

PREDICTING SURFACE WATER LEVEL ESTIMATION BY USING DATA MINING TECHNIQUES

¹ KusumaKavya M , ² DivyaVani K , ³ Sathiyavathi R

Dept. of IT, School of Computing, Sathyabama Institute of Science and Technology, Chennai, India

Abstract— To auspicious get exact pixel water surface extent data through remote detecting is to a great degree noteworthy to the environmental reclamation in inland waterway bowls and for the exact administration of water assets. In regard to the insufficient extraction of water surface extent data show in pixels in the greater part of the ebb and flow water data models, a straightforward model Enhanced Water Index (EWI) in view of Modified Standardized Difference Water Index (MNDWI) has been presented. EWI, which is arranged toward the sub-pixel level examination of water surface extent mapping of inland stream bowl, has been advanced in light of the examination of run of the mill ghostly marks for example, forsake, soil, and vegetation alongside MNDWI in agreement with the Landsat TM band highlights. The examination is done by utilizing strategies for pixel-based EWI esteem with various water extents which are dissected through the presentation of the straight crossover reenactment between the water body and the comparing foundation. In conclusion, the impact of EWI demonstrate has been tried in the medium and lower ranges. The amendment coefficient for sub-pixel level water surface extent anticipated by the EWI show and the test information. Results demonstrated that the model could viably remove the data about pixel water surface extent in inland stream bowls. This investigation demonstrates that EWI show has awesome potential in its application for water extent mapping applications.

Key words—Enhanced Water Index (EWI), Modified Standardized Difference Water Index (MNDWI), percent surface water estimation, remote detecting of condition.

I.Introduction

INFILTRATION and flood of regular surge in inland waterway bowls are huge to the reclamation of biological condition on the two sides of the waterway, and can have a direct effect on the examples, process, and elements of the biological framework. Most inland waterway bowls convey in parched zones, and need enough water assets for vegetation rebuilding as a result of fast rural advancement, consequently cause the absence of water asset for common vegetation rebuilding [2, 3].

Satellite-based convenient dissemination data of surge along inland stream is noteworthy to chiefs for administration of water asset and biological frameworks. In the course of recent decades, the majority of the looks into on water data extraction through remote detecting underlined on the otherworldly mark investigation of the water body what's more, related foundation data, and the single band-based or on the other hand multiband-blend based water data extraction demonstrate [1]– [3], [6], [9]. This approach has generally been utilized both at home and abroad because of its expediency, accommodation, and effectiveness [2], [4], [5], [7], [8], [9]. Be that as it may, this sort of inquire about depends on recognizing the water

body data from the non-water-body data using pixel scales instead of from the purpose of the extent of the water region, which antagonistically impacts precise water data extraction exactness to a critical degree [5], [8].

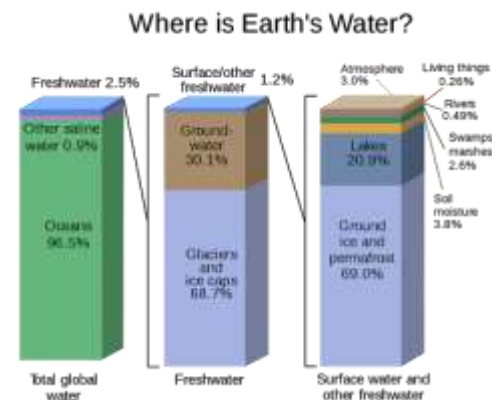


Fig.1 Detailed Division of water resources.

The primary reason for this investigation is to acquire the extraction display which can adequately gain the water proportion pixel. The accompanying issues might be given careful consideration all together to build up the pixel-situated water surface extent: (1) the model can viably recognize water surface from foundation data, for example, betray, soil and vegetation; (2) the models must be distinctive in the pressure of various foundations (vegetation and soil, and so forth.) for a superior separation of water bodies [3, 4]. To precisely acquire pixel-arranged water zone extent, the examination in light of Modified Normalized Difference Water File (MNDWI) principally took after the accompanying focuses: (1) the impact of water pixels with various water bodies and foundation highlights is broke down in view of Landsat TM band highlights what's more, numerical reproduction [3]; (2) the examination of model application what's more, approval in the medium and lower compasses of the zone.

Aside from the water record (WI)- based division, the edge implanted marker-based watershed calculation permits to get exact protest limits in high-determination satellite pictures [4]. The computerized lake water mapping at sub-pixel level is conceivable in the satellite pictures having fine spatial determination [1]. For the most part, the free water surface location is an essential to remove the shoreline, which prompts the precise estimation of wetland flow. In this view, the stochastic entropies what's more, separations fill in as a productive edge-recognition technique for the manufactured gap radar symbolism [5]. Besides, there are some powerful calculations exist for water surface limit extraction, i.e., the dark level co-occurrence lattice [6], and watchful edge administrator [7]. Plainly, the WSB-DA is concerned just with free water surface, i.e.,

skimming water without event of soil (sloppy condition) and vegetation cover compose. In this way, this work intends to build up a semi-automated free water surface limit extraction calculation. The proposed calculation is chiefly in light of WI included decorated top change (TCT), and multispectral colorimetric technique, i.e., shade immersion esteem (HSV) change.

II. LITERATURE SURVEY

One of critical advance in information mining is setting up the info information which might be unstructured, meager, and may contain commotion, for example, inadequate exchanges, records excess, missing esteems, poor picture, and so on. Handling the various types of information to separate the suitable highlights for facilitate investigation is a urgent assignment. There are distinctive programming frameworks that can be used to process the information, however the use of programming depends on the sort of information.

Range estimation of the run of the mill basic spatial highlights is done according to each sort on the two sides of the Waterway bowl, and the subtle elements are appeared. Water Body Spectrum: It is obvious that the reflectance for water body is by and large weaker—around 6%– 8% in the scope of 400– 660 nm and the rest of the piece of the range is likewise indicating debilitating pattern as the wavelength increments. These discoveries contrast from some authentic records to a specific degree, principally as a result of the streaming condition of the surge test and a lot of silt in that. Along with the expansion of such silt, the reflectance is improved, furthermore, the reflection pinnacle of the range bends pushes toward the long waves [4].

Spectrum Analysis on Bare Sand (Desert), Bare Soil, and Vegetation: The range of green and prospering vegetation presents reflection pinnacles of shifting degrees at 530, 1640, what's more, 2200 nm, and ingestion valleys of shifting degrees at 660, 1420, and 1900 nm and also the high reflection district in the vicinity of 740 and 1300 nm. The variety pattern of the wilted vegetation is like that of the green vegetation, while the relating reflection pinnacles and retention valleys are clear than those of the green vegetation. The reflection pinnacles and retention valleys of wilted vegetation in the reflecting locale in the vicinity of 720 and 1300 nm are lower than that of the green vegetation while in alternate locales it is higher than that of the green vegetation. The variety slant of the forsake range is fundamentally the same as that of the soil range, and their reflectance, which is in the vicinity of 570 and 1390 nm. The reflectance of the abandon (which is between 1300 also, 2500 nm) is higher than that of the dirt and that of the shriveled vegetation, while the reflectance of the dirt, which lies in the vicinity of 1900 and 2500 nm, is higher than that of the shriveled vegetation.

The inland stream bowl requires higher accuracy in the surge scope display. The flood and invasion zones of the expansive zone are to various degrees, shaped by leave surge spoke to as blended pixels covering a vast territory in remote detecting

pictures, what's more, the natural condition and water assets exactness administration require more exact spatial water surface dissemination data.

The hidden surface highlights of abandon bowls fit in with the prerequisites of the EWI demonstrate. As the inland stream bowl landscape as a rule has inconspicuous vacillations, what's more, as the surge flood zone of extensive zones in deserts comprises of sand soil, exposed soil, and flushing and wilted vegetation as commonplace fundamental spatial highlights in this manner the water body foundation is generally basic, and the EWI display is proper, and can adequately pack the foundation data and feature water body-related data.

It is apparent by recreation and tests led amid this inquire about that the EWI demonstrate is solid in inland stream bowl to a specific degree. Through numerical esteem reproduction, it can be demonstrated that the EWI display presents water content plainly in pixels, and the River bowl test demonstrates that this display is equipped for effective expectation of the water content extents in pixels. In any case, when far reaching examination is performed on the on location examination, reenactment investigation, and remote detecting mapping comes about, it is obvious that the profundity, stream, and silt substance of the water body can likewise impact the outcome.

The combination of panchromatic and other groups can prompt a superior outcome, yet in this investigation, it didn't upgrade the calculation yield. The absence of adequate phantom data makes PAN, a confined band for this examination. Quite, amid the WSB-DA advancement, the "ultra-blue" band misuses the property of being slightest consumed by water. A straightforward visual translation does not uncover any impressive contrast between ultra and ordinary blue groups. Curiously, the histogram evening out technique encourages the representation of these distinctions.

The wetland locales, which have been tried in this investigation, are not shallow, but rather perhaps impacted by dregs transitions, phytoplankton wealth, turbidity varieties, and so on. These impacts make a mind boggling biological system for every wetland. At once, it is genuinely hard to think about every one of these issues in the calculation improvement process. At last, the proposed calculation gives a nonstop smooth single line, which speaks to the free water surface limit for regions under thought.

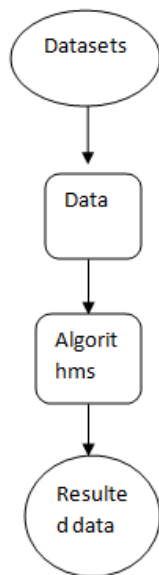
A. Drawbacks in existed system:

The extraction of water surface limit just, which is as of now non-existent for to a great degree wet terrains. Furthermore, sensible intermediary for the foundation of free water surface limit identification.

- It isn't guaranteeing careful estimations of datasets.
- It isn't investigating all the datasets.
- It might misfortune a few information.
- It doesn't give potential in its application to water surfaces.

III. PROPOSED SYSTEM

In this convenient acquire exact pixel surface water extent data through remote detecting. Remote detecting is greatly noteworthy to the natural rebuilding in inland waterway bowls and for the exact administration of water assets. In regard to the lacking extraction of water surface extent data introduce in pixels in a large portion of the momentum water data models like Enhanced Water Index (EWI) in view of Modified Normalized Difference Water Index (MNDWI).



EWI is arranged towards the sub pixel level examination of water surface extent mapping of inland stream bowl. The examination of run of the mill otherworldly mark, for example, betray, soil, vegetation alongside MNDWI as per the band highlights. The investigation is finished by utilizing techniques for pixel in view of EWI esteem with various water extents which are broke down through the presentation of direct half and half reproduction between water body and comparing foundation. The impact of EWI demonstrate has been tried in the medium and the lower groups.

A. List of Components:

1. Data preprocessing
2. Extraction of datasets

3. Analysis of KNN algorithm
4. Analysis of SVM algorithm

3.1. Data preprocessing:

Datasets of surface water areas have been investigated that they are affiliated with two classes, which has seven features (specialties) difference in groundwater areas

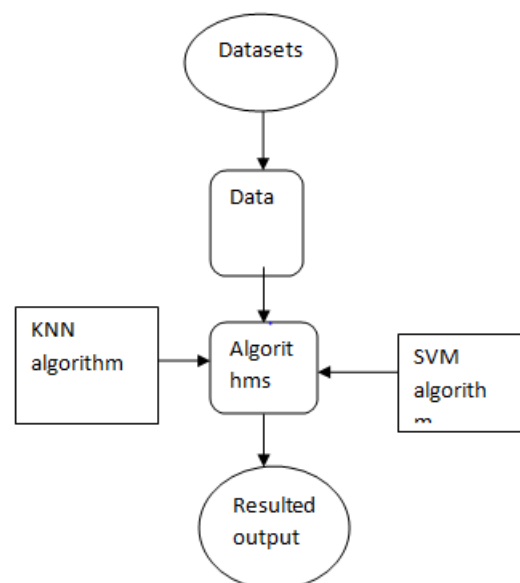
These are the features Height, defect, rain, gradient, temperature, Wada and Lake Our groundwater datasets are collected from these social media data's.

3.2 Extraction of Data Sets:

Datasets of surface water territories have been researched that they are associated with two classes, which has seven highlights (claims to fame) distinction in groundwater regions. These are the highlights Height, imperfection, rain, inclination, temperature, Wada and Lake Our groundwater datasets are gathered from these web-based social networking data's.

3.3 Analysis of KNN algorithm:

The KNN calculation is straightforward. In view of preparing test information, KNN finds the KNNs of the preparation information utilize the k-neighbors classes to dole out the test class case. The scores of comparability of each example of the close to the test occurrence is utilized as the heaviness of the classes of close example. At the point when a few KNNs share a class, at that point the pre-neighbors weights of that class ought to be included, and the consequence of included weights ought to be utilized as the likelihood score of that class in regards to the test example. To locate a positioning for the test example, the scores of the competitor classes ought to be requested.



3.4 Analysis of SVM algorithm:

SVMs created are a directed information mining calculation that can be utilized for the two classifications or relapse challenges. It is as often as possible utilized as a part of characterization issues. SVM have turned into the technique for decision to take care of troublesome characterization issues in a wide range scope of utilization spaces. SVM based on the base limiting basic hazard. In direct order, The SVM deliver a hyper which isolates the preparation information into two gatherings with most extreme edge. A hyper greatest edge design is a hyper plane that isolates two focuses and is an equivalent separation from the two. Numerically, the SVM learns work sign $f(x) = s(w \cdot x + b)$, where w is weighted vector in R . The SVM discover the design $y = w \cdot x + b$ di isolating the R space in two half-spaces with the most extreme edge. Straight SVMs can be summed up for nonlinear issues, the information is mapped into another space H and we run the direct SVM calculation on this new space.

3.5 Advantages :

- This model was able to effectively extract the information about pixel water surface proportion in inland river basins.
- It proves that EWI model has great potential in its application for water proportion mapping applications.
- It will analysis all the datasets and gives accurate analysis.

IV. RESULTS AND DISCUSSION:

An examination is exhibited for surface water quality forecast utilizing different information mining methods at various areas. As per the datasets we have given different index ranges, water type, area, temperature , dimensions as shown in the fig.2

A.Dataset:

id	water type	Surface le	water tex	humidity	area	PH range	compactn	temperat	fractal_Dimension
1 F	23	12	151	954	0.143	0.278	0.242	0.079	
2 M	9	13	133	1326	0.143	0.079	0.181	0.057	
3 F	21	27	130	1203	0.125	0.16	0.207	0.06	
4 F	14	16	78	386	0.07	0.284	0.26	0.057	
5 F	9	19	135	1297	0.141	0.133	0.181	0.059	
6 M	25	25	83	477	0.128	0.17	0.209	0.076	
7 F	18	26	120	1040	0.095	0.109	0.179	0.057	
8 F	15	18	90	578	0.119	0.165	0.22	0.075	
9 F	19	24	88	520	0.127	0.193	0.235	0.074	
10 F	25	11	84	476	0.119	0.24	0.203	0.082	
11 F	24	21	103	798	0.082	0.067	0.153	0.057	
12 F	17	15	104	781	0.097	0.129	0.184	0.061	
13 M	14	15	132	1123	0.097	0.246	0.24	0.078	
14 F	12	22	104	783	0.084	0.1	0.185	0.053	
15 F	12	13	94	578	0.113	0.229	0.207	0.077	
16 F	22	19	97	659	0.114	0.16	0.23	0.071	
17 F	10	16	95	685	0.099	0.072	0.139	0.059	
18 F	15	14	108	799	0.117	0.202	0.216	0.074	
19 F	20	14	130	1260	0.098	0.103	0.138	0.054	
20 M	17	11	87	566	0.098	0.081	0.189	0.058	
21 M	16	14	86	520	0.108	0.127	0.197	0.068	
22 M	17	24	90	274	0.102	0.065	0.182	0.069	

Fig.2 Datasets for different water areas with humidity ,temperature and water texture.

B.Output:

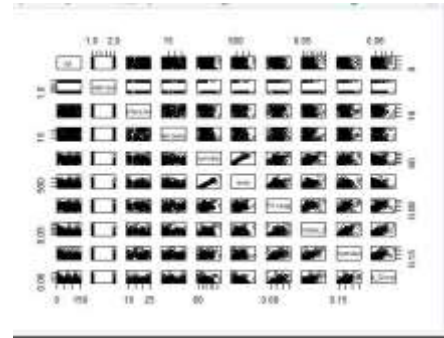


Fig.3 Output for different water index Ranges.

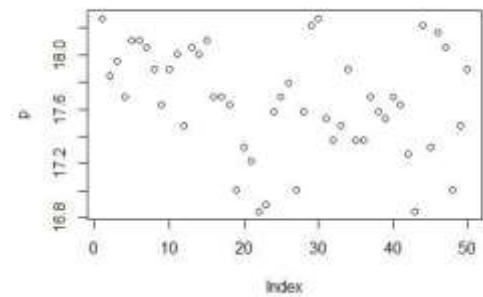


Fig.4 Pictorial Representation for the datasets, Index as on X-axis and PH as on Y-axis.

V. CONCLUSION

In this paper we have given an analysis is presented for surface water quality prediction using various data mining techniques at different locations. According to analyses Back Propagation is used frequently. Basic, quick, and helpful and exact remote detecting mapping technique for water-related data has continuously been in center in water mapping research. Despite the fact that the water data recognizable proof models and strategies at pixel level have been developed and generally connected for a few time, be that as it may, on account of the flooding territory which is exceptionally normal in surge season in inland waterway bowls which are amazingly inadequate in water assets, the extraction of water scope data of high exactness is especially imperative. In this investigation, in view of the MNDWI demonstrate, the water data demonstrate EWI is created with the reason for performing quick high-exactness mapping of the water content extent in pixels. In this way, it is prescribed that the real circumstances, extensive reenactments and pictures are joined

with a specific end goal to decide the EWI max-mean and EWI min-mean of the EWI show exhaustively.

References

- [1] Ajmar A, P. Boccardo, F. Disabato, and T. F. Giulio, "Near real time flood monitoring tool," in Proc. Gi4DM 2010 Conf., Torino, Italy, Feb. 2–4, 2010, pp. 158–163.
- [2] Boland D, Trophic Classification of Lakes Using Landsat-1 (ERTS21) Multispectral Scanner Data [A]. Corvallis, OR, USA: US EPA, Office of Research and Development, Corvallis Environmental Research Laboratory, 1976.
- [3] Chen, Y. Z. Zhang, A. Ekroos, and M. Hallikainen, "The role of remote sensing technology in the EU water framework directive (WFD)," Environ. Sci. Policy, vol. 7, pp. 267–276, 2004.
- [4] Chen, P. Cui, Y. Li, Z. Yang, and Y. Q. Qi, "Changes in glacial lakes and glaciers of post-1986 in the Poiqu River basin, Nyalam, Xizang (Tibet)," Geomorphology, vol. 88, pp. 298–311, 2007.
- [5] Nascimento, M. M. Horta, A. C. Frery, and R. J. Cintra, "Comparing edge detection methods based on stochastic entropies and distances for PolSAR imagery," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 7, no. 2, pp. 648–663, Feb. 2014.
- [6] Ouma and R. Tateishi, "A water index for rapid mapping of shoreline changes of five East African Rift Valley lakes: An empirical analysis using Landsat TM and ETM+ data," Int. J. Remote Sens., vol. 27, pp. 3153–3181, 2006.
- [8] Singh, M. Ghosh, S. R. Sharma, and P. Kumar, "Blue-Red-NIR model for chlorophyll-a retrieval in Hypersaline-Alkaline water using Landsat ETM+ sensor," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 7, no. 8, pp. 3553–3559, Aug. 2014.
- [9] NASA, Landsat 7 Science Data User's Handbook, USA: NASA/USGS, 2012.
- [10] Feyisa, H. Meilby, R. Fensholt, and S. R. Proud, "Automated water extraction index: A new technique for surface water mapping using Landsat imagery," Remote Sens. Environ., vol. 140, pp. 23–35, 2014.
- [10] Vorovencii, "Use of the "Tasseled Cap" transformation for the interpretation of satellite images," Cadastre J. RevCAD, vol. 7, pp. 75–82, 2007.