INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

SAR Backscatter, not Optical Indices or Phenology, Drives Cropland Classification Accuracy: A Multi-Sensor Feature Ablation Study in the Abuja Municipal Area Council, Nigeria

Idris Ibrahim[1]*, Salman Salis Khalid[1], Hudu Hamza Musa[2], Nafisah Abdullahi Ahmed[2]

[1]Strategic Space Applications Department, National Space Research and Development Agency, Abuja, Nigeria

[2]Atlantic International Research Center, National Space Research and Development Agency, Abuja, Nigeria

***Corresponding author**

DOI: https://doi.org/10.51583/IJLTEMAS.2026.150600159

Received: 02 July 2026; Accepted: 07 July 2026; Published: 18 July 2026

ABSTRACT

Cropland mapping in rapidly urbanizing tropical regions increasingly relies on multi-sensor Earth observation, yet the individual contribution of different sensor and feature types to classification accuracy is rarely quantified explicitly. Building on a prior comparative study of Random Forest (RF) and Gradient Tree Boosting classifiers using a 20-band Sentinel-1/Sentinel-2 feature stack over the Abuja Municipal Area Council (AMAC), Nigeria, this study conducts a systematic feature ablation to isolate the marginal contribution of spectral indices, NDVI-derived phenological metrics, and Synthetic Aperture Radar (SAR) backscatter to cropland classification accuracy. Six feature-stack combinations, ranging from optical-only to the full fused stack, were classified using an RF classifier with fixed hyperparameters and the same stratified reference dataset (373 validation points) used in the original study. Overall accuracy, Kappa coefficient, and cropland-specific precision, recall, and F1-score were compared across combinations, and McNemar's test was used to assess whether accuracy differences between stacks were statistically significant. Results show that SAR backscatter, not spectral indices or phenological metrics, was the dominant driver of accuracy gains: the optical-plus-SAR combination achieved the highest overall accuracy (74.8%) and Kappa (0.685) of all six combinations tested, exceeding even the full fused stack (73.5%, κ=0.668), while adding spectral indices and phenological metrics without SAR reduced accuracy below the optical-only baseline (67.0% vs. 68.9%). McNemar's tests confirmed the full stack was not significantly different from the optical-only baseline or the optical-plus-SAR combination ($p > 0.05$), but was significantly more accurate than any combination lacking SAR ($p < 0.05$). Gini-based variable importance ranked VH backscatter and VH temporal variance as the two most important predictors among all 20 bands, and this ranking was corroborated by a permutation-importance cross-check on an independently trained model, which agreed with Gini importance on both top-ranked SAR bands despite the two methods diverging further down the ranking indicating SAR's dominance is a robust finding rather than an artifact of Gini importance's known bias toward continuous variables. These findings indicate that, in this tropical savanna setting, radar data delivers the largest return on investment for operational cropland monitoring, while spectral indices and phenological metrics contribute comparatively little value on their own.

Keywords: Feature Ablation, Random Forest, Sentinel-1, Sentinel-2, SAR-Optical Fusion, Variable Importance, Cropland Mapping

INTRODUCTION

Agricultural land constitutes approximately 38% of the Earth's terrestrial surface and underpins food security and ecosystem services, particularly in rapidly urbanizing regions of the developing world (Azadi et al., 2021). Timely and accurate information on cropland extent is essential for land use planning, food security assessment,

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

and climate-resilient urban development (Dubey et al., 2025). Satellite remote sensing, and the Sentinel constellation in particular, has become central to operational land cover monitoring because it provides continuous, repeatable observation independent of ground access constraints (Du et al., 2019).

A growing body of work has demonstrated that fusing optical and radar imagery, together with derived spectral indices and phenological metrics, improves cropland and land cover classification accuracy relative to single-sensor approaches, because each data source captures complementary information: optical bands and indices capture spectral and structural crop characteristics, phenological metrics capture the temporal growth signature that distinguishes crops from other vegetation, and SAR backscatter is sensitive to canopy structure and moisture and is unaffected by cloud cover (Van Tricht et al., 2018; Li & Xiao, 2025). However, most studies that use such fused, multi-sensor feature stacks report classification accuracy only for the full, combined stack, or compare classifier algorithms on that fixed stack, as in a recent comparison of Random Forest (RF) and Gradient Tree Boosting (GTB) classifiers for cropland mapping in the Abuja Municipal Area Council (AMAC), Nigeria (Ibrahim et al., 2026). What remains less well quantified is the marginal contribution of each feature group within such a stack: whether the added processing cost of incorporating SAR time series or phenological metrics is justified by a commensurate gain in accuracy, relative to a simpler optical-only or optical-plus-indices approach and, notably, whether combining every available feature type is even the best strategy, as opposed to a more selective combination.

This distinction matters for practical deployment. In data- and resource-constrained settings, the cost of acquiring, preprocessing, and maintaining multiple sensor time series is non-trivial, and operational agencies benefit from evidence on which feature investments yield the greatest returns. The objective of this study is therefore to quantify, through a systematic feature ablation design, the individual and combined contribution of spectral indices, NDVI-derived phenological metrics, and SAR backscatter to cropland classification accuracy in a tropical savanna, peri-urban environment, using the same study area, satellite data, and RF classifier configuration established in Ibrahim et al. (2026), together with variable importance analysis and formal statistical testing of accuracy differences between feature-stack combinations.

MATERIALS AND METHODS

Study Area

The study area is located within the Abuja Municipal Area Council (AMAC), Federal Capital Territory, Nigeria, covering approximately 1,769 km² (8°45′–9°10′N, 7°10′–7°35′E), within a tropical savanna climate zone characterized by distinct wet and dry seasons. Full details of the study area are provided in Ibrahim et al. (2026).

Satellite Data and Preprocessing

All data processing was implemented in Google Earth Engine (GEE) (Gorelick et al., 2017). The optical dataset comprised Sentinel-2 Level-2A Surface Reflectance imagery (COPERNICUS/S2_SR_HARMONIZED) for the year 2025, cloud-masked using the Scene Classification Layer (SCL) to exclude cloud, cloud shadow, and cirrus pixels.

The radar dataset comprised Sentinel-1 Ground Range Detected Interferometric Wide-mode imagery (COPERNICUS/S1_GRD) with VV and VH polarizations over the same period. Preprocessing steps follow Ibrahim et al. (2026) and were not modified for this study, so that any accuracy differences observed here are attributable to feature composition rather than changes in the underlying preprocessing pipeline.

Feature Groups

The full 20-band feature stack used in Ibrahim et al. (2026) was partitioned into four feature groups to enable systematic ablation, summarized in Table 1: optical spectral bands (OPT, n=6), spectral indices (IDX, n=2), NDVI-derived phenological metrics (PHEN, n=5), and SAR backscatter and texture measures (SAR, n=7).

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Table 1: Summarized Optical Spectral Bands, Spectral Indices, NDVI-derived phenological metrics, and SAR backscatter texture measures.

Feature Group	Bands / Metrics Included	n bands
Optical Spectral (OPT)	Blue (B2), Green (B3), Red (B4), NIR (B8), SWIR1 (B11), SWIR2 (B12)	6
Spectral Indices (IDX)	NDVI, NDBI	2
Phenological Metrics (PHEN)	NDVI_min, NDVI_max, NDVI_mean, NDVI_std, NDVI_amplitude (annual NDVI time-series statistics)	5
SAR / Radar (SAR)	VV backscatter, VH backscatter, VH/VV ratio, VH temporal variance, VH GLCM contrast, VH GLCM entropy, VH GLCM variance	7
Total		**20**

Reference Data

Reference labels were derived from the ESA WorldCover v200 (2021) product, reclassified into five land use/land cover classes: cropland, other vegetation (comprising tree cover, shrubland, grassland, herbaceous wetland, mangroves, and moss/lichen), built-up, bare/sparse vegetation, and permanent water bodies. A stratified random sample was drawn using GEE's stratifiedSample function with numPoints=250; because this parameter specifies the sample size per class rather than in total, the resulting reference dataset comprised approximately 1,250 points across the five classes. Following a 70/30 train-validation split with a fixed random seed (42), this yielded 373 validation points (approximately 69–77 per class), which were held fixed across all six feature-stack combinations so that any accuracy differences reflect feature composition rather than sampling variation. We note this reference dataset is derived from an existing global land cover product rather than independently collected ground truth, and validation accuracy should be interpreted as agreement with WorldCover rather than an assessment against field-verified labels.

Ablation Experimental Design

To isolate the contribution of each feature group, six feature-stack combinations were classified independently using an identical classifier configuration: (1) OPT only, as a baseline; (2) OPT + IDX; (3) OPT + PHEN; (4) OPT + SAR; (5) OPT + IDX + PHEN, representing an optical-only fused stack without radar; and (6) the full stack (OPT + IDX + PHEN + SAR), replicating the feature set used in the original study. This design allows the marginal contribution of indices, phenology, and SAR to be assessed both individually (combinations 2–4 relative to the baseline) and in combination (combinations 5 and 6).

Classification and Accuracy Assessment

A Random Forest classifier was applied to each of the six feature-stack combinations, using 200 trees and a fixed random seed of 42, consistent with the RF configuration in Ibrahim et al. (2026). Gradient Tree Boosting was not re-run in this study, as the objective here is feature contribution rather than algorithm comparison; RF was retained as the sole classifier because it was identified as the more effective algorithm for cropland-class discrimination in the original comparison. To reduce computational overhead, training and validation pixel values for the full 20-band stack were extracted once, and each combination's classifier was trained using the relevant subset of columns rather than re-extracting pixel values from imagery for every combination. For each combination, overall accuracy, the Kappa coefficient, and cropland-class precision, recall, and F1-score were computed from the confusion matrix on the held-out validation set.

Variable Importance Analysis

For the full-stack RF model, Gini-based variable importance (mean decrease in impurity) was extracted per band using GEE's built-in classifier explanation output, providing an initial ranking of feature contributions across all 20 bands (Breiman, 2001). Because Gini importance is known to be biased toward continuous and high-

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

cardinality variables a concern directly relevant here, given that several SAR and phenology-derived bands are continuous while the underlying reflectance bands are as well this ranking was cross-checked against permutation feature importance, computed as the mean decrease in validation accuracy over 30 random shuffles per band (random_state=42), using an equivalent Random Forest model (200 trees) trained in scikit-learn on the same training points, with band values exported from GEE. Bands consistently ranked highly by both methods were treated as robust contributors; disagreement between the two rankings is reported in Section 3.3 rather than resolved in favor of either method.

Statistical Comparison of Feature-Stack Performance

To test whether differences in classification accuracy between feature-stack combinations were statistically significant rather than attributable to sampling variability, McNemar's test (Foody, 2004) was applied pairwise between the full-stack model and each reduced combination, using the paired correct/incorrect classification outcomes for each of the 373 validation points, tracked by a persistent point identifier across all six combinations. A significance threshold of $\alpha = 0.05$ was adopted, with exact binomial McNemar p-values used given the discordant pair counts involved.

Software and Reproducibility

Feature-stack construction, sampling, and classification were performed in Google Earth Engine (Gorelick et al., 2017). McNemar's tests were computed in Python 3 (SciPy) from per-point prediction outcomes exported from GEE. The random seed was fixed at 42 throughout, consistent with the original study, to ensure reproducibility.

RESULTS

Classification Accuracy by Feature-Stack Combination

Table 2 summarizes overall accuracy, Kappa, and cropland-specific precision, recall, and F1-score for each of the six feature-stack combinations.

Table 2: Summaries of Overall Accuracy, Kappa, Cropland Precision, Cropland Recall, and Cropland F1-score

Feature Stack Combination	**Overall Accuracy (%)**	**Kappa**	**Cropland Precision**	**Cropland Recall**	**Cropland F1-score**
OPT only (baseline)	68.90	0.611	0.716	0.753	0.734
OPT + IDX	68.63	0.608	0.733	0.714	0.724
OPT + PHEN	66.76	0.584	0.688	0.714	0.701
OPT + SAR	**74.80**	**0.685**	**0.753**	**0.714**	**0.733**
OPT + IDX + PHEN	67.02	0.588	0.680	0.662	0.671
Full stack (OPT+IDX+PHEN+SAR)	73.46	0.668	0.743	0.714	0.728

Contrary to the assumption that combining all available feature types maximizes accuracy, the optical-plus-SAR (OPT+SAR) combination achieved the highest overall accuracy (74.80%) and Kappa (0.685) of all six combinations tested, exceeding the full stack (73.46%, κ=0.668) by 1.34 percentage points in overall accuracy. Cropland F1-score was highest for the OPT-only baseline (0.734), closely followed by OPT+SAR (0.733) and the full stack (0.728); the differences among these three top performers are small. In contrast, adding spectral indices or phenological metrics to the optical baseline without SAR reduced performance: OPT+IDX (68.63%), OPT+PHEN (66.76%), and particularly OPT+IDX+PHEN (67.02%) all underperformed the optical-only baseline (68.90%) on overall accuracy, and OPT+IDX+PHEN recorded the lowest cropland F1-score of any combination (0.671). SAR was therefore the only feature group whose addition produced a clear, consistent improvement over the optical baseline across all five metrics.

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING,
MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

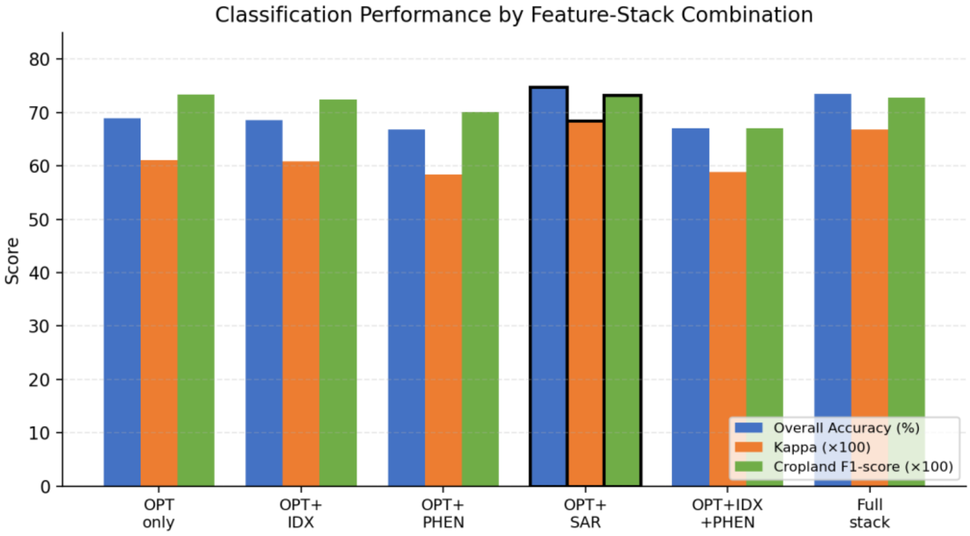

Figure 1. Overall accuracy, Kappa (×100), and cropland F1-score (×100) across the six feature-stack combinations. OPT+SAR (outlined) matches or exceeds the full stack on every metric.

Statistical Significance of Accuracy Differences

McNemar's test results comparing each reduced feature-stack combination against the full stack, based on paired correct/incorrect outcomes across the 373 validation points, are presented in Table 3.

Table 3: McNemar's Significant Testing.

Comparison	**b (Full correct, other wrong)**	**c (Full wrong, other correct)**	**p-value**	**Sig. (α=0.05)**
OPT only vs. Full stack	48	31	0.072	No
OPT + IDX vs. Full stack	45	27	0.045	Yes
OPT + PHEN vs. Full stack	41	16	0.001	Yes
OPT + SAR vs. Full stack	15	20	0.499	No
OPT + IDX + PHEN vs. Full stack	43	19	0.003	Yes

The full stack was not significantly different from the OPT-only baseline (p=0.072) or from OPT+SAR (p=0.499), indicating that neither the addition of SAR to the optical baseline, nor the full combination of all feature groups, produced a statistically distinguishable classification outcome from one another at the point level. By contrast, the full stack was significantly more accurate than OPT+IDX (p=0.045), OPT+PHEN (p=0.001), and OPT+IDX+PHEN (p=0.003). Read together with Table 2, this indicates that SAR, not the full stack's additional complexity is responsible for whatever advantage the full stack holds over the reduced combinations: stacks lacking SAR are significantly worse than the full stack, while OPT+SAR performs statistically indistinguishably from it despite being a simpler, 13-band model rather than the full 20-band stack.

Variable Importance: Gini and Permutation Cross-Check

Gini-based variable importance rankings for the full-stack RF model are presented in Table 4, alongside a permutation-importance cross-check computed on the held-out validation set using an equivalent Random Forest model trained in scikit-learn (n_estimators=200, random_state=42; validation accuracy 71.85%, closely matching the 73.46% overall accuracy obtained from the GEE-native model, with the small difference attributable to implementation differences between the two Random Forest libraries).

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Table 4: Gini Variable Importance

Gini Rank	Feature / Band	Feature Group	Gini Importance (%)	Permutation Rank	Permutation Importance
1	**VH backscatter**	**SAR**	**5.87**	**2**	**0.0086**
2	**VH temporal variance**	**SAR**	**5.74**	**1**	**0.0155**
3	B12 (SWIR2)	OPT	5.38	3	0.0073
4	B11 (SWIR1)	OPT	5.32	14	-0.0025
5	NDVI amplitude	PHEN	5.31	10	-0.0013
6	NDBI	IDX	5.31	15	-0.0030
7	**VV backscatter**	**SAR**	**5.31**	**8**	**0.0029**
8	**VH GLCM variance**	**SAR**	**5.13**	**7**	**0.0038**

The two importance methods agree strongly at the top of the ranking: VH backscatter and VH temporal variance are ranked first and second by both Gini and permutation importance, with only their relative order swapped (Gini: VH=1st, VH_var=2nd; permutation: VH_var=1st, VH=2nd), and B12 (SWIR2) is ranked third by both methods. Across all 20 bands, Gini and permutation rankings are moderately-to-strongly correlated (Spearman's $\rho=0.564$, $p=0.010$), and 5 of the 8 top Gini-ranked bands including four of the five SAR-derived bands in the top 8 also appear in the permutation-importance top 8. This agreement at the top of the ranking indicates that SAR's dominant contribution to classification accuracy is not an artifact of Gini importance's known bias toward continuous, high-cardinality variables (Breiman, 2001), but is corroborated by a bias-resistant, accuracy-based importance measure.

The two methods diverge more noticeably further down the ranking. B11 (SWIR1), NDBI, and NDVI amplitude rank highly under Gini importance (4th, 6th, and 5th, respectively) but drop substantially under permutation importance (14th, 15th, and 10th, respectively, with near-zero or slightly negative permutation importance scores), while B8 (NIR) and B4 (Red) rank comparatively low under Gini (19th and 13th) but rise into the permutation top 5 (4th and 5th). Eleven of the 20 bands recorded negative mean permutation importance, indicating that shuffling their values did not measurably reduce validation accuracy and, in some resampling iterations, coincided with marginally higher accuracy purely by chance. This pattern is consistent with the interpretation offered in Section 3.1: several of the correlated NDVI-derived and index bands (particularly NDBI and NDVI amplitude, both inflated under Gini relative to permutation importance) appear to contribute little unique, non-redundant information once the two dominant SAR bands are already in the model, reinforcing that SAR not the size of the feature stack is driving classification performance in this study.

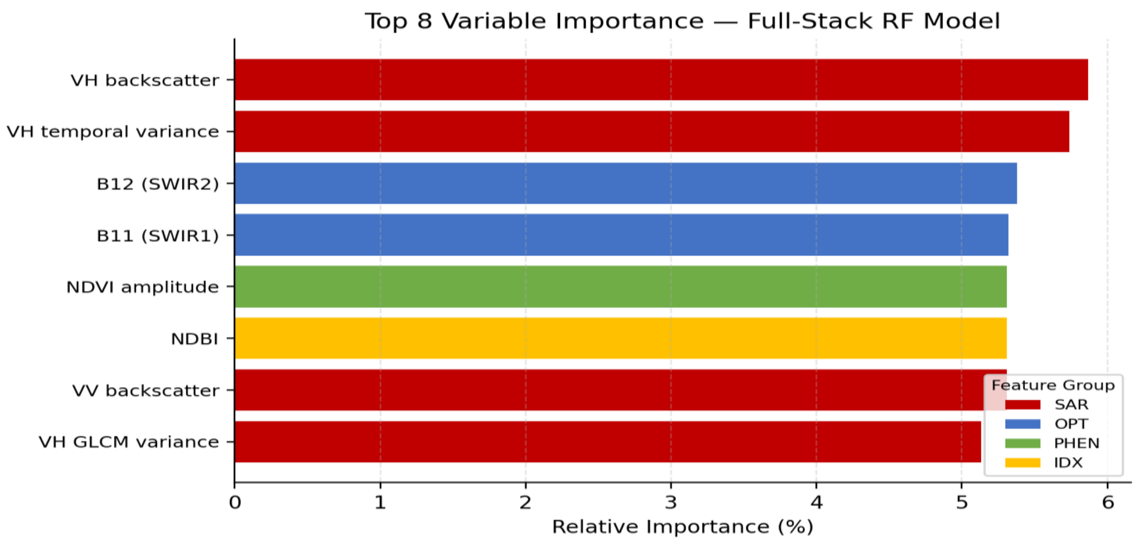

Figure 2. Top 8 Gini variable importance rankings for the full-stack RF model, colored by feature group. SAR-derived bands occupy 4 of the top 8 positions, including ranks 1 and 2 a result corroborated by permutation importance (Table 4).

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

DISCUSSION

The ablation results converge on a consistent conclusion across four independent lines of evidence accuracy comparison, paired significance testing, Gini variable importance, and a permutation-importance cross-check that SAR backscatter, not spectral indices or phenological metrics, is the dominant contributor to cropland classification accuracy in this tropical savanna, peri-urban setting. The agreement between Gini and permutation importance at the top of the ranking (Section 3.3; both methods place VH backscatter and VH temporal variance first and second) is an important robustness check: it rules out the possibility that SAR's apparent dominance is simply an artifact of Gini importance's known bias toward continuous, high-cardinality variables (Breiman, 2001), since permutation importance is not subject to that bias and independently reaches the same conclusion. This is broadly consistent with prior findings that SAR-optical fusion improves crop mapping accuracy, particularly in regions where cloud cover limits the availability of clean optical composites (Van Tricht et al., 2018; Li & Xiao, 2025). What is less expected, and arguably the more actionable finding of this study, is that spectral indices and phenological metrics did not simply contribute less than SAR when added without SAR, they measurably reduced accuracy below the optical-only baseline (OPT+IDX+PHEN: 67.0% vs. OPT-only: 68.9%). A plausible explanation is that, with a training set of moderate size (877 points across five classes), adding seven correlated NDVI-derived and index bands increases the dimensionality of the feature space without providing proportionally more separable information, a pattern consistent with the well-documented risk of Random Forest performance degrading when highly correlated or low-information features dilute the signal from a smaller number of strongly discriminative ones (Belgiu & Drăguţ, 2016). The permutation-importance results in Section 3.3 offer direct support for this explanation: eleven of the twenty bands, including NDBI and NDVI amplitude both inflated in the Gini ranking relative to permutation importance recorded near-zero or negative permutation importance, indicating they contributed little unique predictive signal once the dominant SAR bands were already in the model.

The finding that OPT+SAR statistically matched the full stack (Table 3; McNemar's $p=0.499$) while using 13 of the 20 available bands, rather than requiring the full feature set, has direct practical value: it suggests that, in this setting, the phenological metrics and spectral indices in the original 20-band stack could be omitted from an operational pipeline with no significant loss of classification accuracy, simplifying preprocessing to two sensor sources (Sentinel-2 spectral bands and Sentinel-1 SAR) rather than four derived feature groups. This runs somewhat counter to the common assumption in the multi-sensor fusion literature that combining more complementary feature types monotonically improves accuracy; here, more features did not mean better performance, and a targeted two-group combination outperformed the full stack on every point-estimate metric, even though the difference from the full stack was not statistically significant.

Why Radar, and Why Here: Linking the Result to AMAC's Agro-Ecology

The dominance of SAR, and in particular the high ranking of VH temporal variance (Table 4, rank 2), is plausible in light of AMAC's specific agro-ecological setting, though we note at the outset that this study did not collect field-level crop-type or phenology data, so the explanation below should be read as an interpretation consistent with the pattern of results and the known regional context, not as something the dataset directly confirms. The Federal Capital Territory falls within the Guinea savanna zone of Nigeria's middle belt, where cultivation is dominated by rain-fed smallholder farming of maize, sorghum, millet, and cassava, sown between April and June and harvested between August and October, closely tracking the territory's single rainy season (April–October) (FCTA/ACReSAL, 2025). Crucially, the natural vegetation matrix surrounding cropland in AMAC classified here as 'other vegetation' and comprising the park/grassy savanna, savanna woodland, and shrub savanna that together cover the majority of the territory greens up on largely the same rainfall-driven schedule as the surrounding crops. Because both cover types flush green at the onset of the rains and senesce at their end, their NDVI-derived phenological signatures (green-up timing, peak greenness, senescence) are expected to overlap substantially, which is consistent with phenological metrics contributing comparatively little discriminative power in this study (Table 2; Table 4) despite their demonstrated value in other cropland mapping contexts with more phenologically distinct land cover mosaics (Van Tricht et al., 2018).

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

SAR backscatter, by contrast, is sensitive to canopy structure and volume scattering rather than pigment- or greenness-driven reflectance, and is therefore better positioned to distinguish the relatively low, seasonally cultivated, row-structured canopy of smallholder cereal and tuber fields from the taller, woodier, structurally persistent canopy of savanna woodland and shrubland, even when the two are spectrally similar during the wet season. VH backscatter is additionally sensitive to soil moisture and surface roughness; the tilled, periodically bare soil characteristic of smallholder fields between land preparation and canopy closure (concentrated in the April–June sowing window) contrasts with the more continuous ground cover of uncultivated savanna, plausibly reinforcing the separability captured by the VV and VH bands. The particularly strong contribution of VH temporal variance is consistent with the annual crop cycle itself: cropland fields transition from bare or lightly vegetated soil at planting, through peak canopy structure during the growing season, to harvested or fallow bare ground by the October–November dry-season onset, producing a large within-year range in backscatter that a single natural-vegetation phenology metric would not necessarily capture. Woody savanna vegetation, retaining structural elements (trunks, branches) even through senescence, would be expected to show comparatively lower structural variance across the same period. This structural-and-temporal contrast, rather than a spectral or greenness-based one, offers a physically grounded explanation for why radar not optical indices or phenology emerged as the dominant contributor to cropland discrimination in this landscape, and suggests that the value of SAR here may be tied specifically to the presence of a structurally distinct, seasonally dynamic smallholder cropping system set within a spectrally similar but structurally different natural vegetation matrix, a configuration that may not generalize to landscapes where cropland and surrounding vegetation differ more strongly in spectral phenology, canopy height, or field size.

From a practical standpoint, these findings suggest that operational cropland mapping programs in similar tropical savanna, data-constrained contexts should prioritize investment in radar time series and SAR preprocessing capacity over the added complexity of deriving and maintaining phenological metrics and spectral index stacks, at least as a first-order feature engineering decision. This recommendation is offered with the caveat that it reflects a single growing season and a single geographic setting, and should be tested across additional seasons and regions before being generalized into standard operational practice.

This study has several limitations. First, the feature ablation was conducted using RF only; whether the pattern of SAR dominance holds for GTB or other classifiers, as compared in Ibrahim et al. (2026), remains untested, and it is possible that a different classifier would extract more value from the indices and phenology features that underperformed here. Second, although the validation set (n=373) is larger than initially assumed, McNemar's test power is still limited for comparisons with small numbers of discordant pairs (e.g., OPT+SAR vs. Full stack had only 35 discordant pairs), so the absence of a significant difference between OPT+SAR and the full stack should be interpreted as inconclusive rather than as strong evidence of true equivalence. Third, reference labels were derived from the ESA WorldCover product rather than independently collected ground truth, meaning reported accuracy reflects agreement with an existing global product rather than field-verified classification performance. Fourth, the ablation reflects a single growing season; the relative value of SAR versus optical/phenological features, particularly given SAR's sensitivity to soil moisture, may differ across seasons with different rainfall patterns or during periods of more persistent cloud cover, when SAR's cloud-penetration advantage would be expected to matter even more.

CONCLUSIONS

This study conducted a systematic feature ablation to quantify the individual and combined contribution of spectral indices, NDVI-derived phenological metrics, and SAR backscatter to cropland classification accuracy in the Abuja Municipal Area Council, Nigeria, building on the multi-sensor feature stack established in Ibrahim et al. (2026). Contrary to the assumption that a fully fused feature stack maximizes accuracy, SAR backscatter emerged as the dominant contributor across accuracy comparison, McNemar's significance testing, and Gini variable importance: the optical-plus-SAR combination matched or exceeded the full stack on every accuracy metric, while spectral indices and phenological metrics contributed little value on their own and measurably reduced accuracy when combined without SAR. These findings provide practical guidance for prioritizing sensor investment in operational cropland monitoring in data- and resource-constrained tropical settings favoring SAR-optical fusion over the added complexity of phenological and index-based feature engineering and demonstrate

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

a reproducible, GEE-based ablation framework that can be applied to other study areas, land cover classes, and classifiers in future work.

REFERENCES

1. Azadi, H., Taheri, F., Burkart, S., Mahmoudi, H., De Maeyer, P., & Witlox, F. (2021). Impact of agricultural land conversion on climate change: H. Azadi et al. Environment, Development and Sustainability, 23(3), 3187-3198..
2. Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS journal of photogrammetry and remote sensing, 114, 24-31.Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.
3. Du, J., Watts, J. D., Jiang, L., Lu, H., Cheng, X., Duguay, C., & Tarolli, P. (2019). Remote sensing of environmental changes in cold regions: Methods, achievements and challenges. Remote Sensing, 11(16), 1952.
4. Dubey, A., Singh, P. K., Singh, S., Manzoor, U., Sahoo, C., & Saini, A. (2025). Impact of Climate Change and Land Degradation on Agriculture. In Eco-Resilience: Climate Change, Land Degradation and Sustainable Solutions (pp. 49-95). Cham: Springer Nature Switzerland.
5. FCTA/ACReSAL. (2025). FCT: Agro-Climatic Resilience in Semi-Arid Landscapes. Federal Capital Territory Administration. Retrieved from https://acresal.gov.ng/states/fct/
6. Foody, G. M. (2004). Thematic map comparison. Photogrammetric Engineering & Remote Sensing, 70(5), 627-633.
7. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote sensing of Environment, 202, 18-27.
8. Ibrahim, I., Khalid, S. S., Musa, H. H., & Ahmed, N. A. (2026). A comparative analysis of Random Forest and Gradient Tree Boosting for cropland mapping using multi-sensor Sentinel data: A case study of Abuja Municipal Area Council, Nigeria. International Journal of Latest Technology in Engineering, Management & Applied Science (IJLTEMAS), 15(4), 156–164. https://doi.org/10.51583/IJLTEMAS.2026.150400014
9. Li, Y., & Xiao, X. (2025). Deep learning-based fusion of optical, radar, and LiDAR data for advancing land monitoring. Sensors, 25(16), 4991.
10. Van Tricht, K., Gobin, A., Gilliams, S., & Piccard, I. (2018). Synergistic use of radar Sentinel-1 and optical Sentinel-2 imagery for crop mapping: A case study for Belgium. Remote Sensing, 10(10), 1642.