Assessment of Land Use Land Cover Classification using Support Vector Machine and Random Forest Techniques in the Agusan River Basin through Geospatial Techniques

Juven G. Luzorata
Department of Agricultural and Biosystems Engineering, Center for Resource Assessment Analytics and Emerging Technologies (CREATE), College of Engineering and Geosciences

Aljon E. Bocobo
Department of Agricultural and Biosystems Engineering, Center for Resource Assessment Analytics and Emerging Technologies (CREATE), College of Engineering and Geosciences

Liahona M. Detera
Caraga State University

Neki Jel B. Pocong
Caraga State University

他

https://doi.org/10.5109/7157978

出版情報：Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES). 9, pp.240-246, 2023-10-19. 九州大学大学院総合理工学府
バージョン：

Assessment of Land Use Land Cover Classification using Support Vector Machine and Random Forest Techniques in the Agusan River Basin through Geospatial Techniques

Juven G. Luzorata[1,2*] , Aljon E. Bocobo[1,2], Liahona M. Detera[2], Neki Jel B. Pocong[2],Anamarie P. Sajonia[1,2]
[1]Department of Agricultural and Biosystems Engineering, Center for Resource Assessment Analytics and Emerging Technologies (CREATE), College of Engineering and Geosciences
[2]Caraga State University, Ampayon, Butuan City, 8600, Philippines
Corresponding e-mail: jgluzorata@carsu.edu.ph

Abstract: *The Agusan River basin is a lifeline for residents in Agusan del Norte, Agusan del Sur, and Davao del Norte. However, human activities have caused water contamination and siltation, leading to significant structural and physical changes in the river. This study utilized two machine learning classifiers, Support Vector Machine (SVM) and Random Forest (RF), within the Google Earth Engine (GEE) platform to assess the land use and land cover (LULC) changes from 2000 to 2020. The results unequivocally favored SVM, with higher accuracies of 95.53%, 95.61%, and 92.21% in 2000, 2010, and 2020, respectively. Notably, the study unveiled the substantial impact of LULC changes on critical water quality parameters, including turbidity, total suspended solids, and pH. These findings bear profound implications for the conservation and management of the Agusan River Basin, providing policymakers with invaluable insights for crafting interventions to preserve this invaluable natural resource.*

Keywords: Land Use Land Cover, Google Earth Engine, SVM, RF, Water Quality

1. INTRODUCTION

Land cover describes the physical properties of the earth's surface. The Earth's surface is not uniform at all; there are differences such as water, bare ground, trees, grass, asphalt, and more. Land use refers to how people use the land for social and economic purposes. Urban and agricultural sectors are the two land use types that are the most well-known. The land cover represents all kinds of differences in the Earth's surface in a broad sense.

Land use and land cover (LULCC) change refers to human activities that influence hydrological processes [1] [2]. Natural events like storms, forest fires, and landslides, as well as human-induced factors such as deforestation, climate change, and stochastic events can drive LULC change [3] [4]. LULC change can result in increased storm runoff, reduced vegetation cover, and sediment transport, leading to water quality degradation. Effective management of land use and land cover is crucial in mitigating land degradation, climate change, and extreme rainfall events [4] [5]. Changes in land cover impact the atmosphere, climate, and biology of the Earth, affecting flooding, sedimentation, and stream habitats [6] [7] [8]. Conversion of forests, wetlands, and agricultural land into impervious urban surfaces yields economic benefits but poses environmental costs, increasing runoff and nonpoint source pollution [9] [10].

Remote sensing data has played a significant role in monitoring land use and land cover (LULC) changes. These data provide valuable information about the Earth's surface in terms of spatial, spectral, and temporal resolutions. Numerous studies have investigated the impact of land development on floods using remote sensing [11]. Specifically, the use of Landsat images has facilitated the precise identification and mapping of LULC change on a large scale [12]. Various methods and datasets, such as supervised classification, PCA, hybrid classification, unsupervised classification or clustering, and different classifiers, have been employed for change detection analysis using remote sensing imagery [13] [14] [15] [16]. Among these methods, supervised classification techniques are commonly considered the most effective, although ongoing discussions persist regarding their respective advantages and disadvantages. Among these classifiers are SVM and RF which are machine learning techniques used for classification and prediction.

In the early 1970s, remote sensing techniques were introduced for monitoring water quality. Water bodies across the landscape are impacted by various factors such as suspended sediments (turbidity), algae (chlorophylls and carotenoids), chemicals (nutrients, pesticides, and metals), dissolved organic matter (DOM), thermal releases, aquatic vascular plants, pathogens, and oils. For this particular study, the focus was on assessing water quality based on three parameters - pH, turbidity, and Total Suspended Solids (TSS). Geospatial technology, such as remote sensing and GIS, enables the analysis and visualization of LULC changes and their impact on water quality indicators, advancing our understanding of this relationship [10].

Guppy and Anderson [17] emphasized the importance of water for sustaining life, and Kummu [18] highlighted the crucial role of freshwater resources from rivers in supporting various aspects of everyday life. The increase in water and resource availability and demand, driven by the growing global population, has significantly impacted LULC change in rivers, negatively affecting their physical properties and ecology [19] [20]. Numerous studies have examined the effects of LULC changes on river basins, including impacts on runoff, discharge, water yield, headwater fluvial, morphology, structure, and dam construction [21] [22] [23] [24] [25]. However, the effectiveness of the different models has not been thoroughly compared.

This study aimed to evaluate land use and land cover (LULC) classification using RF and SVM machine learning algorithms within the Google Earth Engine (GEE) platform. The focus was on assessing how LULC changes in the Agusan River may have influenced river water quality and developing effective management practices addressing both concerns. The accuracy of both methods was compared to determine the most precise classifier, followed by mapping and analyzing LULC changes in the study area from 2000, 2010, and 2020 to identify significant patterns impacting water quality and inform management strategies.

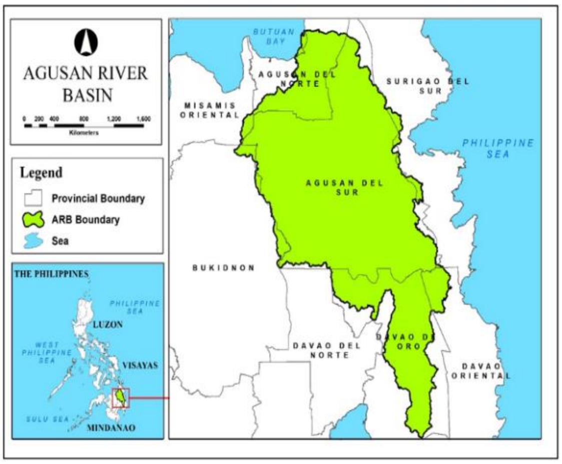

Figure 1. Agusan River Basin – Study Map Area

2. METHODOLOGY

The study utilized the Landsat satellite dataset integrated into the Google Earth Engine image collections, with Landsat-7 for 2000-2012 and Landsat-8 for 2013-2020, to generate LULC maps for 2000, 2010, and 2020. Using SVM and RF classifiers, the study aimed to simulate and evaluate the relationship between water quality parameters such as pH, Turbidity, and Total Suspended Solids (TSS) and LULC changes between 2000-2020, employing the IDW interpolation method for surface mapping. The IDW method was selected for predicting values within the range of observed data, and zonal statistical tools were employed to assess the influence of land use/land cover on water quality by calculating zone values based on another dataset. Water quality and LULC values were obtained through point feature conversion and the "zonal statistics as table" tool [26] [27].

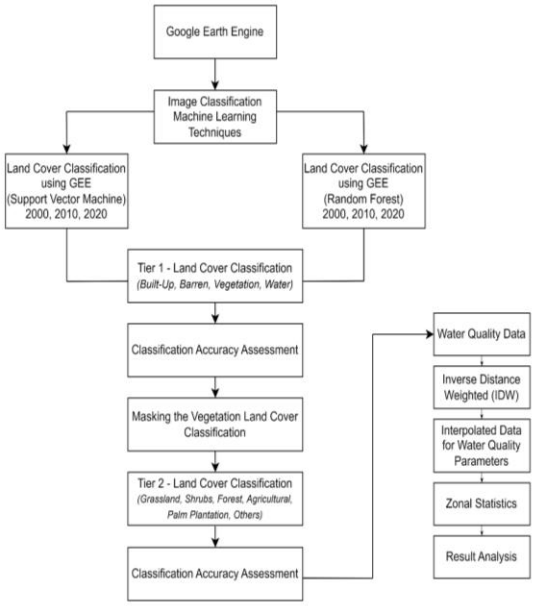

Figure 2. The General Framework of the Study

Google Earth Engine (GEE) was utilized to efficiently classify a vast amount of remote sensing data. The Landsat images from 2000, 2010, and 2020 were accurately classified using SVM and RF algorithms. To ensure data reliability, the accuracy of the classifications was thoroughly checked. Water quality data, specifically pH, turbidity, and total suspended solids (TSS), were collected from the Environmental Management Bureau of the Department of Environmental and Natural Resources (DENR). Additionally, an assessment was conducted to determine the impacts of LULC change on these parameters.

2.1 LULC Sampling Points

In this study, the training and validation points were selected using a technique previously applied and theorized by Kanniah [28] and Lin [29] which involves selecting sample points from high-resolution Google Earth photos. The selection of these sample points was performed using the time-slider feature of Google Earth, ensuring that the selected points were within the study area. To ensure accuracy, ground truth samples were collected using Google Earth imagery, Landsat data composites, and expert knowledge. The sample points were randomly divided into two categories for the classification process: training samples (70%) and validation samples (30%). This approach ensured that the classification was carried out with high accuracy and consistency from both machine learning classifiers.

2.2 Data Collections

Table 1. Data used in the Study

Datasets	Source
Agusan River Basin Boundary	CSU-CreATe (Center Resource Assessment-analytics and Emerging Technologies)
Landsat Images	USGS Earth Explorer (https://earthexplorer.usgs.gov/)
Water Quality Data	DENR – Environmental Management Bureau, Butuan City 8600

2.3 Land Cover Classification using GEE

The Google Earth Engine (GEE) offers convenient online access to a range of research satellite data, including fields operations data, historical land use and land cover (LULC) datasets, and aerial photographs. GEE enables users to access archived satellite images, such as Landsat data, which are provided as a collection by the United States Geological Survey. This eliminates the need to download the satellite images, simplifying the data acquisition process [30].

In this study, Landsat image composites were created every 10 years from 2000 to 2020 using the Google Earth Engine (GEE) platform. Machine learning classifiers, SVM and RF, were employed for accurate land use and land cover (LULC) classification. Ground truth samples were collected using historical imageries from Google Earth and expert knowledge. The classification process aimed to identify four main classes (built-up, barren, vegetation, and water) and several subclasses within the vegetation class (grassland, shrubs, forest, agricultural, palm plantation, and others). The accuracy of the classifications was evaluated using a Confusion Matrix. Landsat-7 ETM and Landsat-8 OLI sensors were used for generating the image composites. The GEE platform facilitated efficient data collection, classification, and analysis, making it a valuable tool for producing accurate LULC maps.

2.4 Spatial Analysis for Water Quality Parameters

To generate surface maps for water quality parameters, the study employed the Inverse Distance Weighted (IDW) interpolation method using ArcGIS 10.8. IDW is a commonly used algorithm for spatially interpolating point data, allowing estimation of values beyond the sampled points [26]. This method assumes that each measurement point has a localized influence that diminishes with distance, with the strongest influences occurring near the observed point [27]. The IDW method was selected because it enables the prediction of values for unsampled locations within the range of observed data. Zonal statistical tools in ArcGIS 10.8 were also utilized to assess the impact of land use and land cover on water quality. These statistical tools calculated zone values based on data from another dataset. To obtain water quality and land use/land cover values in tabular form, the centroid of the union councils was converted into a point feature, and the "zonal statistics as table" tool in ArcGIS 10.8 was employed.

3. RESULTS AND DISCUSSION

3.1 Base Land Cover Map

Using Landsat 7 mages for the years 2000-2012 and Landsat 8 images for the years 2013-2020 with a 30-meter resolution and a supervised SVM classifier and RF classifier, the land cover map of Agusan River Basin (ARB) was generated, as shown in Figure 3. The eight classes were Agricultural (sage dust), Barren (cocoa brown), Built-Up (mars red), Forest (leaf green), Grassland (lemongrass), Palm Plantation (autunite yellow), Shrubs (light apple), and Water (big sky blue). Its total land area was recorded to be 1210948.83 hectares. The three land covers for the years 2000, 2010, and 2020 were the generated LULC using Google Earth Engine (GEE) and ArcGIS software.

Land use and cover maps generated by RF and SVM classifiers show notable changes over time. SVM reveals an increase in agricultural land cover, a decrease in barren and shrubland areas, a rise in palm plantations, and a slight growth in built-up areas. RF, on the other hand, shows a decline in agricultural and forest land cover, an increase in palm plantations, and a slight increase in built-up areas. These changes are influenced by natural factors and human activities such as urbanization, industrialization, and the global demand for palm oil.

3.2 Summary of Overall Accuracy Assessment

SVM algorithm achieved higher overall accuracy (92.21% to 95.61%) compared to Random Forest (RF) for land cover classification. SVM showed promise in accurately classifying Forest and Palm Plantation, while accuracy for Grassland and Agricultural classes was lower. RF demonstrated good overall accuracy (93.94% to 94.7%) but accuracy varies among land cover classes, emphasizing the need for improvements. Considering the user's Accuracy is crucial for identifying classes that require classification enhancement. Forest had the highest accuracy, while Palm Plantation and Agricultural classes showed lower accuracy.

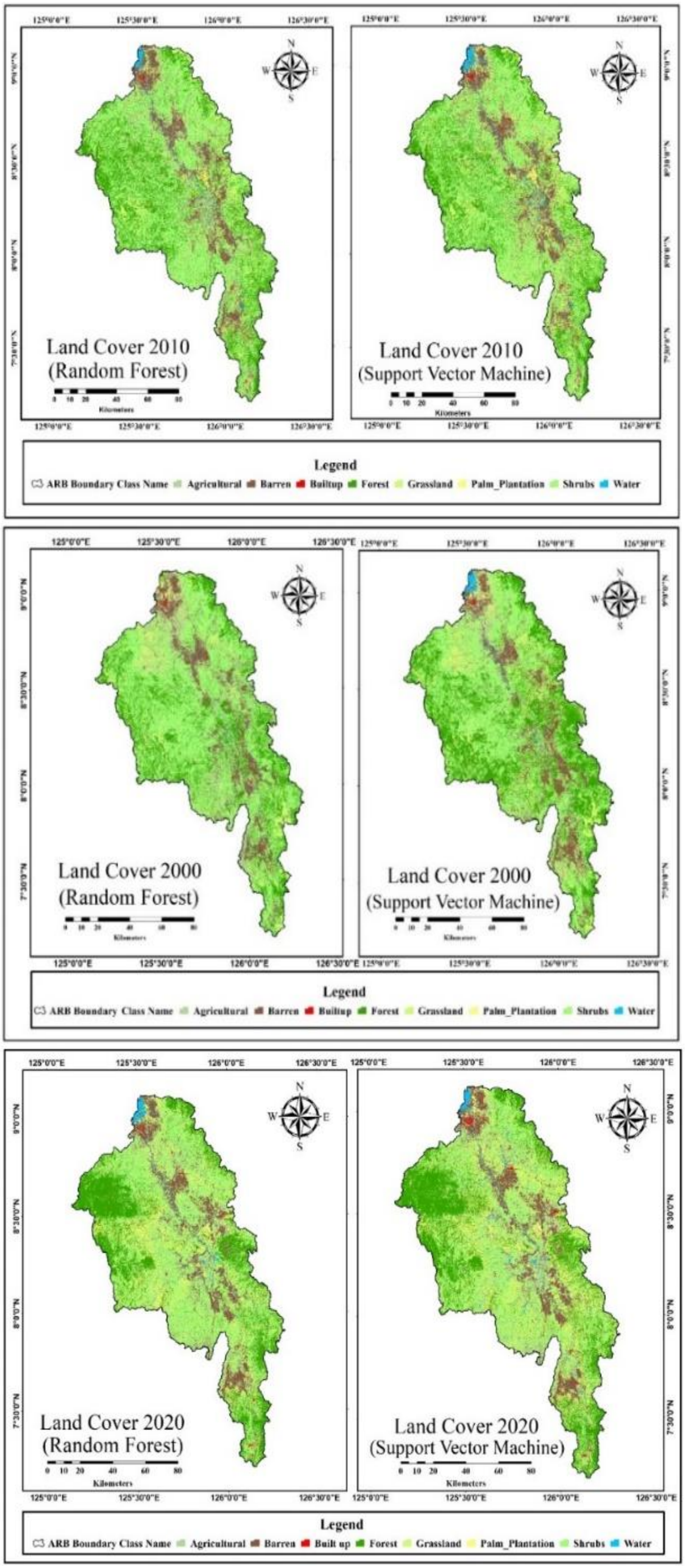

Figure 3. Agusan River Basin Land Cover Map for the years 2000, 2010 and 2020

Improvements are needed for specific classes in both algorithms.

3.3 LULC Change Detection Analysis

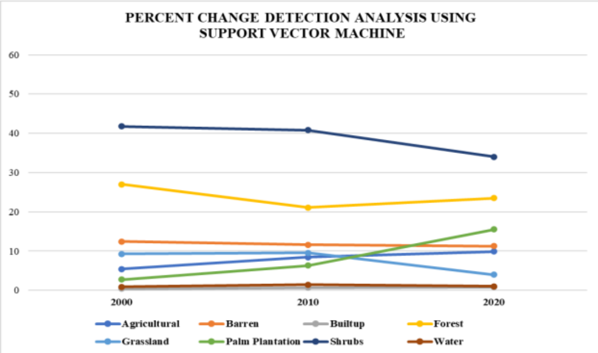

Figure 4. Change Detection Analysis using SVM

Between 2000 and 2020, notable land cover changes occurred in the study area. Shrubs were the largest land cover class in 2000, but by 2010, Agricultural, Palm Plantation, and Built-Up areas increased, while Barren, Forest, and Shrubs decreased. The most significant changes were observed in Agricultural, Palm Plantation, and Grassland. Palm Plantation experienced the largest increase, while Forest and Grassland decreased. Overall, agricultural land expanded, built-up areas increased gradually, and forest cover declined due to potential deforestation. Palm Plantations significantly increased, while shrubs decreased gradually.

Land cover changes in the study area between 2000, 2010, and 2020 indicate fluctuations in forest and shrub cover. Forest cover increased from 19.12% to 22.81% in 2010 but decreased to 21.57% in 2020. Shrubs decreased from 50.11% to 43.13% in 2010 and further to 41.44% in 2020. Agricultural land decreased, while built-up areas and barren land showed slight changes. Grassland decreased significantly, and palm plantation areas expanded. The chart highlights land cover percentage changes over a ten-year interval. These changes reflect natural growth, urbanization, and industrialization, providing insights into land use dynamics in the region.

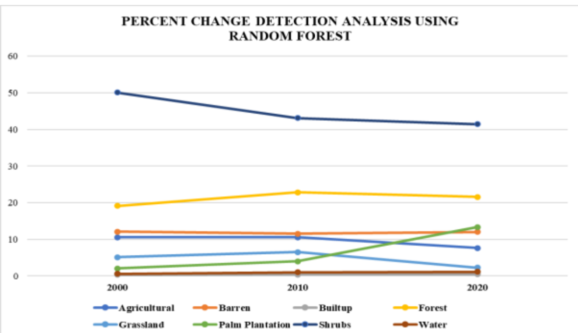

Figure 5. Change Detection Analysis using RF

3.4 Impact of LULC Change on Water Quality Parameters

The pH level of the Agusan River in the Agusan River Basin showed a slight increase from 2010 to 2020, remaining within a stable and neutral range. Despite significant land cover changes, including urban expansion and shifts in vegetation types, the river's pH has maintained stability. Agricultural activities and their associated inputs can impact water pH, but the river system demonstrates resilience and buffering capacity, mitigating these effects to some extent.

Turbidity in the Agusan River Basin consistently decreased from 2010 to 2020, despite significant land use and land cover changes. The decrease in vegetation cover, coupled with sustainable land management practices like erosion control and agroforestry, contributed to lower turbidity levels. The increase in forest and palm plantation land cover likely played a role in preserving riparian buffer zones, which enhance water quality. Vegetation and sustainable practices are crucial in mitigating turbidity in the Agusan River Basin.

Table 2. Impacts of LULC in pH

Land Cover	2010		2020	
	Area (ha)	pH	Area (ha)	pH
Agricultural	3,329,100	7.27	4,530,600	7.54
Barren	4,630,500	7.28	1,578,600	7.52
Built-up	280,800	7.30	3,861,000	7.56
Forest	8,443,800	7.26	13,566,600	7.54
Grassland	3,824,100	7.27	328,500	7.55
Palm Plantation	2,566,800	7.26	6,183,900	7.54
Shrubs	16,166,700	7.27	392,400	7.49
Water	582,300	6.96	9,382,500	7.52

Table 3. Impacts of LULC in Turbidity

Land Cover	2010		2020	
	Area (ha)	Turbidity	Area (ha)	Turbidity
Agricultural	3,329,100	227.72	4,530,600	211.39
Barren	4,630,500	235.00	1,578,600	202.36
Built-up	280,800	225.44	3,861,000	218.93
Forest	8,443,800	221.20	13,566,600	207.37
Grassland	3,824,100	228.24	328,500	188.22
Palm Plantation	2,566,800	231.97	6,183,900	209.98
Shrubs	16,166,700	225.10	392,400	186.15
Water	582,300	185.40	9,382,500	194.49

Land use and land cover (LULC) changes have significantly contributed to the decrease in total suspended solids (TSS) levels in the Agusan River. The expansion of agricultural, built-up, forest, and palm plantation areas, coupled with improved land management practices, has effectively reduced soil erosion and sediment runoff, leading to lower TSS levels. Although certain land cover types have decreased, their past contributions in retaining soil particles and stabilizing the landscape are still evident. Sustainable agricultural practices and effective stormwater management systems have also played a role in minimizing erosion and sedimentation, further reducing sediment reaching the river and contributing to the decrease in TSS levels.

4. CONCLUSION AND RECOMMENDATIONS

This study utilized Support Vector Machine (SVM) and Random Forest (RF) classifiers to assess changes in Land Use Land Cover (LULC) in the Agusan River Basin from 2000-2020. SVM showed effectiveness in accurately classifying narrow rivers, while RF may struggle with such classification due to potential misclassification. LULC changes were found to have significant impacts on water quality parameters such as turbidity, Total Suspended Solids (TSS), and pH. Ongoing monitoring and management strategies are crucial for preserving water quality in the basin. interventions to mitigate human impact, regulate land use, restore degraded areas, and promote sustainable livelihoods for local communities. These findings have important implications for the conservation and management of the Agusan River Basin. Additionally, future research should explore improved classification algorithms, involve community engagement, and address climate change adaptation within the context of conservation and management efforts in the basin.

REFERENCES

[1] J. Liu, T. Liu, A. Bao, P. Maeyer, X. Feng, E. Carreira and X. Chen, "Assessment of Different Modelling Studies on the Spatial Hydrological Processess in an Arid Alpine Catchment," *Water Resources Management,* vol. 30, no. 5, pp. 1757-1770, 2016.

[2] S. Sharma, S. Chaudry and P. Attri, "Remote Sensing & GIS based Approaches for LULC Change Detection," *International Journal of Current Engineering and Technology* , vol. 5, no. 5, 2015.

[3] P. K. Srivastra, D. Han, M. A. Rico- Ramirez, M. Bray and T. Islam, "Selection of classification techniques for land use/land cover change investigation," *Advances in Space Research,* vol. 50, no. 9, pp. 1250-1265, 2012.

[4] Holman- Dodds, K. A. Jennifer, A. Bradley and K. W. Potter, "Evaluation of Hydrologic Benefits of Infilltration Based Urban Storm Water Management," *Journal of the American Water Assessment (JAWRA),* vol. 39, no. 1, pp. 205-215, 2003.

[5] Z. M. Hassan, R. Shabbir, S. Ahmad, A. H. Malik, N. Aziz, A. S. Butt and S. Erum, "Dynamics of land use and land cover change (LULCC) using geospatial techniques: a case study of Islamabad Pakistan," *SpringerPlus,* vol. 5, no. 1, 2016.

[6] C. Lin, C.-C. Wu, K. Tsogt, Y.-C. Ouyang and C.-I. Chang, "Effects of atmospheric correction and pansharpening on LULC classification accuracy using WorldView-2 imagery," *Information Processing in Agriculture,* vol. 2, no. 1, pp. 25-36, 2015.

[7] R. M. Jacobson, S. R. Femmer and R. McKenney, "Land-use changes and the physical habitat of streams - a review with emphasis on studies within the U.S. Geological Survey Federal-State Cooperative Program," *U.S. Geological Survey circular,* 2001.

[8] G. Brierly, "Landscape memory: the imprint of the past on contemporary landscape forms and processes," *Area,* vol. 42, no. 1, pp. 76-85, 2010.

[9] J. E. Grove, A. Troy, J. O'Neil-Dunne, W. R. Burch, M. L. Cadenasso and S. T. A. Pickett, "Characterization of Household and its Implication for the Vegetation of Urban Ecosystems," *Ecosystems,* vol. 9, no. 4, pp. 578-597, 2006.

[10] Z. Tang, B. Engel, B. C. Pijanowski and K. E. Lim, "Forecasting land use change and its environmental impact at a watershed scale," *Journal of Environmental Management,* vol. 76, no. 1, pp. 35-45, 2005.

[11] N. S. Reynard, C. Prudhomme and S. M. Crooks, "Potential Effects of Changing Clmate and Land Use," *Climate Change,* vol. 48, pp. 343-359, 2001.

[12] M. C. Hansen and T. R. Loveland, "A review of a large area monitoring of land cover change using Landsat data," *Remote Sensing of Envirnment,* vol. 122, pp. 66-74, 2012.

[13] J. S. Rawat and M. Kumar, "Monitoring land use/cover change using remote sensing and GIS techniques: a case study of Hawalbagh block, district Almora, Uttarakhand, India," *The Egyptian Journal of Remote Sensing and Space Science,* vol. 18, no. 1, pp. 77-84, 2015.

[14] A. S. Butt, R. Shabbir, S. Ahmad and N. Aziz, "Land use change mapping and analysis using Remote Sensing and GIS: A case study of Simly watershed, Islamabad, Pakistan," *The*

Egyptian Journal of Remote Sensing and Space Science, vol. 18, no. 2, pp. 251-259, 2015.

[15] A. Mahmud and A. S. Achide, "Analysis of land use/land cover changes to monitor urban sprawl in Keffi-Nigeria," *Environmental Research Journa*, vol. 6, no. 2, pp. 129-134, 2012.

[16] N. M. Gazzaz, M. K. Yussof, M. F. Ramili, A. Z. Ariz and H. Juahir, "Characterization of spatial patterns in river water quality using chemometric pattern recognition techniques," *Marine Pollution Bulletin*, vol. 64, no. 4, pp. 688-698, 2012.

[17] L. Guppy and K. Anderson, "Water Crisis Report," United Nations University Institute for Water, 2017.

[18] M. Kummu, H. De Moel, P. Ward and O. Varis, "How Close Do We Live to Water? A Global Analysis of Population Distance to Freshwater Bodies," *PLOS ONE*, vol. 6, no. 6, p. e2058, 2011.

[19] A. Chin, "Urban transformation of river landscapes in a global context," *Geomorphology*, vol. 79, no. 3-4, pp. 460-487, 2006.

[20] S. Zhao, C. Peng, H. Jiang, D. Tian, X. Lei and X. Zhou, "Land use change in Asia and the ecological consequences," *Ecological Research*, vol. 21, no. 6, pp. 890-896, 2006.

[21] F. Wang, G. Quansheng, Y. Qibiao, W. HuaXin and X. Xu, "Impacts of land-use and land-cover changes on river runoff in Yellow River basin for period of 1956–2012," *Chinese Geographical Science*, vol. 27, no. 1, pp. 13-24, 2017.

[22] P. Petchprayoon, P. D. Blanken, C. Ekkawatpanit and K. F. A. Hussein, "Hydrological impacts of land use/land cover change in a large river basin in central-northern Thailand," *International Journal of Climatology*, vol. 30, no. 13, pp. 1917-130, 2010.

[23] G. Xiaoli, X. Wang, H. Yan, Q. Zhang and G. Jin, "Land Use/Land Cover Change Induced Impacts on Water Supply Service in the Upper Reach of Heihe River Basin," *Sustainability*, vol. 7, no. 1, pp. 366-383, 2014.

[24] C. P. Harden, "Human impacts on headwater fluvial systems in the northern and central Andes," *Geomorphology*, vol. 79, no. 3-4, pp. 249-263, 2006.

[25] S. Yousefi, H. Moradi, S. Keesstra, S. Keesstra, O. Navratil and J. Hooke, "Effects of urbanization on river morphology of the Talar River, Mazandarn Province, Iran," *Geocarto International*, vol. 34, no. 3, pp. 276-292, 2019.

[26] S. Samanta, D. Pal, D. Lohar and B. Pal, "Interpolation of climate variables and temperature modeling," *Theoretical and Applied Climatology*, vol. 107, no. 1-2, pp. 35-45, 2012.

[27] B. Gossweiler, I. Wesström, I. Messing, A. M. Romero and A. Joel, "Spatial and Temporal Variations in Water Quality and Land Use in a Semi-Arid Catchment in Bolivia," *Water*, vol. 11, no. 11, p. 2227, 2019.

[28] K. D. Kanniah, "Quantifying green cover change for sustainable urban planning: A case of Kuala Lumpur, Malaysia," *Urban Forestry & Urban Greening*, vol. 27, pp. 287-304, 2017.

[29] X. Lin, M. Xu, C. Cao, R. Singh, W. Chen and H. Ju, "Land-use/land-cover changes and their influence on the ecosystem in Chengdu City, China during the Period of 1992-2018," *Sustainability*, vol. 10, no. 10, p. 3580, 2018.

[30] B. Feizizadeh, T. Blaschke, M. Gholamalifard and D. Tiede, "Change detection and analysis of land use/land cover using remote sensing and GIS in Iran," *Land Use and Land Cover Mapping in Europe*, pp. 129-143, 2021.

[31] N. S. Kudnar, "GIS-based assessment of morphological and hydrological parameters of Wainganga River Basin, Central India," *Modeling Earth Systems and Environment*, vol. 6, no. 3, pp. 1933-1950, 2020.