

Ecological Legacies of Pre-Columbian Settlements Evident in Palm Clusters of Neotropical Mountain Forests

Sebastian Fajardo^{1*†}, Sina Mohammadi^{1*†}, Jonas Gregorio de Souza², César Ardila³, Alan Tapscott², Shaddai Heidgen², Maria Isabel Mayorga Hernández⁴, Sylvia Mota de Oliveira⁵, Fernando Montejo³, Marco Moderato², Vinicius Peripato⁶, Katy Puche³, Carlos Reina³, Juan Carlos Vargas⁷, Frank W. Takes¹, Marco Madella^{2,8,9}

¹ Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Leiden, The Netherlands 2300 RA

² Department of Humanities, CASEs group, Universitat Pompeu Fabra, Barcelona, Spain 08023

³ Subdirección de Gestión del patrimonio, Instituto Colombiano de Antropología e Historia, Bogotá, Colombia 111711

⁴ Facultad de Artes, Universidad Nacional de Colombia, Bogotá, Colombia 111321

⁵ Naturalis Biodiversity Center, Leiden, The Netherlands, 2300 RA

⁶ Division of Earth Observation and Geoinformatics (DIOTG), General Coordination of Earth Sciences (CG-CT),

National Institute for Space Research (INPE), São José dos Campos, SP, Brazil

⁷ Departamento de Antropología, Universidad del Magdalena, Santa Marta, Colombia 470004

⁸ ICREA, Barcelona, Spain 08010

⁹ GAES, The University of the Witwatersrand, Johannesburg, South Africa 2050

*Corresponding author. Sebastian Fajardo: s.d.fajardo.bernalliacs.leidenuniv.nl;

Sina Mohammadi: s.mohammadi@liacs.leidenuniv.nl

†These authors contributed equally to this work.

Ancient populations markedly transformed Neotropical forests, yet the spatial extent of their ecological influence remains underexplored at high resolution. Here we present a deep learning and remote sensing based approach to estimate areas of pre-Columbian forest modification based on modern vegetation. We apply this method to high-resolution satellite imagery from the Sierra Nevada de Santa Marta, Colombia, as a demonstration of a scalable approach, to evaluate palm tree distributions in relation to archaeological infrastructure. Palms were significantly more abundant near archaeological sites with large infrastructure investment. The extent of the largest palm cluster indicates that ancient human-managed areas linked to major infrastructure sites may be up to two orders of magnitude bigger than indicated by current archaeological evidence alone. Our findings suggest that pre-Columbian populations influenced vegetation, fostering conditions conducive to palm proliferation, leaving a lasting ecological footprint. This may have lowered the logistical costs of establishing infrastructure-heavy settlements in less accessible locations.

Introduction

The activities of ancient populations transformed the composition of forests in many Neotropical regions. Ancient human management included the cultivation, tending, and dispersal of tree species for food and technological materials (1, 2), as well as the use of fire for forest clearance and enriching nutrient soil composition (3, 4). Forest tree inventory data, reported archaeological sites, and classic machine learning models have been used to examine the relationship between plant species and their association with past human activities in large regions (1, 5, 6). These studies show that the impacts of past activities on modern forest structures near the location of ancient settlements are long-lasting, to the extent that they may influence our current estimates of forest composition in tropical forests, such as the Amazon (7, 8) and the eastern flank of the tropical Andes (6).

Archaeological sites and forest tree inventories used to estimate the level of ancient human management on modern forest composition are limited by sparse data on the variability of pre-Columbian settlement patterns and the small number of forest tree inventories that represent

the vast areas once occupied by these communities. Both limitations result in coarse resolution estimates, making it difficult to accurately assess the long-term effects of various past human activities on local forest composition (9). This challenge is especially evident in montane forests, where persistent fog and light limitations (10, 11), low soil fertility (11, 12), and abrupt terrain (13) create conditions that require land-use practices beyond conventional agrarian expectations. Yet, archaeological evidence demonstrates that pre-Columbian communities frequently occupied these environments across Colombia (14, 15) and other Tropical Andean forests (6). While the distribution of tree species exploited for food is relatively well documented (16), the management of plants used for raw materials—and their relationship to ancient settlements of different sizes and spatial arrangements—remains poorly understood. Understanding the extent, intensity, and cultural sources of ancient human management on the composition and dynamics of Neotropic montane forests, within a continuous area broad enough to capture both ecological and archaeological variability, is essential for the integration of human legacy strategies into ecological models and incorporating current forest composition into automatic methods for detecting archaeological sites using remote sensing data.

In this study, we propose a method called “PABAM” (Past Areas of Bio-cultural Activity and Management). This method, combined with legacy archaeological data on settlement locations, enables us to explore the relationship between archaeological sites and palm trees with a star-shaped crown (e.g. *Dictyocaryum* sp. and *Bactris* sp.) associated with multi-purpose exploitation by indigenous populations (17, 18). We tested the method in the montane forests region of the Sierra Nevada de Santa Marta (SNSM) in Colombia. By combining deep learning to identify the locations of individual palms based on the star-shaped crown, alongside archaeological legacy data and pedestrian surveys, we aimed to understand the association between the presence of these plants due to human management and the archaeological sites. While the PABAM method is tested in the SNSM, it is designed to be broadly applicable to other Neotropical—and potentially global—regions with similar vegetation and archaeological complexity. This integrated approach provides deeper insights into the formation of cultural landscapes in Neotropical forests. The long-term human history of Neotropical forests reflects the recognition that pre-colonial Indigenous societies shaped resource-rich, enduring landscapes across the region; we hypothesize that the present-day distribution of palms is partially linked to these past human activities. To test our

hypothesis, we analyzed palm concentrations around archaeological sites using a Inverse Distance Weighting (IDW) score and conducted an elevation analysis comparing site-associated palms to regional baselines. These analyses allowed us to assess whether palm presence is related to archaeological site locations and elevation patterns indicative of ancient human management.

Our main contributions are as follows. 1) A method to analyze tropical vegetation and human activity. We propose a multi-component machine learning method to investigate the spatial relationship between vegetation patterns and archaeological sites. 2) Release of a manually annotated palm tree dataset. We are releasing a high-quality palm tree dataset covering 69.5 km^2 , manually labeled by human annotators (Figure 1). We believe such a dataset to be a valuable resource for future research, particularly in the area of domain adaptation in other computer vision applications. Researchers can use our well-annotated dataset as a source domain to develop and test methods for adapting models to new regions. 3) Release of ground-surveyed archaeological site locations. We are also releasing the locations of archaeological sites in the area of study, along with associated levels of certainty, verified through ground surveys and reviews of legacy data. This dataset can support future studies exploring the relationship between the presence of archaeological sites, vegetation patterns, and past human activity.

Results

Palm Tree Detection Model Performance

We evaluated the performance of the model on the test dataset using different confidence thresholds of 0.2, 0.3, and 0.4. Precision and recall metrics are reported in Figure 2. In our application, minimizing false positives was more important than maximizing recall because high false positive rates in palm tree detection could produce artificial clusters and misleading density estimates. Since clusters should reflect true patterns with high precision, we prioritized precision over recall. Therefore, a confidence threshold of ≥ 0.4 was selected to reduce false positives. The final palm tree bounding box predictions for the entire study area, based on the ≥ 0.4 threshold, are shown in Figure 2.

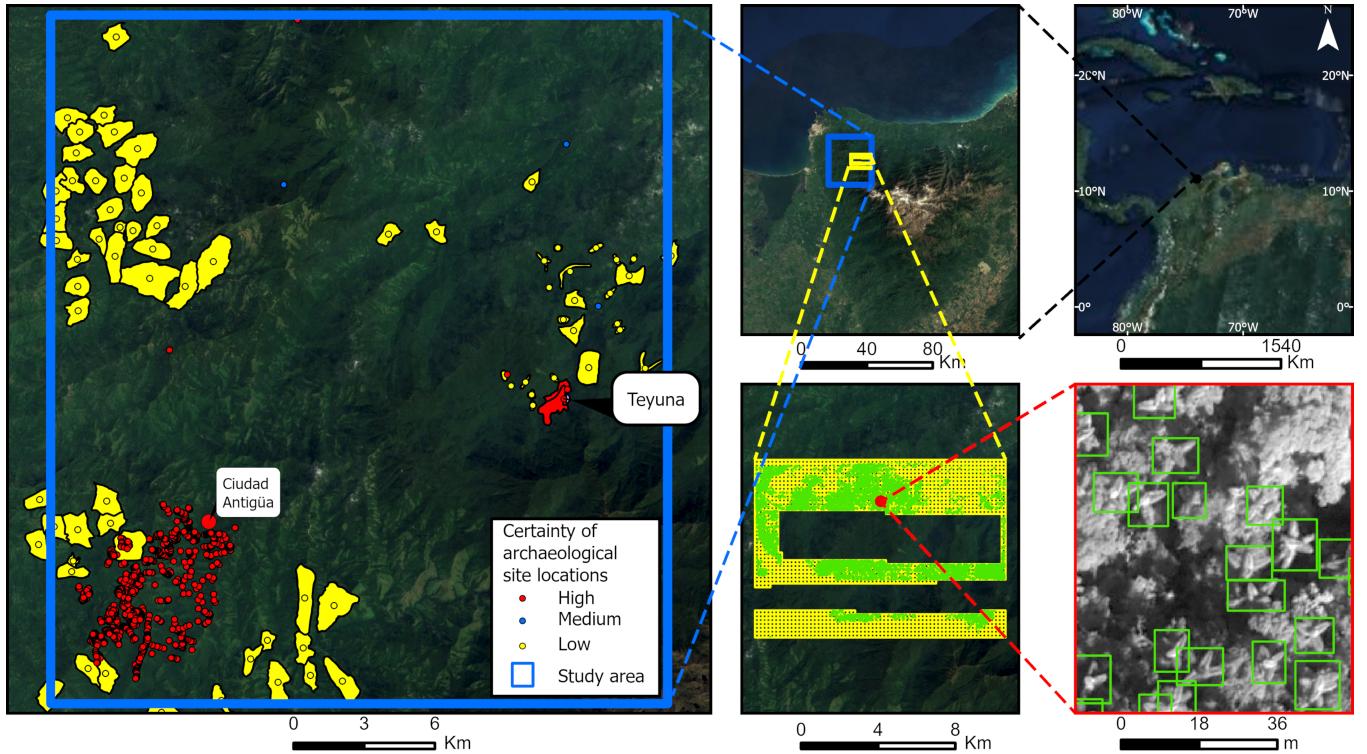


Figure 1: Study area at different resolutions. Regional and global context of the study area (top-center and top-right). Archaeological site locations with varying levels of location certainty (left). Annotated 200×200 m grid in yellow and palm tree labels in green (bottom-center). Detailed view of annotated palm trees on high-resolution imagery (bottom-right). Bottom-right panel includes material © CNES (2024), Distribution Airbus DS. All other panels include Sentinel-2 cloudless - <https://s2maps.eu> by EOX IT Services GmbH (Contains modified Copernicus Sentinel data 2016 & 2017).

Palm Distribution and Archaeological Sites

To evaluate the relationship between palm tree distribution and past human presence, we conducted a spatial clustering analysis of palm trees across the study area. Clusters 1, 2, and 15, which are located in the southwest, west, and east portions of the study area (Figure 2), exhibited spatial overlap with known archaeological features. Because the areas in the western portion of the study area have undergone more modern human intervention, we chose to focus our analysis in the less-disturbed eastern part of the area of study.

Notably, the largest archaeological site (Teyuna), marked by significant infrastructure invest-

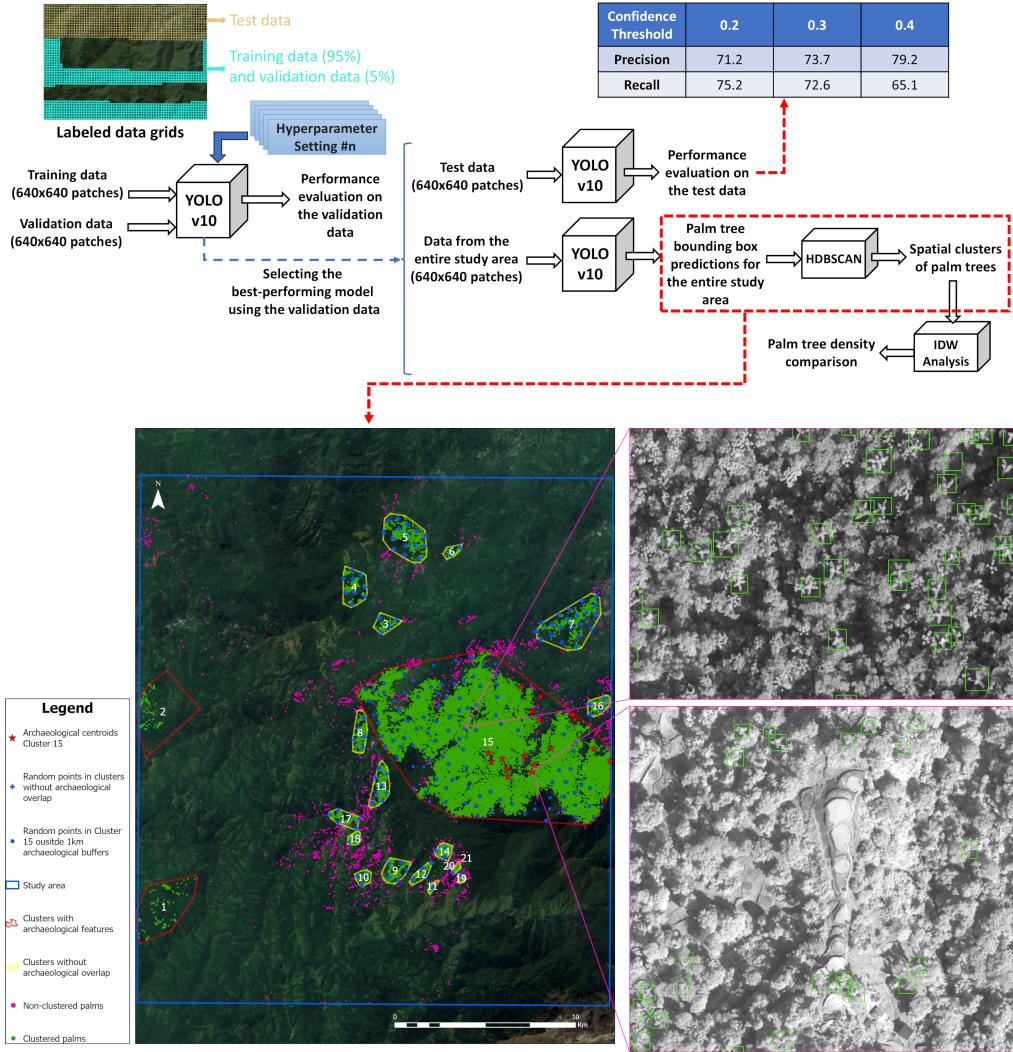


Figure 2: PABAM methodology, which incorporates the YOLOv10 architecture. After selecting the best-performing model based on validation data, we applied it to the test set for performance evaluation, as well as to the entire study area for palm tree detection. Detected palm trees were clustered using the HDBSCAN algorithm, followed by an IDW analysis to compare palm tree densities. Each cluster with its corresponding cluster ID is shown on the bottom-left frame. Right panels include material © CNES (2024), Distribution Airbus DS. Left panel includes Sentinel-2 cloudless - <https://s2maps.eu> by EOX IT Services GmbH (Contains modified Copernicus Sentinel data 2016 & 2017).

ment, is located within the largest cluster of palms (Cluster 15), spanning approximately 100 km² across both well-lit ridges and steep shaded slopes in the eastern part of the study area (Figure 2).

This cluster also contains 17 archaeological zones recorded as points or polygons. We calculated the centroids of these sites to serve as reference points for spatial analysis. This resulted in 17 reference points used in our study. Archaeological sites within the largest palm cluster are located in areas with significantly higher IDW values than random locations within palm clusters lacking archaeological features (Figure 3A). To estimate the extent of ancient human management influence in Cluster 15, we compared palm tree densities at 100 randomly selected points within the cluster, but outside a 1 km buffer around archaeological sites, with densities at the 17 archaeological points in the same cluster. The observed means overlap at the 80% bootstrapped confidence interval, indicating that the differences in palm density are likely not statistically significant (Figure 3A).

Beyond the localized analysis of archaeological centroids near Teyuna, we expanded the IDW analysis to include four additional archaeological centroids situated in geographically distinct regions of the study area. These centroids were spatially independent from both the Teyuna-associated centroids and the control points used in the initial analysis. One of these centroids, in the southwest of the study area corresponds to El Congo-Ciudad Antigua (19), a previously known archaeological site. The remaining two centroids were estimated based on nearby archaeological sites reported in legacy data recorded in our surveys. The IDW values for the centroids in the northwest, southwest (Ciudad Antigua), and south-center regions were 0.005, 0.002, and 0 respectively. In contrast, the Teyuna centroid yielded a IDW value of 3.75. These results suggested that palm tree density around Teyuna is substantially greater than in the other regions.

Modern Buildings and Palm Tree Distribution

The Spearman correlation between the number of modern human buildings and the number of palm trees detected in each 200x200 m cell grid and the detected palms (Figure 3B) was low ($\rho = 0.21$, $p\text{-value} < 0.001$), suggesting weak relationship between the presence of modern buildings and the distribution of palm trees in the study area. This pattern is mainly driven by small modern nucleated settlements and low density dispersed farms populations located at lower elevations on the western side of the study area, away from zones from the largest clusters of detected palms.

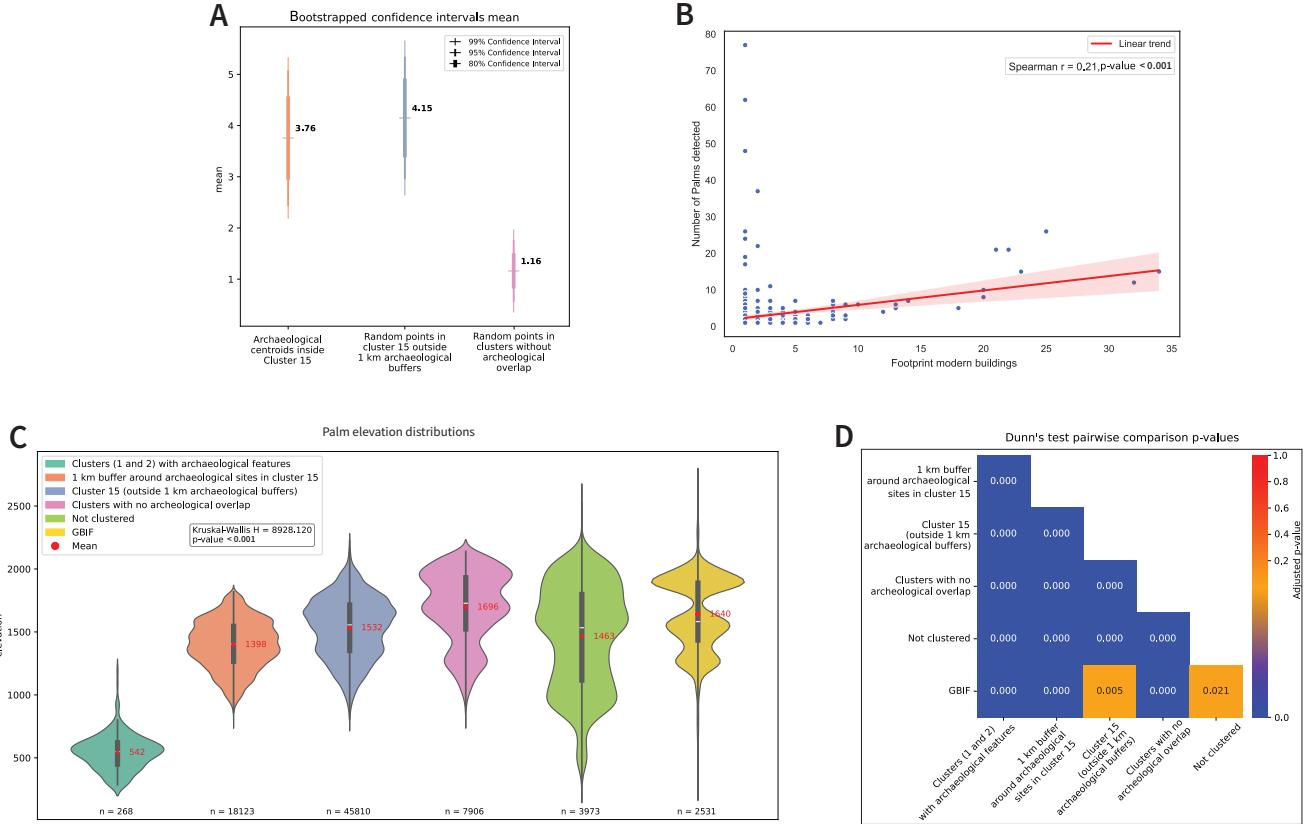


Figure 3: Human activity and palm distribution. (A) Bootstrapped confidence intervals for observed means (Bootstrap samples = 1000; sample size = 17). The subsets shown are IDW means for archaeological centroids inside Cluster 15 (left), 100 random points in Cluster 15 outside 1 km archaeological buffers (center), and 100 random points in clusters without overlap with archaeological features (right). Means are shown as text labels to the right of each bullet graph. (B) Spearman correlation between modern buildings (20) and palm detections in each 200x200 m grid inside the study area. (C) Elevation distributions across subsets of detected palms and the GBIF baseline for the distribution of *Dictyocaryum* in South America. Violin plots (left) display Gaussian kernel density estimates for each subset, with overlaid box plots indicating the median (white line) and mean (red dot). (D) The heatmap presents pairwise Dunn test results between subsets following a significant Kruskal-Wallis test.

Elevation Analysis

The centroids of archaeological locations within Cluster 15 span elevations ranging from 819 to 1570 m, with a mean elevation of 1113 m. The violin plots in Figure 3C show that palm trees located

within a 1 km radius of archaeological sites in Cluster 15 tend to occur at lower elevations compared to those outside this buffer zone but within Cluster 15. Furthermore, both elevation distributions—palms near archaeological sites and those outside the buffer in Cluster 15—show more frequently values at lower than those of other palm clusters without archaeological overlap. In addition, all median elevation values of clustered palms that overlap with archaeological sites are below the GBIF baseline median. In contrast, only the clusters of palms without archaeological overlap exhibit median elevation values higher than that of the GBIF baseline. All elevation distributions of the detected palm subsets fall within the range of the Global Biodiversity Information Facility (GBIF) baseline, suggesting that the palm detection is consistent capturing the overall variability of palm distribution.

Discussion

Ecological Implications of Palm Distribution and Elevational Patterns

Previous studies typically examined broad geographical regions using sparse sampling (1,6), making it difficult to capture finer-scale patterns of human-environment interaction. Our fine-scale detection analysis across a contiguous, yet regionally extensive, area reveals a spatial correlation between the pre-Columbian populations and the densest cluster of palm trees in the Neotropic forest of the SNSM. This correlation is not consistent across all sites but is particularly pronounced around Teyuna, where pre-Columbian populations invested significantly in infrastructure development. Settlement data from the southwestern part of our study area indicate that most archaeological sites are relatively small (≤ 1 ha), with a maximum surface of around 5 ha (21, 22). This surface is considerably smaller than the main site of Teyuna (~ 33 ha). The area around Teyuna contains a significantly higher number of individuals from the detected plant taxa than would be expected under a random distribution model, suggesting that pre-Columbian activities shaped this pattern. While Teyuna is occupied from around 400 CE (23), paleoecological reconstructions show an increase in palm pollen (24, 25), coinciding with the first centuries (1100 and 1250 CE) of the Tairona period. Given that the estimated ancient human management area at the Teyuna site alone is around 100 ha (19), incorporating the large surrounding cluster of detected palms suggests a

much broader area of influence. This palm cluster spans approximately 100 km² suggesting that the extent of activity around Teyuna, including paths, cultivation, resource collection, and general forest management could be up to two orders of magnitude larger than what archaeological evidence alone indicates. This finding complements the temporal data from the palynological analysis and is aligned with the period of major infrastructural investment in the SNSM. Similar dynamics have been observed in many parts of the Amazon, where archaeological earthworks coincide with the hyperdominance of useful species (1, 5), and in the highlands of Colombia, where activities around pre-Hispanic population centers likely impacted high-altitude lakes (26).

Teyuna and adjacent sites located within the largest identified palm cluster exhibit significantly higher palm densities compared to randomly selected points inside clusters without evidence of prehispanic occupation within the same region. This pattern suggests that pre-Columbian activities that managed the forest, and possibly opened the canopy, favored these palms. Along with canopy openness, the typically short dispersal distances in palms combined with the low density of other species probably contributed to the higher local density. In the case of *Dictyocaryum lamarckianum*, its stilt roots provide an increased mechanical support on steep hillsides, again offering an advantage to occupy the cleared landscape. Following initial settlement and palm proliferation, the technological utility of palms likely enhanced the value of the area and influenced subsequent settlement decisions by offering valuable resources—such as materials for construction, containers, thatch, and edible parts like palm hearts (17, 18)—making these areas more suitable for long-term habitation and infrastructure investment.

Building on the extended analysis across multiple archaeological sites, and given Teyuna's historical prominence as a major pre-Columbian center (19, 23), it is plausible that anthropogenic influences—such as the planting, cultivation, or preferential retention of palms—were more intensive in this region. These practices may have constituted a form of low-intensity management or stewardship of plant communities (2). Our results suggest that the legacy effects of pre-Columbian human activity on palm distribution are not uniform across the landscape. Instead, they appear to be spatially variable, with a more pronounced presence persisting near major centers such as Teyuna.

Elevation analysis also supports the idea that prehistoric human activity contributed to localized ecological modifications that promoted palm persistence or proliferation at lower altitudes than the GBIF baseline data. The presence of Pre-Columbian Indigenous populations possibly facilitated the

creation of micro-environmental conditions, such as access to light, modified hydrology, reduced interspecific competition, and/or direct management practices, that favored the survival and expansion of clusters of *Dictyocaryum* palms at lower altitude sites. This presence of palms supports the hypothesis that pre-Columbian people management of the forest has left a durable ecological imprint. Although such occurrences are unlikely to act as standalone predictors for the presence of archaeological sites, they offer valuable supplementary information when integrated with broader environmental and multispectral and hyperspectral datasets. The palm distribution detected in our area of work of the SNSM is an example of long-term human-ecological interactions and niche construction (27), where sometimes subtle but persistent anthropogenic influences shape the distribution of key plant species in time and space. Although Teyuna exhibits a particularly strong signature of ecological legacy, the methodology developed here is not limited to this region. Future applications across other Neotropical and tropical environments can further test the model's potential to detect hidden landscapes of ancient human activity.

The current palm distribution is the result of ecological processes that unfolded during the Pre-Columbian occupation of these sites. The composition of the SNSM forest in areas without reported human presence indicates that palms are a key component of the floristic assemblage but their frequency in the vegetation is relatively low (28, 29). Although short-term successional dynamics following forest disturbance are relatively well understood for lowland Amazonia (30)—for instance, the dominance of early successional species depending on disturbance scale—such processes in Neotropical montane environments remain less studied. Our data suggest that patterns of certain plant frequencies that coincide with nearby archaeological sites are most likely the result of human agency rather than only forest regeneration dynamics. More research using plant proxies for plants such as phytoliths or environmental DNA extracted from archaeological deposits and soils from nearby areas is needed to understand these dynamics.

Potential applications, limitations, and future work directions

Our approach could be extended to analyze other Neotropical forests, such as the Amazon, as well as more arid landscapes with sparse vegetation—where plant species may serve as especially clear ecological markers of past human activity. These markers may reflect lasting ecological legacy

of indigenous land use (31) or post-abandonment conditions shaped by earlier human presence (32). Combining automatic species detection and botanical field surveys with archaeological and paleoenvironmental data can enhance our understanding of pre-Columbian indigenous settlements in the Neotropics, and contextualize them within interconnected local and global dynamics of past human–environment relationships (33). Beyond archaeological applications, the model have potentials for efficient palm tree distribution mapping (34) for bio-economic development, land-use planning, and ecological monitoring. Our method, using satellite imagery and deep learning, provides a scalable approach to detect such patterns on wider surfaces and more remote regions. By automatically mapping the distribution and dominance of specific plant species—particularly those with cultural or ecological ties to past human groups—this approach can aid in identifying affiliation of current indigenous groups to land, complementing traditional field-based evidences.

Although the palm detection model used in this study successfully captured the overall distribution and density patterns of palm trees throughout the canopy of the study area, it occasionally failed to accurately detect individual trees. These errors are primarily caused by visual similarities between palm trees and certain human-made structures, such as crossroads or buildings, that can exhibit textures or shapes resembling the star-like appearance of a palm tree crown. These misclassifications mainly occur in heavily modified landscapes; in well-preserved forests—where this approach is intended—the risk of confusing natural features with roads or buildings is minimal. To address this issue, future work can focus on integrating semi-supervised learning into the detection framework to enhance palm tree detection performance, even in urbanized landscapes. By combining limited labeled data with large unlabeled image collections, these methods can help models learn more robust and generalizable features. Techniques such as pseudo-labeling (35) and consistency regularization (36) may be particularly effective in leveraging the variability present in unlabeled data, including differences in appearance, lighting, and environmental conditions. This strategy is expected to improve detection accuracy and adaptability to unseen regions and palm tree varieties.

The computer vision model detected a specific crown pattern at the top of the canopy. Palms with different crown patterns or located lower in the canopy were less detected. Some areas in the study region—particularly in the southwest—show signs of modern human intervention, while the area around Teyuna appears more preserved. To further enhance the efficiency and accuracy of palm

detection, future research could focus on integrating active learning into the training pipeline (37). This would involve iteratively training a model on a small labeled set, using the model to identify the most informative unlabeled samples (e.g., those with high prediction uncertainty), acquiring annotations for these selected samples, and retraining the model with the expanded labeled set. This iterative process has the potential to significantly improve performance while minimizing annotation costs.

We believe that our palm tree detection dataset will be a valuable asset for future research, particularly in the context of domain adaptation. By using our dataset as a well-annotated source, researchers can investigate methods to adapt models to new regions, thereby significantly reducing the need for extensive labeling efforts in those target domains.

Conclusion

Using our scalable methodology for identifying potential archaeological features through their ecological signatures we find that palm density is indicative of past human activity, suggesting that a synergetic relation between initial settlement and subsequent palm proliferation influenced decisions on where to continue developing settlements. The abundance of palm trees around archaeological centroids—especially at Teyuna—and the observation that site-associated palms occurred at lower elevations than both regional baselines and non-site-associated palms, indicate that pre-Columbian populations actively shaped environments in ways that promoted palm proliferation in areas where plant species competition should be intense. This in turn, may have lowered the logistical costs of establishing infrastructure-heavy settlements in otherwise less accessible locations.

These findings contribute to our understanding of ancient human-environment interactions in Neotropical ecosystems and suggest that pre-Columbian populations left lasting ecological imprints that persist centuries after site abandonment. The integrative approach demonstrated here, combining automated vegetation detection with archaeological data, has particular relevance for tropical regions where traditional archaeological methods face significant challenges due to dense vegetation and limited accessibility. The approach is scalable to any large areas with submetric resolution satellite imagery available. This work underscores the importance of interdisciplinary collaboration between ecologists, archaeologists, and computer scientists to uncover the complex

legacies of ancient human activities in shaping contemporary forest compositions.

Materials and Methods

Pre-Columbian Settlements and Forest Management in the Sierra Nevada de Santa Marta

Pre-Columbian communities of the Tairona cultural tradition transformed the SNSM landscape. Human occupation, dated from 200 to 1600 CE (19, 23), is divided into two archaeological periods: Neguanje (200–1100 CE) and Tairona (1100–1600 CE). Settlements occur on ridgelines, valley slopes (19, 23, 38), riverbanks, bays, inlets, flat coastal areas, and deeper canyons (23, 39). Long-distance mobility between sites within the SNSM has been documented both historically (40) and in the present, with a 60 km round-trip between two Indigenous communities completed in 18 hours (41). Entire valleys, typically 30 km long, can be traversed in three days from sea level to the paramo (~ 3000 m elevation) (42). Spanning from coastal lowlands to 2000 m elevation, the Tairona sites feature terraced platforms with curved stone walls, rammed earth foundations, rainwater drainage systems, and circular stone floors for wood-and-thatch structures connected by paved paths (23, 38, 43–46). All this represent a significant infrastructure investment given the rugged terrain and the logistical challenges of transporting large amount of stones.

Small settlements (<3 ha) are scattered across the landscape (22). Larger sites (≥ 3 ha) are less common but present in both the Neguanje and Tairona periods. However, estimating the size of settlements in the Neguanje phase is challenging due to the superimposing Tairona structures that cover the earlier deposits (23). The largest known settlement is Teyuna (Ciudad Perdida or Buritaca-200), occupied in the Neguanje period from around 400 CE (23). During the Tairona period, it expanded to a settlement of 33 ha with typical regional architecture (19, 23, 46). A remote sensing analysis estimates the footprint of Teyuna at approximately 100 ha (Figure 1), supporting the role of the settlement as a regional center within an extensive and interconnected system, but also suggesting a broader anthropogenic landscape surrounding the site. This aligns with the models of “low-density agricultural urbanism” seen in Mesoamerica and Southeast Asia (19) and with evidence from other Neotropical forests such as the Amazon (16).

Palynological analyses (25, 45) in a radius of 6 km from Teyuna, along with phytolith studies at the site (23), indicate that landscape transformation began between 600 and 900 CE, likely for small-scale cultivation of maize (*Zea mays*) and manioc (*Manihot* sp.). Phytoliths from *Bactris* sp. and other palms (Arecaceae) were identified in Teyuna, suggesting the exploitation of such plant resource (23). Also, an increase in palm pollen between 1100 and 1250 CE in the lowlands (24), and at 900-1300 m elevations (25), coincided with the Tairona occupation period. Among palms, *Dictyocaryum lamarckianum*—common in the SNSM at elevations of 900-2000 m—is used for construction and may represent one of the sources of phytoliths within the recovered assemblages (23, 47). Palynological evidence also suggests rapid forest regeneration after site abandonment in the 16th century, implying that Tairona settlements had an ecological footprint on modern forest composition (45).

We selected a study area of approximately 765 km² in Colombia, which is known to contain Tairona occupation. We collected archaeological legacy data (Supplementary Material, Data S1 and S2), harmonizing geospatial location of historical maps and archaeological reports to estimate areas with possible archaeological sites with low and medium certainty (46). High-certainty locations were identified through systematic pedestrian surveys, as detailed in previous studies (21, 46). Figure 1 presents the study area at various zoom levels for better spatial context.

Within this region, we used high spatial resolution satellite imagery from the Pléiades Neo constellation acquired between 2022 and 2024 to annotate the locations of star-shaped palm trees. We utilized the panchromatic band, which offers the highest spatial resolution (30 cm), making it particularly well-suited for detecting individual palm trees. A total of 37 scenes covering the study area were used for analysis. A 69.5 km² non-contiguous subset was annotated with the locations of star-shaped palm trees and used to train and test the object detection model (Supplementary Material, Data S3 and S4). To streamline the annotation process and minimize the risk of missing any palm trees, the area was divided into 200x200 m grid. Human annotators systematically scanned each grid and marked individual palm trees bounding boxes. In total, 24,430 palm trees were annotated. (Figure 1, bottom-center). Since orbital optical imagery is typically acquired at nadir, detection is biased toward the canopy layer. In the montane forest of the study area, the most likely palm detected under these conditions is *Dictyocaryum* sp. This genus is known for its tall, solitary growth form and prominent canopy presence in Andean cloud forests (17, 48). This makes

it more readily distinguishable than understory palms such as *Bactris* sp. or lowland species such as *Phytelephas* sp.

Past Areas of Bio-cultural Activity and Management (PABAM) Methodology

We propose a multi-component machine learning method to estimate areas of ancient human management using known archaeological sites and vegetation patterns as inputs. In this study, we used star-shaped palm tree crowns as indicators of plants exploited for raw materials in the past and archaeological sites categorized by varying levels of location certainty.

First, we trained the extra-large variant of the “You Only Look Once” (YOLO) object detection architecture, a family of deep learning models designed for fast and accurate object localization and classification in images (49). Within its 10th variant, YOLO introduces several improvements in efficiency and accuracy, and the x configuration represents the most accurate variant, optimized for detecting small or complex objects. We applied this model (YOLOv10x) to 30 cm resolution Pléiades Neo satellite imagery to identify palm trees (Figure 2), taking advantage of its ability to capture fine-scale canopy features. The labeled palm area was divided into two spatially distinct regions (Figure 2). The lower region, comprising 47% of the annotations, was used for training and validation, while the upper region, containing the remaining 53%, was reserved for testing. This spatial separation aimed to reduce the risk of spatial bias during classification. Each region was further subdivided into 640×640 pixel patches for model input. Within the training and validation region, 95% of the patches were used for training and the remaining 5% were used for validation to guide hyperparameter tuning and model selection. The best-performing model based on precision and recall scores was applied to the test patches for evaluation, as well as to the full study area, for palm tree detection (Supplementary Material, Data S5).

Next, we applied Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) (50) to identify spatial clusters of palm trees using 100 as minimum number of points required to form a cluster. This algorithm handles clusters of varying densities, it does not require a parameter for neighborhood radius, and it is robust to noise and outliers. We linked these clusters to past human activity by associating them with the centroids of known archaeological sites and low-location-certainty archaeological zones within the convex hull boundaries of each palm cluster.

To compare palm concentrations, we generated centroids for archaeological sites and low-location-certainty archaeological zones. We also generated random centroids in the largest palm cluster and in areas showing clustered palms without reported archaeological sites. Palm concentration differences were measured using a Inverse Distance Weighting (IDW) score, which models the density of nearby palm trees by assigning higher weights to those located closer to each centroid. Palm trees within a 1000 m radius of each centroid were included in the IDW calculations, ensuring that the analysis captured localized patterns while excluding influence from distant and ecologically distinct areas. This radius is an conservative estimate of human mobility based on recently reported walking distances and travel times between Indigenous settlements (41, 42). The score is calculated as $G = \sum_i \frac{1}{d_i^w}$, where G is the IDW score, d_i is the distance from the selected point to palm tree i (only if $d_i \leq 1000$ meters), and w is a decay factor that gives more weight to closer trees. We use $w = 1$ in our analysis. Finally, we used bootstrapping techniques to compare palm concentrations around archaeological centroids with those in control zones—palm areas without known archaeological features. For this, we generated 100 control points, randomly sampled from palm cluster regions without known archaeological features. We also generated an additional set of 100 control points within the largest palm cluster but outside 1 km buffers around the archaeological sites. For both control sets and the archaeological centroids within the largest palm cluster, we created 1000 bootstrapped datasets. Each dataset consisted of 17 observations sampled with replacement, matching the number of archaeological centroids in the largest cluster.

We further evaluated the consistency of the palm detection model by comparing the elevation of detected palms with the elevation distribution of *Dictyocaryum* sp. occurrences recorded in the Global Biodiversity Information Facility (GBIF) database (51). Elevation data for both sets were extracted from the Shuttle Radar Topography Mission (SRTM) dataset at 1 arc second (~ 30 m) resolution (52). An interquartile range-based outlier filtering on the SRTM elevation was applied to the GBIF baseline. GBIF records served as proxy for the species' natural elevational range, likely reflecting areas less affected by pre-Columbian human activity. We then compared GBIF data with the elevations of palm trees detected by our deep learning model within a 1000 m radius of archaeological centroids in the largest palm cluster. We also examined the elevation data for all palm trees located in areas showing clustered and non-clustered palms without archaeological sites.

To assess the potential relationship between the number of modern human activities and the

number of palm trees detected within the study area, we conducted a Spearman rank correlation between the number of modern buildings detected in a 200x200 m cell grid and the detected palms (Figure 3B). The analysis used a dataset containing automatic detected building from satellite imagery between 2014 and 2024 (20) and palm trees detected per spatial unit (Figure 3B).

References and Notes

1. C. Levis, *et al.*, Persistent effects of pre-Columbian plant domestication on Amazonian forest composition. *Science* **355** (6328), 925–931 (2017), publisher: American Association for the Advancement of Science, <https://www.jstor.org/stable/24918447>.
2. S. D. Coelho, *et al.*, Eighty-four per cent of all Amazonian arboreal plant individuals are useful to humans. *PLOS ONE* **16** (10), e0257875 (2021), publisher: Public Library of Science, doi:10.1371/journal.pone.0257875, <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0257875>.
3. A. Leal, R. Gassón, H. Behling, F. Sánchez, Human-made fires and forest clearance as evidence for late Holocene landscape domestication in the Orinoco Llanos (Venezuela). *Vegetation History and Archaeobotany* **28** (5), 545–557 (2019), doi:10.1007/s00334-019-00713-w, <https://doi.org/10.1007/s00334-019-00713-w>.
4. S. Y. Maezumi, *et al.*, Fire-human-climate interactions in the Bolivian Amazon rainforest ecotone from the Last Glacial Maximum to late Holocene. *Frontiers in Environmental Archaeology* **2** (2023), publisher: Frontiers, doi:10.3389/fearc.2023.1208985, <https://www.frontiersin.org/journals/environmental-archaeology/articles/10.3389/fearc.2023.1208985/full>.
5. V. Peripato, *et al.*, More than 10,000 pre-Columbian earthworks are still hidden throughout Amazonia. *Science* **382** (6666), 103–109 (2023), publisher: American Association for the Advancement of Science, doi:10.1126/science.adc2541, <https://www.science.org/doi/10.1126/science.adc2541>.
6. R. K. Sales, *et al.*, Potential distributions of pre-Columbian people in Tropical Andean landscapes. *Philosophical Transactions of the Royal Society B: Biological Sciences* **377** (1849), 20200502 (2022), publisher: Royal Society, doi:10.1098/rstb.2020.0502, <https://royalsocietypublishing.org/doi/10.1098/rstb.2020.0502>.
7. C. H. McMichael, K. J. Feeley, C. W. Dick, D. R. Piperno, M. B. Bush, Comment on “Persistent effects of pre-Columbian plant domestication on Amazonian forest composition”. *Science*

- 358** (6361), eaan8347 (2017), publisher: American Association for the Advancement of Science, doi:10.1126/science.eaan8347, <https://www-science-org.ezproxy.leidenuniv.nl/doi/10.1126/science.eaan8347>.
8. C. N. H. McMichael, F. Matthews-Bird, W. Farfan-Rios, K. J. Feeley, Ancient human disturbances may be skewing our understanding of Amazonian forests. *Proceedings of the National Academy of Sciences* **114** (3), 522–527 (2017), publisher: Proceedings of the National Academy of Sciences, doi:10.1073/pnas.1614577114, <https://www.pnas.org/doi/10.1073/pnas.1614577114>.
 9. C. R. Clement, *et al.*, The domestication of Amazonia before European conquest. *Proceedings of the Royal Society B: Biological Sciences* **282** (1812), 20150813 (2015), publisher: Royal Society, doi:10.1098/rspb.2015.0813, <https://royalsocietypublishing.org/doi/10.1098/rspb.2015.0813>.
 10. M. G. Letts, M. Mulligan, The impact of light quality and leaf wetness on photosynthesis in north-west Andean tropical montane cloud forest. *Journal of Tropical Ecology* **21** (5), 549–557 (2005), doi:10.1017/S0266467405002488.
 11. T. J. Fahey, R. E. Sherman, E. V. Tanner, Tropical montane cloud forest: environmental drivers of vegetation structure and ecosystem function. *Journal of Tropical Ecology* **32** (5), 355–367 (2016), doi:10.1017/S0266467415000176.
 12. Instituto Geográfico Agustín Codazzi (IGAC), *Estudio general de suelos y zonificación de tierras del Departamento de Magdalena* (Imprenta Nacional de Colombia Bogotá, Colombia) (2009).
 13. S. D. Crausbay, P. H. Martin, Natural disturbance, vegetation patterns and ecological dynamics in tropical montane forests. *Journal of Tropical Ecology* **32** (5), 384–403 (2016), doi:10.1017/S0266467416000328.
 14. C. H. Langebaek, ¿ Cuántos eran?; Dónde estaban?; Qué les pasó? Poblamiento indígena en la Colombia prehispánica y su transformación después de la Conquista, in *Colombia. Preguntas y respuestas sobre su pasado y su presente* (Ediciones Uniandes), pp. 27–52 (2010).

15. A. Botiva Contreras, *et al.*, *Colombia prehispánica: regiones arqueológicas* (Colcultura, Instituto Colombiano de Antropología) (1989).
16. J. Iriarte, *et al.*, The origins of Amazonian landscapes: Plant cultivation, domestication and the spread of food production in tropical South America. *Quaternary Science Reviews* **248**, 106582 (2020), doi:10.1016/j.quascirev.2020.106582, <https://www.sciencedirect.com/science/article/pii/S0277379120305448>.
17. N. Y. Paniagua-Zambrana, R. W. Bussmann, Dictyocaryum lamarckianum (Mart.) H. Wendl., in *Ethnobotany of the Andes*, N. Y. Paniagua-Zambrana, R. W. Bussmann, Eds. (Springer International Publishing, Cham), pp. 1–17 (2020), doi:10.1007/978-3-319-77093-2_99-1, https://doi.org/10.1007/978-3-319-77093-2_99-1.
18. M. Moraes, N. Paniagua-Zambrana, R. Cámara-Leret, H. Balslev, M. Macia, Palmas útiles de Bolivia, Colombia, Ecuador y Perú, in *Cosecha Palmas en el Noroeste Suramérica Bases Científicas Para Su Manejo Conservación*, H. Balslev, M. Macia, H. Navarrete, Eds. (Pontificia Universidad Católica del Ecuador, Quito, Ecuador), pp. 87–102 (2015).
19. D. Rodríguez Osorio, S. Giraldo, E. Mazuera, A. Burbano, E. Figueredo, Beyond Visualization: Remote Sensing Applications in Prehispanic Settlements to Understand Ancient Anthropogenic Land Use and Occupation in the Sierra Nevada de Santa Marta, Colombia. *Latin American Antiquity* **34** (4), 821–841 (2023), doi:10.1017/laq.2022.91.
20. Microsoft, Microsoft Global ML Footprints, <https://github.com/microsoft/GlobalMLBuildingFootprints> (2023), accessed: 2025-05-30.
21. J. C. Vargas, W. Londoño, M. Soto, *Arqueología de la microcuenca El Congo, Sierra Nevada de Santa Marta* (Editorial Unimagdalena, Santa Marta, Colombia) (2022).
22. L. M. Soto Rodríguez, J. C. Vargas, Network analysis in Tairona chiefdoms of the Río Frío basin, Sierra Nevada de Santa Marta, Colombia. *Journal of Archaeological Science* **179**, 106230 (2025), doi:<https://doi.org/10.1016/j.jas.2025.106230>, <https://www.sciencedirect.com/science/article/pii/S0305440325000792>.

23. S. Giraldo, *Lords of the Snowy Ranges: Politics, place, and landscape transformation in two Tairona towns in the Sierra Nevada de Santa Marta, Colombia*, Phd thesis, University of Chicago, Chicago, Illinois (2010).
24. T. van der Hammen, G. W. Noldus, Paleoecology of the Ciénaga Grande (Great Lagoon) of Santa Marta, in *La Sierra Nevada de Santa Marta (Colombia) Transecto Buritaca la Cumbre*, T. van der Hammen, P. M. Ruiz, Eds. (J.Cramer, Berlin-Stuggart, Germany), pp. 581–588 (1984).
25. L. F. Herrera, Buritaca 200: survey of the archaeology and a palynological investigation, in *La Sierra Nevada de Santa Marta (Colombia) Transecto Buritaca la Cumbre*, T. van der Hammen, P. M. Ruiz, Eds. (J.Cramer, Berlin-Stuggart, Germany), pp. 531–546 (1984).
26. M. I. Vélez, *et al.*, Humans and climate in ritualized landscapes, the case of Lake Tota in the eastern highlands of Colombia. *The Holocene* **34** (11), 1587–1597 (2024), doi:10.1177/09596836241266408.
27. K. N. Laland, J. Odling-Smee, M. W. Feldman, Niche construction, biological evolution, and cultural change. *Behavioral and brain sciences* **23** (1), 131–146 (2000).
28. R. P. Fortier, *et al.*, Diversity, composition, and structure of a 1-hectare tree plot in the cloud forest of the Sierra Nevada de Santa Marta, Colombia. *Annals of Forest Science* **81** (1), 24 (2024), doi:10.1186/s13595-024-01239-6, <https://doi.org/10.1186/s13595-024-01239-6>.
29. A. Dechner, M. Diazgranados, Floristic composition and structure of forested areas along the low basin of the San Salvador river, north side of Sierra Nevada de Santa Marta, Colombia. *Univ. Sci.* **12** (2), 99–124 (2007).
30. N. Norden, *et al.*, Successional dynamics in Neotropical forests are as uncertain as they are predictable. *Proceedings of the National Academy of Sciences* **112** (26), 8013–8018 (2015), doi:10.1073/pnas.1500403112, <https://www.pnas.org/doi/abs/10.1073/pnas.1500403112>.
31. B. M. Pavlik, *et al.*, Plant species richness at archaeological sites suggests ecological legacy of Indigenous subsistence on the Colorado Plateau. *Proceedings of the National Academy*

- of Sciences* **118** (21), e2025047118 (2021), doi:10.1073/pnas.2025047118, <https://www.pnas.org/doi/abs/10.1073/pnas.2025047118>.
32. Y. Sapir, Y. Sapir, A. Faust, Using floristic characteristics of contemporary vegetation for identifying archaeological sites: Tel ‘Eton archaeological site as a test case. *Israel Journal of Plant Sciences* **66** (1–2), 60 – 68 (2019), doi:10.1163/22238980-00001029, https://brill.com/view/journals/ijps/66/1-2/article-p60_60.xml.
 33. G. Hambrecht, *et al.*, Archaeological sites as Distributed Long-term Observing Networks of the Past (DONOP). *Quaternary International* **549**, 218–226 (2020), long-term perspectives on circumpolar social-ecological systems, doi:<https://doi.org/10.1016/j.quaint.2018.04.016>, <https://www.sciencedirect.com/science/article/pii/S1040618217311746>.
 34. M. P. Ferreira, *et al.*, Individual tree detection and species classification of Amazonian palms using UAV images and deep learning. *Forest Ecology and Management* **475**, 118397 (2020), doi:<https://doi.org/10.1016/j.foreco.2020.118397>, <https://www.sciencedirect.com/science/article/pii/S037811272031166X>.
 35. H. Zhou, F. Jiang, H. Lu, SSDA-YOLO: Semi-supervised domain adaptive YOLO for cross-domain object detection. *Computer Vision and Image Understanding* **229**, 103649 (2023).
 36. S. Zhang, H. Tuo, J. Hu, Z. Jing, Domain adaptive YOLO for one-stage cross-domain detection, in *Asian conference on machine learning* (PMLR) (2021), pp. 785–797.
 37. A. C. Rodríguez, S. D’Aronco, K. Schindler, J. D. Wegner, Mapping oil palm density at country scale: An active learning approach. *Remote Sensing of Environment* **261**, 112479 (2021).
 38. M. Serje, Arquitectura y urbanismo en la cultura Tairona. *Bol. Mus. Oro* **19** (1), 87–96 (1987).
 39. A. Dever, *Social and economic development of a specialized community in Chengue, Parque Tairona, Colombia*, Ph.D. thesis, University of Pittsburgh, Pittsburgh, Pennsylvania (2007).
 40. N. Gutiérrez Montoya, *Las sociedades originarias de la Sierra Nevada de Santa Marta Ocupación espacial y desarrollo arquitectónico Siglos x al xvi* (ediciones doce calles, Madrid) (2022).

41. Y. Yépez, Y. Gutiérrez, *Relato del espacio, territorio y lugar de los asentamientos Kogui dentro de la Línea Negra: Desarrollo cartográfico de ocupación e integración*, Ba thesis, Universidad de la Costa C.U.C., Barranquilla, Atlántico (2020).
42. G. Reichel-Dolmatoff, Cultural Change and Environmental Awareness: A Case Study of the Sierra Nevada de Santa Marta, Colombia. *Mountain Research and Development* **2** (3), 289–298 (1982), <http://www.jstor.org/stable/3673093>.
43. G. Cadavid, L. F. Herrera, Manifestaciones culturales en el área Tairona. *Inform. Antropol.* **1**, 5–54 (1985).
44. A. M. Groot, Arqueología y conservación de la localidad precolombina de Buritaca 200 en la Sierra Nevada de Santa Marta. *Inform. Antropol.* **1**, 55–102 (1985).
45. L. F. Herrera, Agricultural activity in the Sierra Nevada de Santa Marta (Colombia): historical perspective, in *La Sierra Nevada de Santa Marta (Colombia) Transecto Buritaca la Cumbre*, T. van der Hammen, P. M. Ruiz, Eds. (J.Cramer, Berlin-Stuggart, Germany), pp. 501–530 (1984).
46. M. I. Mayorga Hernández, *El agua como elemento constructor de territorio y arquitectura en Colombia. Caso de estudio Teyuna, Sierra Nevada de Santa Marta*, Phd thesis, Universidad de Granada, Granada, Spain (2024), <https://hdl.handle.net/10481/89459>.
47. P. M. van der Hammen, Thomas; Ruiz, *La Sierra Nevada de Santa Marta (Colombia) Transecto Buritaca – La Cumbre*, vol. 2 (Schweizerbart Science Publishers, Stuttgart, Germany) (1986), http://www.schweizerbart.de//publications/detail/isbn/9783443650018/Studies_on_Tropical_Andean_Ecosystems_.
48. A. Henderson, G. Galeano, R. Bernal, *Field guide to the palms of the Americas* (Princeton University Press) (1995).
49. A. Wang, *et al.*, YOLOv10: Real-Time End-to-End Object Detection, in *Advances in Neural Information Processing Systems*, A. Globerson, *et al.*, Eds. (Curran Associates, Inc.), vol. 37 (2024), pp. 107984–108011, https://proceedings.neurips.cc/paper_files/paper/2024/file/c34ddd05eb089991f06f3c5dc36836e0-Paper-Conference.pdf.

50. R. J. G. B. Campello, D. Moulavi, J. Sander, Density-Based Clustering Based on Hierarchical Density Estimates, in *Advances in Knowledge Discovery and Data Mining*, J. Pei, V. S. Tseng, L. Cao, H. Motoda, G. Xu, Eds. (Springer Berlin Heidelberg, Berlin, Heidelberg) (2013), pp. 160–172.
51. GBIF, Dictyocaryum H.Wendl. Occurrence Download, <https://www.gbif.org/occurrence/download/0025145-250310093411724> (2025), doi:10.15468/DL.ND44PT, <https://www.gbif.org/occurrence/download/0025145-250310093411724>, global Biodiversity Information Facility. Accessed: 2025-03-24.
52. NASA JPL, NASA Shuttle Radar Topography Mission Global 1 arc second [Data set], <https://doi.org/10.5067/MEASURES/SRTM/SRTMGL1.003> (2013), nASA Land Processes Distributed Active Archive Center. Accessed: 2025-06-20.

Acknowledgments

Funding: This research is part of the "Mapping the Archaeological Pre-Columbian Heritage in South America – MAPHSA" Project funded by Arcadia – a charitable fund of Lisbet Rausing and Peter Baldwin. This research was also supported by a Leiden-Delft-Erasmus Global Support Grant (Grant No. LDE-127). Pleiades Neo satellite imagery data provided by the European Space Agency.

Author contributions: Conceptualization: S.F., M.Ma., S.M., J.G.S; Data curation: S.F., M.I.M., F.M., K.P., C.R., J.C.V.; Formal analysis: S.F., S.M.; Funding acquisition: S.F., M.Ma., J.G.S.; Investigation: S.F., S.M.; Methodology: S.F., S.M.; Project administration: S.F., M.Ma., M.M., J.G.S., J.C.V., F.W.T.; Resources: C.A., F.M., F.W.T.; Supervision: S.F., M.Ma., F.W.T.; Visualization: S.F., S.M.; Writing – original draft: S.F., S.M.; Writing – review & editing: All authors

Competing interests: There are no competing interests to declare

Data and materials availability: All data supporting the findings of this study are either cited in the references or included in the supplementary materials.

Supplementary materials

Data S1

Data S2

Data S3

Data S4

Data S5

Supplementary Materials for

Ecological Legacies of Pre-Columbian Settlements Evident in

Palm Clusters of Neotropical Mountain Forests

Sebastian Fajardo^{1*†}, Sina Mohammadi^{1*†}, Jonas Gregorio de Souza², César Ardila³,
Alan Tapscott Baltar², Shaddai Heidgen², Maria Isabel Mayorga Hernández⁴,
Sylvia Mota de Oliveira⁵, Fernando Montejo³, Marco Moderato²,
Vinicius Peripato⁶, Katy Puche³, Carlos Reina³,
Juan Carlos Vargas⁷, Frank W. Takes¹, Marco Madella^{2,8,9}

*Corresponding author. Sebastian Fajardo: s.d.fajardo.bernalliacs.leidenuniv.nl;

Sina Mohammadi: s.mohammadi@liacs.leidenuniv.nl

†These authors contributed equally to this work.

This PDF file includes:

Other Supplementary Materials for this manuscript:

Captions for Data S1 to S5

Caption for Data S1. ArchaeologicalCentroids.txt This geojson dataset contains the approximate central location of reported archaeological sites or zones inside the area of study. The dataset includes the ID_site field, which provides a unique identifier or commonly used name for each site, and the ID_source_field, referencing the source from which the site's ID information was derived. Additionally, the Location_source field indicates the main source of the geolocation. When this field is marked as *N/A*, it denotes that the coordinates were established as part of the present study. This dataset is provided in GeoJSON format and uses the WGS84 coordinate reference system. The file is shared as a .txt but contains a valid GeoJSON. File can be changed to .geojson to open it directly in GIS software or GeoJSON-compatible tools.

Caption for Data S2. Study_Area.txt This geojson dataset includes the area of study. the polygon includes the perimeter in kilometers and the area in square kilometers. It is provided in GeoJSON format and uses the WGS84 coordinate reference system. The file is shared as a .txt but contains a valid GeoJSON. File can be changed to .geojson to open it directly in GIS software or GeoJSON-compatible tools.

Caption for Data S3. Manual_Labels_Palms.txt This geojson dataset includes manually annotated polygon features representing individual palms. Each polygon is assigned a unique ID. It is provided in GeoJSON format and uses the WGS84 coordinate reference system. The file is shared as a .txt but contains a valid GeoJSON. File can be changed to .geojson to open it directly in GIS software or GeoJSON-compatible tools.

Caption for Data S4. Labeled_area.txt This geojson dataset includes the area where palms were manually annotated. It is divided in cells of 200 x 200 m. Each polygon is assigned a unique ID, with their area in square meters and their perimeter in meters. The file is provided in GeoJSON format and uses the WGS84 coordinate reference system. The file is shared as a .txt but contains a valid GeoJSON. File can be changed to .geojson to open it directly in GIS software or GeoJSON-compatible tools.

Caption for Data S5. detected_palms_04_threshold.txt This geojson dataset consists of polygon features representing automatically detected palm clusters with a probability ≥ 0.4 provided by our

deep learning model. Each polygon is associated with a unique identifier. This dataset is provided in GeoJSON format and uses the WGS84 coordinate reference system. The file is shared as a .txt but contains a valid GeoJSON. File can be changed to .geojson to open it directly in GIS software or GeoJSON-compatible tools.