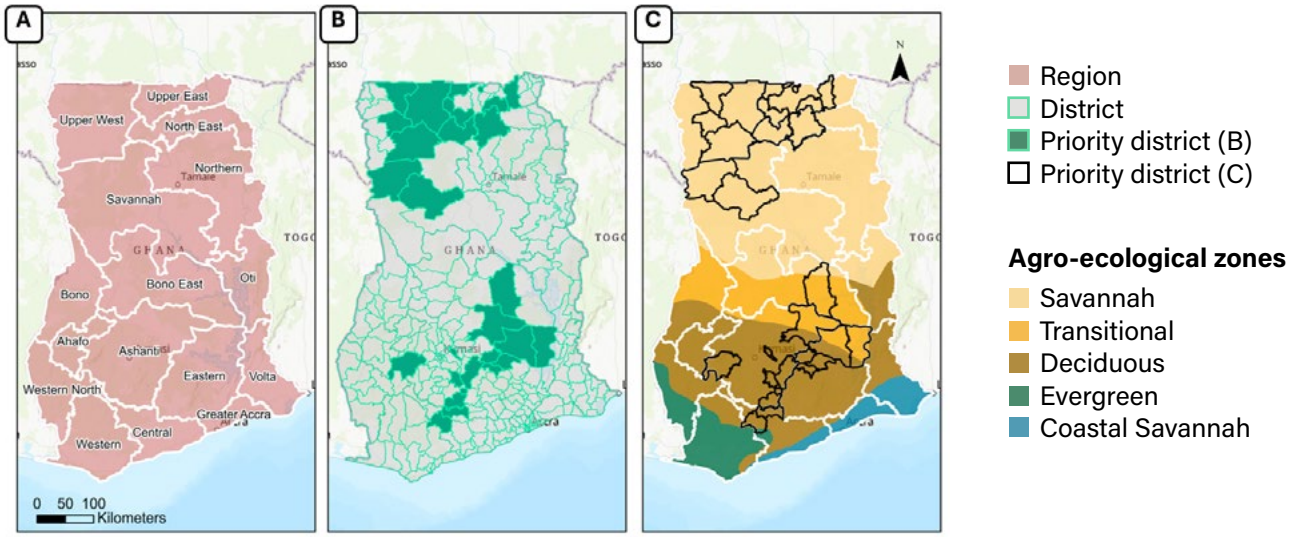
We conducted this study across 26 administrative districts, spanning 8 of Ghana's 16 regions. The 26 administrative districts form part of an ongoing effort by Ghana's EPA to monitor restoration progress and streamline integrated landscape management plans. The target districts were selected as priority areas based on their locations within biological corridors, as well as land degradation and illegal mining pressures (Pers. Comm. 2023a). The following districts were included: Adansi South, Asante Akim South, Assin North, Atwima Mponua, Bawku West, Bosome Freho, Builsa South, Daffiama Bussie Issa, Juaben Municipal, Kassena Nankana West, Kwahu Afram Plains North, Kwahu Afram Plains South, Kwahu East, Kwahu South, Kwahu West, Mamprugu Moagduri, Sawla-Tuna-Kalba, Sekyere Afram Plains North, Sene West, Sissala East, Sissala West, Talensi, Twifo Atti-Morkwa, Wa East, West Gonja, and West Mamprusi Municipal.

Ghana's landscape consists of five general agro-ecological zones: Savannah, Transitional, Deciduous, Evergreen, and Coastal Savannah (Abbam et al. 2018) (Figure 1). This study focuses on two main groups of districts: the southern group and the northern group. The southern group sits between latitudes 5°N and 8°N and stretches from the Deciduous zone to the Savannah-forest Transitional zone. The northern group sits between latitudes 8°N and 11°N and stretches from the Savannah-forest Transitional zone to the Savannah zone. These two district groups cover several forest and game reserves, protected areas, and wildlife corridors, as well as a highly dynamic and evolving agricultural landscape.

Southern Ghana, dominated by deciduous and transitional forests, is well-suited to cocoa trees, which prefer a high average annual rainfall, moderate temperatures, and low climatic variability (Osei-Gyabaah et al. 2023). Other tree crops, such as oil palm, cashew, and shea, are commonly grown in the cocoa belt, thus may be intercropped with cocoa. Oil palm and rubber monoculture plantations are generally concentrated in the South and Southwest, due to the favorable climatic conditions.

Moving northward, the climate becomes drier, transitioning into savannah woodland. The Savannah zone is characterized by lower rainfall, higher temperatures, and greater climatic variability (Abbam et al. 2018), which offer less favorable conditions for cocoa trees. Cashew, acacia, baobab, and shea are common drought-resistant trees in northern Ghana that are often used for agroforestry (Moomen et al. 2024). Agriculture in the region is mainly rain-fed and faces challenges during the dry season from November to March, also called the Harmattan. Strong winds from the Sahel bring dust, dry conditions, and extreme temperature fluctuations, increasing wildfire risks and disrupting farming.

Figure 1 | **Study area map**



Notes: Study area (priority districts) in Ghana: A) Ghana's 16 official regions are outlined in white. B) The 26 priority districts are illustrated with green fill, while all other districts are outlined in light green. C) The agro-ecological zones of Ghana, including the Savannah, Transitional, Deciduous, Evergreen, and Coastal Savannah zones.

Source: Adapted from Abbam et al. 2018.

Land use definitions

We distinguished four land use classes: monoculture, agroforestry, natural, and other (background) (Table 1). This approach differs from traditional land cover or commodity-focused mapping exercises by focusing on a tree-based system definition, rather than an individual crop or simple canopy presence. We focus on the classification of tree-based systems and do not include food crops, such as maize, rice, cassava, or plantain, in the agricultural classes (monoculture or agroforestry). Photo interpreters distinguished the classes using key characteristics, including tree cover, canopy structure, and spatial arrangement.

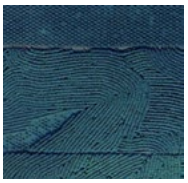
Training data

Collection and photointerpretation

To pinpoint landscapes suitable for collecting training data for our designated classes, we referenced existing forest and tree crop datasets. We referenced four different vector datasets and one field survey in this process.

First, the Spatial Database of Planted Trees (SDPT) (Richter et al. 2024) is a living database of spatial information about the locations of planted forests and tree crops throughout the world. It includes data that national governments, non-governmental organizations, researchers, or a combination of sources provided. In Ghana, SDPT contains comprehensive information about oil palm and rubber plantations, which was used to sample training points for the monoculture class. Second, the West Africa Cocoa (WAC) dataset contains cocoa

Table 1 | **Definitions used for the mapped land use classes and example tree crops**

LAND USE CLASS	DEFINITION	LABELING CRITERIA	EXAMPLE
Other (background) 	Areas containing less than 10% tree cover.	Surfaces appear built, bare, or sparsely vegetated.	Typically represented by human settlement areas, bare land, grasslands, water, and mining areas.
Monoculture 	Agricultural tree systems with a single canopy stratum and no productive or managed understory.	Crown shape and size is relatively consistent, and a distinct planting pattern can be detected.	Typically represented by oil palm and rubber plantations.
Agroforestry 	Agricultural tree systems with natural or planted trees that form 2 distinct canopy strata.	Crown shape and size vary. Area shows signs of disturbance (harvest, regrowth) across years.	Typically represented by shadow systems, like cocoa grown under shade trees or shea serving as shade trees for crops.
Natural 	Primary and secondary woody vegetation that has not been deliberately planted or managed by humans.	No planting pattern or distinct spatial organization. Areas do not exhibit disturbance across years.	Typically represented by forest reserves and protected areas.

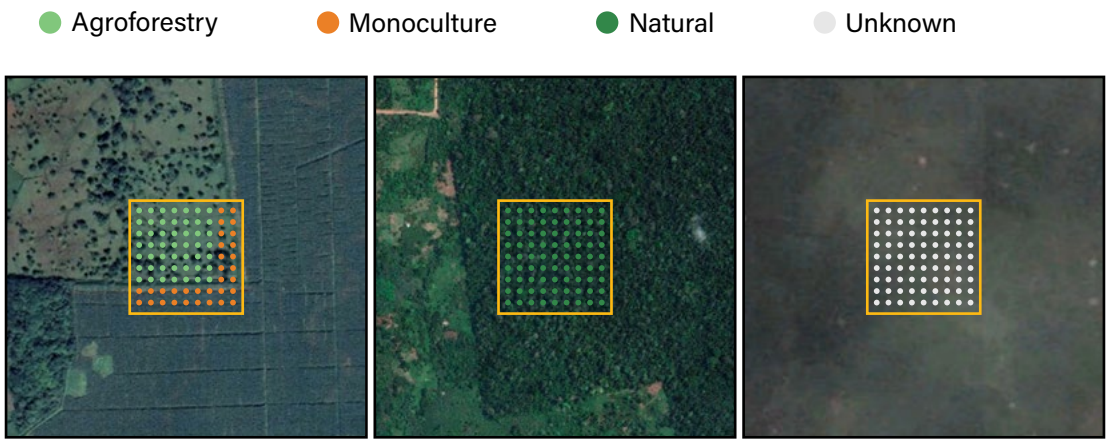
Source: Ashiagbor et al. 2020; WRI authors.

polygons across Côte d'Ivoire and Ghana. Leading cocoa and chocolate companies sourcing directly from these locations supplied the polygons (Schneider et al. 2023). Since cocoa in Ghana is often grown in shaded agroforestry systems, WAC polygons were used to sample training points for the agroforestry class. Third, the World Database on Protected Areas (WDPA) contains spatially referenced information about legally protected areas, including nature reserves, wilderness areas, national parks, and gazetted forests (Bingham et al. 2019). WDPA forest reserve polygons in Ghana were used to sample training points for the natural tree class. Fourth, Ghana's Resource Management Support Centre (RMSC) provided field-collected samples that contained locations for shea, mango, cashew, and citrus tree crops (Pers. Comm. 2023b). RMSC collected the field data between May and July 2022 across five basins in Ghana: Black Volta, Pra, Sene, Tano, and White Volta.

To derive training data, we first calculated the centroid of each polygon in our reference forest and tree crop datasets. Then we performed a secondary photointerpretation step to ensure the samples were assigned an accurate label for 2020, our year of interest. Since the reference vector datasets came from various years, the photointerpretation step was critical to address any temporal inconsistencies (land use change) and ensure that labels aligned with our land use definitions. As an example, WDPA forest reserve polygons were referenced to gather training samples for the natural class; however, Ghana permits forest plantations to be established within reserves, meaning a teak or eucalyptus plantation could likely be present in these areas (Forestry Commission 2021). The photointerpretation step ensured we accurately labeled these samples in the monoculture, rather than natural, land use class.

Photointerpretation surveys were prepared for per-pixel annotation using Collect Earth, an online tool (Saah et al. 2019) (Figure 2). This study's authors performed the annotation. We developed a labeling guide in collaboration with government stakeholders, photointerpretation specialists, and subject matter experts. We drew on existing literature, field data, and spatial datasets to ensure consistency and accuracy. Only 10-m resolution Sentinel imagery was available on the Collect Earth platform, so annotators also referenced high-resolution imagery, accessed through Google Earth Pro, to support the photointerpretation exercise. Google Earth contains a large collection of imagery (satellite, aerial, three-dimensional, and street view) with varying resolutions, ranging from 30 m per pixel to 10–15 centimeters (cm) per pixel in some areas (Google n.d.). The Collect Earth surveys overlaid a 14 x 14 gridded plot on the centroids. Thus each point was spaced at 10-m intervals to create a 140 x 140-meter plot (Figure 2). Annotators classified each Sentinel-2 pixel in the plot into a land cover class according to the definitions outlined in Table 1. Annotators used information about crown shape and size, spatial patterns, and contextual indicators—like proximity to infrastructure and disturbance over time—to inform their classifications. Since some samples remained extremely challenging to classify due to poor image availability and quality,

Figure 2 | **Photointerpretation schema**



Notes: Example of photointerpretation plots across typical scenes of agroforestry, monoculture, and natural systems. The scale, number, and design of gridded plots are illustrated for visual effects and are not consistent with the schema used in this study.

Source: WRI authors.

an "unknown" label category was included in the Collect Earth survey to capture these instances. High-resolution imagery was considered unavailable if there were no images for the plot between 2017–2022, or if images could not be interpreted due to haze or cloud cover. In these scenarios, the entire plot was marked "unknown" and dropped from the training batch.

Quality assessment

Improving label quality is one of the most important means to increase model accuracy (Estes et al. 2022). Considering the complexity of photointerpretation and likeness between certain land use classes, we conducted a data and label assessment to minimize annotation error that may have been introduced in the training dataset. Cleanlab is a tool that algorithmically detects data and label issues in machine learning datasets (Northcutt et al. 2021). Cleanlab automatically detects label, outlier, and duplicate issues using a confident learning framework. Label issues allow users to identify where there might be discrepancies in their annotations. Confident learning can detect these issues by identifying examples where the predicted class consistently disagrees with the annotator's given label, even when the model is confident. Duplicate issues refer to two or more samples (pixels) that exhibit extreme similarity, relative to the rest of the dataset. Near duplicates can negatively impact model generalization and lead to overfitting, because duplicative examples can be unintentionally emphasized during model training. Outlier issues refer to samples that are extremely different from the rest of the dataset, such as rare or anomalous instances. In our case, a large portion of the outlier issues belonged to pixels in the other (background) class, which is reasonable considering this is the only class that includes non-vegetation pixels and encompasses urban land covers. The Cleanlab assessment was performed at the pixel scale, meaning each pixel was individually scored for label, outlier, and duplicate issues. We reviewed the assessment results at the plot level (for all 196 pixels in each plot) to determine if the entire plot introduced confusion into the model and should be removed. This analysis allowed us to identify 28 plots (totaling 5,488 pixels) where all 196 pixels in the plot were affected by label issues. We also identified 27 plots where label issues affected a majority of the pixels (at least 180 of 196). We chose to drop all 55 plots from the training data. We decided against dropping samples with outlier issues, as they had the potential to contain important information for the other (background) class. The Cleanlab assessment also flagged six full plots with near duplicate issues. We chose to drop all these plots, except the example with the lowest near-duplicate score. The near-duplicate score assigned to the sample was determined by its proportional distance to its nearest neighbor (Northcutt et al. 2021). In most cases, the flagged plots came from the same training data batches, signaling consistency in the type of data that introduced confusion or error.

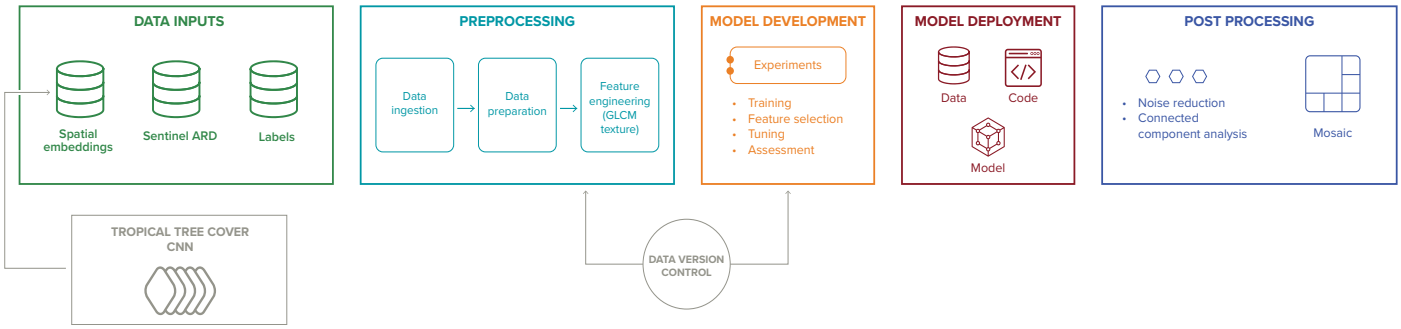
Final training dataset

To produce the final training dataset, we removed 242 plots that were labeled "unknown," 118 plots that had no analysis-ready Sentinel imagery for 2020, and 60 plots that were flagged during the Cleanlab assessment. The final training dataset was comprised of 976 plots, containing 191,296 pixels. The proportion of points in each class was 18.6 percent monoculture, 37.8 percent agroforestry, 16.9 percent natural, and 26.6 percent other (background).

Feature generation

The predictive features for the classification task included Sentinel-1 radar and Sentinel-2 optical imagery for 2020, spatial embeddings extracted from the CNN underpinning the TTC dataset, and texture features. Texture features were extracted from Sentinel-2 imagery using a GLCM analysis

Figure 3 | **Machine learning pipeline**



Notes: The machine learning pipeline takes Sentinel ARD, spatial embeddings, and labels as input, then performs preprocessing, classification, and postprocessing steps. A model development stage is also represented, including hyperparameter tuning, feature selection, and evaluation using a version control system to manage experiments.

Source: WRI authors.

method. Each training sample consisted of 13 optical and radar Sentinel bands, 65 embeddings, and 16 GLCM texture features combined as a 94-band input. These 94 input features were fed into a classification pipeline to assign a land use class to each pixel. Figure 3 shows the machine learning pipeline, including pre- and post-processing steps. This section further details how the 94 input features were produced. The authors performed all data processing, preparation, and classification steps using Python and associated geospatial libraries.

Sentinel analysis ready data

Sentinel-1 and Sentinel-2 images were acquired through the Sentinel Hub Application Programming Interface. The 13 Sentinel analysis ready data (ARD) bands contained four Sentinel-2 bands at a 10-m resolution, six Sentinel-2 bands at a 20-m resolution, two Sentinel-1 bands at a 10-m resolution, and a digital elevation model (DEM) from the Shuttle Radar Topography Mission (Farr and Kobrick 2000). We selected the best monthly 10- and 20-m bottom-of-atmosphere (L2A) Sentinel-2 image based on cloud cover. We then accessed the image angle from the Sen2Cor scene classification data (Louis et al. 2016). Finally, we acquired the monthly composite of vertical transmit/vertical receive (VV) and vertical transmit/horizontal receive (VH) ground range detected (GRD) Sentinel-1 observations for all of Ghana throughout 2020. Additionally, we generated monthly ARD composites of Sentinel-2 and Sentinel-1 imagery from the raw imagery available prior to analysis. We created all ARD using Python's scientific computing libraries: NumPy 1.18.0 (Harris et al. 2020) and SciPy 1.7.3 (Virtanen et al. 2020).

The Sentinel-1 ARD composites were comprised of the monthly median of the decibel gamma naught VV and VH GRD data, processed with no speckle filtering. Speckle filtering is typically not applied to temporally aggregated GRD data, since this filtering is a composite of multiple acquisitions and already spatially aggregated; applying this step would overcorrect for noise.

While Brandt et al. (2023) described the ARD creation process in more detail, we provide an overview of how to prepare Sentinel-2 ARD composites using the following steps applied independently to 6 x 6 kilometer tiles:

- Clouds and cloud shadow identification following Candra et al.'s (2020) and Qiu et al.'s (2019) methods
- Bidirectional reflectance distribution factor correction
- Cloud and shadow removal and gap-filling with relative radiometric normalization and inverse distance weighted interpolation
- Temporal gap filling using the Whittaker smoother (Eilers 2003)
- Super-resolution of 20-m bands to 10-m resolution with a CNN (Lanaras et al. 2018)

The analysis-ready mosaics of Sentinel-2 imagery contained a single cloud-free image for each month of the year. We converted the 12 cloud-free images into a single median image, which showed good results for tropical regions in Brandt et al. (2023). During model training, we augmented the training dataset to improve the classifier's ability to generalize across temporal and geographic scales. This involved calculating a single median image on a different random subset of monthly images for each training sample. For example, one sample (a median image) could be derived from images in January, March, April, and December, while another sample could be derived from images in May, June, July, and October. This method allowed us to train the model on many median composites of images containing varied seasonal characteristics. This standard augmentation practice reduced model overfitting to specific characteristics of the median image. For model inference, we directly applied the trained model to the median of all monthly images.

We designed a haze correction method during ARD processing to combat the effects of Harmattan haze, which generated artifacts along the boundaries of neighboring tiles during model inference. We found the model was particularly sensitive to haze-affected imagery, resulting in class confusion or inconsistent predictions at tile borders. To address this issue, we smoothed boundaries between tile neighbors using Wang et al.'s (2020) methods. This strategy was only applied to tiles that showed visual inconsistency in the final map product.

Feature extraction

We performed feature extraction to access high-dimensional representations, or spatial embeddings, from the TTC model. These intermediate frozen layers of the neural network were isolated and used in a separate supervised classification task. While the exact semantics of these features were not individually interpretable, we hypothesized they could capture critical information about canopy texture, density, or shape that could inform related tasks. Tapping into the expertise of an existing tree cover model allowed us to leapfrog over the initial stage of detecting trees, and instead focus on classification.

The TTC model is a U-Net that takes a bottleneck approach, whereby a 10-m resolution input is encoded to coarser resolution features and then reconstructed stepwise into a 10-m resolution output. The model generates features for 16 separate layers at spatial resolutions of 10, 20, 40, 80, and 160 m. For the purposes of transfer learning, we selected the high resolution (10 m) feature layers from both the U-NET's encoder and decoder. A total of 65 spatial embeddings were extracted for input into the land use classification pipeline.

Texture features

The feature generation process involved computing GLCM texture images. A GLCM is a statistical method that characterizes texture that considers the spatial relationship of pixels. Image texture is derived from the frequency with which pairs of pixels, with specific values and in specific spatial relationships, occur in an image (MATLAB Help Center 2025). In the context of land use classification, texture features can improve classification accuracy because they focus on spatial relationships in a landscape, as opposed to the spectral properties of a single ground unit. GLCM texture-based mapping has proved effective in several efforts to map heterogenous agricultural landscapes, detect shrub crops under forest canopies, and distinguish crops from native forest (Abu et al. 2021; Ashiagbor et al. 2020; Numbisi et al. 2019; Pereira et al. 2022; Sari et al. 2022). Selecting appropriate texture features and parameters to perform the GLCM calculation can be difficult because of the distinct composition of LULC classes and the general lack of guidelines in the field (Mishra et al. 2019). As such, the existing literature informed our parameter selection (Table 2). The authors wrote the GLCM calculations in Python to derive four texture statistics (dissimilarity, correlation, homogeneity, and contrast) for each Sentinel-2 band (blue, green, red, and near infrared [NIR]). We selected a 5 x 5 moving window with one pixel displacement to the right (zero degree angle). These parameters, including selecting the moving window size, balanced computational efficiency with the need to capture spatial patterns and distinguish texture characteristics within plots.

Model architecture and selection

To select the appropriate algorithm for this supervised per-pixel classification task, we tested several commonly used machine learning approaches for creating land use maps in the remote sensing field (Khatami et al. 2016). The tested algorithms included Random Forest, support vector machine, XGBoost, CatBoost, and Light Gradient Boosting Machine (LightGBM). Baseline comparisons of the receiver operating characteristic (ROC) curve and the precision recall (PR) curve were used to determine which model achieved the highest performance. Generally speaking, the performance did not vary drastically across machine learning models with the subset of training data available at the time of testing, which was approximately 50 percent of the final training dataset used in this study. CatBoost classifier (Dorogush et al. 2018), a gradient boosting algorithm, was ultimately selected due to its comparatively higher scores across ROC and PR curves. CatBoost is known for its ability to handle imbalanced classes and resist overfitting better than existing gradient-boosted decision trees like XGBoost and LightGBM (Bentéjac et al. 2021; Dorogush et al. 2018). We identified the optimal combination of features through a feature selection exercise, which is discussed in the following section. We assessed model performance throughout the research period to understand the impact of additional training batches or adjustments in the pre-processing pipeline. We gleaned insights through a review of multiple accuracy metrics, but we chose balanced accuracy as the main metric for evaluation given the imbalanced nature of the training dataset.

Table 2 | **Literature review of GLCM texture properties used for land use classification**

SOURCE	REGION	COMMODITY	TEXTURE FEATURE	WINDOW	ANGLES
(Numbisi et al. 2019)	West Africa	Cocoa	4 GLCM features (contrast, entropy, correlation, variance)	5 x 5	(0°, 45°, 90°, and 135°)
(Abu et al. 2021)	West Africa	Cocoa	18 GLCM features. GLCM features were not in the top 10 selected.	3 x 3	(0°, 45°, 90°, and 135°)
(Descals et al. 2019)	Sumatra	Oil palm	GLCM sum average (band 11), correlation (band 4), correlation (enhanced vegetation index [EVI])	10 x 10 and 30 x 30	unspecified
(Pereira et al. 2022)	Guinea Bissau	Cashew	18 GLCM features (sum average [SAVG] contrast, dissimilarity, inertia)	unspecified	unspecified
(Sari et al. 2022)	Vietnam	Oil palm, rubber	4 GLCM features (mean, variance, homogeneity, and contrast)	7 x 7	unspecified
(Maskell et al. 2021)	Vietnam	Coffee	8 GLCM features	5 x 5 (tested 3, 5, and 7)	(0°, 45°, 90°, and 135°)

Notes: A literature review of GLCM texture properties helped identify the correct parameters for land use classification.
Sources: Abu et al. 2021; Descals et al. 2019; Maskell et al. 2021; Numbisi et al. 2019; Pereira et al. 2022; Sari et al. 2022.

Tuning and feature selection

We performed a feature selection and hyperparameter optimization exercise to narrow down the most influential features for land use classification. Training with only the most important features reduced overfitting and training time, while improving accuracy by removing misleading data and noise.

We used CatBoost's tree-based feature importance method to incrementally evaluate and remove features from the training set until the accuracy metric no longer improved. Recursive feature elimination was performed on the 81 embeddings and texture features, not including the 13 Sentinel ARD bands. The Sentinel bands were retained in the experiment based on their demonstrated predictive value (Brandt et al. 2023). This exercise narrowed the feature set down to 40 features that explained 85–90 percent of the model's performance. Identifying the best set of hyperparameters for a given dataset can be challenging, so we performed an objective search across a variety of values within a search space by using Scikit-learn's RandomizedSearchCV (Pedregosa et al. 2011). The random search enabled us to discover hyperparameter combinations that allowed us to find a good model configuration as the training dataset grew over time.

Accuracy assessment

To evaluate the final classification model, we labeled an independent set-aside validation dataset and calculated a confusion matrix to obtain accuracy metrics by land use class. We collected 1,208 validation points using a stratified random sample by land cover class. To determine the sample size for each category, we considered the mapped proportion of each class and the minimum count of points to produce a sufficiently precise accuracy estimate (FAO 2016). A cushion was added to account for the possibility of dropped samples during photointerpretation due to low image availability or quality. A minimum distance of 10 m was enforced between training and validation plots.

We labeled validation points using the same photointerpretation criteria as training points. The only difference in this process was the use of a consensus labeling approach, whereby two annotators labeled all samples, and a third party reviewed any disagreements. This approach's goal was to minimize label errors and reduce individual annotator bias, given the annotator's role in labeling both training and validation datasets. Similar to the process for labeling training data, an "unknown" label category was included in the validation survey to capture instances where high resolution imagery was unavailable between 2017–2022. If two annotators agreed on an "unknown" label, it was dropped from the validation dataset.

Results

Map accuracy and area assessments

After fitting and optimizing the CatBoost classifier, we evaluated it against the set-aside validation dataset. The model achieved an overall accuracy of 65.27 percent plus or minus 2.69 percent. The user's and producer's accuracy for each land use class was estimated at the 95 percent confidence interval (Table 3). The CatBoost classifier's accuracy was assessed using a confusion matrix, which highlighted where misclassifications occurred for each class (Figure 4).

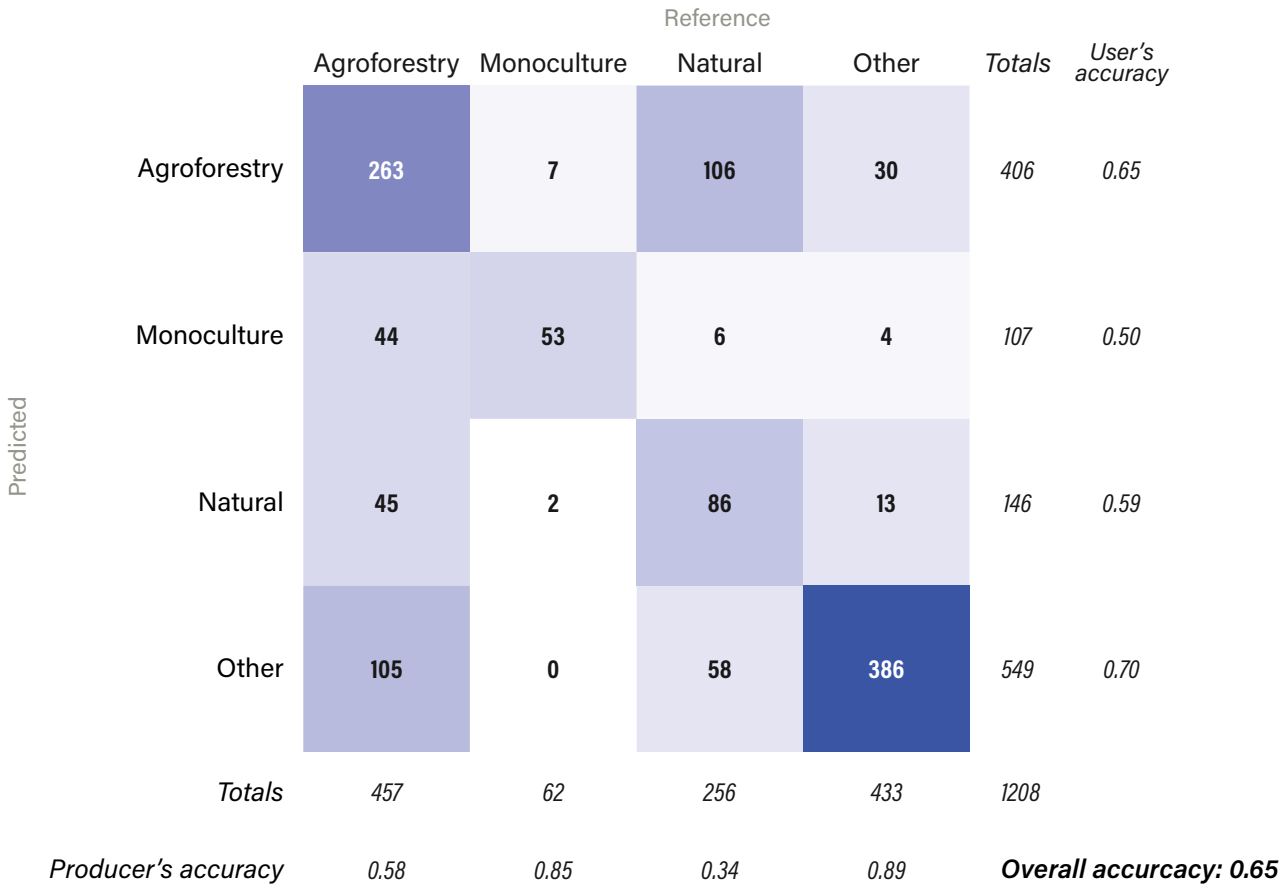
We used the final CatBoost model to generate land use maps and area assessments for all 26 priority districts. Figure 5 shows the map results for the Twifo Atti-Morkwa district in Central Region, Ghana, calling out detail in a selected region with various production systems. The map communicates the predominance of the agroforestry class, which borders two protected areas. The inset illustrates the landscape's heterogeneity, where high-resolution image interpretation indicates large industrial oil palm plantations sit alongside monoculture smallholdings intercropped with cocoa and other commercial tree crops.

Table 3 | **Accuracy assessment of the land use map**

LAND USE CLASS	USER'S ACCURACY	PRODUCER'S ACCURACY	OVERALL ACCURACY
Other	70.24% ± 3.76%	89.02% ± 2.86%	65.27% ± 2.69%
Monoculture	49.73% ± 9.73%	85.01% ± 8.54%	
Agroforestry	64.75% ± 4.75%	57.52% ± 4.53%	
Natural	58.70% ± 7.97%	33.59% ± 5.96%	

Source: WRI authors.

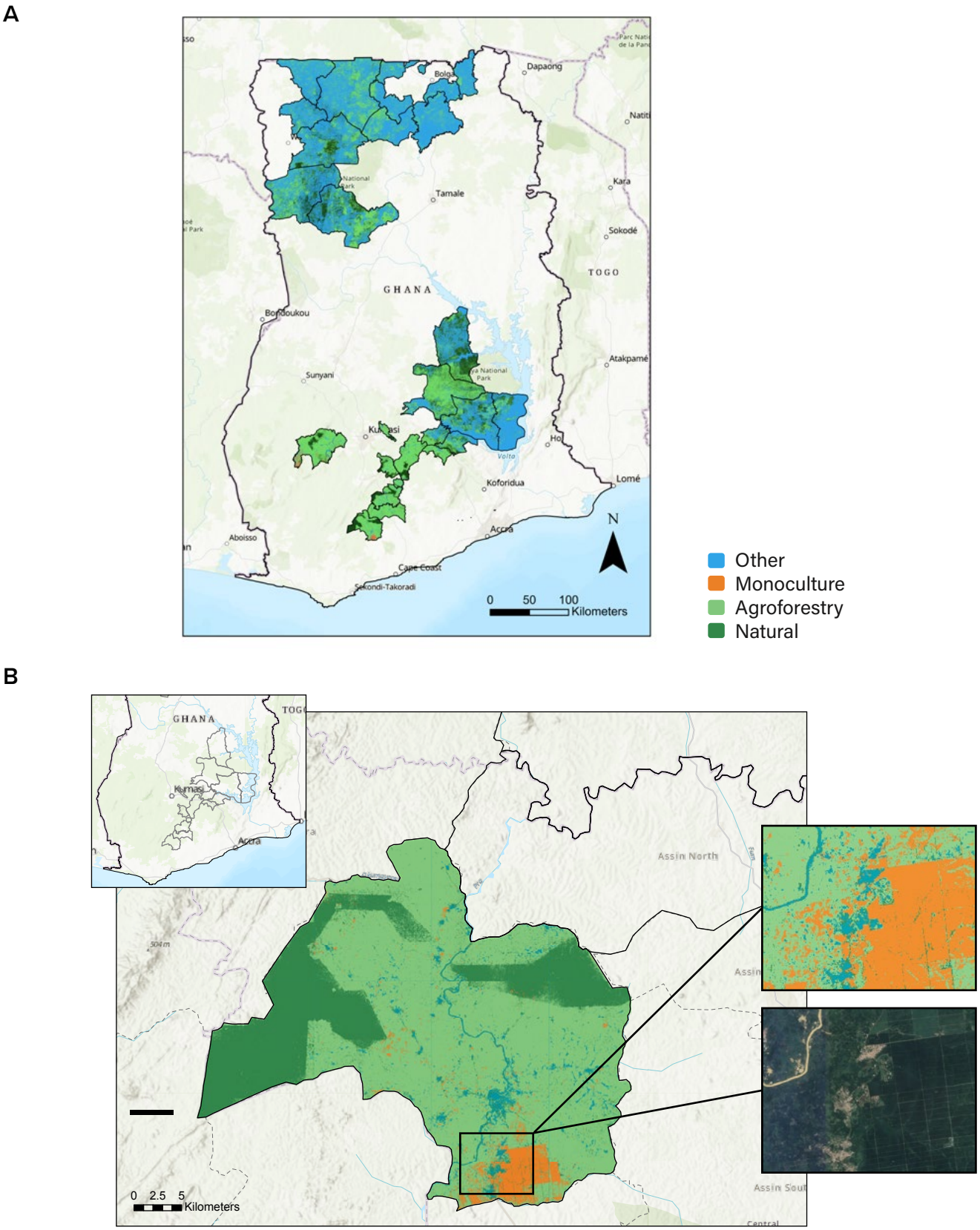
Figure 4 | **Confusion matrix**



Notes: The confusion matrix corresponds to the final CatBoost model's performance on the independent validation dataset. The results illustrate which classes saw the most confusion.

Source: WRI authors.

Figure 5 | **Map results**



Notes: A) Pixel-based land use map for 26 districts using the CatBoost model. B) District-scale map results for Twifo Atti-Morkwa, home to Twifo Oil Palm Plantations, illustrating agroforestry, monoculture, natural, and other (background) land use classes.

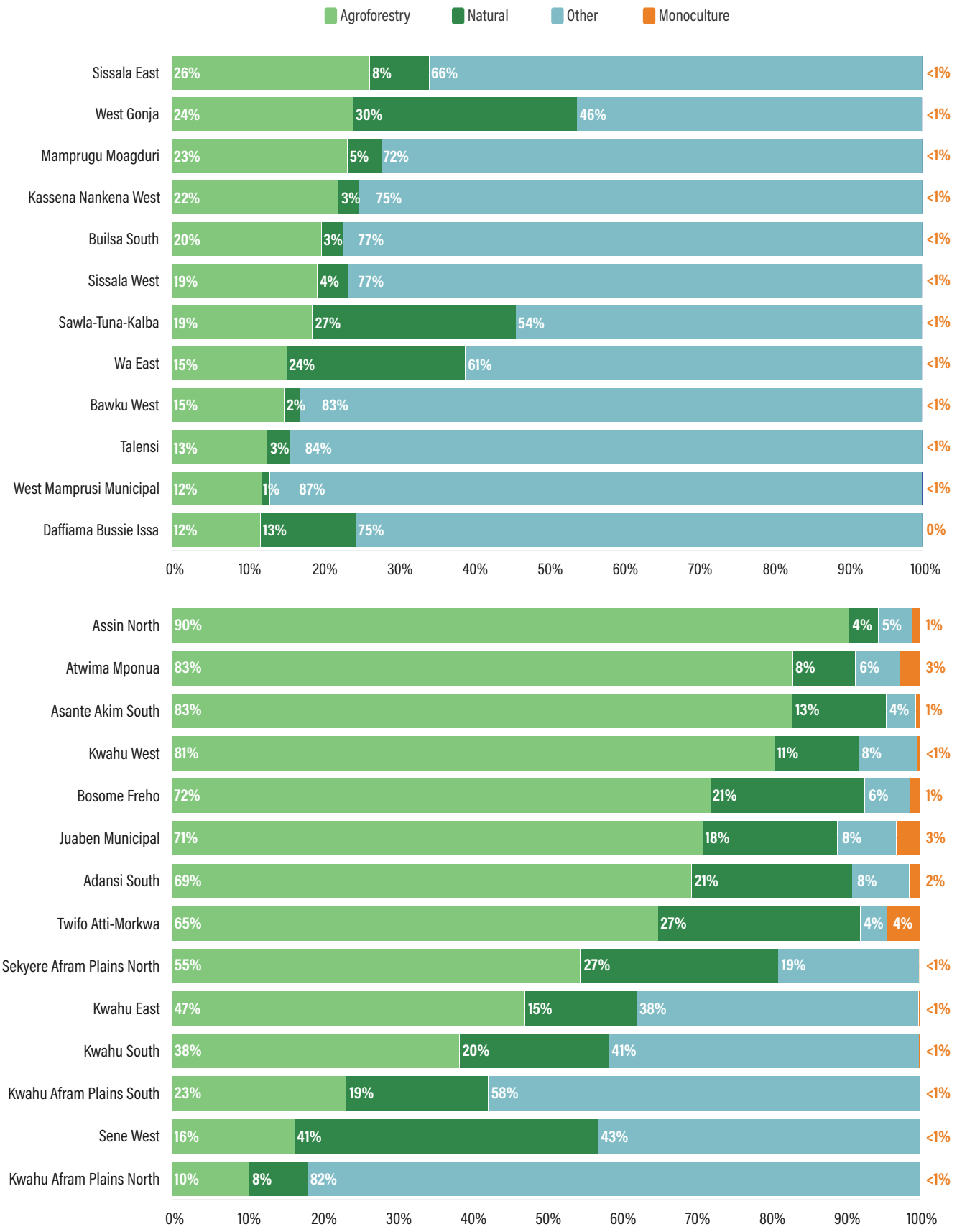
Source: WRI authors.

Table 4 | **Area extent (ha) for each priority district in Ghana and total extent for each region**

ZONE	DISTRICT	MONOCULTURE (HA)	AGROFORESTRY (HA)	NATURAL (HA)
North	West Mamprusi Municipal	10	32,540	2,860
	Wa East	0	68,180	106,320
	Talensi	0	11,100	2,640
	Bawku West	0	16,630	2,380
	Sissala West	10	41,450	8,660
	Builsa South	0	25,810	3,710
	Daffiama Bussie Issa	0	17,880	19,310
	Sissala East	10	139,900	42,230
	Kassena Nankana West	10	20,160	2,530
	West Gonja	0	117,770	145,220
	Sawla-Tuna-Kalba	0	81,810	118,550
	Mamprugu Moagduri	20	51,560	10,150
South	Twifo Atti-Morkwa	4,220	63,140	26,360
	Sene West	70	54,660	135,930
	Sekyere Afram Plains North	150	196,600	95,710
	Adansi South	1,070	53,780	16,630
	Kwahu South	110	31,160	16,240
	Kwahu Afram Plains South	70	73,350	60,140
	Kwahu Afram Plains North	80	24,300	18,950
	Juaben Municipal	550	12,400	3,140
	Bosome Freho	740	41,720	11,940
	Atwima Mponua	5,130	159,430	16,010
	Assin North	730	66,770	2,960
	Asante Akim South	660	97,540	14,790
	Kwahu West	160	32,930	4,570
	Kwahu East	120	29,990	9,590
NORTH	TOTAL	70	624,780	464,570
SOUTH	TOTAL	13,840	937,760	432,960

Source: WRI authors.

Figure 6 | **Proportional distribution of each land use system per district (top: northern districts, bottom: southern districts)**



Source: WRI authors.

With the land use map, we calculated the total error-adjusted area assessments for each class in our study (Table 4). We detected a total of 13,907 ha of monoculture area, 1,562,542 ha of agroforestry area, and 897,523 ha of natural area in the 26 priority districts. We presented the results by zone, with 11 districts clustered in the northern Savannah agro-ecological zone, and the remaining 15 districts clustered in the South across Transitional and Deciduous agro-ecological zones. Districts with the largest extent of monoculture area also contain industrial-scale plantation systems, like teak and Cedrela plantations in the Jimira Forest Reserve in Atwima Mponua (Forestry Commission 2021) and the Twifo Oil Palm Plantations in Twifo Atti-Morkwa. Districts with a higher proportion of agroforestry extent are largely located in Ashanti, where cocoa production is concentrated (Abu et al. 2021) (Table 4 and Figure 6). As expected, districts in the northern Savannah zone have low-to-no monoculture production systems and a higher proportion of area that belongs to the other (background) class (Figure 6).

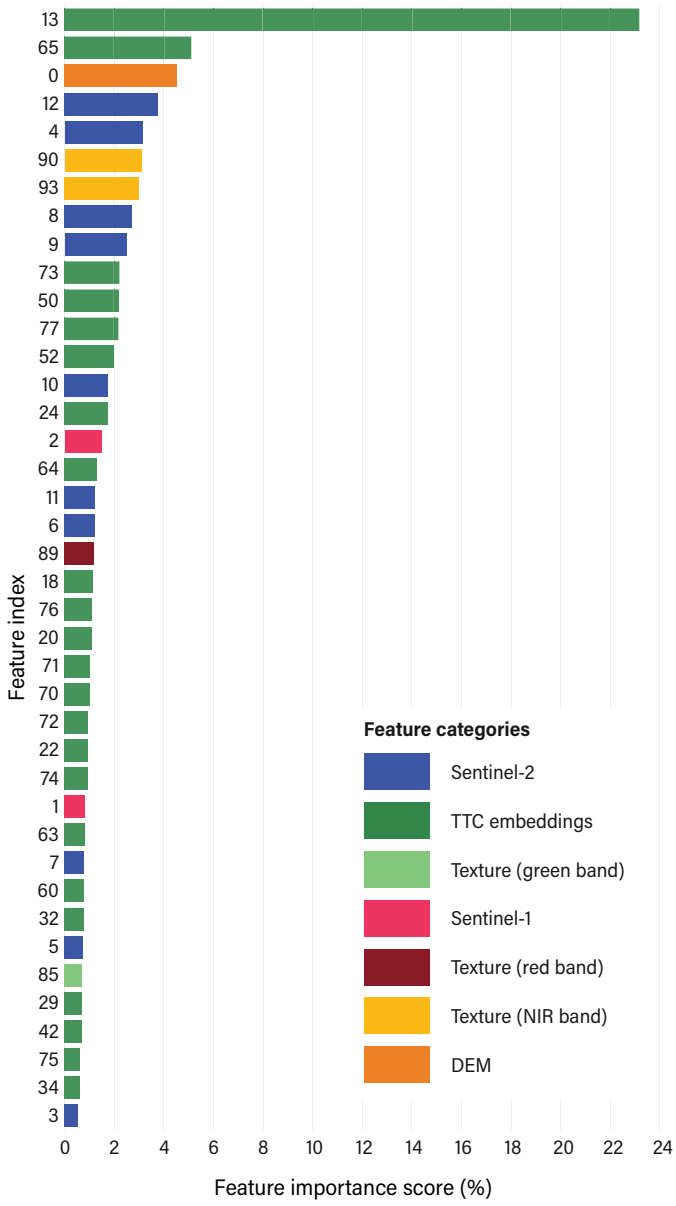
Contributions of transfer learning and texture

This study follows existing literature, which recognizes texture properties for their ability to separate tree crops from surrounding land cover and within production systems (Maskell et al. 2021). Feature importance scores highlighted these texture features' contribution to the classifier's ability to learn the differences between tree systems. In a pixel-based classification exercise that tries to identify systems, a texture feature that considers relationships with neighbors is, understandably, helpful for the interpretation. Given their high feature importance scores, dissimilarity, and contrast, texture features derived from the NIR band were particularly relevant to the classification (Figure 7). Texture features often exhibit high correlation with one another (Hall-Beyer 2017), which is a reason the feature selection exercise was helpful in improving accuracy. Figure 7 shows the feature importance scores of the top 40 features used in the model, color coded by category (Sentinel, embeddings, and texture). The sum of importances for all features is normalized to equal 100, with higher scores indicating larger effects on the model's predictive ability (Prokhorenkova et al. 2019). Figure 7 shows how embeddings and texture features explained approximately 90 percent of the model's performance.

Transfer learning performance gains

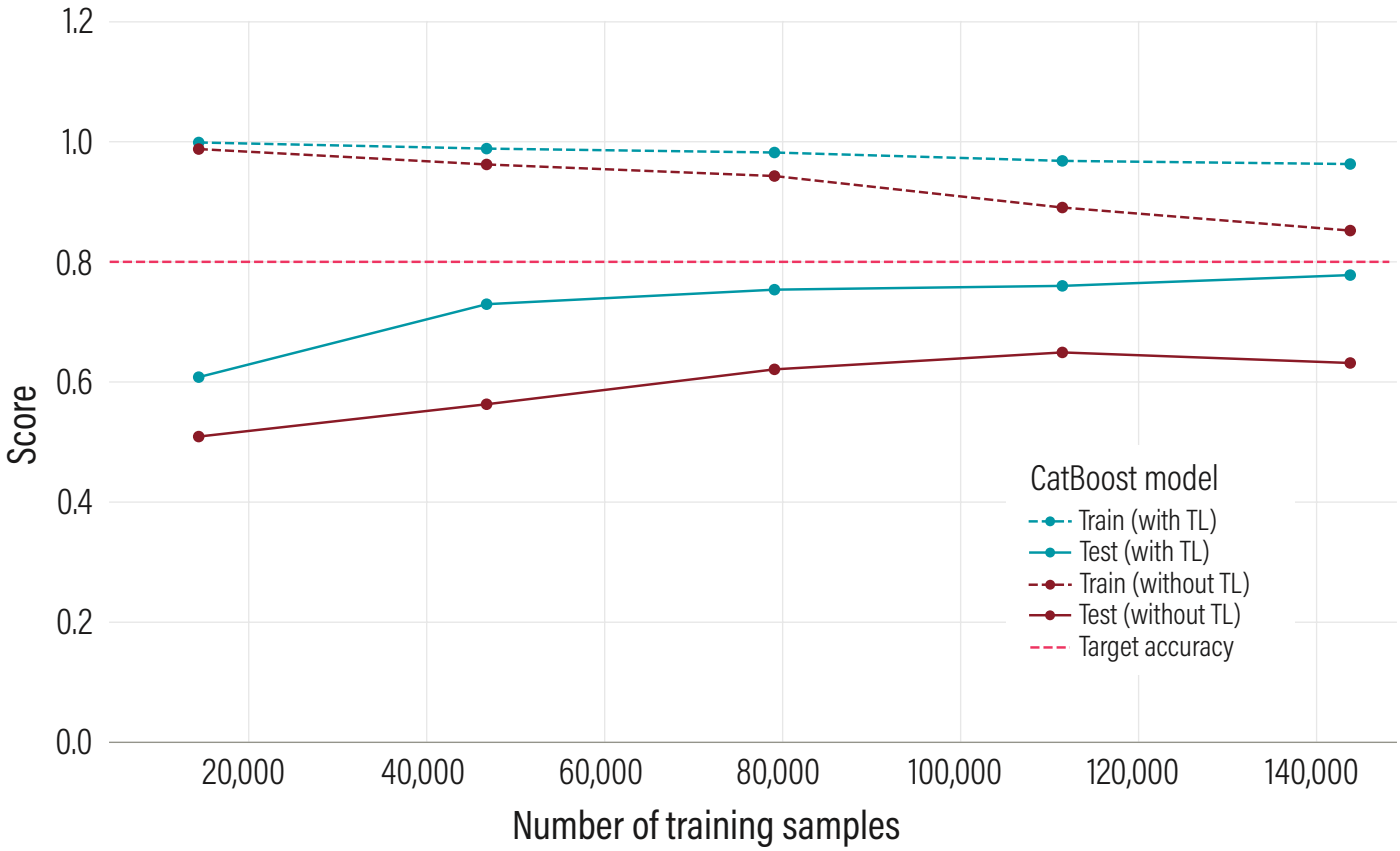
We tested our hypothesis about the value of a transfer learning-based classification approach using two comparative modeling exercises. The first experiment explored the performance gains attributable to incorporating extracted spatial embeddings and texture features, particularly given the limited size of our training dataset. In other words, we

Figure 7 | **Feature importance for the top 40 selected features**



Notes: This bar chart illustrates feature importance scores for the top 40 features. Features are named based on the relative location, or index, in the 94-band input array and color coded by category. Features are sorted by descending importance.
Source: WRI authors.

Figure 8 | **Learning curve comparison**



Notes: TL = transfer learning in Figure 8. A learning curve comparison illustrates how a CatBoost model trained with transferred features (with TL) achieves higher performance with the same number of training samples, evaluated using five-fold cross-validation.

Source: WRI authors.

Table 5 | **Performance comparison with deep learning network**

MODEL	FEATURE COUNT	INPUT SIZE	BALANCED ACCURACY
CatBoost classifier	40	14 x 14 patches	0.8605
CNN	13	14 x 14 patches	0.7379

Source: WRI authors.

wanted to understand whether spatial embeddings' contribution substantially accelerated model learning. Due to the time-intensive process of acquiring training data, we needed to forecast how many labeled training data were required for acceptable accuracy. Learning curves illustrate the relationship between learning and experience and can indicate how a training dataset's size can affect model performance. In the early research stages, comparing learning curves confirmed that adding extracted TTC and texture features could improve model accuracy with a low number of training samples.

Figure 8 shows the results of the learning curve comparison for two CatBoost classifiers. The first classifier was trained using only Sentinel imagery, and the second was trained using Sentinel imagery, GLCM texture features, and the extracted embeddings from the TTC model. A target accuracy of 80 percent is shown as a general illustrative goal. The value of the model that uses transfer learning is evident as more samples are introduced. The results of the test set confirmed that training with transfer learning (blue) allowed the model to approach the hypothetical target accuracy faster, and with fewer training samples, compared to a baseline model (red).

We performed this experiment to validate our early hypothesis that transfer learning improves performance with fewer labeled samples. The baseline model without transfer learning exhibited a slight downward trend in accuracy as training samples increased. This finding likely reflects the model's sensitivity to label noise and class confusion in the more heterogenous dataset, as well as transfer learning's potential stabilizing effect (blue lines in Figure 8).

Our second experiment involved benchmarking the CatBoost model, trained with extracted features and GLCM texture features, against a standard deep learning approach. For this experiment, we implemented a CNN with a U-Net architecture, similar to the one Brandt et al. (2023) used. We trained the CNN using the same features, labels, and train-test split to ensure a direct comparison. We evaluated the model over 10 epochs using pixel-wise cross entropy loss and standard accuracy metrics.

To compare the CatBoost classifier's performance with a standard deep learning algorithm, we looked at the balanced accuracy score for the entire validation set. Balanced accuracy was selected due to the slight class imbalance. Early comparative tests confirmed the transfer learning approach, incorporating extracted features and texture properties, offered a performance advantage for this task (Table 5). These exercises supported the study hypothesis that a transfer learning approach could improve performance when given limited labeled data (training samples) and a heterogenous class structure.

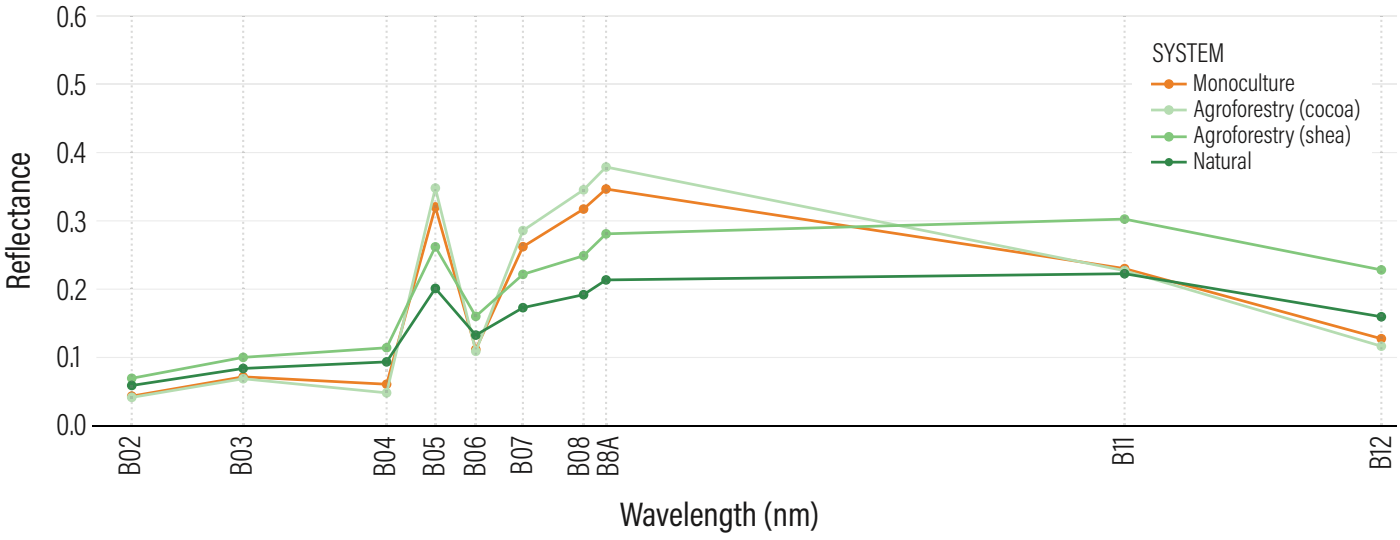
Discussion

Classifier performance

The results of the classification exercise and associated model comparisons demonstrate how transfer-learned features can help distinguish between tree-based systems in Ghana.

The model experienced confusion between the natural and agroforestry classes, which was expected due to the similar spectral signatures and spatial organization of shaded cocoa and open canopy forests. This finding could also be a manifestation of the fact that individual trees of the same species belonged to both of these land use classes. Most false positives for the natural class are classified as agroforestry and vice versa. The inclusion of diverse agroforestry systems within a single class, specifically shaded cocoa and cultivated shea, is likely responsible for low precision. While priority districts in eastern and western Ghana are predominantly comprised of cocoa agroforestry, agroforestry parkland systems (cropland areas with dispersed trees) characterize much of the savannah region of the northern priority districts. Low producer's accuracy for the natural class, often misclassified as agroforestry, likely stems from confusion in the North, where shea trees grow in both wild and agroforestry systems. In fact, studies analyzing shea tree populations across different land-use types (Jepsen et al. 2024) have found a higher density of shea trees in uncultivated bushland areas compared to agricultural land. Local farmers in Ghana even describe shea trees as wild, since they often stem from natural regeneration rather than intentional planting efforts. The confusion between natural

Figure 9 | **Sentinel-2 spectral signature analysis**



Notes: nm = nanometers in Figure 9. This figure shows the Sentinel-2 spectral signature of four representative tiles of the target land use classes. The curves illustrate variation in average surface reflectance across Sentinel-2 bands (B02-B12).

Source: WRI authors

and agroforestry systems in the final map product is caused by the presence of shea trees in uncultivated bushlands, as well as highly managed agricultural systems.

Initial exploratory analyses of the Sentinel-2 optical data hinted at this separability problem between classes and across regions in Ghana. Figure 9 shows the spectral signature of Sentinel-2 ARD for four tiles. Each tile represents a target land use class, including two subsets of the agroforestry class (cocoa and shea agroforestry systems). The graph shows the mean surface reflectance value across the 10 Sentinel-2 bands (B02-B12). Recognizing some inherent limitations to this analysis (Sentinel-2 images could come from varying seasons with varying quality, and no tile is solely comprised of a single land use class), we were still confident these results showed enough distinction in spectral signature alone to perform a decent classification. At the same time, this exercise underscored separability challenges for certain agroforestry system types. Despite being in the same target class, the spectral signatures of the two agroforestry systems are relatively unique. Moreover, the spectral signatures of an agroforestry shea system and a natural parkland show relative similarities. The low degree of separability for these two northern Savannah zone systems, which make up the largest proportional land area in the final map, are the reason we see so much confusion between the natural and agroforestry classes in the confusion matrix (Figure 9). One step to improve class-specific precision and recall is creating training data for more narrowly defined subclasses that reflect the distinct structural and spatial characteristics of different agroforestry systems.

Of all the land use classes, the monoculture class had the lowest accuracy and highest margin of error. Monoculture samples only represented 18 percent of the training data. To account for class imbalance in the training data, we assigned higher weight to underrepresented classes when training the CatBoost model. However, the low proportion of training data could have affected the model's ability to learn this minority class, especially for smallholder oil palm and other monoculture tree crops. The high producer's accuracy relative to the other two tree classes (agroforestry and natural) indicates most of the positives were true positives (correctly identified). The low user's accuracy of the monoculture class stems from confusion with the agroforestry class. Judging by the large industrial monoculture farms identified in the map (Figure 5), this error is less likely to occur at the industrial plantation scale. It is more likely to occur in irregularly shaped, smallholder farms that make up 93 percent of oil palm concessions by area in Ghana (Meijaard et al. 2018).

Sources of error in ground truth labeling

We took multiple steps to mitigate the risk of poor label quality, such as developing a rigorous photointerpretation protocol, gathering input from experts in Ghana, and using algorithmic label review with Cleanlab. To gather sufficient training and reference data, this study relied on photointerpretation and image-drawn labels of complex agricultural systems at small scales. Despite these measures, training data labels were subjected to the annotators' (in this case, this study's authors) expertise and judgment. These data labels also relied on the availability of high-resolution imagery to support visual interpretation of tree dynamics in the study year. If a high-resolution image was not available for 2020, an image from an earlier or later year was used to determine the appropriate pixel label. This approach often resulted in a temporal mismatch, compounded by the effect of using different source imagery to inform the labels and perform the classification. For this study, we used Google Earth Pro 10–15 cm resolution data to inform the annotation and Sentinel-2 data at 10-m resolution to produce maps (Estes et al. 2022). In the context of Ghana's rapidly evolving agricultural frontier, this problem is particularly relevant because a land cover change could occur in the time between when the two images were taken. This means even accurately drawn labels have the potential to introduce error during model training and assessment (Estes et al. 2022).

Errors observed in the other (background) class, where the reference label indicated agroforestry or natural systems, are likely attributable to annotation ambiguity or interpretation. This result could happen in scenes where it was unclear whether a pixel intersected a tree or sparse shrubbery. These discrepancies could also reflect seasonal differences in the imagery used during photointerpretation.

Variability in the agroforestry class

Compounding the difficulty in visually distinguishing natural and cocoa agroforestry systems, a key constraint to accurate, operational land use mapping lies in fundamental inconsistencies in definitions (Njomaba et al. 2025; Rosenstock et al. 2019). Varied interpretations of what exactly is an agroforestry system limit the availability of harmonized reference data that can train machine learning models. These differing interpretations also hinder comparative analyses across remote sensing products. Since clear definitions form the foundation of technical methodologies, varying definitions can result in over or underestimations of extent.

In one of the first remote sensing-derived efforts used to quantify the extent of agroforestry at a global scale, Zomer et al. (2014) defined agroforestry as agricultural land with tree cover greater than 10 percent. While an important advancement, this approach depends on the accuracy of an underlying agricultural land cover classification. The consequence is an underrepresentation of agroforestry systems found on land that meet the forest definition, such as in Cameroon, where cocoa agroforests can exhibit over 80 percent tree cover (Rosenstock et al. 2019). Lesiv et al. (2017) conducted another

global-level effort to map agroforestry alongside other forest management classes. The researchers defined agroforestry as managed forests on "other landscapes." This agroforestry definition includes trees in cropland, pasture, or urban areas, as well as shifting cultivation and fruit trees such as olives, apples, nuts, and cocoa (Lesiv et al. 2017). In their research, den Herder et al. (2017) mapped the distribution of agroforestry across the European Union, noting how their comparatively higher estimates of agroforestry extent may have arisen from underlying assumptions and criteria. In contrast to studies that use land cover criteria as part of their agroforestry definition, our study recognizes that agroforestry can occur across nearly all land cover types (Daniel et al. 2018). We also use a broader definition focused on canopy strata and planting patterns. Since this study focuses exclusively on the classification of spectral and textural features in satellite imagery, we exclude an important component of the agroforestry definition that considers intentionality: agroforestry systems are intentionally established and managed (Terasaki Hart et al. 2023). This variability in how agroforestry is defined across remote sensing products highlights the extent to which underlying definitions greatly influence area estimates (den Herder et al. 2017). While this study's objective does not propose a national legend or set of land cover definitions, we hope these findings shed light on the opportunities and limitations in designing a system that uses transfer learning to augment remote sensing machine learning classification techniques.

Conclusion

This study introduces a transfer learning approach to classify tree-based systems, leveraging extracted spatial embeddings from a high-performing neural network to improve classification accuracy in label-scarce environments. We apply a CatBoost classifier to a combination of Sentinel imagery, GLCM texture features, and extracted spatial embeddings to classify four land use classes: natural, agroforestry, monoculture, and other (background). Through comparative modeling and feature selection exercises, we demonstrate that including spatial embeddings and texture features improves model performance. We demonstrate this method for 26 priority districts in Ghana, resulting in a 10-m resolution land use map for 2020. Area assessments reveal a total of 13,907 ha of monoculture area, 1,562,542 ha of agroforestry area, and 897,523 ha of natural area in the 26 districts.

Using embeddings from the TTC algorithm in the prediction pipeline allowed us to leapfrog over standard approaches to develop 10-m resolution land use maps for Ghana. We confirmed this hypothesis through performance comparisons between our final CatBoost model and a CNN, as well as a CatBoost classifier trained without embeddings. In piloting the transfer learning method in Ghana, we hope to contribute to discourse around developing national land use definitions and their associated remote sensing characteristics. This study presents an approach that uses spatially aware texture features to perform a system classification that is agnostic to species. In doing so, we aim to identify a broader set of structural and spatial patterns through this system-level framing to address the absence of studies that focus on detection beyond a single commodity type. We propose this approach for its relevance and potential application outside of the immediate study area used in this paper. Additional training and reference labels in the target expansion areas are needed to expand this method to new geographies. Further research could focus on using an automated approach to gather high-quality training samples, such as using a label propagation algorithm. This study derived texture features directly from spectral bands. Future work could explore texture analyses from vegetation indices, such as the normalized difference vegetation index.

We also build on Brandt et al.'s (2023) novel contributions by capitalizing on their preprocessing pipeline used to produce cloud-free, analysis-ready composites of Sentinel imagery. As the TTC data expand in temporal coverage (2017–2024) in the coming years, the decision to use a methodologically compatible processing pipeline enables future harmonization with TTC change detection applications. Currently, large-scale tree cover datasets, like those of Brandt et al. (2023) and Hansen et al. (2013), do not distinguish between natural and agricultural trees, which limits their ability to identify drivers of tree cover gain or loss. Our findings suggest transfer-learned spatial features from the TTC model offer value beyond their original predictive task and also represent a scalable path forward for broader restoration monitoring efforts. We hope this initial investigation offers a useful starting point for remote sensing and machine learning practitioners who seek to apply transfer learning when designing datasets that monitor the extent of agricultural and natural trees.

References

Abbam, T., F.A. Johnson, J. Dash, and S.S. Padmadas. 2018. "Spatiotemporal Variations in Rainfall and Temperature in Ghana over the Twentieth Century, 1900–2014." *Earth and Space Science* 5 (4): 120–132. https://doi.org/10.1002/2017EA000327.

Abramowitz, J., E. Cherrington, R. Griffin, R. Muench, and F. Mensah. 2023. "Differentiating Oil Palm Plantations from Natural Forest to Improve Land Cover Mapping in Ghana." *Remote Sensing Applications: Society and Environment* 30 (April): 100968. https://doi.org/10.1016/j.rsase.2023.100968.

Abu, I.-O., Z. Szantoi, A. Brink, M. Robuchon, and M. Thiel. 2021. "Detecting Cocoa Plantations in Côte d'Ivoire and Ghana and Their Implications on Protected Areas." *Ecological Indicators* 129 (October): 107863. https://doi.org/10.1016/j.ecolind.2021.107863.

Alem, A., and S. Kumar. 2022. "Transfer Learning Models for Land Cover and Land Use Classification in Remote Sensing Image." *Applied Artificial Intelligence* 36 (1). https://doi.org/10.1080/08839514.2021.2014192.

Ashiagbor, G., E.K. Forkuo, W.A. Asante, E. Acheampong, J.A. Quaye-Ballard, P. Boamah, Y. Mohammed, and E. Foli. 2020. "Pixel-Based and Object-Oriented Approaches in Segregating Cocoa from Forest in the Juabeso-Bia Landscape of Ghana." *Remote Sensing Applications: Society and Environment* 19 (August): 100349. https://doi.org/10.1016/j.rsase.2020.100349.

Asigbaase, M., S.B. Ndego, and B. Effah. 2025. "Shade Tree Selection in Cocoa Agroforestry: Ghanaian Farmers' Preferences, Ecological Insight and Drivers of Local Ecological Knowledge." *Ecology and Evolution* 15 (7): e71685. https://doi.org/10.1002/ece3.71685.

Benefoh, D.T., G.B. Villamor, M. van Noordwijk, C. Borgemeister, W.A. Asante, and K.O. Asubonteng. 2018. "Assessing Land-Use Typologies and Change Intensities in a Structurally Complex Ghanaian Cocoa Landscape." *Applied Geography* 99 (October): 109–19. https://doi.org/10.1016/j.apgeog.2018.07.027.

Bentéjac, C., A. Csörgő, and G. Martínez-Muñoz. 2021. "A Comparative Analysis of Gradient Boosting Algorithms." *Artificial Intelligence Review* 54: 1937–1967. https://doi.org/10.1007/s10462-020-09896-5.

Bingham, H.C., M. Deguignet, E. Lewis, J. Stewart, D. Juffe-Bignoli, B. MacSharry, A. Milam, and N. Kingston. 2019. *User Manual for the World Database on Protected Areas and World Database on Other Effective Area-Based Conservation Measures: 1.6*. Cambridge, UK: UNEP-WCMC (UN Environment Programme World Conservation Monitoring Centre). https://app.ibat-alliance.org/pdf/wdpa_manual.pdf.

Brandt, J., J. Ertel, J. Spore, and F. Stolle. 2023. "Wall-to-Wall Mapping of Tree Extent in the Tropics with Sentinel-1 and Sentinel-2." *Remote Sensing of Environment* 292 (July): 113574. https://doi.org/10.1016/j.rse.2023.113574.

Candra, D.S., S. Phinn, and P. Scarth. 2020. "Cloud and Cloud Shadow Masking for Sentinel-2 Using Multitemporal Images in Global Area." *International Journal of Remote Sensing* 41 (8): 2877–2904. https://doi.org/10.1080/01431161.2019.1697006.

Chamberlin, J. 2008. "It's a Small World after All: Defining Smallholder Agriculture in Ghana." IFPRI (International Food Policy Research Institute) Discussion papers, 823. https://ideas.repec.org//p/fpr/ifprid/823.html.

Daniel, J., K. Tenneson, M. Suber, R. Mulia, P.V. Thanh, J. Arango, and T. Rosenstock. 2018. "Open-and Crowd-Sourced MRV for Agroforestry? Preliminary Results and Lessons Learned from a Pilot Study Using Collect Earth to Identify Agroforestry on Multiple Land Uses in Viet Nam and Colombia." Info Note. Wageningen, Netherlands: CGIAR CCAFS (Research Program on Climate Change, Agriculture and Food Security). https://www.researchgate.net/publication/332529146_Open-and_crowd-sourced_MRV_for_agroforestry_Preliminary_results_and_lessons_learned_from_a_pilot_study_using_Collect_Earth_to_identify_agroforestry_on_multiple_land_uses_in_Viet_Nam_and_Colombia.

den Herder, M., G. Moreno, R.M. Mosquera-Losada, J.H.N. Palma, A. Sidiropoulou, J.J.S. Freijanes, J. Crous-Duran, J.A. Paulo, M. Tomé, A. Pantera, V.P. Papanastasis, K. Mantzanas, P. Pachana, A. Papadopoulos, T. Plieninger, and P.J. Burgess. 2017. "Current Extent and Stratification of Agroforestry in the European Union." *Agriculture, Ecosystems & Environment* 241 (April): 121–32. https://doi.org/10.1016/j.agee.2017.03.005.

Descals, A., Z. Szantoi, E. Meijaard, H. Sutikno, G. Rindanata, and S. Wich. 2019. "Oil Palm (*Elaeis guineensis*) Mapping with Details: Smallholder versus Industrial Plantations and Their Extent in Riau, Sumatra." *Remote Sensing* 11 (21): 2590. https://doi.org/10.3390/rs11212590.

Dorogush, A.V., V. Ershov, and A. Gulin. 2018. "CatBoost: Gradient Boosting with Categorical Features Support." arXiv:1810.11363. arXiv (October). https://doi.org/10.48550/arXiv.1810.11363.

Eilers, P.H.C. 2003. "A Perfect Smoother." *Analytical Chemistry* 75 (14): 3631–36. https://doi.org/10.1021/ac034173t.

Estes, L.D., S. Ye, L. Song, B. Luo, J.R. Eastman, Z. Meng, Q. Zhang, D. McRitchie, S.R. Debats, J. Muhando, A.H. Amukoa, B.W. Kaloo, J. Makuru, B.K. Mbatia, I.M. Muasa, J. Mucha, A.M. Mugami, J.M. Mugami, F.W. Muinde, F.M. Mwawaza, J. Ochieng, C.J. Oduol, P. Oduor, T. Wanjiku, J.G., Wanyoike, R.B. Avery, K.K. Caylor. 2022. "High Resolution, Annual Maps of Field Boundaries for Smallholder-Dominated Croplands at National Scales." *Frontiers in Artificial Intelligence* 4 (February). https://doi.org/10.3389/frai.2021.744863.

Fagan, M.E., D.-H. Kim, W. Settle, L. Ferry, J. Drew, H. Carlson, J. Slaughter, J. Schaferbien, A. Tyukavina, N.L. Harris, E. Goldman, and E.M. Ordway. 2022. "The Expansion of Tree Plantations across Tropical Biomes." *Nature Sustainability* 5: 681–688. https://doi.org/10.1038/s41893-022-00904-w.

FAO (Food and Agriculture Organization of the United Nations). 2016. *Map Accuracy Assessment and Area Estimation: A Practical Guide*, by Y. Finegold, A. Ortmann, E. Lindquist, R. d'Annunzio, and M. Sandker. Rome: FAO. https://openknowledge.fao.org/handle/20.500.14283/i5601e.

Farr, T.G., and M. Kobrick. 2000. "Shuttle Radar Topography Mission Produces a Wealth of Data." *Eos, Transactions American Geophysical Union* 81 (48): 583–85. https://doi.org/10.1029/EO081i048p00583.

Filella, G.B. 2018. "Cocoa Segmentation in Satellite Images with Deep Learning." Bachelor Thesis. Zurich, Switzerland: Institute of Geodesy and Photogrammetry, ETH Zurich. https://ethz.ch/content/dam/ethz/special-interest/baug/igp/photogrammetry-remote-sensing-dam/documents/pdf/Student_Theses/BA_BonetFilella.pdf.

Forestry Commission. 2021. *Ghana Forest Plantation Strategy – 2020 Annual Report*. Accra, Ghana: Forest Services Division of the Forestry Commission. https://fcghana.org/2872/.

Google. n.d. "View a Map over Time." Google Earth Help. https://support.google.com/earth/answer/6327779. Accessed August 1, 2025.

Hall, J.B., and M.D. Swaine. 1981. *Distribution and Ecology of Vascular Plants in a Tropical Rain Forest*. Dordrecht, Netherlands: Springer. https://link.springer.com/book/10.1007/978-94-009-8650-3.

Hall-Beyer, M. 2017. *GLCM Texture: A Tutorial v. 3.0*. Technical Report. Department of Geography. Calgary, Canada: University of Calgary. https://doi.org/10.13140/RG.2.2.12424.21767.

Hamrouni, Y., E. Paillassa, V. Chéret, C. Monteil, and D. Sheeren. 2021. "From Local to Global: A Transfer Learning-Based Approach for Mapping Poplar Plantations at National Scale Using Sentinel-2." *ISPRS Journal of Photogrammetry and Remote Sensing* 171 (January): 76–100. https://doi.org/10.1016/j.isprsjprs.2020.10.018.

Hansen, M.C., P.V. Potapov, R. Moore, M. Hancher, S.A. Turubanova, A. Tyukavina, D. Thau, S.V. Stehman, S.J. Goetz, T.R. Loveland, A. Kommareddy, E. Egorov, L. Chini, C.O. Justice, and J.R.G. Townshend. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." *Science* 342 (6160): 850–53. https://doi.org/10.1126/science.1244693.

Harris, C.R., K.J. Millman, S.J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith, R. Kern, M. Picus, S. Hoyer, M.H. van Kerkwijk, M. Brett, A. Haldane, J. Fernández del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T.E. Oliphant. 2020. "Array Programming with NumPy." *Nature* 585: 357–62. https://doi.org/10.1038/s41586-020-2649-2.

Jepsen, T., G. Stopponi, and N.O.G. Jørgensen. 2024. "Shea Tree (*Vitellaria paradoxa C.F. Gaertn.*) Agroforestry Systems in Northern Ghana: Population Structure, Management of Trees and Impact of Below Canopy Microclimate." *Agroforestry Systems* 98: 1493–1506. https://doi.org/10.1007/s10457-024-01019-1.

Khatami, R., G. Mountrakis, and S.V. Stehman. 2016. "A Meta-Analysis of Remote Sensing Research on Supervised Pixel-Based Land-Cover Image Classification Processes: General Guidelines for Practitioners and Future Research." *Remote Sensing of Environment* 177 (May): 89–100. https://doi.org/10.1016/j.rse.2016.02.028.

Lanaras, C., J. Bioucas-Dias, S. Galliani, E. Baltsavias, and K. Schindler. 2018. "Super-Resolution of Sentinel-2 Images: Learning a Globally Applicable Deep Neural Network." *ISPRS Journal of Photogrammetry and Remote Sensing* 146 (December): 305–19. https://doi.org/10.1016/j.isprsjprs.2018.09.018.

Lesiv, M., S. Fritz, I. McCallum, N.E. Tsendbazar, J.F. Pekel, M. Herold, M. Buchhorn, B. Smets, and R. Van De Kerchove. 2017. "Evaluation of ESA CCI Prototype Land Cover Map at 20m." Working Paper. Laxenburg, Austria: IIASA (International Institute for Applied Systems Analysis). http://pure.iiasa.ac.at/14979/.

Louis, J., V. Debaecker, B. Pflug, M. Main-Knorn, J. Bieniarz, U. Mueller-Wilm, E. Cadau, and F. Gascon. 2016. "Sentinel-2 Sen2Cor: L2A Processor for Users." Paper presented at the Living Planet Symposium, Prague, Czech Republic, May 9 –13. https://elib.dlr.de/107381/.

Ma, Y., S. Chen, S. Ermon, and D.B. Lobell. 2024. "Transfer Learning in Environmental Remote Sensing." *Remote Sensing of Environment* 301 (February): 113924. https://doi.org/10.1016/j.rse.2023.113924.

Maskell, G., A. Chemura, H. Nguyen, C. Gornott, and P. Mondal. 2021. "Integration of Sentinel Optical and Radar Data for Mapping Smallholder Coffee Production Systems in Vietnam." *Remote Sensing of Environment* 266 (December): 112709. https://doi.org/10.1016/j.rse.2021.112709.

MATLAB Help Center. 2025. (Computing platform.) *Texture Analysis Using Gray-Level Co-Occurrence Matrix*. The MathWorks, Inc. Accessed February 3, 2025. https://www.mathworks.com/help/images/texture-analysis-using-the-gray-level-co-occurrence-matrix-glcm.html.

Meijaard, E., J. Garcia-Ulloa, D. Sheil, S.A. Wich, K.M. Carlson, D. Juffe-Bignoli, and T.M. Brooks. 2018. *Oil Palm and Biodiversity: A Situation Analysis by the IUCN Oil Palm Task Force*. Gland, Switzerland: IUCN (International Union for Conservation of Nature). https://portals.iucn.org/library/sites/library/files/documents/2018-027-En.pdf.

Mishra, V.N., R. Prasad, P.K. Rai, A.K. Vishwakarma, and A. Arora. 2019. "Performance Evaluation of Textural Features in Improving Land Use/Land Cover Classification Accuracy of Heterogeneous Landscape Using Multi-Sensor Remote Sensing Data." *Earth Science Informatics* 12: 71–86. https://doi.org/10.1007/s12145-018-0369-z.

Moomen, A.-W., L.L. Yevugah, L. Boakye, J.D. Osei, and F. Muthoni. 2024. "Review of Applications of Remote Sensing towards Sustainable Agriculture in the Northern Savannah Regions of Ghana." *Agriculture* 14 (4): 546. https://doi.org/10.3390/agriculture14040546.

Naudts, K., Y. Chen, M.J. McGrath, J. Ryder, A. Valade, J. Otto, and S. Luyssaert. 2016. "Europe's Forest Management Did Not Mitigate Climate Warming." *Science* 351 (6273): 597–600. https://doi.org/10.1126/science.aad7270.

Njomaba, E., F. Mushtaq, R.K. Nagbija, S. Yakalim, B.E. Aikins, and P. Surovy. 2025. "Adopting Land Cover Standards for Sustainable Development in Ghana: Challenges and Opportunities." *Land* 14 (3): 550. https://doi.org/10.3390/land14030550.

Northcutt, C., L. Jiang, and I. Chuang. 2021. "Confident Learning: Estimating Uncertainty in Dataset Labels." *Journal of Artificial Intelligence Research* 70: 1373–1411. https://doi.org/10.1613/jair.1.12125.

Numbisi, F.N., F.M.B. Van Coillie, and R. De Wulf. 2019. "Delineation of Cocoa Agroforests Using Multiseason Sentinel-1 SAR Images: A Low Grey Level Range Reduces Uncertainties in GLCM Texture-Based Mapping." *ISPRS International Journal of Geo-Information* 8 (4): 179. https://doi.org/10.3390/ijgi8040179.

Osei-Gyabaah, A.P., M. Antwi, S. Addo, and P. Osei. 2023. "Land Suitability Analysis for Cocoa (*Theobroma cacao*) Production in the Sunyani Municipality, Bono Region, Ghana." *Smart Agricultural Technology* 5 (October): 100262. https://doi.org/10.1016/j.atech.2023.100262.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay. 2011. "Scikit-Learn: Machine Learning in Python." *Journal of Machine Learning Research* 12: 2825–30.

Pereira, S.C., C. Lopes, and J.P. Pedroso. 2022. "Mapping Cashew Orchards in Cantanhez National Park (Guinea-Bissau)." *Remote Sensing Applications: Society and Environment* 26 (April): 100746. https://doi.org/10.1016/j.rsase.2022.100746.

Pers. Comm. (Personal communication). 2023a. Correspondence between Jessica Ertel, Data Scientist at WRI, and William Odoi, Assistant Programme Officer at Environmental Protection Authority, Ghana. July 11.

Pers. Comm. (Personal communication). 2023b. Correspondence between Jessica Ertel, Data Scientist at WRI, and William Odoi, Assistant Programme Officer at Environmental Protection Authority, Ghana. August 29.

Prokhorenkova, L., G. Gusev, A. Vorobev, A.V. Dorogush, and A. Gulin. 2019. "CatBoost: Unbiased Boosting with Categorical Features." arXiv:1706.09516. arXiv (January). https://doi.org/10.48550/arXiv.1706.09516.

Qiu, S., Z. Zhu, and B. He. 2019. "Fmask 4.0: Improved Cloud and Cloud Shadow Detection in Landsats 4–8 and Sentinel-2 Imagery." *Remote Sensing of Environment* 231 (September): 111205. https://doi.org/10.1016/j.rse.2019.05.024.

Richter, J., E. Goldman, N. Harris, D. Gibbs, M. Rose, S. Peyer, S. Richardson, and H. Velappan. 2024. "Spatial Database of Planted Trees (SDPT Version 2.0)." Technical Note. Washington, D.C.: WRI (World Resources Institute). https://doi.org/10.46830/writn.23.00073.

Rosenstock, T.S., A. Wilkes, C. Jallo, N. Namoi, M. Bulusu, M. Suber, D. Mboi, R. Mulia, E. Simelton, M. Richards, N. Gurwick, and E. Wollenberg. 2019. "Making Trees Count: Measurement and Reporting of Agroforestry in UNFCCC National Communications of Non-Annex I Countries." *Agriculture, Ecosystems & Environment* 284 (November): 106569. https://doi.org/10.1016/j.agee.2019.106569.

Saah, D., G. Johnson, B. Ashmall, G. Tondapu, K. Tenneson, M. Patterson, A. Poortinga, K. Markert, N.H. Quyen, K.S. Aung, L. Schlichting, M. Matin, K. Uddin, R.R. Aryal, J. Dilger, W.L. Ellenburg, A.I. Flores-Anderson, D. Wiell, E. Lindquist, J. Goldstein, N. Clinton, and F. Chishtie. 2019. "Collect Earth: An Online Tool for Systematic Reference Data Collection in Land Cover and Use Applications." *Environmental Modelling & Software* 118 (August): 166–71. https://doi.org/10.1016/j.envsoft.2019.05.004.

Sari, I.L., C.J. Weston, G.J. Newnham, and L. Volkova. 2022. "Developing Multi-Source Indices to Discriminate between Native Tropical Forests, Oil Palm and Rubber Plantations in Indonesia." *Remote Sensing* 14 (1): 3. https://doi.org/10.3390/rs14010003.

Schneider, M., C. Winchester, E. Goldman, and Y. Shao. 2023. "Mapping Cocoa and Assessing Deforestation Risk for the Cocoa Sector in Côte d'Ivoire and Ghana." Technical Note. Washington, DC: WRI. https://doi.org/10.46830/writn.21.00011.

Terasaki Hart, D.E., S. Yeo, M. Almaraz, D. Beillouin, R. Cardinael, E. Garcia, S. Kay, S.T. Lovell, T.S. Rosenstock, S. Sprenkle-Hyppolite, F. Stolle, M. Suber, B. Thapa, S. Wood, and S.C. Cook-Patton. 2023. "Priority Science Can Accelerate Agroforestry as a Natural Climate Solution." *Nature Climate Change* 13: 1179–90. https://doi.org/10.1038/s41558-023-01810-5.

Virtanen, P., R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, I. Polat, Y. Feng, E.W. Moore, and J. VanderPlas. 2020. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python." *Nature Methods* 17: 261–72. https://doi.org/10.1038/s41592-019-0686-2.

Wang, N., F. Chen, B. Yu, and Y. Qin. 2020. "Segmentation of Large-Scale Remotely Sensed Images on a Spark Platform: A Strategy for Handling Massive Image Tiles with the MapReduce Model." *ISPRS Journal of Photogrammetry and Remote Sensing* 162 (April): 137–47. https://doi.org/10.1016/j.isprsjprs.2020.02.012.

Zomer, R.J., A. Trabucco, R. Coe, F. Place, M. van Noordwijk, and J. Xu. 2014. "Trees on Farms: An Update and Reanalysis of Agroforestry's Global Extent and Socio-Ecological Characteristics." Working Paper 179. Bogor, Indonesia: ICRAF (World Agroforestry Centre Southeast Asia Regional Program) doi:10.5716/WP14064.PDF.

Acknowledgments

We are pleased to acknowledge our institutional strategic partners that provide core funding to WRI: the Netherlands Ministry of Foreign Affairs, Royal Danish Ministry of Foreign Affairs, and Swedish International Development Cooperation Agency.

We are deeply grateful to Daniel Kwaku Owusu for sharing critical data points, contextual knowledge, and local expertise on agricultural practices and restoration efforts in Ghana. We thank him for his support in photo interpretation to develop the validation dataset. His contributions played an important role in strengthening our final model evaluation.

About the authors

Jessica Ertel is a Data Science Manager for WRI's Global Restoration Initiative.

John Brandt is a Data Science Lead for WRI's Global Restoration Initiative.

Rhiannon Rognstad is the Head of Data Science for WRI's Data Lab.

Erin Glen is a GIS Research Associate with WRI's Land & Carbon Lab.

About WRI

World Resources Institute works to improve people's lives, protect and restore nature, and stabilize the climate. As an independent research organization, we leverage our data, expertise, and global reach to influence policy and catalyze change across systems like food, land and water; energy; and cities. Our 2,000+ staff work on the ground in more than a dozen focus countries and with partners in over 50 nations.

About the Environmental Protection Authority (EPA) of Ghana

The mission of the EPA of Ghana is to co-manage, protect and enhance the country's environment, in particular, as well as seek common solutions to global environmental problems. The accomplishment of the mission is to be achieved inter alia through research, scientific, technological and innovative approaches, good governance and partnerships.