Journal of Agrometeorology
(A publication of Association of Agrometeorologists)
ISSN : 0972-1665 (print), 2583-2980 (online)
Vol. No. 28 (3) : 404-414 (September - 2026)
https://doi.org/10.54386/jam.v28i3.3442
https://journal.agrimetassociation.org/index.php/jam

Research paper

Machine Learning-Based Multi-Decadal Analysis of Paddy Field Dynamics in the Cauvery Delta, India: Implications for Sustainable Agriculture

GUNAVATHI SUNDARAM* and SELVAKUMAR RADHAKRISHNAN

School of Civil Engineering, SASTRA Deemed to be University, Thanjavur, Tamil Nadu, India
**Corresponding author email: gunavathi.sundaramc@gmail.com*

ABSTRACT

Paddy cultivation forms the backbone of agriculture in the Cauvery Delta of Tamil Nadu, India; however, long-term monitoring of its spatial dynamics remains limited. This study employed multi-temporal Landsat imagery and machine learning techniques to analyse the spatiotemporal dynamics of paddy cultivation in Thanjavur district over a 23-year period (1995–96, 2007–08, and 2018–19), while an additional reference year (2015–16) was used for independent validation. Three machine learning classifiers, namely Random Forest (RF), Gradient Tree Boosting (GTB), and Support Vector Machine (SVM), were evaluated using Recall, Precision, Overall Accuracy (OA), F1 Score, and Kappa Coefficient (KC). The RF classifier achieved the highest classification accuracy during 1995–96 (OA = 0.84, KC = 0.68), 2007–08 (OA = 0.90, KC = 0.81), and 2015–16 (OA = 0.90, KC = 0.80), whereas GTB showed superior performance during 2018–19 (Recall = 0.85, OA = 0.82, F1 Score = 0.83, KC = 0.65). The classification results were validated using independently sourced Ground Control Points (GCPs) for 2015–16 and official Agriculture Department acreage statistics, demonstrating strong agreement with the reference data. The results revealed a 6.72% decline in paddy cultivation area (approximately 228 km²) between 1995–96 and 2018–19, accompanied by an expansion of non-paddy and urban land uses. Spatially, paddy cultivation became increasingly concentrated in the western and central parts of the district along the Cauvery River, primarily due to groundwater depletion, irrigation constraints, rapid urbanisation, and increasing climate variability. The findings demonstrate the effectiveness of ensemble-based machine learning approaches for large-scale agricultural monitoring and provide a transferable framework for assessing long-term land-use changes, thereby supporting food security, climate adaptation, and sustainable water resource management in monsoon-dependent deltaic regions.

Keywords: Machine learning; Paddy mapping; Land use change; Cauvery Delta; Food security; Sustainable agriculture

Rice (Oryza sativa L.) is the staple food for more than half of the global population, with Asia contributing nearly 90% of total production (Park et al. 2018). However, increasing population pressure, soil degradation, groundwater depletion, inefficient irrigation, and climate change are intensifying food security challenges worldwide (Zhang and Lin 2019). Global rice demand is projected to reach 915 million tonnes by 2030, while paddy cultivation currently occupies approximately 11% of global cropland, consumes nearly 30% of freshwater resources, and contributes substantially to methane emissions (Li et al., 2025). These challenges highlight the need for accurate and continuous monitoring of paddy cultivation systems.

Remote sensing has become an essential tool for agricultural monitoring due to its capability to provide repetitive and large-scale observations. While coarse-resolution datasets such as MODIS and AVHRR have been widely used for regional assessments of rice cultivation (Gumma et al., 2016), their ability to capture field-level changes is limited. In contrast, Landsat imagery offers a long-term archive and moderate spatial resolution (30 m), making it suitable for monitoring agricultural dynamics over extended periods (Wulder et al., 2019). Traditional paddy mapping approaches based on spectral indices and crop phenology have shown promising results (Sakamoto et al., 2009), but their performance is often affected by cloud cover, spectral confusion, and temporal variability (Dong & Xiao, 2016). Recent advances in machine learning algorithms, such as Random Forest (RF), Support Vector Machine (SVM), and Gradient Tree Boosting (GTB), have

Article info - DOI: *https://doi.org/10.54386/jam.v28i3.3442*
Received:01 April 2026; Accepted: 27 July 2026; Published online : 7 September 2026

significantly improved classification accuracy and enabled efficient processing of large geospatial datasets through cloud-computing platforms such as Google Earth Engine (DeVries et al., 2020).

Despite recent advances in machine learning-based paddy mapping, most existing studies have focused on single-year assessments, short-term monitoring, or algorithm development using limited temporal datasets. Comprehensive multi-decadal investigations that systematically compare multiple machine learning algorithms, incorporate independent validation, and evaluate long-term spatial dynamics of paddy cultivation remain limited, particularly in South Asian deltaic agroecosystems. The Cauvery Delta of Tamil Nadu, one of India's most productive rice-growing regions, is increasingly affected by monsoon variability, declining river flows, groundwater depletion, irrigation constraints, and rapid urbanisation, necessitating continuous monitoring of agricultural land-use changes.

To address these gaps, this study presents a comprehensive machine learning-based framework for multi-decadal paddy mapping in Thanjavur district using Landsat imagery spanning 1995–96, 2007–08, and 2018–19. Unlike previous studies, the proposed framework integrates (i) a comparative evaluation of RF, GTB, and SVM algorithms across multiple time periods, (ii) independent validation using a reference year (2015–16) supported by field-based Ground Control Points (GCPs) and official Agriculture Department records, and (iii) spatial change detection to quantify long-term paddy field transitions. This integrated approach provides new insights into the temporal evolution of paddy cultivation in the Cauvery Delta and establishes a transferable framework for agricultural monitoring in climate-sensitive deltaic regions.

MATERIALS AND METHODS

Study Area

The present study focuses on Thanjavur district in the Cauvery Delta region of Tamil Nadu, India, covering approximately 3,600 km² between 9°50′–11°25′ N and 78°45′–79°25′ E. The climate is tropical semi-humid with bimodal rainfall from the southwest (342 mm) and northeast (545.7 mm) monsoons, averaging 887.7 mm annually. Temperatures range from 24°C to 33°C, while increasing monsoon variability poses risks to agriculture and water availability.

The Cauvery River and its tributaries support intensive paddy cultivation on fertile alluvial and sandy clay loam soils, making the region known as the "Granary of South India". However, climatic fluctuations, recurrent droughts, and rainfall variability threaten agricultural sustainability and water security. The study covered all 14 administrative blocks across the district (Fig. 1) (Source: Sundaram and Radhakrishnan, 2023).

Data Acquisition and Pre-processing

The methodological framework adopted in this study is illustrated in Fig. 2. The Samba season, the principal rice-growing season in Thanjavur district, was selected because it accounts for the largest paddy cultivation area. According to the regional crop calendar, the Samba season typically extends from August to February, with sowing/transplanting occurring between August and 15 September, the vegetative growth stage extending from 16 September to 15 December, and harvesting taking place between 16 December and 15 February. Landsat 5, 7, and 8 Level-1 Terrain Precision (L1TP) images for 1995–96, 2007–08, 2018–19, and a reference year (2015–16) were acquired from the USGS Earth Explorer for both pre- and post-harvest periods with less than 10% cloud cover (Table 1). Cloud and shadow masking were applied to the image using the pixel_qa band, followed by clipping to the study area boundary using QGIS. The scan line errors in the Landsat 7 ETM+ imagery were further corrected using gap-filling interpolation algorithms. Radiance measurements were converted to TOA reflectance to account for differences in sun zenith angle and sensor properties. To achieve uniform spatial coverage, adjacent path/row scenes were mosaicked, and spectral bands were combined into multispectral composites.

Training Sample Generation and Classification

The Normalized Difference Vegetation Index (NDVI) () was calculated for both pre- and post-harvest images, and the NDVI difference (pre–post) was used to distinguish paddy and non-paddy areas using a threshold value (>0.65) (Chen et al., 2011). Representative paddy and non-paddy sample points were manually digitised using Landsat imagery supported by high-resolution Google Earth imagery and expert visual interpretation. Paddy sample points were identified using NDVI-based extraction, whereas non-paddy sample points were selected through visual interpretation of high-resolution Google Earth imagery. A total of 860 representative sample points were collected. The complete sample dataset was subsequently partitioned into 70% training and 30% testing subsets using a random sampling procedure in Google Earth Engine, resulting in 600 training samples (300 paddy and 300 non-paddy) and 260 testing samples (130 paddy and 130 non-paddy). The training samples were used to develop the classification models, while the testing samples were used for independent accuracy assessment (Fig. 3).

Supervised classification was performed on the pre-harvest imagery using RF, SVM, and GTB algorithms. RF generates multiple decision trees using bootstrapped training samples and assigns the final class based on majority voting, thereby reducing variance and overfitting (Breiman, 2001). SVM identifies an optimal separating hyperplane in a transformed feature space and is effective for nonlinearly separable remote-sensing data (Srivastava et al., 2012). GTB constructs decision trees sequentially, with each tree reducing the residual errors of the preceding trees (Friedman, 2002). For RF, the parameters *ntree*, *mtry*, *nodesize*, and *maxnodes* were optimized, while the kernel type, cost (C), and gamma (γ) were optimized for SVM. For GTB, the parameters *n.trees*, *shrinkage*, *interaction.depth*, and *n.minobsinnode* were optimized. Hyperparameter tuning was performed using grid search with 10-fold cross-validation within the **mlr** framework in R, separately for each decadal dataset, and the optimal parameter set was selected based on the highest mean test accuracy and the minimum misclassification error. The optimized configurations for

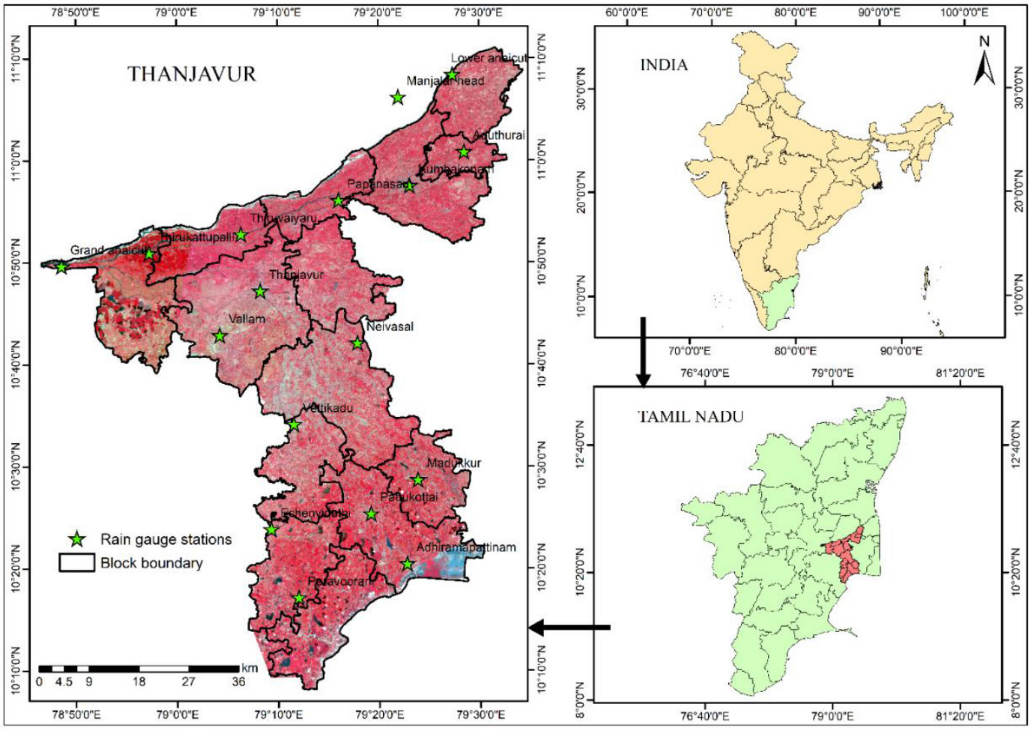


Fig. 1: Location map of the study area with block boundaries

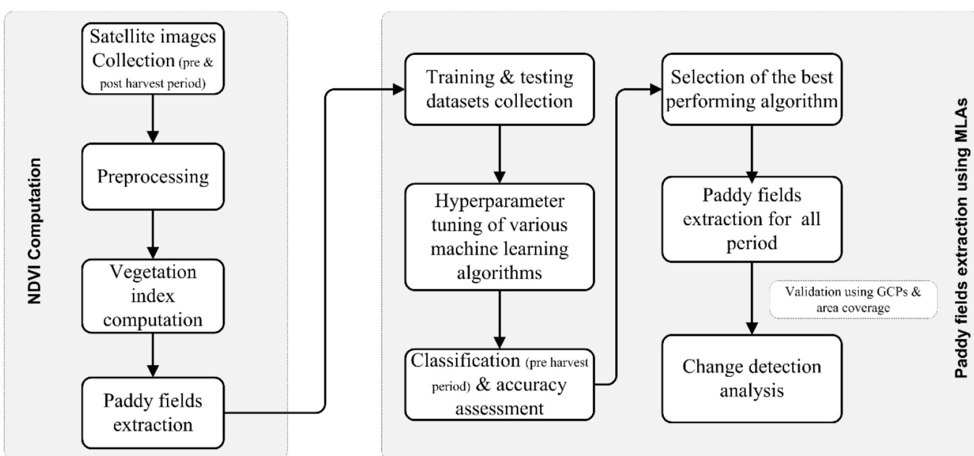


Fig. 2: Methodological flowchart of the study

each classifier and study year are presented in Table 2.

The performance of the optimized models was subsequently evaluated using Overall Accuracy, Precision, Recall, F1 Score, and Kappa Coefficient (Table 3). Overall accuracy is one of the most commonly used measures in classification, reflecting the proportion of correctly classified samples out of the total test samples. Precision indicates the proportion of correctly identified instances among all the samples predicted for a given class, while recall refers to the proportion of correctly identified instances out of all actual samples belonging to that class. In practice, precision and recall often show a trade-off, where an increase in one can lead to a decrease in the other. The F1 score, which combines both precision and recall, is expressed as their harmonic mean, with a value of 1 representing the best performance and 0 the weakest. Additionally, the Kappa Coefficient (KC) serves as a statistical indicator for evaluating the level of agreement between the classified results and the reference data. A KC value approaching 0 signifies poor or slight agreement, whereas a value nearing 1 denotes strong or almost perfect agreement between the two datasets (Keovongsa et al., 2024; W. Zhang et al., 2020) (Equation1-5).

where TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative.

Table 1: Details of Landsat image acquisition (path/row and dates) for pre- and post-harvest periods during the study years

Year	Path /Row	Date of image acquisition	Year	Path /Row	Date of image acquisition
1995 -1996	142/052	26/12/1995	2015 - 2016	142/052	18/01/2016
	142/053	26/12/1995		142/053	18/01/2016
	143/052	17/12/1995		143/052	22/01/2016
	142/052	18/03/1996		142/052	22/03/2016
	142/053	18/03/1996		142/053	22/03/2016
	143/052	04/03/1996		143/052	14/04/2016
2007 -2008	142/052	20/12/2007	2018 - 2019	142/052	10/01/2019
	142/053	20/12/2007		142/053	10/01/2019
	143/052	24/11/2007		143/052	01/01/2019
	142/052	08/03/2008		142/052	15/03/2019
	142/053	08/03/2008		142/053	15/03/2019
	143/052	31/03/2008		143/052	25/04/2019

Table 2: Optimised hyperparameters of the machine learning algorithms corresponding to each study year

Hyperparameters	1995-1996	2007-2008	2015-2016	2018-2019
Gradient boosting				
n.trees	480	330	144	421
shrinkage	0.32	0.09	0.07	0.21
interaction.depth	8	4	10	9
n.minobsinnode	5	5	5	4
Mmce.test.mean	0.21	0.15	0.19	0.14
Tpr.test.mean	0.20	0.12	0.16	0.12
Tnr.test.mean	0.22	0.18	0.24	0.17
Acc.test.mean	0.79	0.85	0.81	0.86
Random forest				
ntree	39	108	54	143
mtry	4	3	8	5
nodesize	3	3	3	2
maxnodes	15	15	15	14
Mmce.test.mean	0.23	0.17	0.26	0.17
Tpr.test.mean	0.14	0.09	0.12	0.19
Tnr.test.mean	0.32	0.24	0.40	0.15
Acc.test.mean	0.77	0.83	0.74	0.83
Support vector machine				
kernel	Radial	Radial	Radial	Radial
cost	1	10	10	10
gamma	1	1	1	1
Mmce.test.mean	0.21	0.16	0.17	0.14
Tpr.test.mean	0.17	0.07	0.16	0.11
Tnr.test.mean	0.28	0.25	0.18	0.19
Acc.test.mean	0.80	0.84	0.83	0.87

$$OA = \frac{TP + TN}{TN + TP + FN + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$F1score = \frac{2 * Recall * Precision}{Recall + Precision} \tag{4}$$

$$KC = \frac{Prob.\ of\ correct\ classification - prob.\ of\ chance\ agreement}{1 - prob.\ of\ chance\ agreement} \tag{5}$$

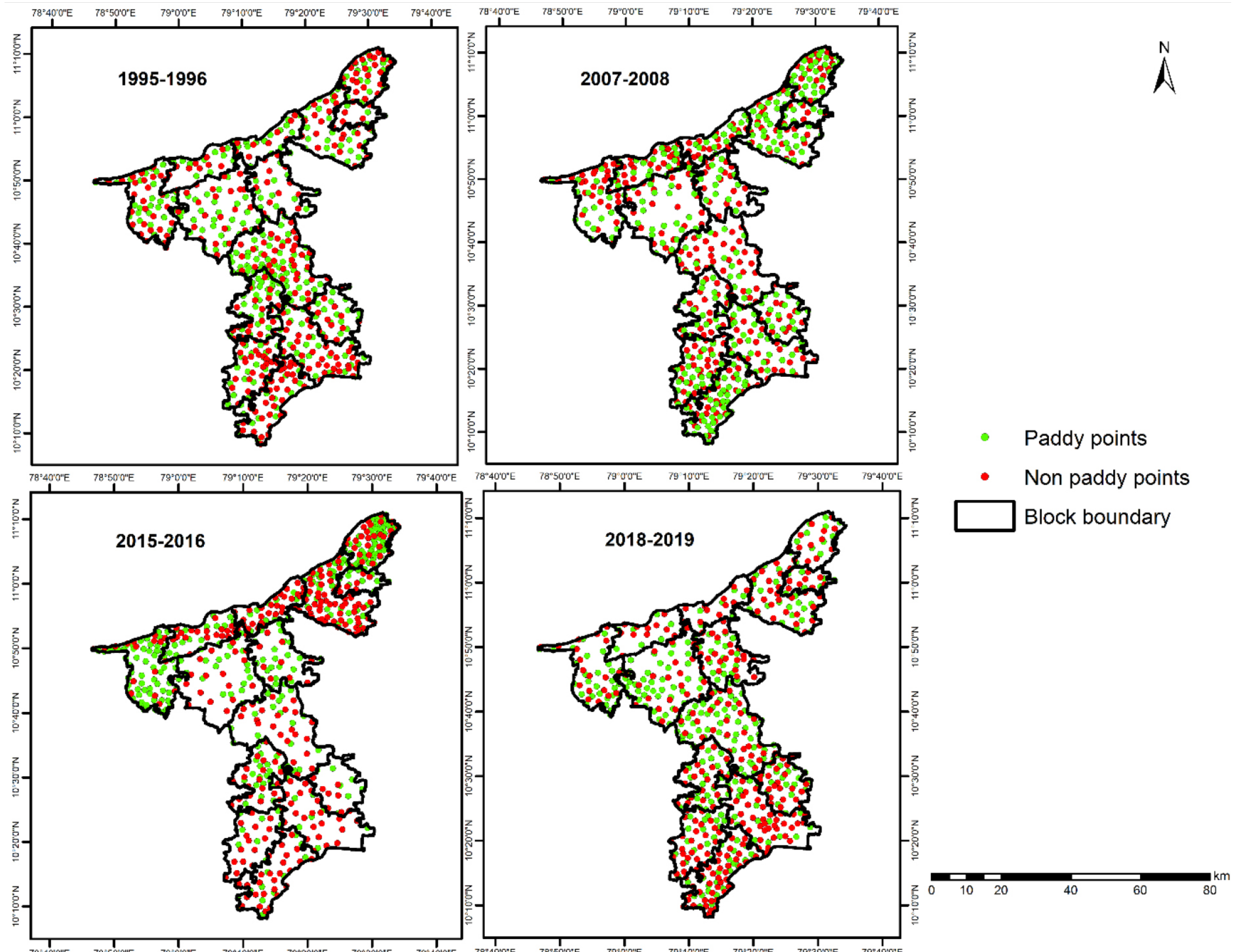


Fig. 3: Spatial distribution of manually digitized paddy and non-paddy sample points used for supervised classification during (a) 1995–96, (b) 2007–08, (c) 2015–16, and (d) 2018-19.

Validation and Change Detection

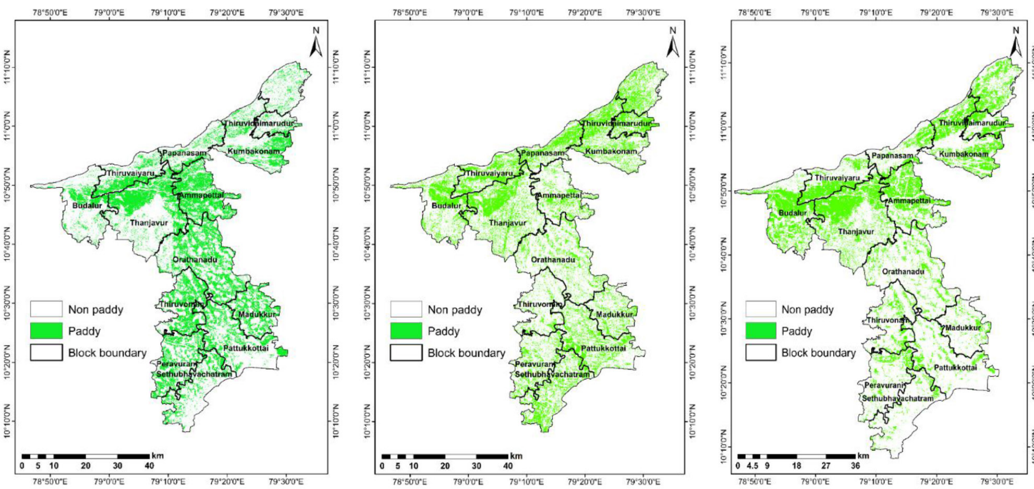
Classification results were validated using two independent approaches. First, spatial accuracy was assessed using Ground Control Points (GCPs) obtained from published field surveys for the reference year 2015–16. Second, classified paddy area estimates were compared with official acreage statistics obtained from the Agriculture Department, Thanjavur. Finally, change detection analysis was conducted to evaluate spatiotemporal changes in paddy cultivation between 1995–2007, 2007–2018, and 1995–2018. The workflow integrated QGIS, R and Google Earth Engine (GEE). QGIS was used for image preprocessing, NDVI computation, and sample extraction, while hyperparameter optimization was performed in R using the *mlr* package. The optimized parameters were subsequently applied in GEE for supervised classification and accuracy assessment.

RESULTS

Selection of the Best-Performing Machine Learning Algorithms

To evaluate model performance, five metrics such as Recall, Precision, OA, F1 Score, and KC were used (Table 4). In 1995–96, RF achieved the highest OA (0.84), F1 Score (0.84), and KC (0.68), outperforming GTB and SVM, although SVM had higher Recall (0.92) and GTB comparable Precision (0.81). In 2007–08, RF again performed best (Recall = 0.92, Precision = 0.89, OA = 0.90, F1 = 0.91, KC = 0.81), while GTB showed strong Precision (0.89) and SVM remained competitive (OA = 0.88, KC = 0.76).

For 2015–16, RF maintained superior performance with Recall = 0.91, Precision = 0.89, OA = 0.90, F1 Score = 0.90, and KC = 0.80, indicating consistent robustness. GTB also performed well (OA = 0.85, KC = 0.71), while SVM showed moderate accuracy (OA = 0.88, KC = 0.75). By 2018–19, GTB achieved the most balanced results (Recall = 0.85, OA = 0.82, F1 = 0.83, KC = 0.65), whereas



(a) RF (1995–96) (b) RF (2007–08) (c) GTB (2018–19)

Fig. 4: Machine learning-based paddy cultivation maps for (a) 1995–96, (b) 2007–08, and (c) 2018–19, showing the spatiotemporal decline in paddy cultivation

Table 3: Representation of the confusion matrix

	Classification			
	Classes	Paddy	Non-paddy	Total
Reference	Paddy	TP	FN	TP+FN
	Non-paddy	FP	TN	FP+TN
	Total	TP+FP	FN+TN	TN+TP+FN+FP

RF retained higher Precision (0.83) but lower Recall (0.79), and SVM showed moderate consistency. Although GTB outperformed the other classifiers in 2018–19, the lower KC is likely due to increased spectral heterogeneity, variations in crop phenology, and mixed-pixel effects in the Landsat imagery, which reduced the separability between paddy and non-paddy classes. Nevertheless, the satisfactory OA (0.82) and F1 Score (0.83) demonstrate the robustness of GTB for regional-scale paddy mapping.

Overall, RF was the most reliable classifier during 1995–96, 2007–08, and 2015–16, while GTB performed better under the more heterogeneous conditions observed in 2018–19.

Paddy Field Extraction Using the Selected Algorithms

The 1995–96 image was classified using the RF algorithm, achieving an overall accuracy of 83.85%. The optimized hyperparameters were: ntree = 39, mtry = 4, nodesize = 3, and maxnodes = 15. Paddy fields covered 1,205.91 sq. km (35.5%), mainly in the central and southern regions with a south–north spread, indicating stable cultivation supported by favorable water availability. Non-paddy areas (2,190.66 sq. km) were distributed across the western, central, northeastern, and southeastern zones (Fig. 4a).

The 2007–08 image was classified using the RF algorithm, achieving an overall accuracy of 90.30%. The tuned hyperparameters were: ntree = 108, mtry = 3, nodesize = 3, and maxnodes = 15. Paddy area declined to 1,087.84 sq. km (32.03%), a 3.47% decrease, and was concentrated along the Cauvery belt from west to northeast. Non-paddy areas increased to 2,308.73 sq. km, reflecting expansion that was likely driven by water stress and land-use changes (Fig. 4b).

The 2018–19 image was classified using the GTB algorithm, yielding an overall accuracy of 82.24%. The optimized hyperparameters were: n.trees = 421, shrinkage = 0.205, interaction. depth = 9, and n.minobsinnode = 4. Paddy fields covered 977.52 sq. km (28.78%), mainly in the western, central, and northeastern regions, while non-paddy areas (2,419.05 sq. km) dominated the southern parts. Overall, the paddy cultivation area declined by 6.72% (228.39 sq. km) from 1995–96, indicating a gradual shift influenced by water scarcity, land conversion, and climate variability (Fig. 4c).

Validation of Classification Results

The 2015–16 image was classified using the RF algorithm, achieving an overall accuracy of 0.90. The optimized hyperparameters were: ntree = 54, mtry = 8, nodesize = 3, and maxnodes = 15. Paddy fields covered 964.14 sq. km (28.40%), while non-paddy areas extended over 2,432.43 sq. km. Paddy cultivation was predominantly distributed in the northern, northeastern, and central parts of Thanjavur district, with scattered patches in southern blocks. In contrast, non-paddy areas were more extensive across the western and southern regions, indicating a fragmented and reduced paddy distribution compared to earlier periods (Fig. 5).

The 2015–16 period was selected for validation due to the availability of reliable ground-truth data. An independent set of 28 Ground Control Points, derived from field observations reported by Ajith et al. (2017), was used for validation and was not included in the training or testing datasets. These GCPs represent paddy field locations distributed across major paddy-growing areas of Thanjavur district. Each point was overlaid on the classified map to assess whether it was correctly identified as paddy. Although the validation

Table 4: Performance metrics of machine learning algorithms for different study periods, with the highest values highlighted in bold

Year	MLA	Recall	Precision	OA	F1 Score	KC
	GTB	0.81	0.81	0.81	0.81	0.62
1995-1996	RF	0.88	0.81	0.84	0.84	0.68
	SVM	0.92	0.76	0.82	0.83	0.63
	GTB	0.84	0.89	0.86	0.86	0.72
2007-2008	RF	0.92	0.89	0.90	0.91	0.81
	SVM	0.91	0.86	0.88	0.89	0.76
	GTB	0.85	0.86	0.85	0.85	0.71
2015 - 2016	RF	0.91	0.89	0.90	0.90	0.80
	SVM	0.88	0.87	0.88	0.88	0.75
	GTB	0.85	0.81	0.82	0.83	0.65
2018-2019	RF	0.79	0.83	0.82	0.81	0.63
	SVM	0.85	0.79	0.81	0.82	0.62

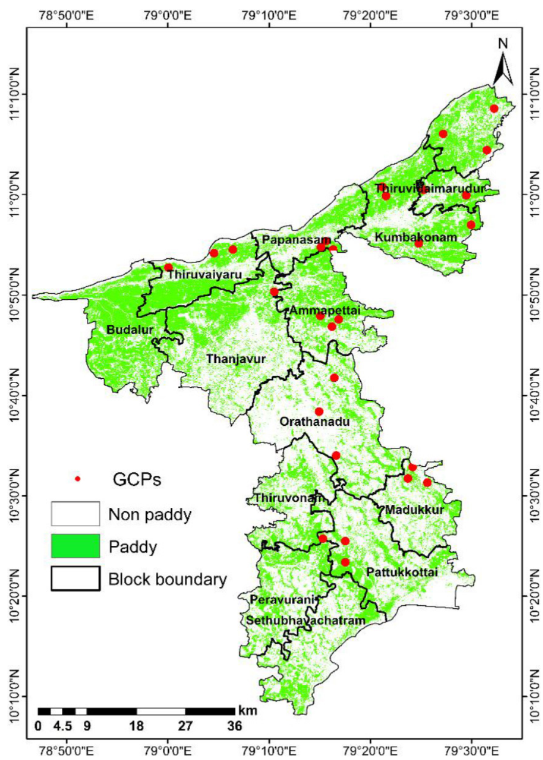


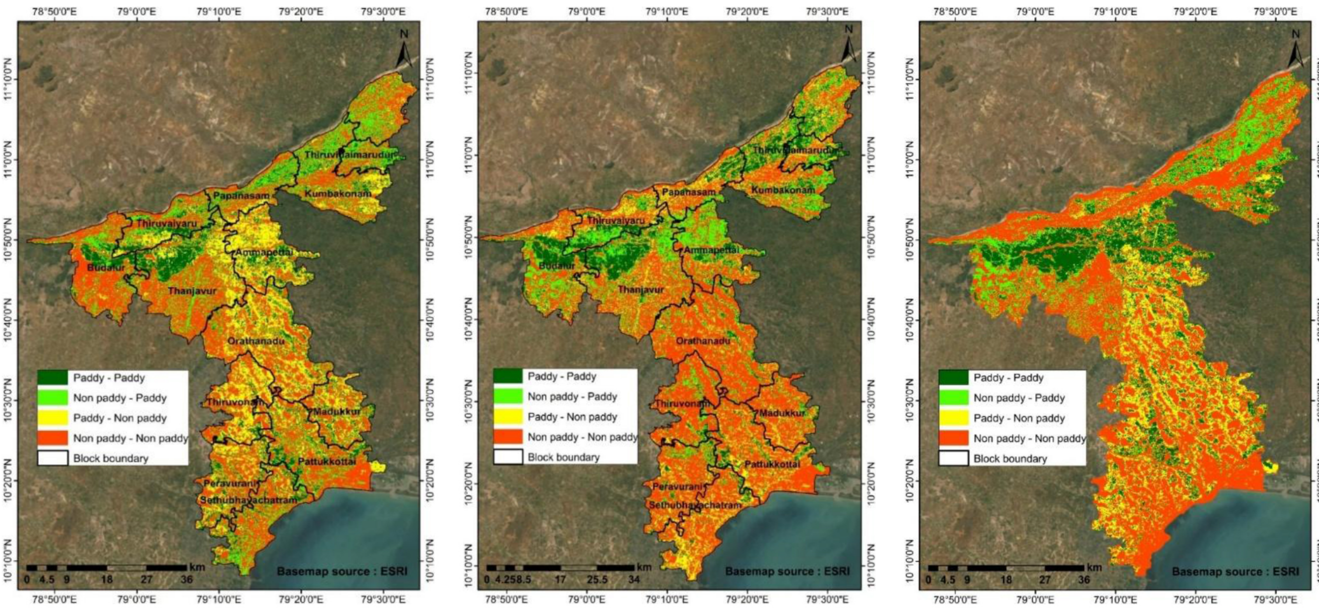
Fig. 5: Supervised classification map showing the spatial distribution of paddy and non-paddy fields in Thanjavur district for the 2015–16 reference year. The locations of the Ground Control Points (GCPs) are overlaid to illustrate the samples used for validation

dataset was limited to 28 independent GCPs and comprised only paddy field locations, all GCPs were correctly identified as paddy in the classified map, indicating strong agreement between the classification results and field observations.

Consequently, the GCP-based validation primarily evaluated the model's ability to correctly identify paddy areas under actual field conditions. However, classifier performance was independently assessed using testing datasets containing both paddy and non-paddy samples, and the resulting confusion matrix metrics accounted for potential misclassification between the two classes. Furthermore, the classified paddy areas were compared with official Agriculture Department acreage statistics, providing an additional assessment of area-level consistency and strengthening the reliability of the classification results.

Table 5 shows a close agreement between the ML-derived paddy area estimates and the Agriculture Department (Thanjavur) records for most years. The lower similarity observed for 1995–96 (63.89%) is primarily due to the absence of season-specific Samba acreage records, necessitating comparison with annual paddy acreage data that include the Samba, Kuruvai, and Thaladi seasons. Consequently, the classified area underestimated the reference area by 682.09 sq. km, corresponding to a 36.13% error. This temporal mismatch likely contributed to the lower agreement.

In contrast, the discrepancies for 2007–08, 2015–16, and 2018–19 were substantially lower, with absolute errors of 66.36 sq. km (5.75%), 18.80 sq. km (1.91%), and 28.24 sq. km (2.81%), respectively. The agreement improved to 94.25% in 2007–08 and exceeded 97% in both 2015–16 and 2018–19, indicating that when



a) 1995-2007 b) 2007-2018 c) 1995-2018

Fig. 6: Change detection maps illustrating paddy and non-paddy land-use transitions during (a) 1995–2007, (b) 2007–2018, and (c) 1995–2018

comparable Samba-season reference data were available, the ML-based approach produced paddy area estimates that closely matched the official records. These results demonstrate the reliability of the proposed methodology for regional-scale paddy area estimation.

Change Detection Analysis

Between 1995 and 2007, change detection in Thanjavur district showed clear spatial variability in paddy dynamics. The paddy–paddy class (447.2 sq. km) was concentrated in the western and parts of the southern regions, indicating persistent paddy cultivation supported by irrigation. Paddy expansion was evident in the non-paddy–paddy class (640.6 sq. km) across the northern, northeastern, and southern areas, while a larger extent of paddy–non-paddy conversion (758.7 sq. km) occurred in the central, southwestern, southeastern, and northeastern regions, reflecting shifts to other land uses. The non-paddy–non-paddy class dominated (1550 sq. km), indicating widespread persistence of non-paddy land uses (Fig. 6a).

During 2007–2018, the paddy–paddy class (455.2 sq. km) remained concentrated in the western, northeastern, and parts of the southern regions, suggesting continued cultivation in well-irrigated zones. The non-paddy–paddy class (522.3 sq. km) in the western, central, and northeastern areas indicates localized expansion, whereas paddy–non-paddy conversion (632.7 sq. km) in the southern, southeastern, northern, and northeastern regions indicate a continuing decline. The non-paddy–non-paddy class further increased to 1786.4 sq. km, reinforcing the dominance of non-paddy land uses (Fig. 6b).

Over the longer period (1995–2018), persistent paddy areas (paddy–paddy) covered 566.8 sq. km, mainly in the western, central, southern, and northeastern regions. Conversion from non-paddy to paddy (410.7 sq. km) was limited to parts of the western and northeastern zones, while substantial paddy loss (paddy–non-paddy: 639.1 sq. km) occurred across the central, southern, and southeastern regions. The non-paddy–non-paddy class remained dominant (1780.1 sq. km), indicating a gradual yet consistent contraction of paddy cultivation alongside expansion of alternative land uses (Fig. 6c).

DISCUSSION

The classification results demonstrated the effectiveness of machine learning algorithms for mapping paddy cultivation across multiple time periods in Thanjavur district. Among the evaluated classifiers, RF achieved the highest accuracy during 1995–96, 2007–08, and 2015–16, whereas GTB performed comparatively better during 2018–19. The superior performance of RF can be attributed to its ability to effectively handle noisy and high-dimensional remote sensing data while minimizing overfitting through ensemble learning. In contrast, GTB likely benefited from its boosting mechanism, which was more effective at capturing complex class boundaries in increasingly fragmented and heterogeneous agricultural landscapes resulting from urban expansion and land-use change. These findings highlight the importance of selecting classification algorithms according to landscape characteristics and temporal conditions.

The temporal analysis of paddy field dynamics in Thanjavur district for 1995–96, 2007–08, and 2018–19 revealed a steady decline in paddy area, despite its status as the "Rice Bowl of Tamil Nadu." Similar trends have been reported across the Cauvery Delta Region. Remote sensing evidence from the Thirumanimuttar sub-basin indicates significant reduction in paddy fields (Vijayakumar et al. 2015). Analyses of Season and Crop Report data (1996–2016) for the Cauvery Delta also show a shift from paddy to less water-intensive crops such as pulses and cotton, along with a decline in sugarcane and banana (Kavipriya et al., 2019). A negative

Table 5: Comparison of paddy area estimates from Agriculture Department data and ML-based classifications.

Year	Reference Paddy Area (sq. km)	ML-Derived Paddy Area (sq. km)	Percent Similarity (%)	Percent Error (%)	Absolute Error (sq. km)
1995-96	1,888	1,205.91	63.89	36.13	682.09
2007-08	1,154.20	1,087.84	94.25	5.75	66.36
2015-16	982.94	964.14	98.09	1.91	18.80
2018-19	1,005.76	977.52	97.19	2.81	28.24

growth trend across Tamil Nadu during 2001–2020 further confirms continued conversion of paddy lands (Narmadha and Karunakaran 2022).

Water scarcity is a major driver in the Cauvery Delta. Groundwater depletion (1990–2019) has increased irrigation stress (Prayag et al. 2023), while deterioration of tank irrigation systems in Tamil Nadu has reduced surface water availability (Neelakantan et al., 2017). Erratic rainfall and inadequate irrigation infrastructure have also altered cropping patterns (Stella et al. 2023). In addition, institutional and market factors such as limited procurement, varietal market constraints, and reliance on groundwater have influenced farmers to shift toward less water-demanding crops (Nair et al. 2024).

Land-use conversion has further accelerated this decline, with agricultural lands in Tamil Nadu increasingly diverted to non-agricultural uses (Govindaprasad and Manikandan 2016). In the Cauvery Delta, urban expansion and resource degradation (Arunpandiyan et al. 2015), along with conversion of coastal paddy fields to shrimp aquaculture causing salinization (Pratheepa et al. 2023), have intensified the loss of paddy areas.

In this context, the System of Rice Intensification (SRI) represents a viable adaptation strategy in the Cauvery Delta, improving water-use efficiency, yield, and profitability (Devi et al. 2023). Overall, the decline in paddy cultivation reflects the combined influence of hydrological, land-use, and socio-economic factors, highlighting the need for integrated management strategies to sustain agriculture in the region.

CONCLUSION

This study demonstrates the effectiveness of machine learning classifiers for long-term monitoring of paddy cultivation dynamics in the Thanjavur district, using multi-temporal Landsat imagery spanning over two decades. Ensemble-based methods-particularly Random Forest and Gradient Tree Boosting exhibited superior classification performance, achieving high overall accuracies (≥0.82) and Kappa Coefficients (≥0.64), underscoring their reliability for spatiotemporal agricultural mapping. The proposed methodological framework effectively integrates NDVI-based seasonal differencing with advanced machine learning algorithms within the Google Earth Engine environment, providing a scalable and transferable model for continuous agricultural monitoring and change detection.

The results reveal a 6.72% reduction (228.39 km²) in paddy area between 1995 and 2018, reflecting the progressive transformation of Thanjavur's agrarian landscape. This decline is primarily driven by groundwater depletion, declining irrigation efficiency, climate variability, and land conversion pressures. These findings highlight the escalating vulnerability of deltaic agricultural systems, where hydrological, climatic, and anthropogenic stressors converge to threaten both ecological stability and rural livelihoods.

Rice cultivation in the Cauvery Delta requires integrated land and water management strategies, including rehabilitation of traditional irrigation systems, groundwater recharge enhancement, water-use efficiency technologies, and climate-resilient rice varieties. These measures strengthen climate adaptability and food and water security in the region. Future research should use multi-source datasets like Sentinel-1 Synthetic Aperture Radar and high-resolution optical imagery to overcome limitations and develop near-real-time decision-support systems for adaptive agricultural management. This approach can guide policymakers, planners, and local communities towards data-driven, climate-resilient, and sustainable agricultural development in monsoon-dependent rice ecosystems.

ACKNOWLEDGEMENT

The authors express their gratitude to the School of Civil Engineering, SASTRA Deemed to be University. They also extend special appreciation to all those who contributed directly or indirectly to this project.

Author's certificate: The manuscript or its part is not under consideration for publication elsewhere and the same has been approved by all co-authors.

Conflict of Interests: The authors declare that there is no conflict of interest regarding the publication of this article.

Funding: The researchers received no funding or support from any external organizations for the submitted work.

Data Availability: The satellite imagery used in this study is publicly available at USGS Earth Explorer.

Author's contribution: **Gunavathi Sundaram:** Methodology, Analysis and investigation, Writing - original draft preparation; **Selvakumar Radhakrishnan:** Supervision.

Publisher's Note: The periodical remains neutral concerning jurisdictional claims in published maps and institutional affiliations.

REFERENCES

Ajith, K., Geethalakshmi, V., Ragunath, K. P., Pazhanivelan, S., & Panneerselvam, S. (2017). Rice Acreage Estimation in Thanjavur, Tamil Nadu Using Lands at 8 OLIIMAGES and GIS Techniques. *International Journal of Current Microbiology and Applied Sciences*, *6*(7), 2327–2335. https://doi.org/10.20546/IJCMAS.2017.607.275

Arunpandiyan, M., Aarthi, M., Vidyalakshmi, R., Savaridhos, R. J., & Devi, P. (2015). Impacts of urbanization on water resources and vegetation on the delta region of Tamilnadu using remote sensing and GIS. *Res. Rev.: J. Ecol. Environ. Sci*, *3*(4), 43-47

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

Chen, J., Huang, J., & Hu, J. (2011). Mapping rice planting areas in southern China using the China Environment Satellite data. *Mathematical and Computer Modelling*, *54*(3–4), 1037–1043. https://doi.org/10.1016/J.MCM.2010.11.033

Devi, G. K., Narayanamoorthy, A., Jothi, P., & Sujitha, K. S. (2023). Can the System of Rice Intensification Save Water and Increase Yield of Paddy?: A Study of Tank Irrigated Area. *Journal of Land and Rural Studies*, *11*(2), 232102492311640–232102492311640. https://doi.org/10.1177/23210249231164032

DeVries, B., Huang, C., Armston, J., Huang, W., Jones, J. W., & Lang, M. W. (2020). Rapid and robust monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine. *Remote Sensing of Environment*, *240*, 111664. https://doi.org/10.1016/J.RSE.2020.111664

Dong, J., & Xiao, X. (2016). Evolution of regional to global paddy rice mapping methods: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *119*, 214–227. https://doi.org/10.1016/J.ISPRSJPRS.2016.05.010

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, *38*(4), 367–378. https://doi.org/10.1016/S0167-9473(01)00065-2

Govindaprasad, P. K., & Manikandan, K. (2016). Farm Land Conversion and Food Security: Empirical Evidences from Three Villages of Tamil Nadu. *Indian Journal of Agricultural Economics*, *71*(4). https://doi.org/10.22004/AG.ECON.302235

Gumma, M. K., Thenkabail, P. S., Teluguntla, P., Rao, M. N., Mohammed, I. A., & Whitbread, A. M. (2016). Mapping rice-fallow cropland areas for short-season grain legumes intensification in South Asia using MODIS 250 m time-series data. *International Journal of Digital Earth*, *9*(10), 981–1003. https://doi.org/10.1080/17538947.2016.1168489

Kavipriya, J., Mahandrakumar, K., Amarnath, J. S., & Prahadeeswaran, M. (2019). An Extend of Farm Diversification in Cauvery Delta Zone-A Spatial and Temporal Analysis. *Int.J.Curr. Microbiol.App.Sci*, *8*(5), 1257–1263. https://doi.org/10.20546/ijcmas.2019.805.143

Keovongsa, I., Fitriyah, A., Okura, F., Noda, K., Yoshida, K., Keokhamphui, K., & Kato, T. (2024). Analysis of Paddy Field Changes (1989–2021) Using Landsat Images and Flooding-Assisted MLC in an Urbanizing Tropical Watershed, Vientiane, Lao PDR. *Sustainability 2024, Vol. 16, Page 9776*, *16*(22), 9776. https://doi.org/10.3390/SU16229776

Li, S., Zhang, Y., Zhao, J., van Groenigen, K. J., Shen, X., Zhang, H., Gu, J., Zhang, W., Hui, D., Chen, Y., & Liu, L. (2025). Water-saving irrigation practices in rice paddies reverse the impact of root aerenchyma on methane emissions. *Agriculture, Ecosystems & Environment*, *378*, 109309. https://doi.org/10.1016/J.AGEE.2024.109309

Nair, J., Thomas, B. K., & Bahinipati, C. S. (2024). Cropping decisions under water stress: Evidence from Cauvery Delta Region, India. *World Water Policy*, *10*(3), 711–729. https://doi.org/10.1002/WWP2.12177

Narmadha, N., & Karunakaran, K. R. (2022). Agricultural Diversification in Tamil Nadu - An Economic analysis. *Economic Affairs (New Delhi)*, *67*(3), 183–188. https://doi.org/10.46852/0424-2513.3.2022.5

Neelakantan, T. R., Suribabu, C. R., & Selvakumar, R. (2017). Opportunity to restore irrigation tanks in the Cauvery delta by mining and deepening. *Environment, Development and Sustainability*, *19*(4), 1463–1472. https://doi.org/10.1007/S10668-016-9814-7/METRICS

Park, S., Im, J., Park, S., Yoo, C., Han, H., & Rhee, J. (2018). Classification and Mapping of Paddy Rice by Combining Landsat and SAR Time Series Data. *Remote Sensing 2018, Vol. 10, Page 447*, *10*(3), 447. https://doi.org/10.3390/RS10030447

Pratheepa, C. M., Raj, R., & Sinha, S. (2023). The socio-ecological contradictions of land degradation and coastal agriculture in south India. *Environment and Planning E: Nature and Space*, *6*(1), 391–411. https://doi.org/10.1177/25148486221079720

Prayag, A. G., Zhou, Y., Srinivasan, V., Stigter, T., & Verzijl, A. (2023). Assessing the impact of groundwater abstractions on aquifer depletion in the Cauvery Delta, India. *Agricultural Water Management*, *279*, 108191. https://doi.org/10.1016/J.AGWAT.2023.108191

Sakamoto, T., Van Phung, C., Kotera, A., Nguyen, K. D., & Yokozawa, M. (2009). Analysis of rapid expansion of inland aquaculture and triple rice-cropping areas in a coastal area of the Vietnamese Mekong Delta using MODIS time-series imagery. *Landscape and Urban Planning*, *92*(1), 34–46. https://doi.org/10.1016/J.LANDURBPLAN.2009.02.002

Srivastava, P. K., Han, D., Rico-Ramirez, M. A., Bray, M., & Islam, T. (2012). Selection of classification techniques for land use/land cover change investigation. *Advances in Space Research*, *50*(9), 1250–1265. https://doi.org/10.1016/J.ASR.2012.06.032

Stella, S., Fmm, M., & Rekha, M. (2023). An Analysis of Factors Affecting Agriculture Productivity in Thanjavur District of the Cauvery Delta Region. *International Journal of Economics, Business and Management Research*, *07*(01), 69–78. https://doi.org/10.51505/IJEBMR.2023.7107

Sundaram, G., & Radhakrishnan, S. (2023). Performance Evaluation of Bias Correction Methods and Projection of Future Precipitation Changes Using Regional Climate Model over Thanjavur, Tamil Nadu, India. *KSCE Journal of Civil Engineering*, *27*(2), 878–889. https://doi.org/10.1007/s12205-022-0151-0

Vijayakumar, N., Gurugnanam, B., & Arulbalaji, K. (2015). Land use Land cover change detection in Thirumanimuttar Sub Basin, Cauvery River, Tamilnadu. *International Journal of Science, Engineering and Technology Research*, *4*(4), 680-683

Wulder, M. A., Loveland, T. R., Roy, D. P., Crawford, C. J., Masek, J. G., Woodcock, C. E., Allen, R. G., Anderson, M. C., Belward, A. S., Cohen, W. B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosilla, T., Hipple, J. D., Hostert, P., Hughes, M. J., … Zhu, Z. (2019). Current status of Landsat program, science, and applications. *Remote Sensing of Environment*, *225*, 127–147. https://doi.org/10.1016/J.RSE.2019.02.015

Zhang, M., & Lin, H. (2019). Object-based rice mapping using time-series and phenological data. *Advances in Space Research*, *63*(1), 190–202. https://doi.org/10.1016/J.ASR.2018.09.018

Zhang, W., Liu, H., Wu, W., Zhan, L., & Wei, J. (2020). Mapping Rice Paddy Based on Machine Learning with Sentinel-2 Multi-Temporal Data: Model Comparison and Transferability. *Remote Sensing 2020, Vol. 12, Page 1620*, *12*(10), 1620. https://doi.org/10.3390/RS12101620