A COMPARATIVE ANALYSIS OF THE SEMI-AUTOMATIC AND AUTOMATIC APPROACHES TO THE IDENTIFICATION OF CROPS AND AGRICULTURAL AREAS IN THE DEPARTMENT OF LOUETSI-WANO (LÉBAMBA, GABON).

Abstract:

Mapping of food-crop farming systems increasingly relies on artificial intelligence algorithms. In our previous work [13-14], we used a series of automatic and semi-automatic algorithms to map agricultural plantations and to discriminate between cassava and plantain crops. This paper presents a systematic comparative analysis of these two approaches. Starting from the methodology, we examine the statistical robustness of the models (automatic and semi-automatic) and their sensitivity to sample size. The results show that tree-based ensemble methods (RF, XGBoost) consistently outperform neural architectures and the parametric classifier when the training sample is small, raising the Kappa index for cassava and plantain discrimination from 0.63 to 0.89 within the same geographic context and the same three-level hierarchical protocol. The analysis concludes that, for crops with small training samples, the choice of algorithm family matters more than the data source alone, and it suggests ways to strengthen the methodology of future agricultural mapping studies in tropical forest areas.

Key words:-

Remote sensing; machine learning; Random Forest; convolutional neural network; agricultural mapping; Lébamba; Gabon.

Introduction:-

Mapping agricultural systems is a key issue for food security and spatial planning, particularly in countries where agriculture relies on small, scattered family farms [8,17,18]. Since the advent of satellites and UAVs, it has become possible to map agricultural land. Furthermore, the use of classification algorithms is becoming increasingly crucial for identifying these systems. Several researchers are utilising algorithms integrated into GIS software [11]. However, others are using slightly more advanced techniques known as machine learning to improve the accuracy of classifications, map agricultural systems effectively and distinguish between different cropping systems [2,9,15]. In Gabon, the remoteness of rural areas and the lack of reliable data from an agricultural census complicate the mapping of cultivated areas and the identification of crop types, despite the authority's stated commitment to achieving food self-sufficiency. Whilst studies have succeeded in mapping land use at a national level [16], these remain focused on monitoring forest cover and distinguish only a generic 'cultivated land' category, without precisely locating or delineating small family farm plots.

In response to this shortcoming, researchers have used semi-automatic and automatic algorithms to map agricultural plantations and attempt to distinguish between cassava and plantain crops in the municipality of Lébamba, using datasets combining Sentinel-2 satellite imagery, field GPS surveys and very high-resolution drone [13-14]. These studies share the same study area, the same three-level hierarchical classification framework and the same statistical evaluation framework (confusion matrix, Kappa coefficient according to Landis and Koch, 1977), but differ fundamentally in the family of algorithms employed. In their first study, they compared Convolutional Neural Networks (CNNs), Support Vector Machines (SVMs) and Maximum Likelihood (ML). In the second, they compared Random Forests (RF), Extreme Gradient Boosting (XGBoost) and a multilayer perception (RNA). This unique setup – the same object, the same terrain, the same data sources and distinct algorithms – offers a rare opportunity to conduct a controlled comparative analysis of the effect of algorithmic choice on the quality of agricultural mapping in a tropical context.

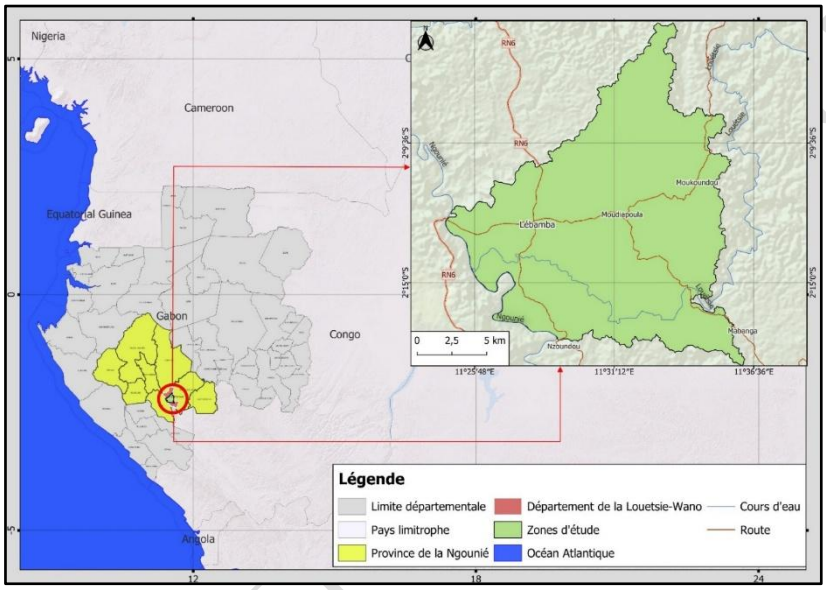
In this paper, our overall objective is to ascertain whether the observed performance differences are due more to the choice of algorithmic family, the data source used (GPS, drone, fusion), or the size of the training sample available

UNDER PEER REVIEW IN IJAR

for each class. To this end, our comparative analysis focuses on three areas: (i) the nature and volume of the data used in each study; (ii) the operating principles of the algorithms tested on each dataset; (iii) the performance metrics reported for each of the three levels of the hierarchical classification (overall accuracy, Kappa coefficient, accuracy per class).

Joint study area:-

Both corpora cover the commune of Lébamba and its surroundings, in the department of Louétsie-Wano (Ngounié Province, southern Gabon), an area defined by the Ngounié and Louétsie rivers, roughly bounded by 2°15’00’’ to 2°9’36’’ South latitude and 11°25’48’’ to 11°36’36’’ East longitude. The landscape, predominantly forested, is interspersed with savannahs, forest clearings and small-scale agricultural plots used mainly for growing cassava and plantains.



Source: Authors

Methodology:-

Data sources:-

To ensure the accuracy of our analysis, we are using same assumptions as in our previous work in the same region. This involves the same field dataset, comprising a Sentinel-2 Level 2A image (April 2024, resampled to 10 m and 20 m resolutions), 246 GPS readings associated with the crops observed, spread across 51 reference plots, and very high-resolution drone imagery (flown at an altitude of 80 m, GSD resolution of 2.2 cm/pixel) used to delineate plot boundaries. The only difference between the two datasets lies in the classification algorithms tested: Convolutional Neural Networks, Support Vector Machines and Maximum Likelihood for the first study, compared with Random Forest, Extreme Gradient Boosting and a multi-layer perceptron-type Artificial Neural Network for the second (FA). The fact that the same dataset was used allows for a direct comparison of performance between the two sets of algorithms.

Comparison of algorithms

The work by [13] is based on different types of architecture and methods:

- The RNC learns non-linear hierarchical representations through successive convolutions;
- The SVM separates classes using an optimal hyperplane in a feature space (radial basis function kernel);
- Maximum Likelihood classifies each pixel according to a multivariate Gaussian distribution based on Bayes’ theorem.

UNDER PEER REVIEW IJAR

Comparative analysis of these three algorithms led to the selection of the RNC algorithm, as it produced the best results. In the second phase, [14]considered three further algorithms:

- FA aggregates the predictions of multiple decision trees trained on bootstrapped subsets [3];
- XGBoost sequentially constructs trees that correct the residuals from previous trees (regularised gradient boosting);
- RNA (multi-layer perceptron) stacks layers of neurons with non-linear activation (ReLU) trained via backpropagation.

The conclusions of this second study show that the FA algorithm performs better.

The Convolutional Neural Network is based on a sequence of mathematical operations, namely: convolution, non-linear activation, pooling, a fully connected layer, classification, followed by backpropagation learning [10]. The table below summarises the main formulae presented by the authors and indicates, for each one, the page of the article on which it appears.

UNDER PEER REVIEW IN IJAR

Table 1. Key concepts used in processing data relating to the Convolutional Neural Network algorithm

Element	Formula
Convolution	$a_{ij} = \sigma\left((W * X)_{ij} + b\right)$
Multi-map convolution	$C_q^l = \left(\sum_{p=1}^{n} S_p^{l-1} * k_{p,q}^l + b_q^l\right)$
Detailed convolution	$C_q^l = \left(\sum_{p=1}^{n}\sum_{u=-x}^{x}\sum_{v=-x}^{x} S_p^{l-1}(i-u, j-v)\cdot k_{p,q}^l(u,v) + b_q^l\right)$
Pooling	$S_q^l(i,j) = \frac{1}{4}\sum_{u=0}^{z}\sum_{v=0}^{z} C_q^l(2i-u, 2j-v)$
Fully connected layer	$\widehat{output} = \sigma(W \times f + b)$
Softmax	$\hat{y}(i) = \frac{e^{\widehat{output}}}{\sum_{labels} e^{output}}$
Loss	$L = \frac{1}{2}\sum_{i=1}^{N}(\hat{y}(i) - y(i))^2$
ReLU / Leaky ReLU	$f(x) = \max(0,x)$; αx si $x < 0$
Adam Optimizer	$g_t = \frac{\partial f(x,w)}{\partial w}$ $p_t = m_1 \cdot p_{t-1} + (1-m_1)\cdot g_t$ $q_t = m_2 \cdot q_{t-1} + (1-m_2)\cdot g_t^2$ $\hat{p}_t = \frac{p_t}{1-m_1^t}$ $\hat{q}_t = \frac{q_t}{1-m_2^t}$ $w_t = w_{t-1} - \alpha \cdot \frac{\hat{p}_t}{\sqrt{\hat{q}_t} + \epsilon}$

UNDER PEER REVIEW IN IJAR

On the other hand, the Random Forest comprises a large number of decision trees, each constructed from a random vector independent of the others, and the class selected is the one that receives the most votes [3]. [3] formalizes the accuracy of such a forest using three concepts, namely: the margin, the strength of the individual trees, and the correlation between these trees. The table below summarises the main formulae from this article and indicates the page on which each appears.

Table 2. Key concepts used in processing data relating to the Random Forest algorithm

Statement	Mathematical formulation
Random forest definition	$\{h(x, \Theta_k),\ k = 1,2, \dots\}$
Margin function of a set of classifiers	$mg(X,Y) = \text{av}_k\, I(h_k(X) = Y) - \max_{j \neq Y} \text{av}_k\, I(h_k(X) = j)$
Generalisation error	$PE^* = P_{X,Y}(mg(X,Y) < 0)$
Convergence of the error (as the number of trees increases)	$PE^* \to P_{X,Y}\big(P_\Theta(h(X,\Theta) = Y) - \max_{j \neq Y} P_\Theta(h(X,\Theta) = j) < 0\big)$
Margin function of the forest	$mr(X,Y) = P_\Theta(h(X,\Theta) = Y) - \max_{j \neq Y} P_\Theta(h(X,\Theta) = j)$
Strength of the classifiers	$s = E_{X,Y}\, mr(X,Y)$
Chebyshev's inequality	$PE^* \leq \dfrac{\text{var}(mr)}{s^2}$
Variance of the margin	$\text{var}(mr) = \bar{\rho}\,(E_\Theta\, sd(\Theta))^2 \leq \bar{\rho}\, E_\Theta\, \text{var}(\Theta)$
Average correlation between trees	$\bar{\rho} = \dfrac{E_{\Theta,\Theta'}\big(\rho(\Theta,\Theta')\, sd(\Theta)\, sd(\Theta')\big)}{E_{\Theta,\Theta'}\big(sd(\Theta)\, sd(\Theta')\big)}$
Upper bound on the variance	$E_\Theta\, \text{var}(\Theta) \leq 1 - s^2$
Upper bound on the generalisation error	$PE^* \leq \dfrac{\bar{\rho}\,(1 - s^2)}{s^2}$
c/s² ratio	$c/s^2 = \bar{\rho}/s^2$

UNDER PEER REVIEW IN IJAR

Statement	Mathematical formulation
Margin in the case of two classes	$mr(X,Y) = 2P_\Theta(h(X,\Theta) = Y) - 1$
Relative strength for class j	$s_j = E_{X,Y}\left(P_\Theta(h(X,\Theta) = Y) - P_\Theta(h(X,\Theta) = j)\right)$
Bound for multiple classes	$PE^* \leq \sum_j \frac{\mathrm{var}\left(P_\Theta(h(X,\Theta) = Y) - P_\Theta(h(X,\Theta) = j)\right)}{s_j^2}$
Number of variables drawn at each node (Forest-RI)	$F = \mathrm{int}(\log_2 M + 1)$
Forest-RI computation time ratio	$\frac{F \cdot \log_2(N)}{M}$
Root mean square error of a numerical predictor	$E_{X,Y}(Y - h(X))^2$
Convergence in regression	$E_{X,Y}(Y - \mathrm{av}_k\, h(X,\Theta_k))^2 \rightarrow E_{X,Y}(Y - E_\Theta\, h(X,\Theta))^2$
Average error of a tree	$PE^*(\mathrm{tree}) = E_\Theta\, E_{X,Y}(Y - h(X,\Theta))^2$
Forest error in regression	$PE^*(\mathrm{forest}) \leq \bar{\rho}\, PE^*(\mathrm{tree})$
Weighted correlation of residuals	$\bar{\rho} = \frac{E_{\Theta,\Theta'}\left(\rho(\Theta,\Theta')\, sd(\Theta)\, sd(\Theta')\right)}{(E_\Theta\, sd(\Theta))^2}$
Proportion of out-of-sample (OOB) votes	$Q(x,j) = \frac{\sum_k I\left(h(x,\Theta_k) = j;\ (y,x) \notin T_{k,B}\right)}{\sum_k I\left((y,x) \notin T_{k,B}\right)}$

NB : Θ_k : random vector that governs the construction of the tree k ; h(x,Θ) : classifier (tree); I(·) : indicator function; av_k: average on the trees; ρ ¯: average correlation between trees; s : strength of trees ; M: number of input variables; N :number of observations; T_(k,B) : bootstrap sample of the tree k.

Both datasets employ the same three-level top-down classification architecture (overall land cover, vegetation component mask, and fine-scale taxonomic discrimination between cassava and banana) and validate their results using the confusion matrix and the interpretation scale for the Kappa coefficient developed by Landis and Koch (1977).

UNDER PEER REVIEW IN IJAR

Comparative results:-

Level 1: Identification of agricultural crops within the overall land use pattern :-

At Level 1, both datasets achieve high overall accuracies, but with significant discrepancies for the 'agricultural plantation' target class. In Dataset 1, the best result is achieved by the RNC using GPS data (overall accuracy 98.85 per cent, Kappa 0.97), although with a omission error of 58.65 per cent for the agricultural class. In corpus 2, the Random Forest model using GPS data achieves a higher overall accuracy (99.75 per cent, Kappa 0.9940), with an F1-score of 0.82 for the agricultural class, which is significantly better than the balance between accuracy and recall observed in corpus 1.

Table 3. Best performance at Level 1 by corpus and by data source

Corpus	Best model	Source	Overall accuracy (%)	Kappa coefficient
1	RNC	GPS	98,85	0,97
1	RNC	UAV	98,73	0,97
1	RNC	Combination	98,28	0,96
2	FA	GPS	99,75	0,9940
2	FA	UAV	99,71	0,9933
2	FA	Combination	99,72	0,9936

Level 2: isolation of agricultural parcels within vegetation classes:-

Level 2 confirms the same hierarchy. The RNC corpus achieved its best result with the GPS + drone combination (98.28 per cent, Kappa 0.96), whilst the FA corpus reached a near-perfect maximum with drone data (99.88 per cent, Kappa 0.9964).

Table 4. Best performance at Level 2 by corpus and by data source

Corpus	Best model	Source	Overall accuracy (%)	Kappa coefficient
1	RNC	GPS	97,11	0,93
1	RNC	UAV	96,34	0,93
1	RNC	Combination	95,16	0,93
2	FA	GPS	99,83	0,9940
2	FA	UAV	99,88	0,9964
2	FA	Combination	99,81	0,9939

Level 3: Distinguishing between cassava and plantains :-

It is at this level that the difference between the two corpora is most revealing. In Corpus 1, GPS sampling alone results in a complete failure to distinguish between the two crops (Kappa of 0.00, with all banana plots being classified as cassava); the addition of drone data partially overcomes this limitation (Kappa of 0.63), whilst the raw fusion of the two sources reduces performance (Kappa of 0.53). In corpus 2, even with the GPS sample alone, the Random Forest achieves a Kappa of 0.7273 (overall accuracy 92.86 per cent); with the drone data, it reaches a Kappa of 0.8941 (overall accuracy 94.74 per cent); the fusion of the sources also performs well (Kappa 0.8712). By contrast, the RNA in the FA corpus exactly replicates the behaviour of the RNC in the RNC corpus when using GPS data alone: a Kappa of zero, with the minority class (banana) being completely absorbed by the majority class (cassava).

UNDER PEER REVIEW IN IJAR

Table 5. Comparison of performance between the best results at Level 3 (cassava/banana discrimination)

Corpus	Best model	Source	Overall accuracy (%)	Kappa coefficient
1	RNC	GPS	87,12	0,00
1	RNC	UAV	81,64	0,63
1	RNC	Combination	76,86	0,53
2	Random forest	GPS	92,86	0,7273
2	Random forest	UAV	94,74	0,8941
2	Random forest	Combination	93,57	0,8712

Discussion:-

An algorithmic family effect that is more decisive than a data source effect:-

A comparison of the two datasets reveals findings consistent with the literature on the robustness of decision-tree-based ensemble methods [3,4,7]: given an identical data collection protocol and study area, switching from a neural network or parametric architecture (RNC, SVM, MV) to an ensemble method (Random Forest, XGBoost) significantly improves crop discrimination performance, particularly at level 3, where the Kappa coefficient increases from 0.63 (the best result in the RNC dataset, obtained using the drone) to 0.89 (the best result in the RF dataset, also obtained using the drone). This discrepancy, observed within the same study area and using the same hierarchical protocol, suggests that the choice of algorithmic family is at least as important as the data source used in determining the final quality of the mapping.

Sensitivity to sample volume as a common explanatory factor :-

The behaviour of the RNA in Corpus 2 closely mirrors that of the RNC in Corpus 1 when the GPS sample is reduced: both neural networks fail to distinguish the minority class (banana) and simply predict the majority class (cassava), which results in both cases in a Kappa coefficient of zero despite a deceptively high overall accuracy. This finding is consistent with the observations made by [1]. This parallelism, observed independently in both studies, confirms that neural networks are particularly sensitive to class imbalance and small training sample sizes [5], whereas tree-based ensemble methods (notably Random Forest) retain a minimal discriminatory capacity even under these degraded conditions [6].

Thematic consistency and cross-validation of the two corpora:-

Beyond the statistical indicators, the two datasets converge thematically: both place plantain cultivation in the wettest and most alluvial areas, and cassava on better-drained land – a distribution consistent with the known ecology of these two species. This qualitative convergence, achieved by two entirely distinct families of algorithms, constitutes a form of independent cross-validation of the agro-ecological reality of the study area, reinforcing the credibility of the entire dataset beyond mere performance scores.

Conclusion :-

A comparative analysis of two studies conducted on the same agricultural plot in the municipality of Lébamba, utilising similar data sources but distinct families of algorithms, shows that the shift from neural and parametric architectures (RNN, SVM, Maximum Likelihood) to ensemble methods based on decision trees (Random Forest, XGBoost) significantly improves the quality of fine-grained discrimination between cassava and plantain crops, particularly when the available training sample is limited. This result suggests that future agricultural mapping work in tropical forest areas should prioritise decision-tree-based ensemble methods as an initial approach and in contexts where data is scarce, whilst reserving neural network architectures for settings with a sufficiently large field sample. A controlled replication testing all six algorithms on a single dataset and using spatially disjoint cross-validation would be the natural next step to consolidate these findings. This comparison has several limitations that should be highlighted. Both studies rely on a random partitioning of training and validation pixels, which introduces spatial

UNDER PEER REVIEW IN IJAR

autocorrelation that may artificially inflate the scores reported in both datasets, as the authors of both papers themselves acknowledge. Finally, one of the main constraints encountered in both datasets was persistent cloud cover in the region, which considerably limited the availability of usable satellite imagery. This situation necessitated an approach based on a single image from April, thereby restricting the analysis to a single (spatial) snapshot. Such a constraint limits the study of the temporal and phenological dynamics of crops, the inclusion of which could, however, improve spectral discrimination in future research.

References :-

1. Anand, R, Mehrotra K.G., Mohan C.K., S. Ranka S. (1993). An improved algorithm for neural network classification of imbalanced training sets. IEEE Transactions on Neural Networks, 4(6), 962-969. https://doi.org/10.1109/72.286891

2. Bolfe, É. L., Parreiras, T. C., Silva, L. A. P. D., Sano, E. E., Bettiol, G. M., Victoria, D. D. C., Sanches, I. D., & Vicente, L. E. (2023). Mapping agricultural intensification in the Brazilian Savanna: A machine learning approach using harmonized data from Landsat Sentinel-2. ISPRS International Journal of Geo-Information, 12(7), Article 263. https://doi.org/10.3390/ijgi12070263

3. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324

4. Genuer, R. (2010). Random Forests: Elements of theory, variable selection and applications [Thèse de doctorat, Université Paris-Sud].

5. Ghosh, K., Bellinger, C., Corizzo, R., Branco, P., Krawczyk, B., &Japkowicz, N. (2024). The class imbalance problem in deep learning. Machine Learning, 113(7), 4845-4901. https://doi.org/10.1007/s10994-022-06268-8

6. Han, S., Williamson, B. D., & Fong, Y. (2021). Improving random forest predictions in small datasets from two-phase sampling designs. BMC Medical Informatics and Decision Making, 21(1), Article 322. https://doi.org/10.1186/s12911-021-01688-3

7. Hu, L., & Li, L. (2022). Using tree-based machine learning for health studies: Literature review and case series. International Journal of Environmental Research and Public Health, 19(23), Article 16080. https://doi.org/10.3390/ijerph192316080

8. Huang, X., Vrieling, A., Dou, Y., Li, X., & Nelson, A. (2025). Divergent crop mapping accuracies across different field types in smallholder farming regions. International Journal of Applied Earth Observation and Geoinformation, 139, Article 104559. https://doi.org/10.1016/j.jag.2025.104559

9. Ibrahim, E. S., Rufin, P., Nill, L., Kamali, B., Nendel, C., & Hostert, P. (2021). Mapping crop types and cropping systems in Nigeria with Sentinel-2 imagery. Remote Sensing, 13(17), Article 3523. https://doi.org/10.3390/rs13173523

10. Indolia, S., Goswami, A. K., Mishra, S. P., & Asopa, P. (2018). Conceptual understanding of convolutional neural network- A deep learning approach. Procedia Computer Science, 132, 679-688. https://doi.org/10.1016/j.procs.2018.05.069

11. Koné, M., Sanogo, T., Sidibé, S., & Mariko, Z. (2026). Performance et limites des workflows de classification d'images satellitaires dans le contexte Ouest Africain : Comparaison de ENVI, ERDAS Imagine et Orfeo ToolBox (OTB). European Scientific Journal, ESJ, 22(2), 128. https://doi.org/10.19044/esj.2026.v22n2p128

12. Landis, J. R., & Koch, G. G. (1977). A one-way components of variance model for categorical data. Biometrics, 33(4), 671-679. https://doi.org/10.2307/2529465

UNDER PEER REVIEW IN IJAR

13. Massala, M. M., Nkoumakali, B., & Obiang, L. O. (2026). Occupation du sol et identification des cultures agricoles à partir des algorithmes d'apprentissage automatique. Perspectives Plurielles, Tome II (Série 1, n° 3), 159-175.
14. Massala, M. M., Tarazewicz, D. M., & Obiang, L. O. (2026). Évaluation comparative des algorithmes d'identification des plantations agricoles et des cultures à partir des algorithmes du Réseau de Neurones Convolutifs (RNC), de Machines à Vecteurs de Support (SVM) et du Maximum de Vraisemblance (MV). Cahier de la Recherche Africaine, 4(7), 205.

15. Mazarire, T. T., Ratshiedana, P. E., Nyamugama, A., Adam, E., & Chirima, G. (2022). Exploring machine learning algorithms for mapping crop types in a heterogeneous agriculture landscape using Sentinel-2 data. A case study of Free State Province, South Africa. South African Journal of Geomatics, 9(2), 333-347. https://doi.org/10.4314/sajg.v9i2.22

16. Nzigou Boucka, F., Obame, C. V., Manfoumbi, F., Nzue Mba, A., Ngui Ondo, M., Ovono, V., & Mambimba Ndjoungui, A. (2021). Cartographie de l'occupation du sol du Gabon en 2015—Changements entre 2010 et 2015. Revue Française de Photogrammétrie et de Télédétection, 223, 118-128. https://doi.org/10.52638/rfpt.2021.567

17. Sun, L., Li, H., Li, S., Zhao, S., Zhang, Y., Mo, Y., Zhang, B., & Wu, Y. (2026). A multi-source remote sensing and machine learning framework for maize mapping and yield estimation in fragmented Loess gully regions. Frontiers in Plant Science, 17, Article 1892566. https://doi.org/10.3389/fpls.2026.1892566

18. Crespin-Boucaud, A., Lebourgeois, V., Lo Seen, D., Castets, M., &Bégué, A. (2020). Agriculturally consistent mapping of smallholder farming systems using remote sensing and spatial modelling. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-3/W11, 35-42. https://doi.org/10.5194/isprs-archives-XLII-3-W11-35-2020

UNDER PEER REVIEW IN IJAR