Land Use and Land Cover Classification using Support Vector Machine Technique: A Case Study of Kaddam Watershed

Sridhar P.[1*], Anil K.[2], Raju J.N.S.S.[3] and Ramana K.V.[4]
1. Centre for Spatial Information Technology, University College of Engineering, Science & Technology Hyderabad, JNTU Hyderabad, INDIA
2. Civil Engineering, Bapatla Engineering College, Bapatla, INDIA
3. Civil Engineering, S.R.K.R. Engineering College, Bhimavaram, INDIA
4. Civil Engineering, Lakireddy Bali Reddy College of Engineering, INDIA
*parupallysridhar@gmail.com

Abstract
Land use and land cover (LU/LC) maps, along with their temporal dynamics, are essential for flood prediction, seasonal water quality monitoring, environmental sustainability planning and ecological assessments. Accurate classification of satellite image datasets presents a challenging task due to the complexity of land use patterns and the need for precise methods. To address these challenges, this study utilized multi-temporal satellite image datasets to perform LU/LC classification and analyze temporal changes within the Kaddam watershed. The Support Vector Machine (SVM) classification technique was employed, using training samples that represent critical land use classes including water bodies, agricultural land, forests, urban areas and barren land. Representative polygons were digitized to train the SVM model and key parameters such as kernel type and gamma values were optimized to enhance classification accuracy.

Performance evaluation was conducted using a confusion matrix to derive metrics such as overall accuracy and Kappa statistics. Ground truth data comparisons further validated the classification results. The high accuracy and robustness of the SVM-based approach demonstrate its potential as a reliable tool for LU/LC classification and its applicability to other regions for effective land use management and planning.

Keywords: Support Vector Machines, LULC, Training, Error matrix.

Introduction

Over the last two decades, numerous research papers have highlighted significant advancements in remote sensing and GIS technologies. These technologies, particularly satellite imagery, provide valuable spatial and temporal datasets that are essential for addressing various domains including environmental monitoring, urban planning, natural resource management, watershed management and disaster management[7,16].

**** Author for Correspondence***

However, effective management, monitoring and planning activities require multi-temporal satellite datasets with varying spatial resolutions to effectively capture changes critical for tackling challenges posed by human activities[9]. The assessment of spatial and temporal changes in regional features can be conducted through either manual or automated methods using specialized software to transform raster to vector data.

Manual methods, although precise, are time-consuming and require significant effort and expertise. Historically, trace films were overlaid to extract features, interpret visual elements and apply classification proficiency to provide information[8]. Land Use Land Cover (LULC) classification techniques for extracting accurate data from remote sensing images have proven highly adaptable. Manual classification, in particular, is a highly effective approach for interpreting remotely sensed data when the analyst has in-depth knowledge of the area being classified. This method leverages the human brain's ability to recognize image features and relate them to real-world objects, often surpassing computers in terms of accuracy.

However, manual interpretation can be time-intensive and subjective, especially as it typically involves images limited to basic red, green and blue colours which may not fully exploit the wealth of spectral information in satellite images. Classification methods can be broadly grouped into supervised and unsupervised approaches, boundary-based and non-boundary classification and hard versus soft classification.

Over recent decades, researchers, planners and scientists have generated LULC maps and analysed statistical and temporal changes using advanced digital techniques and specialized software such as ArcGIS and ERDAS Imagine. The key LULC classification techniques include manual interpretation, numerical approaches and digital methods like NDVI (Normalized Difference Vegetation Index), SAVI (Soil-Adjusted Vegetation Index) and NDWI (Normalized Difference Water Index)[6,15]. Supervised classification methods require prior knowledge of the study area and use algorithms such as Maximum-Likelihood Classifier (MLC), Support Vector Machine (SVM), Random Forest (RF), Decision Tree Classifier, K-Nearest Neighbour (KNN) and Artificial Neural Networks (ANN)[11].

https://doi.org/10.25303/186da0106

In contrast, unsupervised classification techniques, like the ISODATA clustering method, are used when there is no prior knowledge of class labels. The Support Vector Machine (SVM) is a robust supervised classification technique widely recognized for its ability to handle high-dimensional data and complex feature spaces effectively[5]. In the context of remote sensing, SVM excels in classifying satellite imagery, particularly when dealing with medium-resolution datasets, by minimizing misclassification through optimal hyperplane selection. For the present study, SVM was employed to classify multi-temporal satellite images from Landsat and IRS-P6 (LISS-III) into broad land cover categories such as agricultural land, forests, water bodies, barren land and settlements.

Study Area

The Kaddam watershed, located in Telangana, India, serves as a significant sub-watershed of the Godavari River basin, playing a crucial role in regional hydrology due to its undulating terrain, varied slopes and diverse elevations. Geographically, the watershed spans latitudes 18.5° N to 19° N and longitudes 78.5° E to 79° E. It experiences a tropical climate with annual rainfall ranging from 800 to 1000 mm and temperatures fluctuating between 15°C and 45°C. The watershed is predominantly agricultural, supporting crops such as paddy and cotton, while also encompassing deciduous forests that contribute to maintaining ecological balance and influencing hydrological processes. The location map of the study area is presented in fig. 1.

Material and Methods

Land Use and Land Cover (LULC) maps were prepared using satellite imagery and temporal datasets, as detailed in tables 1 and 2. The datasets for this study were obtained from the Bhuvan platform (https://bhuvan-app3.nrsc.gov.in/data/download/index.php; table 1) and Earth Explorer (https://earthexplorer.usgs.gov/; table 2). Access to these datasets required login credentials and the specification of the area of interest (AOI) to download the relevant satellite images. Both platforms have provided free access to a wide range of satellite imagery for the research and scientific community over the past two decades, making them invaluable resources for LULC mapping and related studies.

Preparation of FCC Image: The downloaded multi-temporal satellite datasets, organized by date in separate folders, contain individual spectral bands. Each image is monochrome and requires processing to generate a color image for identifying classes within the dataset. To create a False Color Composite (FCC) image from these bands, the "Composite Bands" tool in the Arc Toolbox within ArcMap is utilized. Individual bands are added to the tool and the output file is named based on the folder's corresponding date. This process is systematically applied to all satellite image datasets, resulting in FCC maps that serve as inputs for generating Land Use and Land Cover (LULC) maps using the Support Vector Machine (SVM) method.

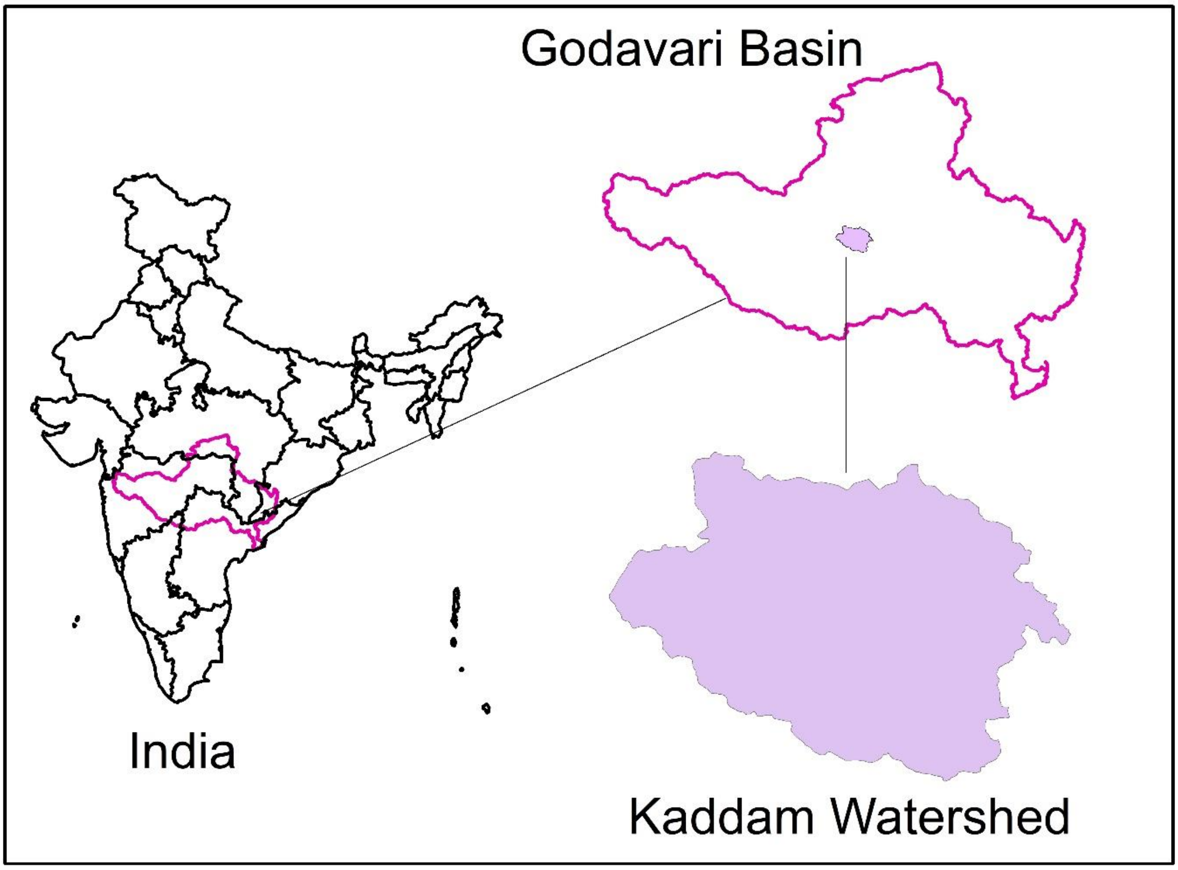

Figure 1: Location Map of the study area

https://doi.org/10.25303/186da0106

Table 1
Satellite data sources of the study area (source: Bhuvan Website)

S.N.	Topo sheet No	Bounding Box	Date of Pass	Spatial Resolution
1	E44A03	78.0E19.25N-78.25E19.5N	14-10-08, 22-12-11 and 25-09-16.	23.5 meters
2	E44A06	78.25E19.5N-78.5E19.75N		
3	E44A07	78.25E19.25N-78.5E19.5N		
4	E44A08	78.25E19.0N-78.5E19.25N		
5	E44A10	78.5E19.5N-78.75E19.75N		
6	E44A11	78.5E19.25N-78.75E19.5N		
7	E44A12	78.5E19.0N-78.75E19.25N		
8	E44A15	78.75E19.25N-79.0E19.5N		
9	E44A16	78.75E19.0N-79.0E19.25N		

Table 2
Satellite data sources of the study area (source: Earth Explorer Website)

S. N.	Data	Path and Row	Date of Pass	Spatial Resolution
1	Land Sat Thematic mapper and ETM7	144/047	21-11-1988 and 19-04-2022.	30 meters

Methodology: The Land Use and Land Cover (LULC) classification for this study was carried out using the Support Vector Machine (SVM) technique, a robust supervised machine learning algorithm widely recognized for its effectiveness in remote sensing applications[5,11]. The process began with the definition of training samples representing different land use and land cover classes such as water bodies, agricultural land, forest and urban areas. Using the Training Sample Manager in ArcMap or ArcGIS Pro, polygons were manually digitized for each class within the area of interest (AOI). These training samples served as the basis for training the SVM classifier, ensuring that the model captured the spectral and spatial characteristics of the study area accurately.

Subsequently, SVM parameters were configured including the selection of an appropriate kernel function (e.g. linear, polynomial, or radial basis function) and adjustments were made to gamma and cost parameters to optimize class separability[12]. This setup ensured that the classifier was tailored to the specific spectral and spatial characteristics of the satellite images used in the study. After configuring the SVM classifier, the model was trained using the defined training samples, resulting in a classified raster where each pixel was assigned to one of the predefined classes.

To refine the results, post-classification processing steps were undertaken, such as applying majority filtering to reduce noise and reclassifying combined or misclassified classes for enhanced clarity[4]. The classification output was evaluated for accuracy using a confusion matrix which compared the predicted classes with reference data to calculate key metrics like overall accuracy, kappa coefficient, user's accuracy and producer's accuracy[3]. This rigorous evaluation provided insights into the reliability of the classification results and highlighted areas for potential improvement.

The final outputs including the FCC (False Color Composite) maps and LULC maps, were generated, offering a detailed spatial representation of land use and land cover patterns in the Kaddam watershed. These outputs serve as a critical tool for understanding spatial and temporal land-use dynamics in the region. A flow chart summarizing the methodology is represented in figure 2.

Results and Discussion

For the present study area, multi-temporal satellite images from Landsat and IRSP6 (LISS3) were classified into six broad land use and land cover categories: Agricultural Land, Cropped Land, Forest, Barren Land, Water Bodies (e.g. tanks and rivers) and Built-up Areas. The Level I classification approach provided a clear distinction of these categories, leveraging the medium resolution of the satellite datasets. From table 3, The land use and land cover (LULC) details provided display changes over time from 1988 to 2022 across various categories.

The analysis of Land Use and Land Cover (LULC) changes in the Kaddam Watershed reveals dynamic trends across various categories, reflecting shifts in land management and environmental conditions over the years. Water bodies exhibited a gradual increase from 18.16 sq km in 1988 to 26.30 sq km in 2022, likely influenced by improved water management practices, reservoir construction and hydrological changes. In contrast, forest cover showed a consistent decline from 1,124.79 sq km in 1988 to 960.95 sq km by 2016, suggesting significant deforestation during this period. However, stabilization after 2016, with a slight increase to 961.77 sq km in 2022, indicates potential conservation efforts. Settlements expanded steadily from 30.03 sq km in 1988 to 34.01 sq km by 2022, highlighting urban growth and population increase. Meanwhile, river areas remained relatively stable, with minor variations due

https://doi.org/10.25303/186da0106

to natural river dynamics or advancements in mapping precision.

Agricultural land showed fluctuations, peaking at 995.36 sq km in 2011 before a sharp decline to 279.49 sq km in 2016, followed by a modest recovery to 289.71 sq km in 2022, possibly reflecting land-use conversions and reclamation efforts. Cropped land demonstrated a dramatic increase from 358.1 sq km in 1988 to 1,315.35 sq km by 2022, driven by intensified agricultural practices and irrigation improvements. Conversely, barren land, which increased initially, dropped significantly from 334.77 sq km in 2008 to just 28.42 sq km in 2022, indicating successful reclamation and repurposing of degraded lands. These results underscore the interplay between human activities, policy interventions and natural dynamics, shaping the LULC patterns within the Kaddam Watershed over time. Temporal and multiple prepared LULC maps of the study area are represented in the figure.3.

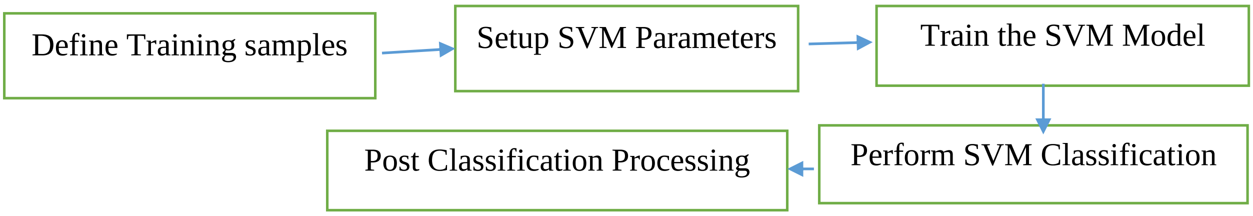

Figure 2: Flowchart of Methodology

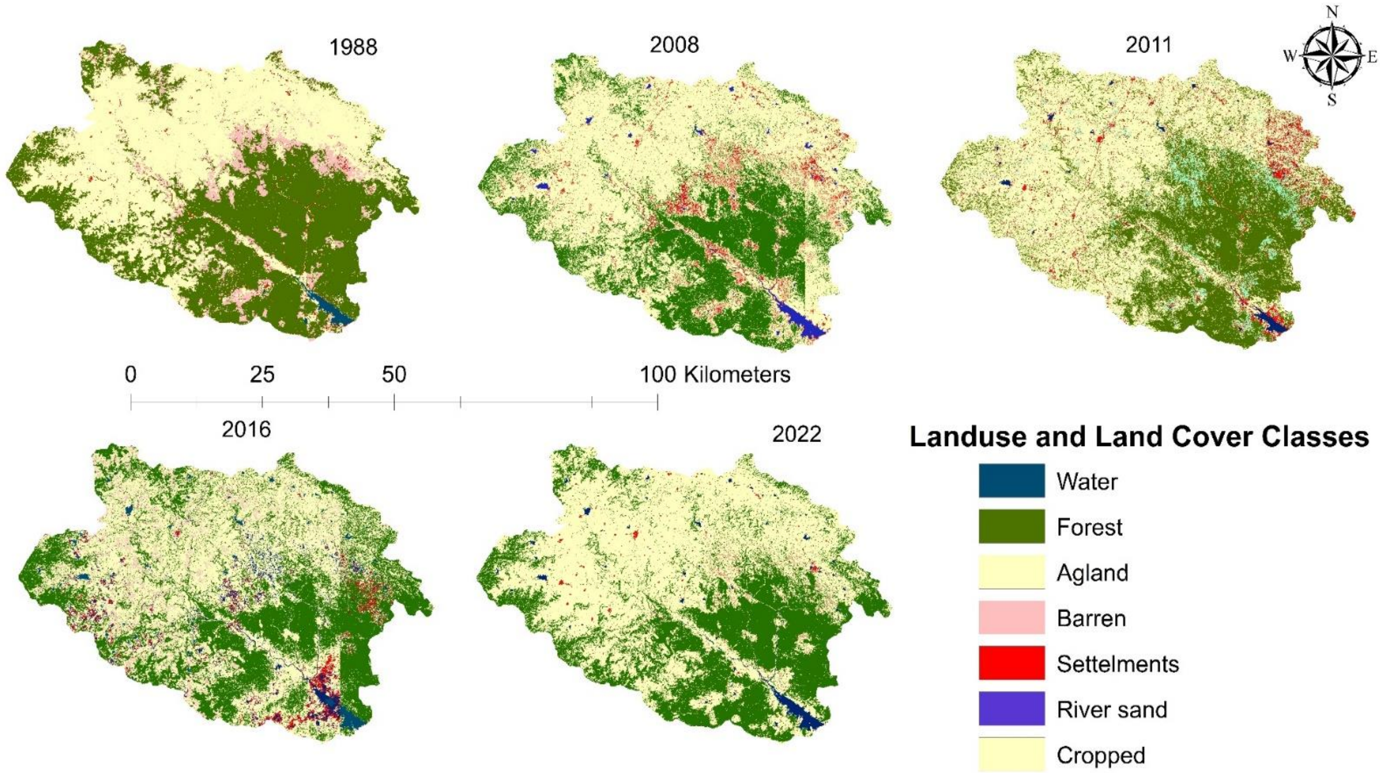

Figure 3: LULC Maps of the study area

Table 3
Land use and Land cover classes derived from satellite images using SVM technique

LULC class	Year-1988	Year-2008	Year-2011	Year-2016	Year-2022
Area in Sq kms					
Water	18.16	17.66	19.66	18.93	26.3
Forest	1124.79	995.35	973.65	960.95	961.77
Agricultural land	886.99	984.91	995.36	279.49	289.71
Barren	237.54	334.76	236.69	68.84	28.42
settlements	30.03	31.35	32.08	33.52	34.01
River	0.58	0.8	0.89	0.64	0.67
Cropped Land	358.1	291.41	398.56	1293.86	1315.35

https://doi.org/10.25303/186da0106

Error Matrix

The classification analysis across multiple years (2008, 2011, 2016 and 2022) highlights the performance of land cover mapping using remote sensing data, with high accuracies and notable insights into classification challenges. In 2022, agricultural land exhibited excellent classification performance, achieving a user accuracy of 96.4% and a producer accuracy of 97.2%. Cropped land stood out with a perfect classification accuracy of 100%, underscoring the effectiveness of the classifier for this category. Barren land also demonstrated strong results, with user and producer accuracies of 96% and 90%, respectively. However, forest and river classes faced some classification challenges. Forest user accuracy dropped to 80%, despite maintaining a producer accuracy of 100%, while the river class showed moderate confusion with barren and forest classes, achieving a user accuracy of 85.2%.

Settlements and water bodies revealed lower accuracies, both at 70%, due to minor overlaps with other categories. The overall Kappa coefficient of 0.903 reflects a strong agreement between classification and reference data, despite some category-specific challenges. In 2016, water bodies were perfectly classified with 100% user and producer accuracies, demonstrating the classifier's robustness in this category. Similarly, forest and cropped land maintained strong accuracies, with user and producer accuracies above 94%. However, the river class showed significant confusion, with a user accuracy of 35.3% and a producer accuracy of 66.7%, indicating challenges in distinguishing this class from others.

Agricultural land, barren land and settlements maintained high classification accuracies above 90%, reinforcing the reliability of the results for these categories. The Kappa coefficient for 2016, at 0.905, highlights consistent classification reliability across most categories, with room for improvement in the river class. Overall, the results from 2008 to 2022 demonstrate robust classification accuracy for most land cover types, with high Kappa coefficients (ranging from 0.898 to 0.938) reflecting strong agreement. Challenges remain for certain classes like river and settlements, likely due to spectral overlaps, but the findings underscore the efficacy of the classification methodology.

Conclusion

The study highlights the effectiveness of Remote Sensing (RS) and GIS as essential tools for monitoring and managing Land Use and Land Cover (LULC) changes. By utilizing Support Vector Machine (SVM) classification, the analysis achieved high classification accuracies across multiple categories, outperforming traditional methods frequently cited in the literature, such as Maximum Likelihood Classifier (MLC) or Decision Tree (DT). SVM's ability to handle medium-resolution satellite datasets and high-dimensional feature spaces resulted in precise categorization of land cover types including agricultural land, forests, water bodies and urban settlements. The SVM-based classification method proved superior in addressing spectral overlaps and minimizing misclassification errors, particularly for challenging classes like water bodies and cropped lands.

For instance, the Kappa coefficients consistently remained above 0.9 across all study years, demonstrating strong agreement between classified and reference data, a benchmark often unmatched by alternative classifiers. Furthermore, the study revealed critical insights such as forest stabilization post-2016, significant expansion of water bodies and urban growth trends, which align with findings from prior research emphasizing the adaptability and accuracy of SVM in diverse geographic contexts. These results underscore the potential of SVM classification as a reliable approach for supporting sustainable land-use planning and ecological management, contributing to informed decision-making and conservation efforts.

References

1. Anil K., Sivaprakasam S. and Sridhar P., Prioritization of Watersheds Using GIS and FAHP (Fuzzy Analytical Hierarchy) Method, Intelligent Computing and Applications, Smart Innovation, Systems and Technologies, **315**, https://doi.org/10.1007/978-981-19-4162-7_15 **(2022)**

2. Anil K., Sivaprakasam S. and Sridhar P., Soil Loss Evaluation in Kaddam Watershed Using Geographical Information Systems and Remote Sensing Techniques, *J. Inst. Eng. India*, Ser. A, https://doi.org/10.1007/s40030-023-00753-6 **(2023)**

3. Congalton R.G., A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, **37(1)**, 35-46 **(1991)**

4. Foody G.M., Status of land cover classification accuracy assessment, *Remote Sensing of Environment*, **80(1)**, 185-201 **(2002)**

5. Huang C., Davis L.S. and Townshend J.R.G., An assessment of support vector machines for land cover classification, *International Journal of Remote Sensing*, **23(4)**, 725-749 **(2002)**

6. Huete A.R., A soil-adjusted vegetation index (SAVI), *Remote Sensing of Environment*, **25(3)**, 295-309 **(1998)**

7. Jensen J.R., Introductory Digital Image Processing: A Remote Sensing Perspective, Upper Saddle River, NJ, Prentice Hall **(2005)**

8. Lillesand T., Kiefer R.W. and Chipman J., Remote Sensing and Image Interpretation, Hoboken, NJ, Wiley **(2015)**

9. Lu D. and Weng Q., A survey of image classification methods and techniques for improving classification performance, *International Journal of Remote Sensing*, **28(5)**, 823-870 **(2007)**

10. McFeeters S.K., The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features, *International Journal of Remote Sensing*, **17(7)**, 1425-1432 **(1996)**

11. Mountrakis G., Im J. and Ogole C., Support vector machines in remote sensing: A review, *ISPRS Journal of Photogrammetry and Remote Sensing*, **66(3)**, 247-259 **(2011)**

12. Pal M. and Mather P.M., Support vector machines for classification in remote sensing, *International Journal of Remote Sensing*, **26(5)**, 1007-1011 **(2005)**

13. Parupalli S., Padma Kumari K. and Ganapuram S., Assessment and planning for integrated river basin management using remote sensing, SWAT model and morphometric analysis (case study: Kaddam river basin, India), *Geocarto International*, **34(12)**, 1332–1362, https://doi.org/10.1080/10106049.2018.1489420 **(2018)**

14. Sridhar P. and Ganapuram S., Morphometric analysis using fuzzy analytical hierarchy process (FAHP) and geographic information systems (GIS) for the prioritization of watersheds, *Arab J Geosci.*, **14**, 236, https://doi.org/10.1007/s12517-021-06539-z **(2021)**

15. Tucker C.J., Red and photographic infrared linear combinations for monitoring vegetation, *Remote Sensing of Environment*, **8(2)**, 127-150 **(1979)**

16. Xie Y., Sha Z. and Yu M., Remote sensing imagery in vegetation mapping: a review, *Journal of Plant Ecology*, **1(1)**, 9-23 **(2008)**.

(Received 06th January 2025, accepted 07th March 2025)