environmental sciences proceedings

MDPI

Proceeding Paper

Satellite-Based Analysis of Lake Okeechobee's Surface Water: Exploring Machine Learning Classification for Change Detection †

Madan Thapa Chhetri [1,*] and Sandip Rijal [2,*]

1 Florida Atlantic University
2 Florida Atlantic University
* Correspondence: mthapachhetr2023@fau.edu (M.T.C.); srijal2023@fau.edu (S.R.)
† Presented at the 5th International Electronic Conference on Remote Sensing, 7–21 November 2023; Available online: https://ecrs2023.sciforum.net/.

Abstract: Water is an essential resource for the survival of living beings. Remote sensing data serves as the best possible way to detect water bodies and monitor change over time. With a surplus amount of remote sensing data, machine learning approaches have become an effective and efficient way to detect and monitor surface water bodies. This research focused on utilizing remote sensing and machine learning approaches to monitor changes in the surface water of Lake Okeechobee, Florida, USA. This investigation used two remotely sensed data, Landsat 7, and Landsat 8, for 2002 and 2022, respectively. Two machine learning algorithms, Support vector Machine (SVM) and Random Forest (RF), were adopted considering their power and robustness, among others, for supervised classification. Both algorithms gave an accuracy of over 92% and a kappa statistic exceeding 0.8. Further, we used image differencing techniques to track changes across two decades. The SVM suggested an increase of 85 km^2, and RF indicated an expansion of 52 km^2 of surface water area. This study explicitly demonstrates how dynamic are the natural resources, especially water sources. Thus, it can be a foundation for research that explores further on environmental assessment and sustainable water resource planning in Lake Okeechobee.

Keywords: remote sensing; change detection; surface water extraction; machine learning; Lake Okeechobee

Citation: Chhetri, M.T.; Rijal, S. Satellite-Based Analysis of Lake Okeechobee's Surface Water: Exploring Machine Learning Classification for Change Detection. *Environ. Sci. Proc.* **2023**, *27*, x. https://doi.org/10.3390/xxxxx

Academic Editor(s): Name

Published: date

1. Introduction

Water bodies are essential for investigating hydrology, ecology, and ecosystem preservation [1]. Identifying water availability and the dynamics of its movement is important for effective water resource management. With advancements in remote sensing, detecting water bodies and assessing changes using satellite-based remote sensing imagery has emerged as a predominant method [2,3]. Change detection involves comparing data sets collected for the same area at different times to identify and understand the changes' nature, extent, and location. Human activities, sudden natural events, or long-term environmental factors can cause these changes. Studying the changing pattern is crucial for planning and formulating action plans, restoration programs, and other hydrological analyses.

With the remote sensing method, monitoring changes in water, land cover/use, vegetation, and forest has been widely used. In recent years, remote sensing data sources have become indispensable tools for comprehensively analyzing surface water dynamics [4–6]. The surface water change detection process typically involves extracting water features from multi-date satellite images, followed by comparative analysis to detect and characterize any observed changes [4,7]. Surface water detection from remote sensing imageries has several challenges. First, misclassifying shadows of hills, melting ice, trees, and wet barren sand as water is the biggest problem when applying the threshold approach [8]. In surface water extraction, classification techniques are generally favored for their higher accuracy when contrasted with single-band methods [4]. Talking about image classification—non-parametric methods, specifically SVM and RF, have gained popularity [9]. SVM and RF demonstrate robustness against noise and overtraining, highlighting their efficacy in handling unbalanced data [10]. These techniques have proven to be highly effective for various classification-related applications.

Lake Okeechobee Watershed Restoration Project is a part of the Comprehensive Everglades Restoration Plan (CERP) implemented from July 2016 to the present has been working to maintain the water flow across South Florida to preserve the ecosystems. Lake Okeechobee is a focal part of South Florida's water supply and flood control systems, requiring a keen understanding of its change dynamics. Over time, the lake has altered for several reasons, mainly caused by anthropogenic activities, flood control, and urbanization. Furthermore, this area is a refuge for many important flora and fauna of sub-tropical regions; thus, a comprehensive assessment of changes in surface water is necessary. This research aims to quantify surface water detection using remote sensing datasets from the Landsat image series (2002 and 2022) employing SVM and RF models. Finally, we will compare the change detection using different machine learning methods and suitable methods will be proposed to detect change in surface water.

2. Materials and Methods

This research utilized satellite imageries from Landsat-7 and 8 of 2002 and 2022, with a cloud cover of less than ten percentage. Data are freely available and downloaded from the United States Geological Survey (USGS Earth Explorer) website. The band specifications of both Landsat-7 and 8 can be found (source). The acquired satellite images were processed and georeferenced using ArcGIS Pro version 3.0.3 [11]. The downloaded layers were merged into a single layer using the Composite Bands tools in ArcGIS Pro for analysis. Subsequently, the image was clipped using the shapefile of Lake Okeechobee (Figure 1). After these preprocessing steps, spectral indices such as Normalized Difference Water Index (NDWI), Modified Normalized Difference Water Index (MNDWI), and Normalized Difference Vegetation Index (NDVI) were calculated and combined with the composite bands for better extraction of surface water. Random samples were collected from the Landsat image, superimposed over a resolution image [12] to collect samples with confidence for water and non-water.

Once training samples were generated for both the imageries, classification was done using classification tools based on SVM and RF models. For the SVM model, the maximum number of samples per class was 100 for classification. For the RF model, the number of trees and maximum tree depth was chosen as 30, and the maximum number of samples per class was selected as 50. After classification was performed for both the imageries, we created confusion matrices for accuracy assessment. Finally, we used a raster calculator for the image differencing approach to analyze the change between the two dated classified imageries.

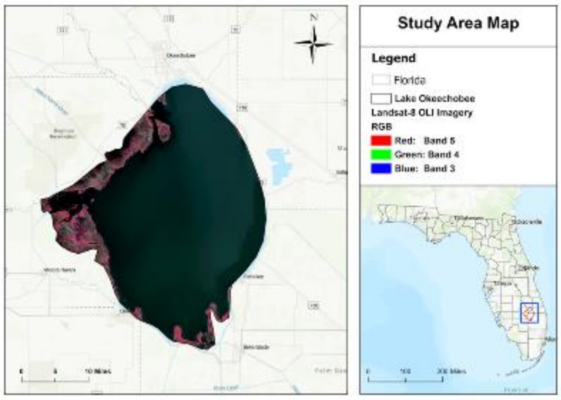

Figure 1. Location of Study Area.

3. Results

3.1. Lake Status

In 2002, within the study area of 1582.8 km², land cover classification using RF indicated a water area of 1292.2 km², while SVM classification estimated a water area of 1267.3 km². The corresponding land areas were 290.6 km² for RF and 315.5 km² for SVM.

Similarly, in 2022, the RF classification identified 1344 km² of water and 238.9 km² of land within the same study area. Meanwhile, the SVM classification classified 1352 km² area as water and 230.9 km² as land. These findings demonstrate the dynamic changes in land and water distribution within the study area over a span of two decades.

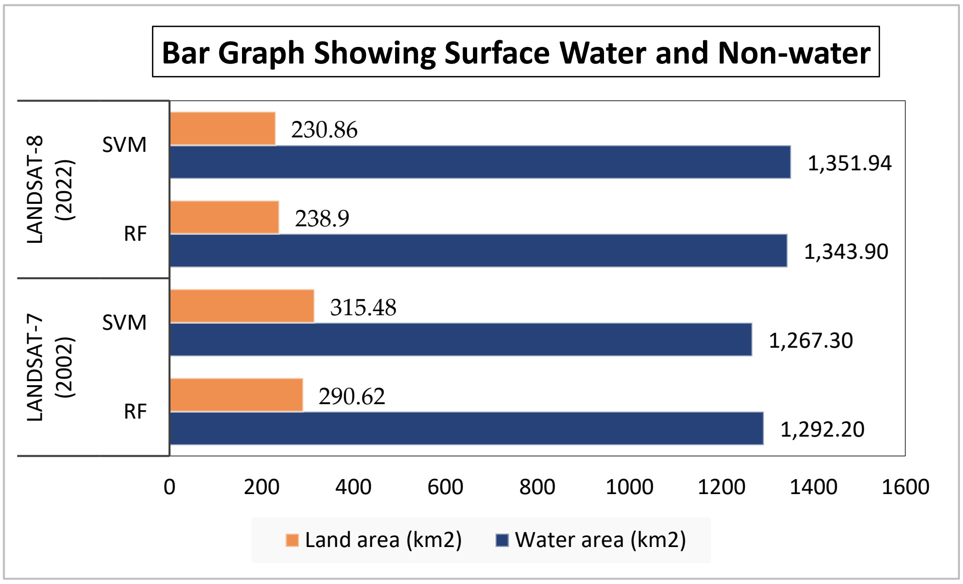

Figure 2. Status of lake water and land in 2002 and 2022 based on Machine learning classification approach.

3.2. Lake Dynamics

Our study generated a consistent increasing trend for surface water area. More specifically, SVM classification indicates an increase of surface water by 87.07 km², accompanied by a decrease of 2.28 km², while RF-based classification suggests an expansion of 60.49 km² and a reduction of 8.65 km².

Table 1. This is a table for changing trends on Lake Okeechobee based on Machine learning approaches.

Status/Methods	SVM (in km^2)	RF (km^2)
Decrease	2.28	8.65
Not Changed	1496.53	1516.74
Increase	87.07	60.49

In addition, the change in the surface water area of Lake Okeechobee is displayed in Figures 3 and 4 below. Here, water surface areas are shown in blue, land surface areas are shown in light grey, increased areas are in green, and decreased areas are in red. The grey color symbolizes no change in water and land surface area.

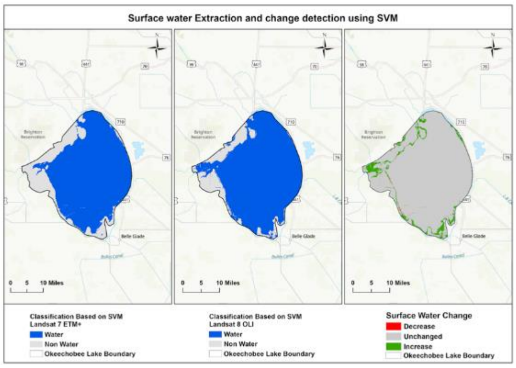

Figure 3. Surface water extraction and change detection analysis using SVM.

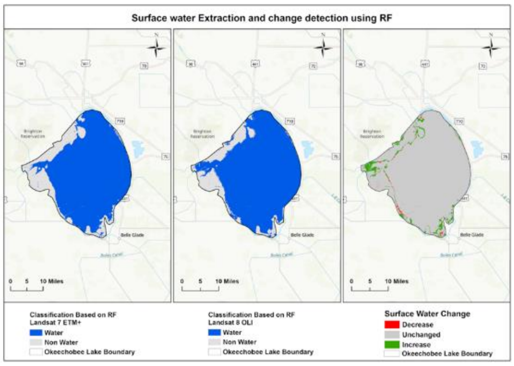

Figure 4. Surface water extraction and change detection analysis using RF.

3.3. Accuracy Assessment

The satellite-based classification and change detection study demonstrated the effectiveness of SVM and RF algorithms in classifying land cover. The classification was performed separately for Landsat-7 and 8 images, using the Machine Learning (SVM and RF) model with a total of 100 water and non-water training samples. The accuracy assessment of the classified images is shown in Table 2. For Landsat-7, SVM achieved an overall accuracy (OA) of 95%, while RF showed slightly higher performance with a 97% OA. Similarly, for Landsat-8, SVM outperformed RF with an impressive OA of 98%, while RF obtained a still commendable OA of 97%. Landsat-8 images yielded better results than Landsat-7 due to their lower cloud cover.

Table 2. This is a table for the accuracy assessment of the classification.

	Landsat-7		Landsat-8	
Parameters	**RF**	**SVM**	**RF**	**SVM**
Overall accuracy (%)	97	95	97	98
Kappa Statistics	0.89	0.83	0.87	0.94

Both SVM and RF proved to be reliable methods for satellite-based classification, providing accurate land cover categorization results in the analyzed imagery.

4. Discussion

Generally, lakes in subtropical regions face significant seasonal and inter-annual variations in rainfall and net evapotranspiration, resulting in water level variations [13]. Moreover, Lake Okeechobee is managed aggressively by lowering or raising the water and controlling in flow with dams and weir time and again. With the operation decision of the U.S. Army Corps of Engineers (USACE) and South Florida Water Management District (SFWMD), they controlled the water level and supply with the intention of maintaining water quality, ecosystem downstream, and flood control. There is Hober Hoover Dike with a levee, which controls high water by releasing it, before rainfall season starts. Less rainfall in 2000 and 2001, right after flowing water from the lake, could cause the water level to decrease significantly in Lake Okeechobee in the early 2000s [14].

Ref. [15] studied the impact of three hurricanes in 2004 and 2005 and found an increased water flow with increased rainfall, redistribution of bottom sediments, and more suspended materials in water. Throughout 2002 to 2022, this lake encountered multiple hurricanes like Hurricane Ian, Irma, Sally, etc. Hurricanes are responsible for rainfall and increased water inflow because the watershed area receives more rain, which is responsible for changing the lake's water surface area.

Evapotranspiration and rainfall act in opposite directions for any hydrological regime. Ref. [16] assume a 10% increase or decrease in rainfall and evapotranspiration to assess the condition of the water level and its impact on supporting vegetation by 2060. Unless in the extreme cases when rainfall decreases and evapotranspiration increases, there could be an unfavorable scenario for vegetation, or else two factors in Lake Okeechobee seem to have a compensating effect, i.e., increased evapotranspiration will be likely to be compensated with increased rainfall under climate change scenario [16]. It suggests that decreasing water surface water is less likely in our study area.

Hober Hoover Dike surrounds Lake Okeechobee with the intention of flood water management; it increases sedimentation compared to outflow from the region [16]. The same volume of water will have increased dispersion within the area because of the raised bottom in this shallow lake.

We understood that the increase in the surface area of Lake Okeechobee cannot be related to one factor as it has a multitude of impacts collectively acting upon it. Increased sedimentation could have raised the bottom of the lake so water spilled over the region [17]; increasing rainfall resulting in increased water levels with or without storms, anthropogenic water control system, etc., could have resulted in surface water dynamics.

This study aims to quantify the change in the surface area of Lake Okeechobee, which we observed through the increase and decrease in surface water area. Water surface area increased by 84.64 km^2 with SVM and 51.7 km^2 with RF within the span of 20 years. Lakes in subtropical regions experience fluctuations in water levels on both an inter-annual and interseasonal basis. Thus, we must consider all the tradeoffs while managing freshwater in Lake Okeechobee.

5. Conclusions

This research demonstrated the significance of combining remote sensing data and machine learning techniques to monitor changes in Lake Okeechobee. Both the RF and SVM work well, performing with an accuracy of over 92% during the classification of land cover in this research. So, it can be suggested that classification objectives are well met with those two machine learning algorithms and are great for flat ground. The active management of water sources and hurricanes are major factors for change in the water surface. Further, it provides valuable insights into the dynamics of water resources and lays the groundwork for future studies and informed environmental and water resource management practices.

Author Contributions: M.T.C. conceived and designed the paper. M.T.C. prepared and collected the data. M.T.C. and S.R. analyzed the data and results. M.T.C. wrote the first draft of the manuscript, and S.R. reviewed and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement:

Informed Consent Statement:

Data Availability Statement: Data will be available on request.

Acknowledgments: Landsat imagery was downloaded from the U.S. Geological Survey (USGS). The research analysis was performed at the Remote Sensing lab at Florida Atlantic University. We acknowledge Caiyun Zhang for her advice and two anonymous editors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ghasemigoudarzi, P.; Huang, W.; De Silva, O.; Yan, Q.; Power, D. A machine learning method for inland water detection using CYGNSS data. *IEEE Geosci. Remote Sens. Lett.* **2020**, *19*, 1–5.
2. Chandrasekar, K.; Sesha Sai, M.V.R.; Roy, P.S.; Dwevedi, R.S. Land Surface Water Index (LSWI) response to rainfall and NDVI using the MODIS Vegetation Index product. *Int. J. Remote Sens.* **2010**, *31*, 3987–4005.
3. Barrett, D.C.; Frazier, A.E. Automated method for monitoring water quality using Landsat imagery. *Water* **2016**, *8*, 257.
4. Zhang, Z.; Lu, H.; Zhao, M. Water body extraction and change detection based on multi-temporal SAR images. In *MIPPR 2009: Remote Sensing and GIS Data Processing and Other Applications*; SPIE: Bellingham, WA, USA, 2009; Volume 7498, pp. 725–731.
5. Tang, Z.; Ou, W.; Dai, Y.; Xin, Y. Extraction of water body based on LandSat TM5 imagery–a case study in the Yangtze River. In Proceedings of the Computer and Computing Technologies in Agriculture VI: 6th IFIP WG 5.14 International Conference, CCTA 2012, Zhangjiajie, China, 19–21 October 2012; Revised Selected Papers, Part II 6; Springer: Berlin/Heidelberg, Germany, 2013; pp. 416–420.
6. Li, W.; Du, Z.; Ling, F.; Zhou, D.; Wang, H.; Gui, Y.; Sun, B.; Zhang, X. A comparison of land surface water mapping using the normalized difference water index from TM, ETM+ and ALI. *Remote Sens.* **2013**, *5*, 5530–5549.
7. Du, Z.; Linghu, B.; Ling, F.; Li, W.; Tian, W.; Wang, H.; Gui, Y.; Sun, B.; Zhang, X. Estimating surface water area changes using time-series Landsat data in the qingjiang river basin, China. *J. Appl. Remote Sens.* **2012**, *6*, 063609.
8. Acharya, T.D.; Subedi, A.; Huang, H.; Lee, D.H. Classification of surface water using machine learning methods from Landsat data in Nepal. *Proceedings* **2018**, *4*, 43.
9. Adam, E.; Mutanga, O.; Odindi, J.; Abdel-Rahman, E.M. Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: Evaluating the performance of random forest and support vector machines classifiers. *Int. J. Remote Sens.* **2014**, *35*, 3440–3458.
10. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.
11. Overview of Georeferencing. Overview of georeferencing-ArcGIS Pro | Documentation. (n.d.). Available online: https://pro.arcgis.com/en/pro-app/latest/help/data/imagery/overview-of-georeferencing.htm (accessed on 10 September 2023).
12. Layer: World Imagery (ID: 0). (n.d.). Services.arcgisonline.com. Available online: https://services.arcgisonline.com/arcgis/rest/services/World_Imagery/MapServer/0 (accessed on 20 September 2023).
13. Abtew, W. Evaporation estimation for Lake Okeechobee in south Florida. *J. Irrig. Drain. Eng.* **2001**, *127*, 140–147.
14. Steinman, A.; Havens, K.; Hornung, L. The managed recession of Lake Okeechobee, Florida: Integrating science and natural resource management. *Conserv. Ecol.* **2002**, *6*.

15. James, T.R.; Chimney, M.J.; Sharfstein, B.; Engstrom, D.R.; Schottler, S.P.; East, T.; Jin, K.R. Hurricane effects on a shallow lake ecosystem, Lake Okeechobee, Florida (USA). *Fundam. Appl. Limnol.* **2008**, *172*, 273.
16. Havens, K.E.; Steinman, A.D. Ecological responses of a large shallow lake (Okeechobee, Florida) to climate change and potential future hydrologic regimes. *Environ. Manag.* **2015**, *55*, 763–775.
17. Gray, P.N.; Florida, A. A brief history of Lake Okeechobee ecosystem responses to water level management. *Water Manag.* **2017**.