environmental sciences proceedings

Proceeding Paper

Extraction of Surface Water Extent: Automated Thresholding Approaches†

Meghaa Sathish Kumar [1,2,*]

[1] Indian Institute of Remote Sensing, Dehradun, Uttarakhand– 248001, India
[2] ITC, University of Twente, 7522 NH Enschede, Netherlands
* Correspondence: meghaasathish@gmail.com
† Presented at the 5th International electronic Conference on Remote Sensing, 7 - 21 November 2023.

Abstract: Inland water bodies play a crucial role in both ecological and sociological contexts. The distribution of these water bodies can change over time due to natural or human-induced factors. Monitoring the extent of surface water is vital to understand extreme events such as floods and droughts. The availability of dense temporal Earth observation data from sensors like Landsat and Sentinel, coupled with advancements in cloud computing, has enabled the analysis of surface water extent over extended periods. In this study, automated thresholding approaches were applied within the Google Earth Engine platform to extract the surface water extent of the Chembarampakkam reservoir in Tamil Nadu, India. Sentinel-2 data spanning from 2019 to 2023 were used to derive two key indices, namely the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI). These indices were then thresholded to determine the presence of water. The performance of two different global thresholding techniques, namely Deterministic thresholding and Otsu thresholding method was compared to achieve better results. To enhance the accuracy of the deterministic technique, an iterative method was implemented. While the threshold values were generally similar for both techniques, the Otsu algorithm slightly outperformed the iterated deterministic technique in water classification. Furthermore, surface water dynamics image was obtained using temporal images, providing insights into the temporal surface dynamism of the reservoir. Overall, this study highlights the significance of surface water monitoring using remote sensing and cloud computing techniques.

Keywords: Google Earth Engine; Thresholding; Otsu; Determinant; Iteration; Surface Water Extent; Chembarampakkam.

Citation: To be added by editorial staff during production.

Academic Editor: Firstname Lastname

Published: date

1. Introduction

Inland water bodies are important both for their ecological and sociological significance. Fresh water bodies like rivers have played a major role in shaping the human settlements since their rich natural resource forms the basis of livelihood and it also provides a medium for transport and access [1]. Surface freshwater is the major source of water to meet agricultural, domestic and industrial water demands [2]. The expansion of human settlements into towns and cities and shift of the economy towards the industrial sector is resulting in poor management of the available freshwater, affecting its availability, quality, ecological balance and eventually affecting the marginal communities whose livelihoods are still dependent on them. Inland surface water bodies are very dynamic both temporally and spatially. Their distribution and course change over time due to natural or human-induced processes [3]. Monitoring their dynamic behavior is crucial for understanding the availability of water stocks and ensuring planned usage of available water resources. Surface water spread also helps in monitoring certain extreme events. For example, an excessive increase in the water spread could indicate the possibility of flooding, while, on the other hand, shrinking suggests the possibility of drought.

Surveying and documenting surface water spread is possible for a smaller spatial and temporal range, but it is not cost efficient, and is time consuming. Recent advances in remote sensing help overcome these challenges and allow us to perform huge computations over a larger spatial and temporal scale in a cost effective manner [4]. Availability of temporally dense earth observation data from sensors like Landsat, Sentinel etc. has made it possible to analyze surface water spread extent for a longer time period [5]. Today, cloud computation has allowed us to process these images in a matter of time and perform temporal analyses with decades of data. Advances in Image Analysis Techniques have led us to extract different levels of information from the imageries. One of these techniques, which suits well for surface water extent extraction is image segmentation of resultant water or vegetation indices images [6], and several previous studies have used the same to map the extent of surface water [7-9]. Image Segmentation is the technique of grouping regions in an image [10]. One of the famous image segmentation techniques is thresholding-based segmentation, where one or more threshold values are used to segment the image. Generally, thresholding is a technique in which the new value of a pixel in the segmented image is decided based on certain criteria set for the old value of the same pixel in the original image [11]. Thresholding-based segmentation is basically grouped into i) global thresholding, ii) local thresholding, and iii) Adaptive thresholding [12].

In this study, an attempt has been made to extract the monthly surface water extent of an inland reservoir using global thresholding in the Google Earth Engine. Generally, determinant and Otsu are two global thresholding techniques that are predominantly used in previous studies on surface water extraction [8][13]. Otsu is a cluster-based thresholding technique [14] where the image is segmented into two classes with a particular grey level as a threshold such that the classes have larger inter-class variance and lesser intra-class variance [11]. Deterministic thresholding uses a single threshold value for segmenting the image into two regions and particularly, using 0 as the deterministic threshold value for extracting water has been used often [6]. But this might lead to either over or under estimation of the water extent while the analysis is made for time-series data or for different water bodies from diverse geographical location [6]. Also, a study on Otsu thresholding [15] shows that, threshold value obtained from the Otsu algorithm is equal the mean value of the average values of both classes. Therefore, this study tries to use this averaging method to improvise the deterministic threshold value for an automated approach and compare it with the threshold generated by the Otsu algorithm. Finally, a layer which pictures the dynamics of the surface water in the reservoir is also produced using the monthly surface water extent layers obtained through thresholding.

2. Data and Method

2.1. Study Area

The Chembarampakkam reservoir is situated in the Adyar river watershed which is one of the many watersheds of the metropolitan city of Chennai in Tamil Nadu. Chennai is one of the most populated cities in the world and this reservoir plays a major role in meeting the city's water needs, importantly potable water needs [16]. The untimely opening of this reservoir due to inadequate monitoring is one of the major reasons that led to flooding in the downstream, which in turn resulted in the death of almost 500 people during the 2015 Chennai Floods [17]. This is the reason why this inland fresh water body was taken as a study area for this study.

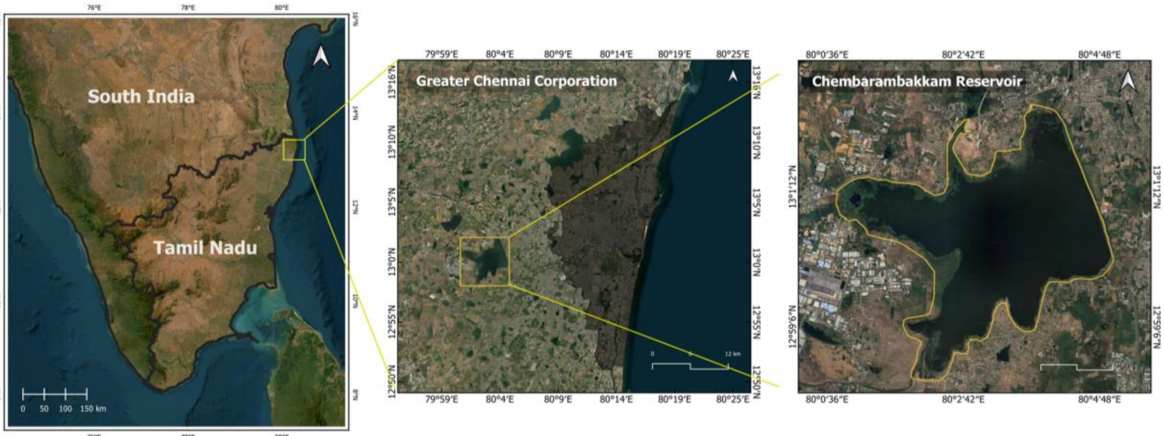


Figure 1. Map of the Chembarampakkam Reservoir situated in Greater Corporation of Chennai in Tamil Nadu .

2.2. Data, Platform and Pre-processing

The study was carried in the Google Earth Engine platform. The sentinel-2 image collection of Level 2 was used for deriving the indices. Initially, the dataset was filtered for cloud cover less than 40 percent. Then, monthly means were generated for the region of interest. The Normalized Difference Water Index (NDWI), and Normalized Difference Vegetation Index (NDVI) were then generated using the following equations (1) & (2) [18,19].

$$NDWI = (Green_{band} - NIR_{band})/(Green_{band} + NIR_{band}) \tag{1}$$

$$NDVI = (NIR_{band} - Red_{band})/(NIR_{band} + Red_{band}) \tag{2}$$

2.3. Methodology

2.3.1. Improvised Determinant Thresholding

Usually, in hydrological studies performed in GEE, 0 has mostly been used as a deterministic threshold value for extracting water extent [6]. However, for better extraction, threshold values can differ for different scenarios. In this study, the averaging technique discussed in [15], was iteratively used to arrive at an improved threshold value. Initially 0 was set as a threshold for the indices, which ranges between -1 to 1, where T is set to 0, and T is the threshold value. Then, the following equation was used to calculate the new threshold.

$$T_{new} = \frac{(Avg_{C1} + Avg_{C2})}{2} \tag{3}$$

Where C_1 and C_2 represent segmented classes with T = 0 as a threshold. Then iteratively new thresholds are derived until T_{old} is much smaller than T_{new} ($T_{old} << T_{new}$). Where T_{old} is the previously used threshold value.

2.3.2. Otsu Thresholding

In Otsu method, a specific gray value of the image is considered as a threshold in such a way the 2 classes segmented will have an increased inter-class variance. The gray value t is obtained in such a way, BSS of the two classes are maximum [20]. Where BSS is 'Between Sum of Squares' and is calculated as,

$$BSS = \sum_{i=1}^{n}(\overline{DN}_i - \overline{DN})^2 \quad (4)$$

Where, i is the number of classes and n is 2 in this case. $\overline{DN}_i$ is the mean value of Digital Numbers in the particular class i and $\overline{DN}$ is the mean value of the overall image.

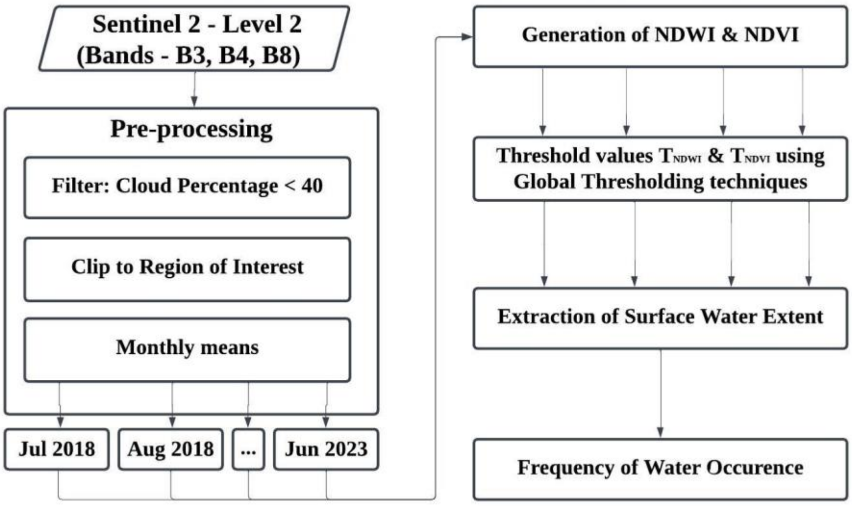


Figure 2. Methodology followed in this study.

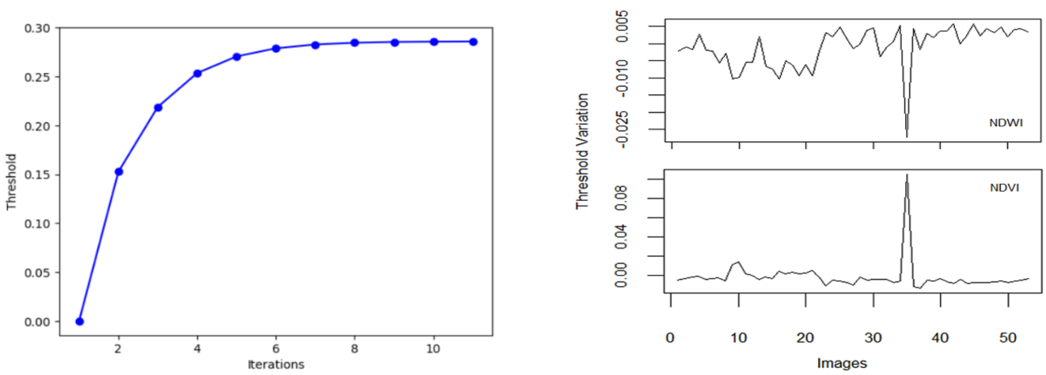
2.3.3. Extraction of Surface Water Extent and Dynamism

Once the thresholds are estimated for both NDWI and NDVI, the surface extent of water is found such that NDWI > T_{NDWI} and NDVI < T_{NDVI} [8]. Where T_{NDWI} and T_{NDVI} are the thresholds of water and vegetation indices respectively. In this way, the surface area where water is present and vegetation is not present is extracted.

For generating the water dynamics layer, the frequency of water occurrence in the surface area is estimated by dividing the number of months water was present in the pixel by the total number of months [8].

3. Results & Discussions

A total of 344 Sentinel 2 Level 2 images for the region of interest, from 2019 January to 2023 June, were collected and processed in the study. In GEE, the process took a matter of seconds and gave the final surface water dynamics layer (figure 4(b)).

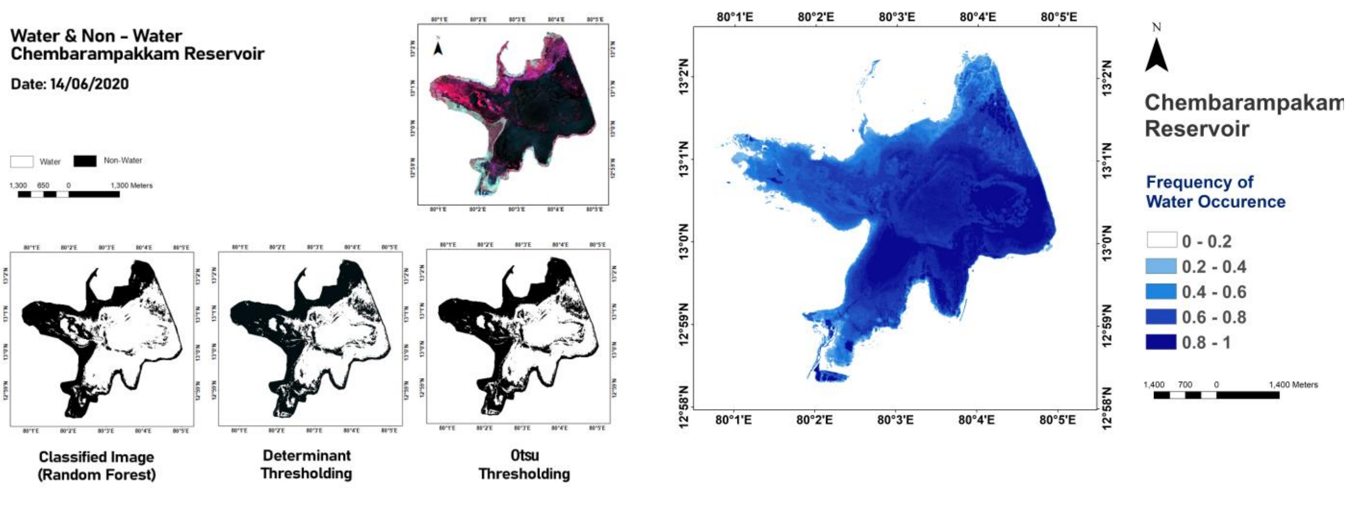


(**a**) (**b**)

Figure 3. (a) Change in threshold values for increasing number of iterations **(b)** Variation of Deterministic threshold value from that of Otsu.

In the iteration process for deterministic thresholding, the threshold values reached negligible variations around 6th iteration as seen in the figure 3 (a). When the same function was repeated for different regions, it successfully worked for areas with significant water cover. But for areas with little or no water cover, the iteration couldn't perform the estimation of threshold value. When the threshold value generated using iteration process was compared with the Otsu threshold value, it was found to be moreover similar with an outlier on just the 35th image of the collection. On average, the Otsu thresholds for NDWI and NDVI were 0.0015 and 0.00108 lower than the iterated thresholds, respectively (refer figure 3(b)). This shows that, this improvised deterministic thresholding approach can be also used in automatic thresholding just as in Otsu approach. For understanding the quality of both the thresholding approaches, it was compared with the classified image generated from a cloud free day (2020-06-14) through supervised classification (figure 4(a)). The classification was carried on with Random Forest classifier with 96% accuracy. With visual interpretation it was evident that both the thresholding technique failed to classify the floating vegetation as water. Since both the threshold images didn't show much variation, the number of misclassified pixels was calculated and was found to be 8.31% and 8.24% for deterministic and Otsu thresholding respectively. Though the difference in threshold value of deterministic and Otsu are not so significant, Otsu proved to perform better.

Finally, the surface dynamism layer (refer figure 4(b)) gives a better understanding of the reservoir's depth, the possible full extent it can reach and zones where the siltation can be removed for enhancing the holding capacity of the reservoir.



(**a**) (**b**)

Figure 4. (a) Comparison of binary images from deterministic and Otsu thresholding with classified image **(b)**The surface water dynamism layer

4. Conclusions

In conclusion, this study has performed two global thresholding techniques: an improved method of deterministic thresholding and Otsu thresholding to automatically generate image specific threshold. From this study, it is evident that Google Earth Engine (GEE) allows us to perform surface dynamism analysis of water bodies from over several years at ease. This study also shows how iterating mean of class averages can lead to improving threshold values which can be used for deterministic thresholding in surface water extraction. Also, it shows that the threshold values estimated for deterministic thresholding and Otsu threshold value are moreover similar. It has been noted that, Otsu slightly outperforms the deterministic thresholding. The surface water dynamism layer gave a better understanding of the reservoir's permanent and temporary water spread. Including Indices that can distinguish Floating Vegetation or algae and Modified NDWI could improve the quality of this study. This study's findings contribute to ongoing hydrological research that aimed at proper water resource management.

Funding: This research received no external funding.

Data Availability Statement: Data were from Google Earth Engine platform and the boundary file for the reservoir was self-digitized.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Fang, Y.; Ceola, S.; Paik, K.; McGrath, G.; Rao, P.S.C.; Montanari, A.; Jawitz, J.W. Globally Universal Fractal Pattern of Human Settlementsin River Networks. *Earth's Futur.* **2018**, *6*, 1134–1145, doi:https://doi.org/10.1029/2017EF000746.
2. Sharma, R.K.; Kumar, S.; Padmalal, D.; Roy, A. Streamflow Prediction Using Machine Learning Models in Selected Rivers of Southern India. *Int. J. River Basin Manag.* **2023**, doi:10.1080/15715124.2023.2196635.
3. Karpatne, A.; Khandelwal, A.; Chen, X.; Mithal, V.; Faghmous, J.; Kumar, V. Global Monitoring of Inland Water Dynamics: State-of-the-Art, Challenges, and Opportunities. In *Computational Sustainability*; Lässig, J., Kersting, K., Morik, K., Eds.; Springer International Publishing: Cham, 2016; pp. 121–147 ISBN 978-3-319-31858-5.
4. Ko, B.C.; Kim, H.H.; Nam, J.Y. Classification of Potential Water Bodies Using Landsat 8 OLI and a Combination of Two Boosted Random Forest Classifiers. *Sensors (Basel).* **2015**, *15*, 13763–13777, doi:10.3390/s150613763.
5. Carroll, M.L.; Loboda, T. V Multi-Decadal Surface Water Dynamics in North American Tundra. *Remote Sens.* **2017**, *9*, doi:10.3390/rs9050497.
6. Huang, C.; Chen, Y.; Zhang, S.; Wu, J. Detecting, Extracting, and Monitoring Surface Water From Space Using Optical Sensors: A Review. *Rev. Geophys.* **2018**, *56*, 333–360, doi:https://doi.org/10.1029/2018RG000598.
7. Sheng, Y.; Song, C.; Wang, J.; Lyons, E.A.; Knox, B.R.; Cox, J.S.; Gao, F. Representative Lake Water Extent Mapping at Continental Scales Using Multi-Temporal Landsat-8 Imagery. *Remote Sens. Environ.* **2016**, *185*, 129–141, doi:https://doi.org/10.1016/j.rse.2015.12.041.
8. Walia, Y.; Gupta, P.K.; Srivastav, S.K.; Gulzat, A.; Saha, S.K. Cloud Based Geo-Processing Platform for Analyzing Large Volume Temporal Satellite Data: A Study in Part of Ghaghara River Basin (India) for Surface Water Spread Analysis. *38th Asian Conf. Remote Sens. - Sp. Appl. Touching Hum. Lives, ACRS 2017* **2017**, *2017-Octob*.
9. Zhou, S.; Kan, P.; Silbernagel, J.; Jin, J. Application of Image Segmentation in Surface Water Extraction of Freshwater Lakes Using Radar Data. *ISPRS Int. J. Geo-Information* 2020, *9*.
10. Gonzalez, R.; Faisal, Z. *Digital Image Processing Second Edition*; 2019;
11. Sahoo, P.K.; Soltani, S.; Wong, A.K.C. A Survey of Thresholding Techniques. *Comput. Vision, Graph. Image Process.* **1988**, *41*, 233–260, doi:https://doi.org/10.1016/0734-189X(88)90022-9.
12. Padmasini, N.; Umamaheswari, R.; Sikkandar, M.Y. Chapter 10 - State-of-the-Art of Level-Set Methods in Segmentation and Registration of Spectral Domain Optical Coherence Tomographic Retinal Images. In; Dey, N., Ashour, A.S., Shi, F., Balas, V.E.B.T.-S.C.B.M.I.A., Eds.; Academic Press, 2018; pp. 163–181 ISBN 978-0-12-813087-2.
13. Das, N.; Bhattacharjee, R.; Choubey, A.; Ohri, A.; Dwivedi, S.B.; Gaur, S. Time Series Analysis of Automated Surface Water Extraction and Thermal Pattern Variation over the Betwa River, India. *Adv. Sp. Res.* **2021**, *68*, 1761–1788, doi:https://doi.org/10.1016/j.asr.2021.04.020.

14. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man. Cybern.* **1979**, *9*, 62–66, doi:10.1109/TSMC.1979.4310076.
15. Xu, X.; Xu, S.; Jin, L.; Song, E. Characteristic Analysis of Otsu Threshold and Its Applications. *Pattern Recognit. Lett.* **2011**, *32*, 956–961, doi:https://doi.org/10.1016/j.patrec.2011.01.021.
16. Brindha, K.; Neena Vaman, K. V; Srinivasan, K.; Sathis Babu, M.; Elango, L. Identification of Surface Water-Groundwater Interaction by Hydrogeochemical Indicators and Assessing Its Suitability for Drinking and Irrigational Purposes in Chennai, Southern India. *Appl. Water Sci.* **2014**, *4*, 159–174, doi:10.1007/s13201-013-0138-6.
17. Veerasingam, S.; Mugilarasan, M.; Venkatachalapathy, R.; Vethamony, P. Influence of 2015 Flood on the Distribution and Occurrence of Microplastic Pellets along the Chennai Coast, India. *Mar. Pollut. Bull.* **2016**, *109*, 196–204, doi:https://doi.org/10.1016/j.marpolbul.2016.05.082.
18. McFEETERS, S.K. The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432, doi:10.1080/01431169608948714.
19. Dutta Gupta, V.; Areendran, G.; Raj, K.; Ghosh, S.; Dutta, S.; Sahana, M. Chapter 26 - Assessing Habitat Suitability of Leopards (Panthera Pardus) in Unprotected Scrublands of Bera, Rajasthan, India. In; Kumar Shit, P., Pourghasemi, H.R., Adhikary, P.P., Bhunia, G.S., Sati, V.P.B.T.-F.R.R. and C., Eds.; Elsevier, 2021; pp. 329–342 ISBN 978-0-12-822931-6.
20. Kolli, M.K.; Opp, C.; Karthe, D.; Pradhan, B. Automatic Extraction of Large-Scale Aquaculture Encroachment Areas Using Canny Edge Otsu Algorithm in Google Earth Engine – the Case Study of Kolleru Lake, South India. *Geocarto Int.* **2022**, *37*, 11173–11189, doi:10.1080/10106049.2022.2046872.