horticulturae

MDPI

Article

Predicting Suitable Regions for Avocado (*Persea americana* Mill.) Tree Cultivation in Tanzania

Ibrahim Juma [1,*,†], Jhon B. Valencia [2,†] and Andrés J. Cortés [3,*,‡]

[1] Department of Food Science and Technology, University of Dar es Salaam, P.O. Box 35134 Dar es Salaam, Tanzania
[2] African Plant Nutrition Institute (APNI), Benguérir 43150, Morocco; jbrayanvalenciag@gmail.com
[3] Department of Plant Breeding, Swedish University of Agricultural Sciences, 23436 Lomma, Sweden
* Correspondence: vuga.ibrahim@udsm.ac.tz (I.J.); andres.cortes@slu.se (A.J.C.)
† These authors contributed equally to this work.
‡ Current address: Facultad de Ciencias Agrarias—Departamento de Ciencias Forestales, Universidad Nacional de Colombia—Sede Medellín, Medellín 050034, Colombia.

Abstract

Avocado cultivation is expanding rapidly in East Africa, driven by growing market demand, yet planning often relies on farmers' experience rather than systematic spatial analysis, raising risks of inefficient land and resource use. Therefore, this study applied four species distribution models (SDMs), Generalized Additive Models (GAM), Boosted Regression Trees (BRT), Maximum Entropy (MaxEnt), and Random Forest (RF), along with an ensemble model to map potential avocado suitability in Tanzania. The models were calibrated using 199 Variance Inflation Factor (VIF)-depurated occurrence records from which climatic, edaphic, and topographic predictor variables were extracted. BRT and RF had the best predictive abilities, with AUC values ranging from 0.77 ± 0.20 to 0.81 ± 0.13. The individual models identified Njombe, Iringa, Songwe, Kigoma, Rukwa, Kagera, and Morogoro as regions with high suitability, with more than 30% of each region's total area predicted to be suitable for avocado production. Moderate suitability (15% to $\leq$30% of the regional area) was recorded for Kilimanjaro, Arusha, Dodoma, Manyara, Mara, Mbeya, Ruvuma, Tanga, and Katavi, whereas negligible suitability was forecasted for most of the remaining regions by the majority of the models. These findings suggest that heavy investments in avocado production and value chain additions should be directed primarily to regions with high suitability in order to use resources efficiently and minimize investment risks. More targeted, site-specific management should be encouraged in moderately suitable regions, with a focus on helping farmers identify and manage the best avocado sites rather than promoting broad expansion across the country. The findings generated by the ensemble model could be incorporated in the Tanzania Agriculture Climate Adaptation Technology Deployment Programme (TACATDP) to enhance sustainable crop investment, lower production risks, and strengthen the resilience of the avocado sector in the country.

Keywords: species distribution modelling (SDM); ecological niche modelling (ENM); agricultural planning; ensemble model; environmental predictors

Check for updates

Academic Editor: Esmaeil Fallahi

Received: 14 November 2025
Revised: 15 December 2025
Accepted: 18 December 2025
Published: 25 December 2025

1. Introduction

Avocado (*Persea americana* Mill.) production has been steadily rising on a global scale, supported by constantly increasing consumer demand, nutritional awareness, and accessibility to high-value international markets [1]. The world avocado production was approximately 10.47 million metric tonnes in 2023 (about 10% increase from the production of the previous year), with the leading countries being Mexico, Colombia, and Peru, which produced 3, 1.09, and 0.87 million tonnes, respectively [2,3]. When the production is accounted at the regional level, 2022 global avocado production was led by the Americas, which produced 72% (about two-thirds) of global avocado production, while production was 14% in Africa and 11% in Asia, and a total of 3% for both Europe and Oceania [4]. The global avocado trade has been growing steadily, too. In 2023, exports totalled around 2.9 million metric tonnes (worth USD 7.26 billion), representing a 10% increase over 2022 levels. Mexico and Peru led the market, exporting 1.22 million tonnes and 598,000 tonnes, respectively, valued at USD 2.8 billion and USD 963 million, respectively [5,6]. The leading avocado importer is the United States, which imported about 1.26 billion kilograms of avocados in 2023, valued at approximately USD 3.09 billion. The second and third top avocado importers were the European Union and the Netherlands, which imported 756.81 and 375.09 million kg of avocados worth USD 2.20 and 1.16 billion, respectively. Other two countries that followed this trend were France and Spain, which imported 189.65 and 240.57 million kg of avocados, valued at USD 584.68 and 558.52 million, respectively [7].

Given the steady growth in avocado demand and cultivation, species distribution modelling (SDM), or ecological niche modelling (ENM), has been used to predict potential suitable areas for establishing new avocado orchards. This approach employs statistical relationships between observed species or crop occurrences and environmental variables to determine potential niches and predict geographical suitability [8]. The application of SDM to perennial crops such as avocados helps map agro-ecological zones and locate underexploited high-potential sites. Additionally, it can establish a correlation between predicted suitability and yield data [9–11]. SDM can therefore enable decision makers to undertake risk-informed planning when integrated with data on biodiversity, disease, water, and infrastructure. SDM has also been utilized to address pollinator overlap and land-use pressures [12–15]. SDM encompasses presence-only approaches, most notably Maximum Entropy (MaxEnt), and presence/absence or pseudo-absence regressions such as Generalized Additive Models (GAM). Also, machine learning models including Boosted Regression Trees (BRT) and Random Forests (RF) are included in SDM [9,10,16]. These algorithms may be combined in ensembles to minimize single-model bias [10,16]. Model development involves gathering crop occurrence records and environmental predictors, curating them, fitting and tuning models with geographically structured cross-validation, and assessing performance before projecting to some climate scenarios [9,10]. Operationally, SDM outputs are frequently integrated with multi-criteria decision analysis (MCDA), specifically, the Analytic Hierarchy Process (AHP) and weighted overlays, to incorporate terrain and soil variability, as well as access infrastructure and market proximity [17,18].

Evidence from species distribution models (SDMs) across multiple regions shows consistent, policy-relevant patterns in avocado suitability. In the Andes, SDMs for 'Hass' avocados revealed new high-suitability zones and showed positive correlations between anticipated suitability and orchard productivity [9]. A continental assessment of Hass avocado across the Americas found that by the mid-21st century, the total area of suitable land is expected to remain stable, with shifts (upslope/poleward reconfiguration) and losses (hot lowland regions) being the main changes [10]. According to SDMs and land-change research for avocado cultivation in Mexico, future warming patterns will shift high suitability toward elevated regions while simultaneously increasing the risk of habitat fragmentation and deforestation, while mismatching pollinator trajectories [12–16]. Studies in Europe and the Eastern Mediterranean highlight frosty and cooler spring nights as climate constraints for avocado cultivation [19,20].

Beyond this global evidence, avocado production in East Africa is organized very differently from the major producing regions of the Americas, where avocados originated [21], and the Mediterranean Europe [22]. In Kenya, Tanzania, and neighboring countries, orchards are largely managed by smallholder farmers in cool, high-elevation tropical zones, where trees are commonly grown under rain-fed conditions, intercropped with other species, and established from a mixture of cultivars and heterogeneous planting material [23–25]. In contrast, most empirical and modelling studies have been developed for large, specialized commercial orchards dominated by grafted clonal 'Hass' in subtropical or Mediterranean climates, where fully irrigated, high-input systems are the norm [22,26,27]. These differences in climate, water management, varietal composition, farm structure, and value-chain development indicate that suitability assessments and agronomic recommendations generated for Latin American and Mediterranean systems are unlikely to be directly applicable to East African conditions without explicit regional testing and adaptation. Yet, despite this clear need for region-specific guidance, there is currently no available SDMs that map suitable areas for avocado growing in East Africa.

Tanzania is an East African country, and its economy depends heavily on agriculture. This sector contributes approximately 23% to its Gross Domestic Product (GDP). The agriculture sector employs about 65% of the nation's workforce and generates approximately 20% of the country's export earnings [23,28]. Within the sector, one of the key horticultural crops gaining recognition in Tanzania is the avocado. This crop is widely cultivated in the Mbeya, Njombe, Songwe, Iringa, Kilimanjaro, Arusha, Tanga, Kigoma, Kagera, and Mara regions [29]. In 2023, the country exported about 25.06 million kilograms of avocados, whose value was USD 37.46 million [30,31]. Export destinations included the Netherlands (11.46 million kg, USD 17.07 million), India (3.23 million kg, USD 4.93 million), the United Arab Emirates (1.95 million kg, USD 3.60 million), the United Kingdom (2.16 million kg, USD 2.47 million), and France (1.11 million kg, USD 1.72 million). Similarly, the country exported about 21,000 kg of avocados to Kenya, which contributed USD 2260 to Tanzania's export earnings, a sign of strengthening East African regional trade [31]. These data show how Tanzania is emerging as a player in the global avocado trade, especially in the European, Asian, and African markets.

Despite Tanzania's growing role in the global avocado trade and the crop's rising economic importance, the expansion of cultivation areas has largely relied on anecdotal knowledge, farmer experience, and historical patterns of success. To the best of our knowledge, there are currently no spatially explicit, model-based frameworks to guide the strategic identification of ecologically optimal regions for avocado production. Avocado is highly sensitive to its growing environment, and orchard performance is strongly governed by the interaction of bioclimatic, topographic, and edaphic factors. Depending on the cultivar, avocado grows in areas with frost-free climates with mean annual temperatures of 12.5 °C to 25.5 °C [32–34] and annual rainfall of about 800 to 2000 mm that is reasonably well distributed throughout the year [23,35,36]. Additionally, the crop thrives at elevations of approximately 700–2500 m above sea level in tropical highland systems [34–38] on gentle to moderate slopes (generally <15%). It performs best on deep (≥0.8–1.0 m), well-drained loamy to sandy-loam soils, alluvial loams, volcanic, or even calcareous soils with slightly acidic to slightly alkaline pH (≈5.5–7.5) and high amounts of organic matter [33,39]. Taken together, these three groups of factors form the central core for delineating geographical regions that are most suitable for establishing new plantations. The absence of planning tools that explicitly integrate these variables creates a gap in geospatial and agro-ecological planning, which in turn imposes a barrier to the effective utilization of the resources, including land, capital, and time, while limiting productivity potential and increasing vulnerability to environmental risks. The lack of predictive spatial tools for avocado in Tanzania limits the

https://doi.org/10.3390/horticulturae12010024

sector's sustainable development and long-term competitiveness, given that the demand for Tanzanian avocados continues growing in domestic and international markets.

Thus, this research addresses this research gap by utilizing SDMs to predict and map the regions that are suitable for growing avocado in Tanzania. The findings from this research will guide policymakers and farmers in locating optimal areas for establishing and/or expanding avocado farming, thereby reducing the risks of crop failure and preventing or minimizing wastage of resources associated with crop failure. In addition to that, the study findings will support the Tanzanian government in formulating targeted agricultural policies and land-use plans, along with informing investors and development partners who seek evidence-based locations for establishing and/or expanding crop-related infrastructure. Ultimately, the research will enhance avocado production while improving national food security, export earnings, and sustainable development in the rural communities, given the fact that this crop is mainly grown by smallholder farmers living in rural areas.

2. Materials and Methods

2.1. Study Sites and Sampling

The study was undertaken in the Morogoro, Iringa, Mbeya, Njombe, Songwe, Rukwa, Ruvuma, Kigoma, Mwanza, Kagera, and Mara regions, which are located in eastern, southern, southwestern, western, northwestern, and northern Tanzania (Figure 1). These regions were selected because they encompass the main avocado-producing zones in the country and capture wide variation in the environmental and management conditions that influence avocado growth. Fieldwork was undertaken across these regions between 2017 and 2024, where 400 local avocado trees were sampled for morphological and genetic diversity analyses. We employed Garmin Epix GPS mapping and multisport watch to determine the latitude and longitude of the collecting sites. These georeferenced occurrence points provided the foundation for SDM and served as the spatial basis for predicting suitable regions for avocado cultivation across Tanzania. The map (Figure 1) and spatial data processing were performed using R statistical software version 4.5.0 [40] in RStudio version 2025.05.01 [41].

2.2. Presence Data and Cleaning

We performed a number of rigorous filtering steps to ensure data quality and minimize spatial bias. They included removing duplicate coordinates to retain only unique records and thereafter excluding records falling outside the boundaries of Tanzania. We filtered out records over water bodies or areas lacking soil information by applying a 1 km resolution iSDA (Innovative Solutions for Decision Agriculture) soil raster tailored to the study region [42]. This dataset derives from iSDAsoil, an open-access, machine learning-based soil mapping product covering Africa at 30 m resolution, providing spatially explicit predictions for key physical and chemical soil properties, with associated uncertainty estimates, for use in environmental and agricultural analyses. To align the spatial resolution of the soil data with the other environmental predictors (e.g., climatic variables), we upscaled the 30 m iSDA rasters to 1 km by calculating the mean value of the cells within each 1 km grid cell, a standard aggregation technique in species distribution modelling [42]. Subsequently, a spatial thinning procedure with 1 km distance was applied to mitigate sampling bias [43]. After these cleaning steps, the dataset was reduced to 199 high-quality, spatially independent records, which were subsequently used for species distribution modelling (Figure 1; Table S1).

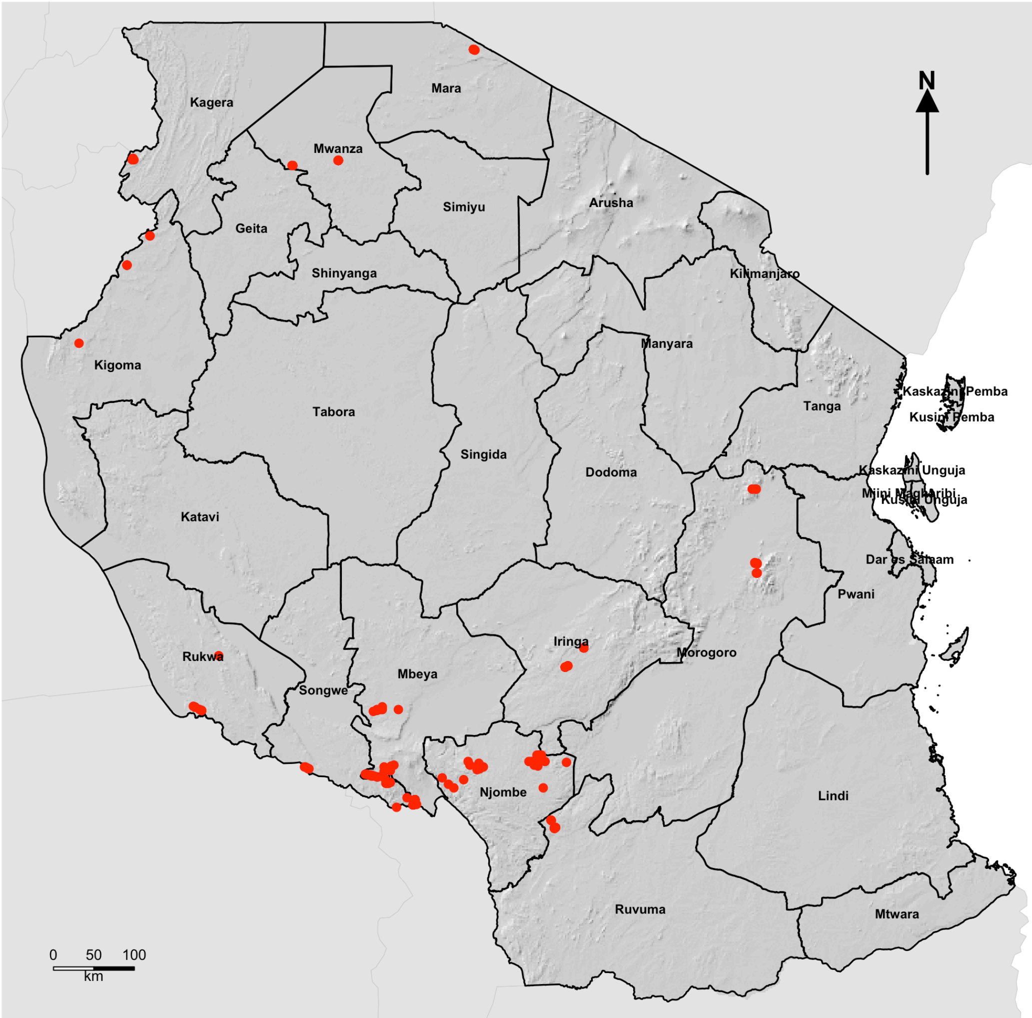

Figure 1. Map of Tanzania showing regions included in avocado tree sampling. Presence locations are marked in red.

2.3. Pseudo-Absence Generation and Data Preparation

Given the limited number of presence records, a distance-constrained pseudo-absence method was employed to reduce spatial bias. Specifically, each presence point was buffered by 20 km, effectively excluding nearby areas from background point generation. The remaining portion of the Tanzanian boundary, defined as the study area minus the buffered zones, was used to randomly sample candidate pseudo-absence points.

To ensure robust environmental representation, we adopted a model-specific sampling strategy as recommended by Barbet-Massin et al. [44]. For machine learning algorithms (RF, BRT), a subset equal to the number of presence records (1:1 ratio) was selected to maximize discriminatory power, whereas for GAM and MaxEnt, a larger background sample (10,000 points) was retained to better define the environmental envelope.

Environmental data was then extracted from the predictor rasters for both presence and pseudo-absence points using the terra package version 1.8-87 [45]. The two datasets were combined into a single response dataset with an indicator variable (1 for presence, 0 for absence), and any remaining missing values were removed, thus ensuring a clean and balanced dataset for model calibration.

2.4. Variable Selection

To model the distribution of avocado in Tanzania, we initially assembled a set of 31 predictor variables (Figure 2). These included 19 bioclimatic variables (bio_01 to bio_19) describing annual and seasonal patterns of temperature and precipitation. A complete list of these variables, including their names and units (e.g., annual precipitation [bio_12]), is provided in Table S2. The dataset also included 2 topographic variables (elevation and slope) and 10 edaphic factors. We were guided by expert recommendations for avocado nutrition in the selection of edaphic/soil variables, and we sourced them from the iSDAsoil database for Tanzania [42]. Specifically, we incorporated a set of chemical soil properties, including aluminum content ("al_0–20 cm"), pH ("ph.h2o_0–20 cm"), total carbon ("c.tot.0–20 cm"), organic carbon ("oc_0–20 cm"), and total nitrogen ("n.tot.ncs_0–20 cm"), alongside key physical properties, sand ("sand.tot.psa_0–20 cm"), silt ("silt.tot.psa_0–20 cm"), clay percentage ("clay.tot.psa_0–20 cm"), bedrock characteristics ("bdr_0–200 cm"), and bulk density ("db.od_0–20 cm").

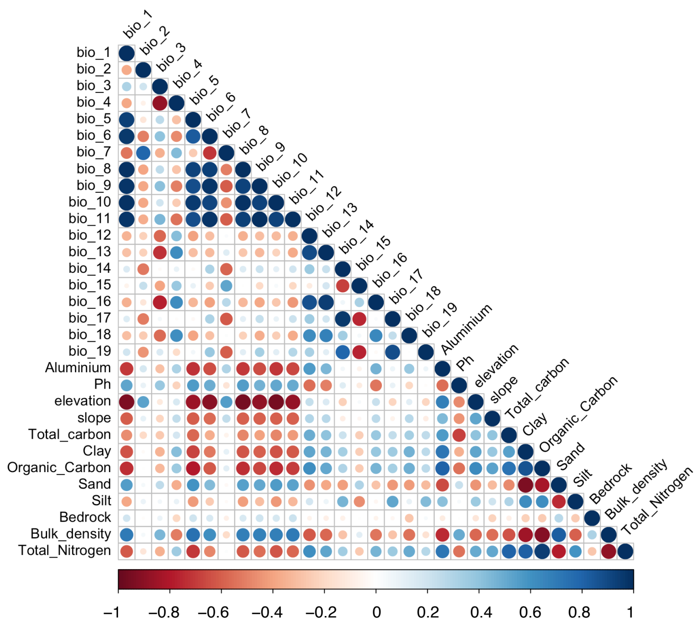

Figure 2. Correlogram among a priori predictor variables selected to model avocado-suitable areas in Tanzania. Bioclimatic variables are abbreviated as in WorldClim [46].

2.5. Variable Reduction

We performed Variance Inflation Factor (VIF) analysis [47] to identify and remove redundant predictors. This diagnostic step was very crucial since it ensured that the final model would not be compromised by correlated variables, enhancing the reliability and interpretability of the model results. Following execution of VIF analysis, we managed to reduce the predictor set from 31 to 11 predictors (Figure 3), which were the most critical variables relevant for avocado distribution in Tanzania. This approach not only minimized the inclusion of redundant predictors in avocado suitability modelling for Tanzania, but also reduced the risk of model overfitting.

2.6. Species Distribution Models

2.6.1. Model Training, Tuning, and Evaluation

To address spatial autocorrelation, the final presence–absence dataset was validated using spatial block cross-validation (5 folds) via the spatial sample package version 0.6.1 [48], ensuring spatial independence between training and testing sets. Four modelling techniques, i.e., Generalized Additive Models (GAM), Boosted Regression Trees (BRT/GBM), Maximum Entropy (MaxEnt), and Random Forest (RF), were applied to predict avocado occurrence. The models were evaluated through assessing the area under the ROC curve (AUC), true skill statistics (TSS), and sensitivity.

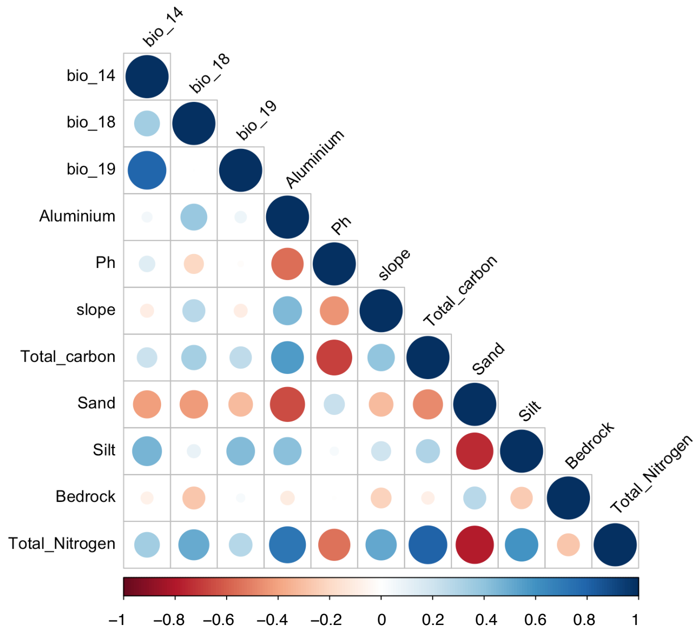

Figure 3. Retained predictor variables (11) after VIF analysis to model avocado distribution in Tanzania, with bioclimatic variables being abbreviated according to WorldClim [46].

2.6.2. Spatial Prediction and Thresholding

After model calibration and evaluation, spatial predictions were generated over the entire study area. Each model produced a continuous probability surface reflecting the suitability for avocado occurrence. Continuous suitability scores were converted to binary maps (suitable/unsuitable) using thresholds derived from the Youden Index (J = sensitivity + specificity − 1) [49,50], which maximizes the sum of sensitivity and specificity and balances omission and commission errors.

Critically, thresholds were calculated and applied individually for each algorithm to account for their distinct probability distributions. As detailed in Table S3, the algorithms required varying cut-offs to achieve optimal performance (e.g., RF required a stricter threshold of ~0.323, whereas GAM utilized a lower threshold of ~0.032).

To develop the ensemble avocado habitat suitability model for Tanzania, we integrated predictions from the four SDMs (GAM, BRT, MaxEnt, and RF) using a spatially conditional mean of their suitability outputs. This robust approach ensures model integrity by calculating the mean of all four models in inland areas, but conditionally excluding the GAM model (averaging only BRT, MaxEnt, and RF) in coastal zones (<10 km) to mitigate known regional extrapolation (Equation (1)).

$$\text{Ensemble}(x) = \frac{1}{4}\sum_{m=1}^{4} s_m(x) \tag{1}$$

where x is the pixel and $s_m(x) \in [0,1]$ is the normalized suitability predicted by model $m \in \{1,2,3,4\}$ that represents BRT, RF, GAM, and MaxEnt, respectively.

The final binary ensemble consensus map was generated using an AUC-weighted threshold of 0.184. This value was calculated as the weighted average of the individual models' optimal Youden Index thresholds, ensuring that the final binary consensus is proportional to the validated predictive power (AUC mean) of each algorithm.

3. Results

3.1. Model Evaluations

We evaluated four SDMs using AUC, TSS, and optimized probability thresholds (Table 1). Performance was good under rigorous spatial cross-validation (AUC ranges 0.72–0.81), indicating robust discrimination between suitable and unsuitable avocado conditions. RF was the top performer (AUC = 0.81 ± 0.13), followed by BRT (0.77 ± 0.20), while MaxEnt and GAM followed with 0.75 ± 0.08 and 0.72 ± 0.15, respectively.

Table 1. Model evaluation metrics (mean ± SD) derived from spatial block cross-validation. Thresholds were determined using the Youden Index (sensitivity + specificity − 1).

Model	AUC	TSS	Threshold
RF	0.81 ± 0.13	0.62 ± 0.27	0.32 ± 0.09
BRT	0.77 ± 0.20	0.53 ± 0.33	0.28 ± 0.25
MaxEnt	0.75 ± 0.08	0.54 ± 0.18	0.09 ± 0.05
GAM	0.72 ± 0.15	0.55 ± 0.25	0.03 ± 0.05

Threshold-based metrics showed that the models behaved quite differently. RF reached the highest TSS (0.62 ± 0.27), but only at a relatively high probability cut-off (0.32 ± 0.09). In contrast, GAM achieved a similar TSS (0.55 ± 0.25) at a very low threshold (0.03 ± 0.05), indicating a more sensitive and less conservative classification of suitable conditions. MaxEnt (0.09 ± 0.05) and BRT (0.28 ± 0.25) fell in an intermediate threshold range. Because TSS = sensitivity + specificity − 1, these values indicate that RF provides the best balance between true positives and true negatives. Although all models showed reasonable discrimination, the TSS values indicate that RF (0.62 ± 0.27) separated presences from background points more effectively than GAM (0.55 ± 0.25), MaxEnt (0.54 ± 0.18), and BRT (0.53 ± 0.33).

Optimal thresholds varied markedly among the algorithms, reflecting their different calibration to environmental gradients. RF required the highest probability cut-off (0.32 ± 0.09) to define suitability, indicating a conservative model that emphasizes high-confidence locations. In contrast, GAM (0.03 ± 0.05) and MaxEnt (0.08 ± 0.05) optimized at very low thresholds, suggesting that they produced broad probability surfaces in which even low values were treated as suitable, likely contributing to the expansive predictions in coastal zones. BRT occupied an intermediate position (0.28) but showed high variability (SD = 0.25), indicating that its suitable–unsuitable boundary was unstable across spatial folds.

3.2. Spatial Suitability Patterns Based on Individual Models

The models differed in their level of conservatism, with RF and BRT and, to a lesser extent, MaxEnt highlighting the Southern and Northern Highlands, as well as the Western and Lake regions, as major avocado production zones (Table 2; Figure 4). RF and BRT were the most robust algorithms, identifying broad areas of suitable habitat across these regions, whereas GAM and, to a lesser degree, MaxEnt produced much more restrictive suitability patterns.

Njombe emerged as the premier region for avocado cultivation, characterized by strong consensus between the tree-based models, with suitability estimated at 84.7% (RF) and 82.0% (BRT). MaxEnt was comparatively conservative in Njombe, identifying 65.1% of the area as suitable, whereas GAM predicted only minimal suitability (8.9%). A similar pattern occurred in Rukwa, where RF (73.6%) and BRT (70.7%) indicated very high potential, but MaxEnt identified just 26.0% of the region as suitable. Kagera and Iringa also stood out as important avocado-growing regions, with suitable land covering 43.9–58.3% of Kagera and 43.3–48.0% of Iringa as estimated by BRT and RF. Other regions that displayed > 30% suitability in at least two models are Songwe (BRT: 36.1%; RF: 35.2%) and Kigoma (BRT: 30.7%; RF: 42.1), with the models effectively agreeing on the available niche for the Songwe region. Morogoro recorded > 30% suitability in only the RF model (43.6%), whereas Kaskazini Unguja scored suitability exceeding 30% in BRT (57.1%) only.

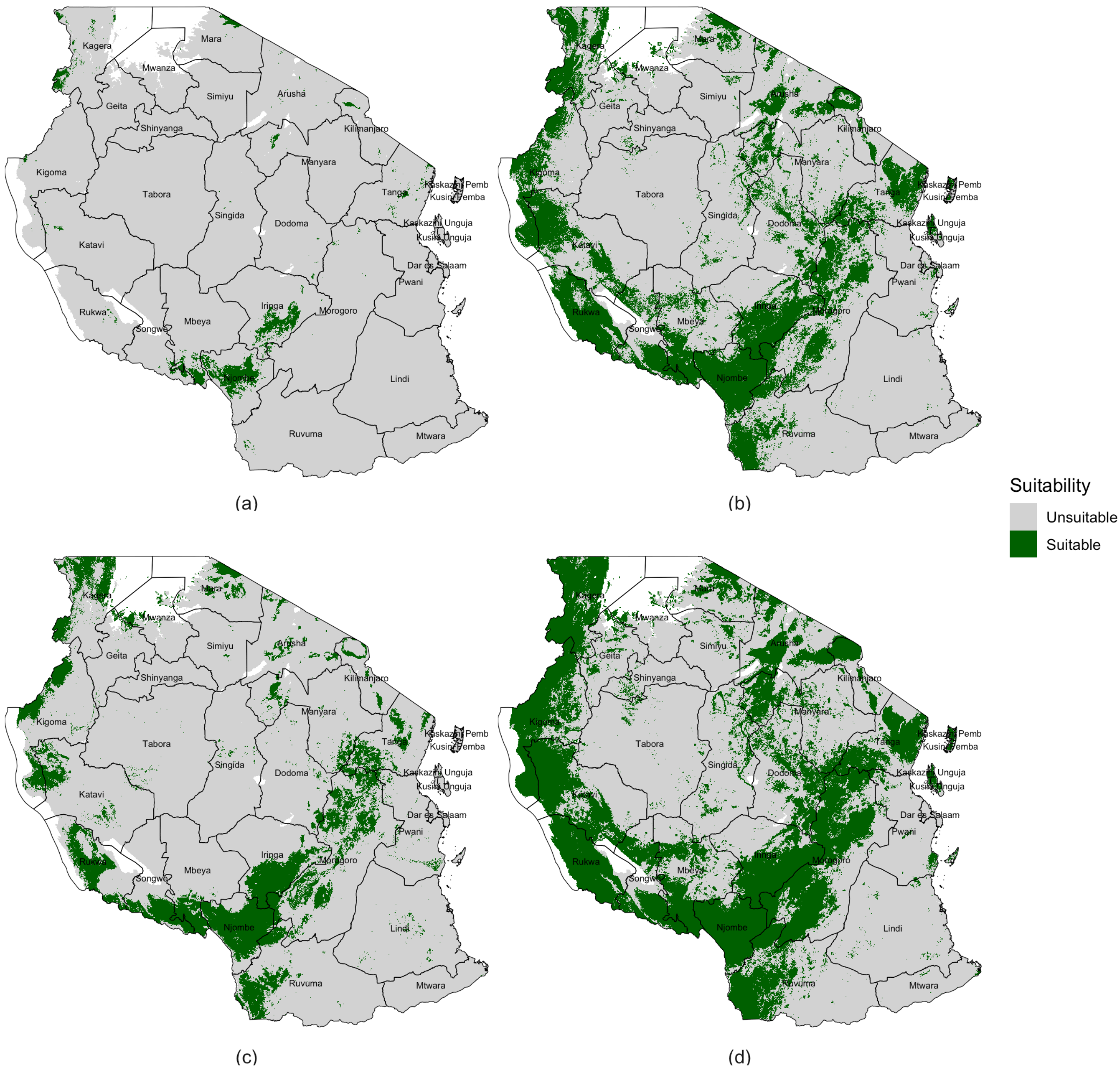

Figure 4. Binary maps of avocado suitable areas (in green) across regions in Tanzania summarizing outputs from (**a**) Generalized Additive Models (GAM), (**b**) Boosted Regression Trees (BRT), (**c**) Maximum Entropy (MaxEnt), and (**d**) Random Forest (RF).

Meanwhile, moderate suitability levels of 15 to ≤30% were assigned across the models for Arusha (GAM: 0.0%, RF: 22.3%), Kaskazini Pemba (GAM/MaxEnt: 0.0%, BRT: 26.2%), Katavi (GAM: 0.0%, RF: 28.8%), Kilimanjaro (GAM: 1.2%, RF: 30.1%), Mara (GAM: 2.3%, RF: 18.6%), Mbeya (GAM: 2.2%, RF: 26.5%), Ruvuma (GAM: 0.0%, RF: 20.7%), and Tanga (GAM: 0.1%, BRT: 29.7%). The tree-based models consistently identified viable niches in these regions that were overlooked by the regression approaches. Areas with moderate suitability can still grow avocado effectively, but suitable land is more limited and/or patchy compared to highly suitable regions, so achieving good results depends more on careful site selection and management.

Table 2. Percentage of land area in each administrative region of Tanzania suitable for avocado cultivation based on the four SDMs (BRT, GAM, MaxEnt, and RF).

Region	BRT	GAM	MaxEnt	RF
Arusha	17.5	0.0	4.1	22.3
Dar es Salaam	1.8	0.0	0.0	0.0
Dodoma	14.0	0.0	1.3	11.5
Geita	1.7	0.0	0.4	5.2
Iringa	43.3	1.6	24.9	48.0
Kagera	43.9	5.2	23.8	58.3
Kaskazini Pemba	26.2	0.0	0.0	4.4
Kaskazini Unguja	57.1	0.0	0.2	8.5
Katavi	26.8	0.0	2.6	28.8
Kigoma	30.7	0.0	11.2	42.1
Kilimanjaro	25.5	1.2	6.6	30.1
Kusini Pemba	13.4	0.0	0.0	1.0
Kusini Unguja	7.5	0.0	0.0	1.2
Lindi	0.1	0.0	0.1	0.0
Manyara	9.7	0.1	2.6	11.6
Mara	15.3	2.3	10.1	18.6
Mbeya	23.3	2.2	10.1	26.5
Mjini Magharibi	24.1	0.0	0.0	0.0
Morogoro	26.1	0.0	11.6	43.6
Mtwara	0.0	0.0	0.0	0.0
Mwanza	8.0	0.0	6.7	9.4
Njombe	82.0	8.9	65.1	84.7
Pwani	0.7	0.0	0.7	1.1
Rukwa	70.7	0.0	26.0	73.6
Ruvuma	15.0	0.0	6.8	20.7
Shinyanga	0.0	0.0	0.0	0.1
Simiyu	0.8	0.0	0.1	1.4
Singida	3.1	0.0	0.2	2.5
Songwe	36.1	0.7	20.3	35.2
Tabora	0.1	0.0	0.0	0.2
Tanga	29.7	0.1	7.6	24.8

Several regions were consistently classified as unsuitable by all models, with suitability values close to zero. The coastal lowlands regions such as Dar es Salaam (0–1.8%), Lindi (0–0.1%), Mtwara (0.0%), and Pwani (0–1.1%) were universally classified as poor avocado-growing zones. Tabora (0.0–0.2%), Shinyanga (0.0–0.1%), Simiyu (0.0–1.4%), and Singida (0.0–3.1%) were also identified with negligible suitability for crop production. In these regions, low suitability reflects the very small proportion of land meeting avocado requirements, rather than the overall magnitude of the climatic factors themselves. Biophysical constraints such as saline soils, high evapotranspiration, and erratic rainfall drive these low suitability scores.

Lastly, in the Zanzibar archipelago, a notable prediction divergence was observed. GAM and MaxEnt forecasted negligible suitability (0–0.2%) across the islands, whereas BRT predicted exceptionally high suitability in Kaskazini Unguja (57.1%), Kaskazini Pemba (26.2%), and Mjini Magharibi (24.1%), and RF produced intermediate values (1.0–8.5%) (Table 2). This pattern suggested that BRT may have been oversensitive to particular environmental gradients in these islands while GAM and MaxEnt imposed much stronger constraints and RF occupied an intermediate position. The archipelago is characterized by a unique environmental mosaic, including low elevation gradients paired with distinct maritime climate regimes and coral-rag soils, which pose challenges for model transferability from mainland training data.

3.3. Prediction Agreement of Avocado Distribution in Tanzania

Since GAM exhibited low performance and/or introduced statistical artifacts compared to the other algorithms (BRT, MaxEnt, and RF), we excluded the results of the GAM from the prediction agreement. This was performed on purpose to ensure that the resulting spatial outputs reflect the most reliable and robust consensus and variability among the three best-performing models.

The consensus between the prediction results generated by the three selected machine learning models (BRT, MaxEnt, and RF) was visually assessed through a model voting map, which illustrated the spatial distribution of suitability agreement (Figure 5). The consensus map depicted 'Model Votes' on a scale of 0 to 3, where a score of 3 indicated that all three algorithms converged on a prediction of high suitability. These areas of highest consensus (dark green) formed a distinct and continuous belt in the Southern Highlands, specifically covering the high-altitude districts of Mbeya, Njombe, and Iringa. Similarly, strong model agreement was found in the Northern Highlands (Kilimanjaro and Arusha) and in the western regions (Kigoma and Kagera). Conversely, most of the central plateau (Dodoma, Singida, and Tabora) received a vote count of 0 (gray), indicating strong consensus among all three models that these regions are unsuitable for avocado cultivation. Intermediate vote counts of 1 or 2 (pink and blue) outlined transitional zones, including parts of Morogoro and the lake-border regions, indicating areas where the models disagreed in their binary classifications.

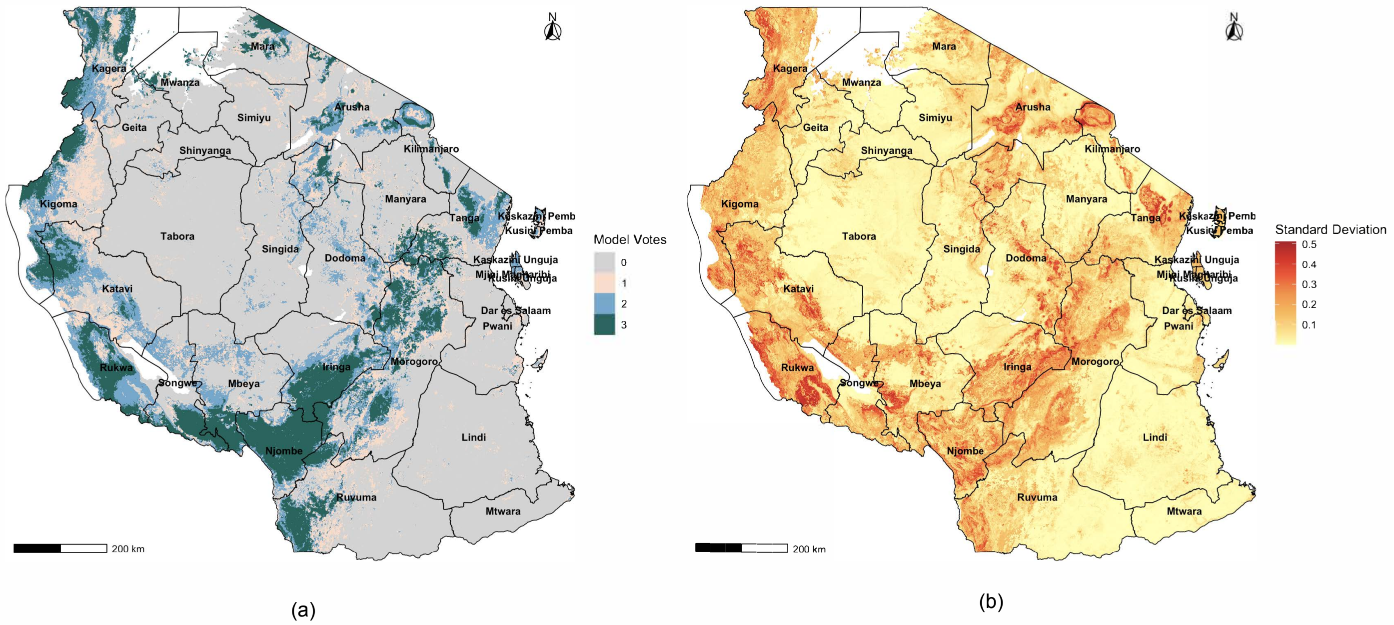

Figure 5. Inter-model uncertainty for avocado suitability across Tanzania generated based on the results from the three most robust models (BRT, MaxEnt, RF), displaying agreement between the three models (**a**) and variability among the three models (**b**). Low standard deviation (SD) values (yellow) indicate high model consensus, while high SD values (dark red) indicate high divergence and higher prediction uncertainty.

Prediction variability was further assessed using a standard deviation map of suitability scores for the three best-performing algorithms, excluding GAM, which highlighted areas of greater statistical uncertainty (Figure 5). The SD values in this analysis ranged from 0.1 to 0.5. Regions with high uncertainty (SD $\approx$ 0.3–0.5; orange to red areas) were concentrated mainly in the Southern Highlands, including Mbeya, Njombe, Iringa, and parts of Ruvuma, as well as in the Northern Highlands around Arusha and Kilimanjaro. Similar pockets of elevated SD were evident in western Tanzania (Kagera, Kigoma, Katavi, and Rukwa). These areas corresponded to environments characterized by sharp elevation gradients, variable rainfall patterns, and heterogeneous land cover. Such complexity leads models to respond differently to environmental predictors, resulting in higher variability among predictions.

Areas with moderate uncertainty (SD $\approx$ 0.2–0.3; light to medium orange) were evident in transitional zones, including parts of Morogoro, the Lake Zone margins (notably Mwanza, Geita, and Mara) and sections of the eastern interior such as Tanga and Manyara. These regions typically lie between highland and lowland systems, where small differences in temperature, moisture availability, or topographic position may shift model outcomes across suitability thresholds.

In contrast, low uncertainty (SD < 0.1–0.2; pale yellow) dominated much of the central plateau (i.e., Dodoma, Singida, and Tabora) and the coastal and southern lowlands (i.e., Dar es Salaam, Pwani, Lindi, and Mtwara). The consistently low SD in these regions pointed to strong agreement among the models and reflected broadly uniform climatic constraints, principally low and highly seasonal rainfall and heat stress, that restricted avocado suitability regardless of the algorithm used.

Taken together, the SD patterns showed more uncertainty in climatically and topographically complex landscapes, many of which are also important agricultural areas, whereas confidence in the predictions was highest in clearly unsuitable zones. As a result, suitability outputs should always be interpreted together with their associated uncertainty when guiding site selection and regional planning for avocado cultivation.

4. Discussion

4.1. Tree-Based Algorithms Perform Better than Regression and Presence-Only Models in Species Distribution Model Evaluation

Model evaluation in SDM can provide information about predictive accuracy, ecological validity, and the transferability of models to different contexts [51,52]. In this study, all four SDMs achieve AUC values between 0.72 and 0.81 (Table 1), indicating an excellent level of discriminatory capacity [53]. The two tree-based algorithms had the highest AUCs (BRT $\approx$ 0.77; RF $\approx$ 0.81; Table 1), which reflects their ability to model nonlinear responses and interactions without parametric constraints. Our results agree with the findings by Mi et al. [54], Fern et al. [55], and Valavi et al. [56], who noted BRT and RF surpassed regression-based or presence-only models in working with complex environmental gradients.

The highest value for RF's TSS (0.615) (Table 1) points to the most consistent equilibrium between omission and commission errors for the model [51]. Comparatively, GAM and MaxEnt scored lower TSS (0.546 and 0.536, respectively; Table 1), which points to a lower balance between sensitivity and specificity. Although GAMs are excellent at fitting smooth responses, they struggled on the environment with complex interactions and strong non-linearities [51,57–59]. MaxEnt, on the other hand, sometimes generated overly generalized suitability estimates over heterogeneous landscapes when they were fed limited data [60,61].

Optimal classification thresholds RF/BRT were higher (0.32–0.28; Table 1) compared to the ones for GAM and MaxEnt (0.03–0.09), in line with recommendations based on the True

https://doi.org/10.3390/horticulturae12010024

Skill Statistic (TSS = sensitivity + specificity − 1) [62]. Conservative thresholds (such as those derived for RF/BRT) are preferred where there are costs associated with false positives, whereas more liberal thresholds that maximize sensitivity (like those for GAM/MaxEnt) are desirable for exploratory screening [50,63]. Because threshold performance can shift across space and scenarios, it can be helpful to recalibrate thresholds (using geographically structured cross-validation), report continuous and binary outputs together, and use a two-tier mapping scheme that is guided by ensemble agreement [51,63,64].

4.2. Tanzania's Southern Highlands, Some Western, Eastern, and Lake Regions Predicted as Highly Suitable Avocado-Growing Zones

The spatial patterns of avocado suitability produced by individual and ensemble models (Table 2; Figure 4) show that Njombe, Iringa, Songwe (Southern Highlands), Morogoro (Eastern Zone), Kigoma, Rukwa (Western Zone), and Kagera (Lake zone) are the most suitable areas for avocado production. The regions identified as suitable for avocado growing coincide with the regions where coffee and tea cultivation are taking place [65–67]. Moderate to high levels of agreement among models in these regions point to the high ecological confidence for avocado growing in these areas under the current climate, likely due to the presence of suitable agro-ecological conditions for avocado production, which are moderate temperatures (≈15–30 °C), high rainfall > 1000 mm yr^{-1}, with good distribution), and well-drained loamy to sandy-loam soils (preferably pH ~ 5–7) [35,68]. Strategic expansion of avocado production should focus on these regions. However, ongoing climate change may shift temperature and rainfall beyond the crop's optimal range, making adaptive management and other climate-resilient practices necessary to sustain productivity.

4.3. Regions with Moderate Suitability

Regions classified as having moderate suitability (15–≤30%; Table 2; Figure 4) are those in which 15–30% of the total land area is predicted to be suitable for avocado cultivation. In these regions, suitable land is present but more limited in extent than in high-suitability regions. Across the four modelling approaches (BRT, GAM, MaxEnt, and RF), Kilimanjaro, Arusha, Dodoma, Manyara, Mara, Mbeya, Ruvuma, Tanga, and Katavi repeatedly fall within this moderate suitability class, although the degree of agreement among models varies.

Across the study area, Kilimanjaro, Arusha, Mbeya, Tanga, and Katavi consistently show moderate suitability for avocado production, especially in the BRT and RF models, which generally predict wider suitable areas. In Arusha (BRT: 17.5%; RF: 22.3%) and Mbeya (BRT: 23.3%; RF: 26.5%), suitability falls firmly within the moderate range, indicating agro-ecological conditions that can support avocado cultivation where management is good and sites are carefully selected. Kilimanjaro (BRT: 25.5%; RF: 30.1%) and Katavi (BRT: 26.8%; RF: 28.8%) are close to the upper boundary of this moderate class, suggesting a relatively large area of suitable land and generally favorable, though spatially variable, conditions.

By contrast, Dodoma, with BRT and RF estimates of 14.0% and 11.5%, respectively, and Manyara, with corresponding values of 9.7% and 11.6%, sit at the lower end of the moderate class and tend to grade into marginal suitability. This pattern is consistent with their relatively dry climates and stronger rainfall seasonality [69,70], both of which can depress avocado yields where irrigation is not available. Mara, with 15.3% and 18.6% of its land classified as suitable by the BRT and RF models, respectively, and Ruvuma, with corresponding values of 15.0% and 20.7%, were identified as moderately suitable by the two models. These results indicate that a meaningful proportion of land has favorable conditions overall, although suitability is shaped in sub-regions limited by local factors such as altitude, soil properties, and temperature variability. Taken together, these patterns reinforce the importance of environmental heterogeneity, especially differences in elevation, rainfall distribution, and land use, in shaping moderate avocado suitability, rather than assuming uniform optimal growing conditions.

Model behavior differs markedly within this suitability class. GAM outputs are comparatively conservative, contributing minimally to moderate suitability across the selected regions and generally producing lower estimates than the other models. This reflects GAM's tendency to prioritize smooth, dominant environmental gradients [71,72]. MaxEnt provides intermediate results, identifying moderately suitable areas in several regions but with lower magnitudes than BRT and RF. BRT and RF, in particular, identify a wider span of conditions as suitable, and this is most evident in regions with complex, heterogeneous landscapes.

Yet, while the SDMs rate Mbeya, Arusha, Tanga, Manyara, and Mara only as moderately suitable, the current distribution of avocado production suggests a different pattern. National reports and previous studies describe several of these regions as among major avocado-producing areas in Tanzania [23,37,73]. One possible explanation is that avocado expansion in Tanzania has been driven largely by farmers' own, often anecdotal, decisions to establish orchards in these regions, many of which the models classify as only moderately suitable. As a result, a substantial share of the country's avocado trees is now concentrated in such areas, whereas production in some high-suitability zones remains underdeveloped simply because these areas have not been exploited to the same extent.

4.4. Regions Identified with Low Suitability

Regions with low suitability (0.0–3.1%) were projected in the coastal regions (Dar es Salaam, Lindi, Mtwara, and Pwani) and in the semi-arid interior regions of Tabora, Shinyanga, Singida, and Simiyu (Table 2; Figure 4). Avocado is extremely sensitive to salinity, with a reduction in yields when the electrical conductivity of irrigation water is greater than around 0.65 dS m^{-1} [74]. In Tanzania, the incursion of saline groundwater along the coast [75] and intense drought and/or heat stress in the semi-arid regions [76] might contribute to the recorded low suitability in these regions.

4.5. Inconsistent Model Predictions in Zanzibar Archipelago

In the Zanzibar archipelago, model outputs exhibit marked inconsistencies. The GAM predict 0.0% suitability across all Zanzibar regions, including Kaskazini Unguja and Pemba (Table 2), indicating extremely restrictive behavior and poor transferability to these unsampled coastal–island environments. GAM's behavior in these areas is consistent with limited extrapolation capacity when local environmental conditions diverge strongly from those represented in the mainland training data.

By contrast, BRT produce relatively high suitability estimates in parts of Zanzibar, particularly in Kaskazini Unguja (57.1%) and Kaskazini Pemba (26.2%), whereas MaxEnt and RF remain restrictive, with suitability values of ≤0.2% and ≤8.5%, respectively (Table 2). This pattern indicates that BRT is more prone to extrapolating suitable conditions into coastal–island environments, while MaxEnt and RF apply more stringent climatic constraints. Such behavior is consistent with the known tendency of BRT to extrapolate in environmental space [58,77] and with the more restrictive effect of regularization and clamping in MaxEnt and, to a lesser degree, RF [56,78,79].

However, the final ensemble model, which employed a spatially conditional exclusion of GAM in coastal zones, indicated negligible suitability in these regions. This confirms that the inflated single-model predictions are modelling artifacts rather than real-world ecological patterns. This discrepancy concurs with findings by Wenger & Olden [80], Owens et al. [81], and Mesgaran et al. [82], who observed that regression-based/correlative SDMs overpredict in novel and/or data-poor environments with covariate ranges that differ from training conditions. Rios et al. [83] reported overpredictions in island systems characterized by abrupt climatic and edaphic gradients. However, Araújo & New [84], Marmion et al. [85], and Thuiller et al. [86] showed that ensemble forecasting approaches, by integrating complementary algorithmic strengths, can improve predictive robustness and minimize single-model bias. Collectively, these findings emphasize the significance of applying extrapolation control and spatially structured cross-validation [56,87,88] when extending SDMs to ecologically unique or under-sampled regions such as the Zanzibar archipelago.

4.6. Regions with High Model Uncertainty

The standard deviation map shows more uncertainty in climatically and topographically complex landscapes including those found in the Njombe, Iringa, Mbeya, Songwe, Arusha, Kilimanjaro Kigoma, Rukwa, and Kagera regions, located in the Southern/Northern Highlands and Western Tanzania (Figure 5). The pronounced agro-ecological variability in these regions creates both optimal and suboptimal microenvironments for avocado cultivation, which are captured differently by the three most robust models (BRT, MaxEnt, and RF; Table 2). Many of these regions are also important agricultural areas for a wide range of horticultural, cereal, legume, and cash crops [89,90].

This sort of inter-model variability has been previously documented by Pearson et al. [91] and Buisson et al. [92], who observed that species distribution model uncertainty is highest at ecotones and along environmental transition areas where predictor gradients are steep. In these marginal areas in our study, RF tends to predict a larger area of suitability than BRT, indicating a less conservative classification of suitable conditions, whereas BRT is more restrictive, limiting suitability to a narrower set of environmental combinations. MaxEnt remains the most restrictive of the three in these areas. Consequently, suitability outputs should always be interpreted together with their associated uncertainty when guiding site selection and regional planning for avocado cultivation.

4.7. Management, Investment, and Policy Implications

The modelling results provide a spatially explicit basis for guiding avocado development in Tanzania by distinguishing areas where investment is likely to be most productive from those where expansion would carry high biophysical and financial risk. The strongest and most consistent signals of suitability arise from the tree-based models (RF and BRT), which also demonstrated the highest predictive performance under spatial cross-validation (AUC $\geq$ 0.77; TSS up to 0.62; Table 1). These outputs, together with the model agreement and uncertainty patterns (Figures 4 and 5), offer a robust framework for management and policy decision-making.

Njombe, Iringa, and Songwe in the Southern Highlands, Kigoma, and Rukwa in the Western Zone, Kagera in the Lake Zone, and Morogoro in the Eastern Zone emerge as the most suitable areas for avocado production. These regions consistently show the highest suitability, especially for the RT/BRT models, and therefore represent the core zones for prioritizing avocado development and investment.

In these high-suitability regions, policy should focus on consolidating and upgrading existing production systems rather than expanding avocado into new areas. Priorities include improving rural roads, packhouses, cold storage, and export logistics, together with stronger extension services on pruning, nutrient management, pest and disease control, and post-harvest handling. A complementary priority is to improve the supply and equitable availability of grafted planting material for commercially preferred cultivars, particularly 'Hass' and 'Fuerte', which are currently concentrated mainly in parts of the Southern and Northern Highlands. Strengthening nursery systems to produce and distribute certified, disease-free grafted saplings using locally sourced seedling rootstocks [93] of these cultivars across avocado-growing regions is essential for improving productivity, market access, and phytosanitary standards nationwide. Land-use planning authorities can further use the high-suitability maps to delineate "avocado development corridors", where orchard expansion and processing facilities are encouraged on already cleared agricultural land, while avoiding affecting forests and other high-conservation-value ecosystems.

Regions with moderate suitability, such as Kilimanjaro, Arusha, Mbeya, Tanga, Katavi, Mara, and Ruvuma, represent transitional agro-ecological zones where avocados can perform well on appropriately chosen sites, but suitable land is more limited and patchier than in the core high-suitability regions. In these areas, priority should be given to assisting farmers in finding and managing the best avocado sites, both in existing orchards and in any new establishments, rather than encouraging broad expansion across the whole landscape. This can be supported through targeted investments in small-scale irrigation and water-harvesting structures on suitable slopes, better soil fertility management (including organic amendments and liming where needed), and stronger local nurseries that can supply certified, disease-free, and locally adapted grafted planting material. Pilot schemes and adaptive trials should be set up in microclimates where model agreement and suitability are highest to test predictions and fine-tune management recommendations before promoting larger-scale investment.

By contrast, the coastal lowlands (Dar es Salaam, Lindi, Mtwara, and Pwani) and the semi-arid interior (Singida, Tabora, Shinyanga, and Simiyu) show negligible suitability across all models and are likely constrained by salinity, high evapotranspiration, and recurrent drought. In these regions, large-scale avocado expansion would entail substantial risk. Policy guidance should therefore discourage major public or private investments in new avocado orchards and instead promote alternative, better-adapted crops or diversified agroforestry systems, with tolerant avocado rootstocks employed only in very localized favorable niches under close technical monitoring [94]. SDM outputs can serve as a screening tool in extension and credit programs to avoid steering smallholders into high-risk avocado ventures in these marginal environments.

At the national level, species distribution model-derived suitability layers should be formally integrated into existing planning instruments such as the Agricultural Sector Development Programme Phase II (ASDP II) and the Tanzania Agriculture Climate Adaptation Technology Deployment Programme (TACATDP). A concrete implementation pathway would include (i) using district-level suitability classes (high, moderate, low, and negligible) to guide the registration and approval of new large- and medium-scale avocado investments; (ii) overlaying suitability maps with infrastructure and market-access data to prioritize locations for public investment in roads, collection centers, and pack-houses; (iii) incorporating suitability information into extension targeting so that training in avocado agronomy and post-harvest handling is concentrated in high- and moderately high-suitability wards; and (iv) utilizing mapped suitability as part of environmental impact assessments to ensure that proposed expansions do not occur in ecologically sensitive or water-stressed areas.

Finally, the spatial patterns identified in this study indicate that current avocado hotspots are concentrated in relatively cool, high-rainfall highlands, which may themselves be vulnerable to future warming and shifts in rainfall regimes. Therefore, highly suitable regions should not be regarded as climate-change safe areas. Climate-smart adaptation packages, combining mulching, shade and windbreak management, efficient irrigation where feasible, soil and water conservation, and careful selection of cultivars and rootstocks [95] should be mainstreamed across all avocado development programs. Linking the present suitability maps with future climate projections would further strengthen decision-

https://doi.org/10.3390/horticulturae12010024

making and help ensure that current investments in avocado orchards remain viable over coming decades.

5. Conclusions

This study presents the first assessment of avocado suitability in Tanzania, combining bioclimate, topography, and soil predictive factors. The implementation of spatial block cross-validation confirmed that our final models possess robust predictive capacity, with AUC values ranging from about 0.72 to 0.81, and a reliable consensus among the tree-based algorithms. The study has shown that Njombe, Iringa, Songwe, Kigoma, Rukwa, Kagera, and Morogoro are the core regions highly suitable for avocado growing. Thus, crop expansion and investment strategies should focus on these areas. Nevertheless, policies and strategies are needed to make sure that avocado expansion and investment in these regions do not cause negative impacts to the environment, like deforestation and loss of biodiversity or ecosystem services.

A second group of regions including Kilimanjaro, Arusha, Manyara, Mara, Mbeya, Ruvuma, Tanga, and Katavi contains a moderate proportion of land classified as suitable and shows greater environmental variation. In these transitional areas, the models indicate viable niches for avocados but also a stronger dependence on the local microclimate, soils, and management practices. Here, investment should aim to raise yields and strengthen resilience in existing orchards, alongside carefully planned expansion into remaining suitable land, so that the available potential is used without undermining environmental integrity.

The coastal lowland regions of Dar es Salaam, Lindi, Mtwara, and Pwani, together with the semi-arid interior regions of Singida, Tabora, Shinyanga, and Simiyu, show low suitability, largely due to soil salinity, extreme temperatures, and limited rainfall, among other constraining factors. In these regions, large-scale avocado expansion would entail substantial biophysical and financial risks and should generally be discouraged.

Finally, the spatially corrected ensemble model consistently observes negligible suitability for the Zanzibar archipelago, largely due to constraints and environmental divergence from the mainland. This result, achieved after addressing individual model anomalies (GAM and BRT overpredictions), confirms that the observed low suitability in the archipelago is a modelling reality for this crop under current conditions. Due to the inherent uniqueness of the environmental factors in these islands, this study continues to advocate for more research in exploring ecological suitability for this crop in Zanzibar regions using hyper-local data and process-based or mechanistic modelling approaches. Once all these findings are incorporated into Tanzania's national agricultural planning and climate adaptation programs, it will stimulate sustainable crop investment, lower production risks, and strengthen the resilience of Tanzania's avocado sector.

6. Future Perspectives

This study provides a first spatially explicit, model-based assessment of avocado suitability in Tanzania, utilizing robust spatial cross-validation. Yet, some limitations point to clear avenues for future research. First, model calibration relied on 199 presence records collected mainly from known production areas between 2017 and 2024. Expanding both the density and geographic coverage of occurrence records, particularly in underexplored regions [96], would strengthen model performance and may reveal additional high-potential areas that are currently underutilized.

Second, we used pseudo-absences and correlative SDMs, which assume equilibrium between avocado distribution and current environmental conditions. Future work should incorporate systematic absence data where available and explore hybrid or mechanistic approaches that explicitly represent physiological thresholds, water balance, and tree growth responses to stress [97]. Linking SDM outputs to yield and fruit-quality data would also allow the transition from simple occurrence-based suitability toward productivity-oriented agro-ecological zoning.

Third, we used 10,000 pseudo-absence points compared with 199 presence records, and this large gap likely contributed to the very low suitability thresholds produced by both the BRT and MaxEnt models. In effect, the sheer number of background points may have dominated the modelling process, making it harder for the models to capture the ecological signal contained in the relatively small set of presence records. Future modelling efforts must systematically experiment with varying background sample sizes ranging, for example, from 2000 to 10,000 points to establish a prevalence ratio that optimizes the trade-off between AUC, TSS, and threshold stability.

Fourth, our predictors were restricted to climate, topography, and soil at a kilometer-scale resolution. Fine-scale microclimatic and edaphic variability [98], as well as management factors such as irrigation, shade, cultivar and seedling rootstock choice [94], pruning, fertilization, and disease pressure were not explicitly represented. This likely contributed to the discrepancies between modeled suitability and the observed distribution of avocado production, particularly in regions where orchards are already well established such as Mbeya, Arusha, and Tanga, but the SDMs suggest only moderate suitability. Future studies should integrate higher-resolution environmental layers and farm-management information, and, where possible, couple SDM with local agronomic trials to better capture the realized niche of avocado under current farming practices.

Fifth, we modeled avocado suitability only under current climate conditions. We did not project suitability under future climate scenarios. Future research should apply the same or improved models under several climate models and emission pathways and for different time horizons. This will help to identify areas where long-lived tree species are less exposed to future climate stress [99], as well as standing adaptation [100], and recommend mobilization trajectories of current germplasm [101].

Lastly, we also did not include socio-economic or land-use constraints in our suitability maps. For example, we did not account for land tenure, processing facilities, or access to domestic and export markets. Future studies should overlay the biophysical suitability and biotic interaction maps [102] with these socio-economic and policy layers. Working directly with farmers, extension officers, and planning authorities would also help to develop simple decision-support tools. This will make it easier to use the model outputs in practical, climate-smart investment and land-use planning for avocados in Tanzania.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/horticulturae12010024/s1, Table S1: Geographic and Administrative Data for *Persea americana* (Avocado) Field Presence Records. Table S2: Full names, definitions, and units of the 19 WorldClim bioclimatic variables used in the study [103]. Table S3: Algorithm-specific classification thresholds (t) derived from cross-validation, associated mean AUC values, and interpretation of threshold stringency.

Author Contributions: Conceptualization, I.J., J.B.V., and A.J.C.; methodology, I.J., J.B.V., and A.J.C.; data analysis, J.B.V.; writing—original draft preparation, I.J.; writing—review and editing, I.J., J.B.V., and A.J.C.; project administration, A.J.C.; funding acquisition, A.J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research and the article processing charge (APC) were funded by Vetenskapsrådet (the Swedish Research Council) through grant 2022-04411, awarded to A.J.C.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials as well as at this link: https://drive.google.com/drive/folders/1BHUZ4rLKjLrv-ahL0SJ9Z3MN-GCy7A2k?usp=sharing (accessed on 17 December 2025). Any additional material not contained within this manuscript or in the weblink will be provided by the authors upon specific request.

Acknowledgments: The authors thank R. Ortiz for support and guidance, as well as J.M. Hussein for accompanying the field expeditions in 2023 and 2024. Additionally, the authors thank University of Dar es salaam, Tanzania, for providing a platform to run this research in Tanzania, as well as the Tanzania Commission for Science and Technology (COSTECH) and the Tanzania Agricultural Research Institute (TARI) for enabling the corresponding research permits. The African Plant Nutrition Institute (APNI) is also acknowledged for hosting the authors and enabling discussions during the institute's annual meeting in late September 2024 in Sea Cliff, Zanzibar, Tanzania.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviation	**Full Form**
SDM	Species Distribution Modelling
SDMs	Species Distribution Models
ENM	Ecological Niche Modelling
FAO	Food and Agriculture Organization
USD	United States Dollar
WITS	World Integrated Trade Solution
GDP	Gross Domestic Product
GPS	Global Positioning System
iSDA	Innovative Solutions for Decision Agriculture
VIF	Variance Inflation Factor
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve
TSS	True Skill Statistic
BRT	Boosted Regression Trees
RF	Random Forest
GAM	Generalized Additive Models
MaxEnt	Maximum Entropy
MCDA	Multi-Criteria Decision Analysis
AHP	Analytic Hierarchy Process
SD	Standard Deviation
ASDP II	Agricultural Sector Development Programme Phase II
USDA FAS	United States Department of Agriculture—Foreign Agricultural Service
TANTRADE	Tanzania Trade Development Authority

References

1. Patra, S.; Maurya, L.L.; Muradi, K.B.; Raghavan, M.; Krishnan, A.G.; Maurya, S.; Bhaskar, J.; Prusty, R.; Mohapatra, S.R.; Panda, A.K. Unlocking the potential of avocado in India: Exploring diversity, cultivation practices, and pathways to progress. *Appl. Fruit Sci.* **2025**, *67*, 154. [CrossRef]
2. FAO (Food and Agriculture Organization of the United Nations). FAOSTAT: Crops and Livestock Products (Avocados, Production Quantity, 1961–2023). Available online: https://www.fao.org/faostat/en/#data/QCL (accessed on 6 December 2025).
3. Statista Research Department. Global Avocado Production from 2000 to 2023 (in Million Metric Tons). Available online: https://www.statista.com/statistics/577455/world-avocado-production/ (accessed on 21 July 2025).
4. Portal del Campo. America Accounted for 72% of Global Avocado Production in 2022. Available online: https://www.portaldelcampo.cl/Noticias/97082_Am%C3%A9rica-concentr%C3%B3-el-72--de-la-producci%C3%B3n-de-palta-a-nivel-mundial-en-2022.html (accessed on 13 August 2025).
5. Tridge. Overview of Global Fresh Avocado Market. Available online: https://www.tridge.com/intelligences/avocado (accessed on 12 August 2025).

6. FreshPlaza. The Americas Lead Global Avocado Production. Available online: https://www.freshplaza.com/north-america/article/9676861/the-americas-lead-global-avocado-production/ (accessed on 2 August 2025).
7. WITS (World Integrated Trade Solution). Avocados, Fresh or Dried Imports by Country in 2023. Available online: https://wits.worldbank.org/trade/comtrade/en/country/ALL/year/2023/tradeflow/Imports/partner/WLD/product/080440?utm (accessed on 6 August 2025).
8. Mod, H.K.; Scherrer, D.; Luoto, M.; Guisan, A. What we use is not what we know: Environmental predictors in plant distribution models. *J. Veg. Sci.* **2016**, *27*, 1308–1322. [CrossRef]
9. Ramírez-Gil, J.G.; Morales-Osorio, J.G.; Peterson, A.T. Potential geography and productivity of 'Hass' avocado crops in Colombia estimated by ecological niche modeling. *Sci. Hortic.* **2018**, *241*, 108–118. [CrossRef]
10. Ramírez-Gil, J.G.; Cobos, M.E.; Jiménez-García, D.; Morales-Osorio, J.G.; Peterson, A.T. Current and potential future distributions of Hass avocados in the face of climate change across the Americas. *Crop Pasture Sci.* **2019**, *70*, 1024–1033. [CrossRef]
11. Grüter, R.; Trachsel, T.; Laube, P.; Jaisli, I. Expected global suitability of coffee, cashew and avocado due to climate change. *PLoS ONE* **2022**, *17*, e0261976. [CrossRef]
12. Sáenz-Ceja, J.E.; Sáenz-Reyes, J.T.; Castillo-Quiroz, D. Pollinator species at risk from the expansion of avocado monoculture in Central Mexico. *Conservation* **2022**, *2*, 407–424. [CrossRef]
13. Arima, E.Y.; Denvir, A.; Young, K.R.; González-Rodríguez, A.; García-Oliva, F. Modelling avocado-driven deforestation in Michoacán, Mexico. *Environ. Res. Lett.* **2022**, *17*, 045012. [CrossRef]
14. Denvir, A. Avocado expansion and the threat of forest loss in Michoacán, Mexico under climate change scenarios. *Appl. Geogr.* **2023**, *156*, 102856. [CrossRef]
15. Ramírez-Mejía, D.; Levers, C.; Kolb, M.; Ghilardi, A.; Godínez-Gómez, O.; Mas, J.-F. Mapping spatiotemporal patterns of avocado expansion and land-use intensity in central Mexico and their effects on landscape connectivity. *Environ. Res. Lett.* **2024**, *19*, 064017. [CrossRef]
16. Charre-Medellín, J.F.; Mas, J.-F.; Chang-Martínez, L.A. Potential expansion of Hass avocado cultivation under climate change scenarios threatens Mexican mountain ecosystems. *Crop Pasture Sci.* **2021**, *72*, 1016–1028. [CrossRef]
17. Selim, S.; Koç-San, D.; Selim, Ç.; San, B.T. Site selection for avocado cultivation using GIS and multi-criteria decision analyses: Case study of Antalya, Turkey. *Comput. Electron. Agric.* **2018**, *153*, 450–459. [CrossRef]
18. Anacona Mopan, Y.E.; Solis Pino, A.F.; Rubiano-Ovalle, O.; Paz, H.; Ramírez-Mejía, I. Spatial analysis of the suitability of Hass avocado cultivation in the Cauca Department, Colombia, using multi-criteria decision analysis and GIS. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 136. [CrossRef]
19. Domínguez, A.; García-Martín, A.; Moreno, E.; González, E.; Paniagua, L.L.; Allendes, G. Identifying optimal zones for avocado (*Persea americana* Mill.) cultivation in the Iberian Peninsula: A climate suitability analysis. *Land* **2024**, *13*, 1290. [CrossRef]
20. Çelik, M.Ö.; Orhan, O.; Kurt, M.A. Predicting climate change impacts on subtropical fruit suitability using MaxEnt: A regional study from southern Türkiye. *Sustainability* **2025**, *17*, 5487. [CrossRef]
21. Berdugo-Cely, J.A.; Cortés, A.J.; López-Hernández, F.; Delgadillo-Durán, P.; Cerón-Souza, I.; Reyes-Herrera, P.; Navas-Arboleda, A.A.; Yockteng, R. Pleistocene-dated genomic divergence of avocado trees supports cryptic diversity in the Colombian germplasm. *Tree Genet. Genomes* **2023**, *19*, 42. [CrossRef]
22. Cárceles Rodríguez, B.; Durán Zuazo, V.H.; Franco Tarifa, D.; Cuadros Tavira, S.; Sacristan, P.C.; García-Tejero, I.F. Irrigation alternatives for avocado (*Persea americana* Mill.) in the Mediterranean subtropical region in the context of climate change: A review. *Agriculture* **2023**, *13*, 1049. [CrossRef]
23. Juma, I.; Fors, H.; Persson Hovmalm, H.; Nyomora, A.; Geleta, M.; Carlsson, A.S.; Ortiz, R.O. Avocado production and local trade in the Southern Highlands of Tanzania: A case of an emerging trade commodity from horticulture. *Agronomy* **2019**, *9*, 749. [CrossRef]
24. SAGCOT Centre Ltd. Avocado Value Chain. Available online: https://sagcot.co.tz/wp-content/uploads/2024/09/AvocadoValuechain.pdf (accessed on 1 December 2025).
25. Kariuki, J.G. Embracing avocado (*Persea americana*) farming among smallholder farmers in rural households in Kenya: Challenges, opportunities and strategies for sustainable growth. *J. Agribus. Rural Dev.* **2025**, *1*, 110–124. [CrossRef]
26. Moreno-Ortega, G.; Pliego, C.; Sarmiento, D.; Barceló, A.; Martínez-Ferri, E. Yield and fruit quality of avocado trees under different regimes of water supply in the subtropical coast of Spain. *Agric. Water Manag.* **2019**, *221*, 192–201. [CrossRef]
27. Hass Avocado Board. Country Profile: Mexico. Available online: https://hassavocadoboard.com/wp-content/uploads/2019/11/hab-marketers-country-profiles-2019-mexico.pdf (accessed on 1 December 2025).
28. World Bank Group. Background Paper for the Country Climate and Development Report (CCDR): United Republic of Tanzania. Available online: https://documents1.worldbank.org/curated/en/099121924163520324/pdf/P18018715c2c9e0651b367172fb62ee6b60.pdf (accessed on 17 December 2025).
29. Mwakalinga, M.M. *Avocado Value Chain Development in Tanzania*; Ministry of Agriculture, United Republic of Tanzania: Dar es Salaam, Tanzania, 2019.

30. World Bank. Tanzania Avocados, Fresh or Dried Exports by Country in 2023. World Integrated Trade Solution (WITS). Available online: https://wits.worldbank.org/trade/comtrade/en/country/TZA/year/2023/tradeflow/Exports/partner/ALL/product/080440 (accessed on 27 July 2025).
31. World Bank. Kenya Avocados, Fresh or Dried Imports by Country in 2023. World Integrated Trade Solution (WITS). Available online: https://wits.worldbank.org/trade/comtrade/en/country/KEN/year/2023/tradeflow/Imports/partner/ALL/product/080440 (accessed on 27 July 2025).
32. Knight, R.J.; Campbell, C.W. Ecological adaptation and the evolution of modern avocado cultivars. *Rev. Chapingo Ser. Hortic.* **1999**, *5*, 49–54. [CrossRef]
33. Whiley, A.W.; Schaffer, B.; Wolstenholme, B.N. (Eds.) *The Avocado: Botany, Production and Uses*; CABI Publishing: Wallingford, UK, 2002.
34. Sseruwagi, P.; Lehmann, E.; Sigombe, P.; Ddamulira, G.; Van Casteren, J.W.; De Bauw, P. Characterizing avocado production systems for Ugandan exports: The need for consolidation and support for sustainable development. *Front. Sustain. Food Syst.* **2025**, *9*, 1500012. [CrossRef]
35. Tanzania Trade Development Authority (TANTRADE). Tanzania Avocado Profile. Available online: https://www.tantrade.go.tz/downloads-prodcut-profile/TANZANIA%20AVOCADO%20PROFILE (accessed on 21 July 2025).
36. Imbert, E. Avocado—Tanzania. In *Hass Avocado Board: Country Profile*; Hass Avocado Board/CIRAD: Irvine, CA, USA, 2016. Available online: https://hassavocadoboard.com/wp-content/uploads/2019/07/hab-marketers-country-profiles-2016-tanzania.pdf (accessed on 1 December 2025).
37. Juma, I.; Nyomora, A.; Hovmalm, H.P.; Fatih, M.; Geleta, M.; Carlsson, A.S.; Ortiz, R.O. Characterization of Tanzanian avocado using morphological traits. *Diversity* **2020**, *12*, 64. [CrossRef]
38. Ramírez-Gil, J.G.; Henao-Rojas, J.C.; Diaz-Diez, C.A.; Peña-Quiñones, A.J.; León, N.; Parra-Coronado, A.; Bernal-Estrada, J.A. Phenological variations of avocado cv. Hass and their relationship with thermal time under tropical conditions. *Heliyon* **2023**, *9*, 19642. [CrossRef] [PubMed]
39. Dubrovina, I.A.; Bautista, F. Analysis of the suitability of various soil groups and types of climate for avocado growing in the state of Michoacán, Mexico. *Eurasian Soil Sci.* **2014**, *47*, 491–503. [CrossRef]
40. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. 2025. Available online: https://www.R-project.org/ (accessed on 17 December 2025).
41. Posit Team. RStudio: Integrated Development Environment for R. Posit Software, PBC. 2025. Available online: http://www.posit.co/ (accessed on 17 December 2025).
42. Hengl, T.; Miller, M.A.; Križan, J.; Shepherd, K.D.; Sila, A.; Kilibarda, M.; Antonijević, O.; Glušica, L.; Dobermann, A.; Haefele, S.M.; et al. African soil properties and nutrients mapped at 30 m spatial resolution using two-scale ensemble machine learning. *Sci. Rep.* **2021**, *11*, 6130. [CrossRef]
43. Inman, R.; Franklin, J.; Esque, T.; Nussear, K. Comparing sample bias correction methods for species distribution modeling using virtual species. *Ecosphere* **2021**, *12*, e03422. [CrossRef]
44. Barbet-Massin, M.; Jiguet, F.; Albert, C.H.; Thuiller, W. Selecting pseudo-absences for species distribution models: How, where and how many? *Methods Ecol. Evol.* **2012**, *3*, 327–338. [CrossRef]
45. Hijmans, R.J. *Terra: Spatial Data Analysis, R Package*, version 1.8-87; R Foundation for Statistical Computing: Vienna, Austria, 2025.
46. WorldClim. Bioclimatic Variables—WorldClim 1 Documentation. Available online: https://www.worldclim.org/data/bioclim.html (accessed on 11 November 2025).
47. Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J.R.G.; Gruber, B.; Lafourcade, B.; Leitão, P.J.; et al. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **2013**, *36*, 27–46. [CrossRef]
48. Mahoney, M.J.; Johnson, L.K.; Silge, J.; Frick, H.; Kuhn, M.; Beier, C.M. Assessing the performance of spatial cross-validation approaches for models of spatially structured data. *arXiv* **2023**, arXiv:2303.07334. [CrossRef]
49. Youden, W.J. Index for rating diagnostic tests. *Cancer* **1950**, *3*, 32–35. [CrossRef]
50. Liu, C.; Berry, P.M.; Dawson, T.P.; Pearson, R.G. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **2005**, *28*, 385–393. [CrossRef]
51. Norberg, A.; Abrego, N.; Blanchet, F.G.; Adler, F.R.; Anderson, B.J.; Anttila, J.; Araújo, M.B.; Dallas, T.; Dunson, D.; Elith, J.; et al. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecol. Monogr.* **2019**, *89*, e01370. [CrossRef]
52. Fiorentino, D.; Núñez-Riboni, I.; Pierce, M.E.; Oesterwind, D.; Akimova, A. Improving species distribution models for climate change studies: Ecological plausibility and performance metrics. *Ecol. Model.* **2025**, *508*, 111207. [CrossRef]
53. Çorbacıoğlu, Ş.K.; Aksel, G. Receiver operating characteristic curve analysis in diagnostic accuracy studies: A guide to interpreting the area under the curve value. *Turk. J. Emerg. Med.* **2023**, *23*, 195–198. [CrossRef]

54. Mi, C.; Huettmann, F.; Guo, Y.; Han, X.; Wen, L. Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. *PeerJ* **2017**, *5*, e2849. [CrossRef]
55. Fern, R.R.; Morrison, M.L.; Wang, H.H.; Grant, W.E.; Campbell, T.A. Incorporating biotic relationships improves species distribution models: Modeling the temporal influence of competition in conspecific nesting birds. *Ecol. Model.* **2019**, *408*, 108743. [CrossRef]
56. Valavi, R.; Guillera-Arroita, G.; Lahoz-Monfort, J.J.; Elith, J. Predictive performance of presence-only species distribution models: A benchmark study with reproducible code. *Ecol. Monogr.* **2022**, *92*, e01486. [CrossRef]
57. Cutler, D.R.; Edwards, T.C., Jr.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random forests for classification in ecology. *Ecology* **2007**, *88*, 2783–2792. [CrossRef]
58. Elith, J.; Leathwick, J.R.; Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* **2008**, *77*, 802–813. [CrossRef]
59. Becker, E.A.; Carretta, J.V.; Forney, K.A.; Barlow, J.; Brodie, S.; Hoopes, R.; Jacox, M.G.; Maxwell, S.M.; Redfern, J.V.; Sisson, N.B.; et al. Performance evaluation of cetacean species distribution models developed using generalized additive models and boosted regression trees. *Ecol. Evol.* **2020**, *10*, 5759–5784. [CrossRef]
60. Morales, N.S.; Fernández, I.C.; Baca-González, V. MaxEnt's parameter configuration and small samples: Are we paying attention to recommendations? A systematic review. *PeerJ* **2017**, *5*, e3093. [CrossRef] [PubMed]
61. Feng, X.; Park, D.S.; Liang, Y.; Pandey, R.; Papeş, M. Collinearity in ecological niche modeling: Confusions and challenges. *Ecol. Evol.* **2019**, *9*, 10365–10376. [CrossRef] [PubMed]
62. Allouche, O.; Tsoar, A.; Kadmon, R. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **2006**, *43*, 1223–1232. [CrossRef]
63. Elith, J.; Leathwick, J.R. Species distribution models: Ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Evol. Syst.* **2009**, *40*, 677–697. [CrossRef]
64. Hao, T.; Elith, J.; Lahoz-Monfort, J.J.; Guillera-Arroita, G. Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. *Ecography* **2020**, *43*, 549–558. [CrossRef]
65. Food and Agriculture Organization of the United Nations (FAO). *Report of the Tea Industry in Tanzania; CCP:TE 16/CRS.7, Intergovernmental Group on Tea, 22nd Session, Naivasha, Kenya*; FAO: Rome, Italy, 2016.
66. Tea Board of Tanzania. *The State of Tea Industry in Tanzania*; Tea Board of Tanzania: Dar es Salaam, Tanzania, 2024. Available online: https://www.teaboard.go.tz/uploads/documents/sw-1733290446-TEA%20PROFILE..pdf (accessed on 12 December 2025).
67. USDA FAS. *Tanzania Coffee Annual (TZ2024-0002)*; U.S. Department of Agriculture, Foreign Agricultural Service: Washington, DC, USA, 2024. Available online: https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Coffee+Annual_Dar+Es+Salaam_Tanzania_TZ2024-0002.pdf (accessed on 12 December 2025).
68. Salazar-García, S.; Garner, L.C.; Lovatt, C.J. Reproductive biology. In *The Avocado: Botany, Production and Uses*; Whiley, A.W., Schaffer, B., Wolstenholme, B.N., Eds.; CABI: Wallingford, UK, 2013; pp. 118–167.
69. Deus, D.; Gloaguen, R.; Krause, P. Water balance modeling in a semi-arid environment with limited in situ data using remote sensing in Lake Manyara, East African Rift, Tanzania. *Remote Sens.* **2013**, *5*, 1651–1680. [CrossRef]
70. John, O. Evaluation of rainfall extreme characteristics in Dodoma Urban, a central part of Tanzania. *Int. J. Environ. Geoinf.* **2022**, *9*, 165–177. [CrossRef]
71. Hastie, T.; Tibshirani, R. Generalized additive models. *Stat. Sci.* **1986**, *1*, 297–310. [CrossRef]
72. Wood, S.N. *Generalized Additive Models*, 2nd ed.; Chapman and Hall/CRC: New York, NY, USA, 2017.
73. Yangaza, I.S.; Nyomora, A.M.; Joseph, C.O.; Sangu, E.M.; Alcaraz, M.L.; Hormaza, J.I. Genetic diversity and population structure of local avocado (*Persea americana* Mill.) from northern Tanzania assessed using SSR markers. *Genet. Resour. Crop Evol.* **2025**, *72*, 4789–4807. [CrossRef]
74. Celis, N.; Suarez, D.L.; Wu, L.; Li, R.; Arpaia, M.L.; Mauk, P. Salt tolerance and growth of 13 avocado rootstocks related best to chloride uptake. *HortScience* **2018**, *53*, 1737–1745. [CrossRef]
75. Van Camp, M.; Mtoni, Y.; Mjemah, I.C.; Bakundukize, C.; Walraevens, K. Investigating seawater intrusion due to groundwater pumping with schematic model simulations: The example of the Dar es Salaam coastal aquifer in Tanzania. *J. Afr. Earth Sci.* **2014**, *96*, 71–78. [CrossRef]
76. Ministry of Agriculture. *National Food Security Bulletin*; Ministry of Agriculture: Dodoma, Tanzania, 2025; Volume 58. Available online: https://www.kilimo.go.tz/uploads/documents/sw-1744942626-FEBRUARY%202025-FOOD%20SECURITY%20BULLETIN-FINAL.pdf (accessed on 12 November 2025).
77. Elith, J.; Kearney, M.; Phillips, S. The art of modelling range-shifting species. *Methods Ecol. Evol.* **2010**, *1*, 330–342. [CrossRef]
78. Phillips, S.J.; Anderson, R.P.; Schapire, R.E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **2006**, *190*, 231–259. [CrossRef]
79. Elith, J.; Phillips, S.J.; Hastie, T.; Dudík, M.; Chee, Y.E.; Yates, C.J. A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* **2011**, *17*, 43–57. [CrossRef]

80. Wenger, S.J.; Olden, J.D. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods Ecol. Evol.* **2012**, *3*, 260–267. [CrossRef]
81. Owens, H.L.; Campbell, L.P.; Dornak, L.L.; Saupe, E.E.; Barve, N.; Soberón, J.; Ingenloff, K.; Lira-Noriega, A.; Hensz, C.M.; Myers, C.E.; et al. Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecol. Model.* **2013**, *263*, 10–18. [CrossRef]
82. Mesgaran, M.B.; Cousens, R.D.; Webber, B.L. Here be dragons: A tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. *Divers. Distrib.* **2014**, *20*, 1147–1159. [CrossRef]
83. Rios, E.B.; Sadler, J.; Graham, L.; Matthews, T.J. Species distribution models and island biogeography: Challenges and prospects. *Glob. Ecol. Conserv.* **2024**, *51*, e02943. [CrossRef]
84. Araújo, M.B.; New, M. Ensemble forecasting of species distributions. *Trends Ecol. Evol.* **2007**, *22*, 42–47. [CrossRef]
85. Marmion, M.; Parviainen, M.; Luoto, M.; Heikkinen, R.K.; Thuiller, W. Evaluation of consensus methods in predictive species distribution modelling. *Divers. Distrib.* **2009**, *15*, 59–69. [CrossRef]
86. Thuiller, W.; Lafourcade, B.; Engler, R.; Araújo, M.B. BIOMOD—A platform for ensemble forecasting of species distributions. *Ecography* **2009**, *32*, 369–373. [CrossRef]
87. Roberts, D.R.; Bahn, V.; Ciuti, S.; Boyce, M.S.; Elith, J.; Guillera-Arroita, G.; Hauenstein, S.; Lahoz-Monfort, J.J.; Schröder, B.; Thuiller, W.; et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **2017**, *40*, 913–929. [CrossRef]
88. Yates, K.L.; Bouchet, P.J.; Caley, M.J.; Mengersen, K.; Randin, C.F.; Parnell, S.; Fielding, A.H.; Bamford, A.J.; Ban, S.; Barbosa, A.M.; et al. Outstanding challenges in the transferability of ecological models. *Trends Ecol. Evol.* **2018**, *33*, 790–802. [CrossRef] [PubMed]
89. United Republic of Tanzania. *Agricultural Sector Development Programme Phase Two (ASDP II)*; Ministry of Agriculture: Dar es Salaam, Tanzania, 2016. Available online: https://faolex.fao.org/docs/pdf/tan169870.pdf (accessed on 9 December 2025).
90. United Republic of Tanzania. *Agriculture Annual Report 2023–2024*; Ministry of Agriculture: Dar es Salaam, Tanzania, 2024. Available online: https://www.kilimo.go.tz/uploads/documents/sw-1747227277-Agriculture%20Annual%20Report%202023%20-%202024%20compressed.pdf?utm (accessed on 9 December 2025).
91. Pearson, R.G.; Thuiller, W.; Araújo, M.B.; Martinez-Meyer, E.; Brotons, L.; McClean, C.; Miles, L.; Segurado, P.; Dawson, T.P.; Lees, D.C. Model-based uncertainty in species range prediction. *J. Biogeogr.* **2006**, *33*, 1704–1711. [CrossRef]
92. Buisson, L.; Thuiller, W.; Casajus, N.; Lek, S.; Grenouillet, G. Uncertainty in ensemble forecasting of species distribution. *Glob. Change Biol.* **2010**, *16*, 1145–1157. [CrossRef]
93. Cañas-Gutiérrez, G.P.; López-Hernández, F.; Cortés, A.J. Whole genome resequencing of 205 avocado trees unveils the genomic patterns of racial divergence in the Americas. *Int. J. Mol. Sci.* **2025**, *26*, 10353. [CrossRef]
94. Reyes, P.H.; Muñoz, L.; Velázquez, V.; Patiño, L.; Delgado, O.A.; Díaz, C.A.; Navas, A.A.; Cortés, A.J. Inheritance of rootstock effects in avocado (*Persea americana* Mill.) cv. Hass. *Front. Plant Sci.* **2020**, *11*, 555071. [CrossRef]
95. Cañas-Gutiérrez, G.P.; Sepulveda-Ortega, S.; López-Hernández, F.; Navas-Arboleda, A.; Cortés, A.J. Inheritance of yield components and morphological traits in avocado cv. Hass from criollo elite trees via half-sib seedling rootstocks. *Front. Plant Sci.* **2022**, *13*, 843099. [CrossRef]
96. Valencia, J.B.; Mesa, J.; León, J.G.; Madriñán, S.; Cortés, A.J. Climate vulnerability assessment of the *Espeletia* complex in páramo sky islands of the northern Andes. *Front. Ecol. Evol.* **2020**, *8*, 565708. [CrossRef]
97. Bedoya-Cañas, L.; López-Hernández, F.; Cortés, A.J. Climate change responses of high-elevation *Polylepis* forests. *Forests* **2024**, *15*, 811. [CrossRef]
98. López-Hernández, F.; Rosero-Alpala, M.G.; Rosero, A.; Cortés, A.J. Projected shifts in Colombian sweet potato germplasm under climate change. *Horticulturae* **2025**, *11*, 1080. [CrossRef]
99. Cortés, A.J.; Restrepo-Montoya, M.; Bedoya-Cañas, L.E. Modern strategies to assess and breed forest tree adaptation to changing climate. *Front. Plant Sci.* **2020**, *11*, 583323. [CrossRef]
100. Cortés, A.J.; López-Hernández, F.; Blair, M.W. Genome—Environment associations: An innovative tool for studying heritable evolutionary adaptation in orphan crops and wild relatives. *Front. Genet.* **2022**, *13*, 910386. [CrossRef]
101. Cortés, A.J. Unlocking genebanks for climate adaptation. *Nat. Clim. Change* **2025**, *15*, 590–592. [CrossRef]
102. Guevara, M.; Osorio, A.N.; Cortés, A.J. Integrative breeding for biotic resistance in forest trees. *Plants* **2021**, *10*, 2022. [CrossRef]
103. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [CrossRef]