IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 6, June 2021

DOI: 10.17148/IARJSET.2021.8612

Waste Water Quality Assessment in an Industrial Belt of West Bengal Using Statistical methods and Geo-Spatial Techniques

K. Das [1], D. Adak [2], B. Raychaudhuri [3], S. Sengupta [4]

Senior Scientific Officer, Department of Science & Technology and Biotechnology, Kolkata, India [1]

Ex.Project Fellow, Department of Science & Technology and Biotechnology, Kolkata, India [2]

Professor, Department of Physics, Presidency University, Kolkata , India [3]

Ex Professor , Indian Institute of Technology, Kharagpur, India [4]

Abstract: Maintaining the quality of surface water is a complex issue since they are used in domestic, agricultural and industrial purposes . Hence, to protect the water bodies proper water management strategies are necessary. Here an attempt has been made to evaluate the water quality in an industrial area of West Bengal. Industrial waste water samples were collected for the parameters such as TSS, BOD, COD, OIL& GREASE and PHENOL and tested at the laboratory .To compare means and test for significance of the parameters ANOVA was applied and based on their significance, Water quality index was calculated to assess the overall water quality status of the region . Landsat 8 satellite image was processed and correlated with water quality data. By correlation - regression analysis parameters were identified at respective spectral regions and their spatial distribution were measured using Inverse Distance Weighted method and mapped at ARC GIS 10.3.

Keywords : ANOVA, Remote Sensing and GIS, WQI, Regression Analysis and Inverse Distance Weighted .

I. INTRODUCTION

Water is an integral and essential part of our very existence as lives of all beings are sustained by water. In all the countries surface water play an important role to meet the industrial, agricultural, domestic and infrastructural needs of human beings as well as it also facilitates their waste discharge including agricultural runoff which ultimately makes it vulnerable to pollution [1-2]. However , the level of treatment required for human consumption, agriculture and industry necessitates an understanding of the quality of source waters [3]. The largest source of water pollution in India is untreated sewage [4] .With the increasing water pollution, degradation of water quality has caused widespread concern not only to the human society but also to natural ecosystem. Therefore, information on water quality , to locate and monitor the pollutants for a sustainable water-use management strategies is very essential [5-7]. Most of the water quality parameters such as TSS,TDS,BOD,COD, Oil& Grease, chlorophyll-a concentration, turbidity, salinity, total phosphorus (TP), Secchi disk depth (SDD), Temperature, pH, etc. are traditionally determined by collecting samples from the field and then analyzing the samples in the laboratory. Although, this method provides more accurate result but it is very labour intensive, costly, time consuming and even lacks in providing spatio-temporal data which is essential for monitoring a large waterbody [8-12]. But with advancement in space science and computerizations, applications of remote sensing techniques along with Geographical Information System (GIS) have become useful tools to overcome these short falls. He et al. and Kallio.K have rightly mentioned that with the development of remote sensing techniques, water quality monitoring based on remote sensing methods has become accessible and very efficient as i) it gives a synoptic view of the entire waterbody ; ii) Provides a comprehensive record of water quality in an area and represents trends over time and iii) Prioritizes sampling locations and field surveying times which are essential for effective water quality monitoring purposes [12-13].

Such an attempt has been made in the present study where Geo-Spatial technology coupled with Statistical methods have been used to evaluate the water quality status in an industrial belt of Hooghly district adjoining to Kolkata. Here besides spectral analysis of Landsat 8 satellite image, a number of statistical and mathematical analysis such as ANOVA, Correlation- Regression, WQI, NDVI etc. have been carried out to find the correlation between spectral regions and the water quality parameters as well as to establish suitable algorithm necessary for systematic water quality monitoring in any region.

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 6, June 2021

DOI: 10.17148/IARJSET.2021.8612

II. STUDY AREA

The study area which belongs to Hooghly district lies between 22°41´32″ - 22°42′18″ N latitude ’and 88°15´35″ - 88°20′07″E longitude particularly on the west bank of the famous river Hooghly and covers an area of 114.42 km^2 with a total population of 3,10,662 . The minimum and maximum temperature ranges between 13° to 36° C. Rainy season commences from the middle of June and continues up to September with July and August are the rainiest months. Mainly alluvial soils are found in the region. The district Hooghly is surrounded by the districts of Bankura and Barddhaman to the North, Nadia and North Twenty Four Parganas to the East, Howrah to the South and Paschim Medinipur to the West. The principal rivers in the district are the Hugli, the Damodar and the Rupnarayan (Figure 1).

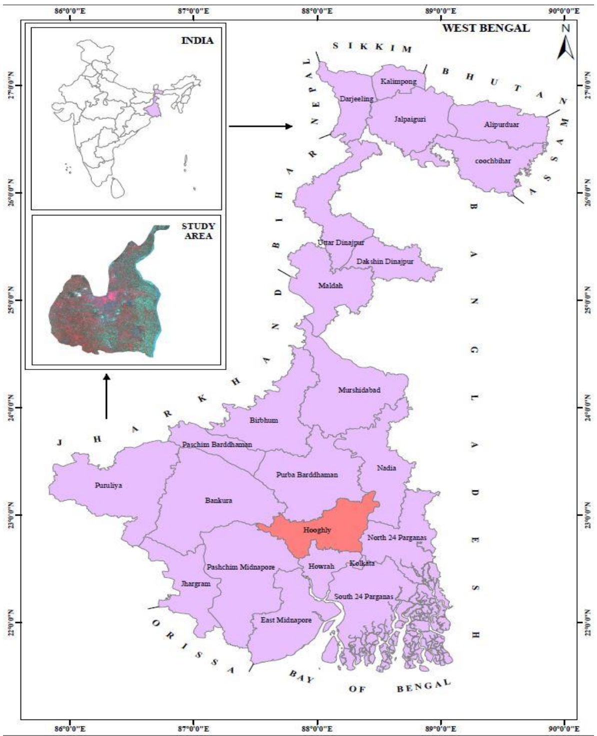


Figure 1: Location Map of the Study Area

III. MATERIALS AND METHODS

A) Sampling and Data Analysis.

Twenty five samples were collected from seven effluent discharge points of various industries located in the study area which discharge their effluents into River Hooghly through various small streams, drains or canals. The collected samples were analyzed in registered laboratory using APHA,2005 standard method [14] and compared with national standard (Table 1). Water quality Parameters which were considered for analysis include Total Dissolved Solids (TDS), Total Suspended Solids (TSS), Oil & Grease, Chemical Oxygen Demand (COD), Biochemical Oxygen Demand (BOD) ,phenol , Sulphide, Arsenic and Chloride.

Table I **:** Concentration of Water Quality parameters

INDUSTRIAL OUTFALL POINTS	TDS (mg/L)	TSS (mg/L)	OIL& GREASE (mg/L)	COD (mg/L)	BOD (mg/L)	PHENOL (mg/L)	SULPHIDE (mg/L)	ARSENIC (mg/L)	CHLORIDE (mg/L)	LATITUDE/ LONGITUDE
1	566	205	24.2	693	143	5.2	3.21	0.044	0.012	88°16'57.895"E 22°42'41.174"N
2	663	305	21.7	536	251	2.03	1.15	0.014	0.011	88°18'50.59"E 22°43'51.896"N
3	2145	55	11.5	126	45	<1.0	4.4	0.026	0.031	88°18'21.725"E 22°43'5.62"N
4	462	156	7.2	179	71	<1.0	1.39	0.22	0.025	88°21'15.856"E 22°43'21.104"N
5	403	66	8.3	137	51	<1.0	1.63	<0.01	0.039	88°18'58.51"E 22°43'44.172"N

6	819	171	12.4	210	113	<1.0	2.36	0.015	0.007	88°18'42.362"E 22°43'40.859"N
7	528	210	11.2	302	128	1.85	4.63	0.17	0.016	88°21'19.559"E 22°42'50.847"N

B) Analysis of Variance (ANOVA).

Water quality results were analyzed using a one-way Analysis of Variance (ANOVA). ANOVA was applied to compare means and test for significance of each water quality parameter. The method enables the difference between two or more sample means to be analyzed. The purpose is to test for significant differences between class means, and this is done by analyzing the variances. The basis of ANOVA is the partitioning of sums of squares between-class and within-class [15-16] It provides an F statistic, the ratio of the variance computed among the means to the variance within the samples. Following central limit theorem, if the group means belong the same population, the variance between the group is supposed to be lower than the variance within, quantifying the spread of sampling stations within groups. A higher value of the F statistic than its critical value F crit implies significant differences between the groups, meaning that the samples do not belong to the same population (Table 2).

Table II **:** Calculation of ANOVA

SUMMARY

Groups	**Count**	**Sum**	**Average**	**Variance**
TDS (mg/L)	25	13495	539.8	203235.6
TSS (mg/L)	25	1799	71.96	5411.957
OIL& GREASE (mg/L)	25	148.9	5.956	36.7584
COD (mg/L)	25	3142	125.68	24390.06
BOD (mg/L)	25	1039	41.56	3018.007
PHENOL (mg/L)	25	22.03	0.8812	0.891169
SULPHIDE (mg/L)	25	30.375	1.215	1.4529
AMMONIUM (mg/L)	25	142	5.68	44.35743
ARSENIC (mg/L)	25	0.606	0.02424	0.001746
CADMIUM (mg/L)	25	0.2915	0.01166	0.000117
LEAD (mg/L)	25	0.597	0.02388	0.000382

ANOVA

Source of Variation	**SS**	**df**	**MS**	**F**	**P-value**	**F crit**
Between Groups	7760182.903	25	310407.3	23.5133	1.3E-73	1.523774
Within Groups	8237643.717	624	13201.35			
Total	15997826.62	649				

C) Water Quality Index (WQI).

Water Quality Index is a mathematical method which provides a single number that expresses overall water quality at a certain location and time, based on various water quality parameters [17-19]. It gives an indication how suitable the water is for various purposes (Table-3).

Table III **:** Status of water quality

Water Quality Index Level	**Water quality status**	**Category**
0-25	Very bad	A
26-50	Bad	B
51-70	Medium	C
71-90	Good	D
91-100	Excellent	E

DOI: 10.17148/IARJSET.2021.8612

Horton initially proposed the application of WQI in 1965 and since then many different methods for the calculation of WQI's have been developed such as National Sanitation Foundation Water quality Index (NSFWQI), Oregon Water Quality Index (OWQI) etc .

In the present study water quality index has been estimated using Weighted Arithmetic Index Model . In this model, different water quality components are multiplied by a weighting factor and are then aggregated using simple arithmetic mean (Table-5). Weighted arithmetic water quality index method classified the water quality according to the degree of purity by using the most commonly measured water quality variables.

Here ,for assessing the quality of water, the quality rating scale (Qi) for each parameter was Calculated by using the following equation;

qn = 100 (Vn – Vio) / (Sn – Vio)

(Let there be *n*water quality parameters and quality rating or sub index *(qn)* corresponding to *nth* parameter is a number reflecting the relative value of this parameter in the polluted water with respect to its standard permissible value).

(Let there be *n*water quality parameters and quality rating or sub index *(qn)* corresponding to *nth* parameter is a number reflecting the relative value of this parameter in the polluted water with respect to its standard permissible value).

qn=Quality rating for the *nth* Water quality parameter.

Vn=Estimated value of the *nth* parameter at a given sampling station obtained from laboratory analysis.

Sn =Standard permissible value of the *nth* parameter.(here recommended CPCB standard are used)

Vio= Ideal value of nth parameter in pure water, (i.e.,0 for all other parameters except the parameter pH and Dissolved oxygen (7.0and 14.6 mg/L respectively) can be obtained from the standard Tables.

Unit weight was calculated by a value inversely proportional to the recommended standard value *Sn* of the corresponding parameter.

Wn =K / Sn

Wn= Unit weight for the *nth* parameters. *Sn*= Standard value for *nth* parameters. K= Constant for proportionality.

$$k = (1/(1/\sum_{i=1}^{n} si))$$

Proportionality constant " K " value using formula
where "si" is standard permissible for nth parameter.
The overall Water Quality Index calculated by aggregating the quality rating with the unit weight linearly.

WQi = ΣQiWi/ Σ Wi

Where, Qi = Quality rating, Wi = Relative weight

D) Image Processing And Generation Of Regression Models.

Landsat 8 (2018) cloud free Satellite Data along wth Survey of India Topographical sheets, Police station maps, hand held GPS etc. have been utilized in this study. At first using Erdas Imagine 9.2 software Geometric Corrections between Satellite Images and Topoghaphical Sheets has been performed considering UTM / WGS 1984 coordinates system and Nearest Neighbour Resampling method which were further rectified with the GPS points collected from the ground. Then base layer database such as study area boundary demarcation, locating industries , analysis of waste water status at various effluents points with land use / land cover pattern of the surrounding areas etc. were generated at ARC GIS 10.3 platform. Afterwards, the DN values of images were converted into reflectance values for performing ratios between different bands and subsequently to find their correlation with water quality parameters correlation analysis were done [20] . Finally, based on the output correlation- regression models were generated using Statistical Package of Social Sciences (SPSS) software.

E) Inverse Distance Weighted (IDW) Interpolation.

Inverse distance weighted (IDW) interpolation explicitly makes the assumption that things that are close to one another are more alike than those that are farther apart. To predict a value for any unmeasured location, IDW uses the measured values surrounding the prediction location. The measured values closest to the prediction location have more influence on the predicted value than those farther away. IDW assumes that each measured point has a local influence that diminishes with distance. It gives greater weights to points closest to the prediction location, and the weights diminish as a function of distance, hence the name inverse distance weighted. The equation which is followed is :-

$$Z_o = \left\{ \frac{\sum_{i=1}^{s} \frac{Z_i}{d_i^k}}{\sum_{i=1}^{s} \frac{1}{d_i^k}} \right\}$$

Where: Z is the estimated value at a point 0, Z is the distance between control point I, is the distance between control point and point o, is the number of control points used in estimation, k is the specified power.

Using this method spatial distribution of various water parameters were estimated at Arc GIS 10.3 platform.

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 6, June 2021

DOI: 10.17148/IARJSET.2021.8612

IV. RESULTS AND DISCUSSIONS

The laboratory tested results of the samples (Table-1) indicates that discharge rate of the parameters like TSS, TDS, Oil & Grease, COD, BOD, Phenol, Sulphide, Arsenic, Chloride were quite high. Accordingly, to observe whether these parameters are statistically significant or not one-way ANOVA (Analysis of Variance) was done (Table -4)

Table IV : Result of ANOVA Single Factor

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	7760182.903	25	310407.3	23.5133	1.3E-73	1.523774
Within Groups	8237643.717	624	13201.35			
Total	15997826.62	649				

The output of Analysis of Variance (ANOVA) indicates that the parameters are statistically very significant as F value is greater than Fcrit value (F=23.5133>$Fcrit$ =1.523774) and based on this significance whether there exists any difference in the physio-chemical characteristic among the parameters at all the stations or not,P value analysis was performed (Table-5).

Table V : P-Value Analysis through Statistical Method

Parameters	p-value	Remarks
TDS (mg/L)	4.37E-07	Significant
TSS (mg/L)	0.000227	Significant
OIL& GREASE (mg/L)	0.000566	Significant
COD (mg/L)	0.000744	Significant
BOD (mg/L)	0.013117	Significant
PHENOL (mg/L)	1.24E-10	Significant
SULPHIDE (mg/L)	3.11E-10	Significant
AMMONIUM (mg/L)	0.000577	Significant
ARSENIC (mg/L)	1.33E-11	Significant
CADMIUM (mg/L)	1.29E-11	Significant
LEAD (mg/L)	1.33E-11	Significant

The result of P-Value analysis again proves that there are significant differences in their physico-chemical characteristics of the parameters as p <0.05 at all the stations. Observing the significance and variations among parameters, calculation of Water Quality Index (WQI) was done to know the overall water quality of the area(Table-6).

Table VI : Calculation of water Quality Index

Parameters	Observed Values	Standard Values(Sn)	1/Si	Unit Weight (Wn)	Quality Rating(qn)	Wnqn
TDS (mg/L)	280	2100	0.00048	3.28017E-06	13.33333333	4.37356E-05
TSS (mg/L)	19	100	0.01	6.88835E-05	19	0.001308787
OIL& GREASE (mg/L)	3.4	10	0.1	0.000688835	34	0.0234204
COD (mg/L)	82	250	0.004	2.75534E-05	32.8	0.000903752
BOD (mg/L)	18	30	0.03333	0.000229612	60	0.013776706
PHENOL (mg/L)	<1.0	1	1	0.006888353	60	0.413301169
SULPHIDE (mg/L)	0.42	2	0.5	0.003444176	21	0.072327705
AMMONIUM (mg/L)	2.02	50	0.02	0.000137767	4.04	0.000556579
ARSENIC (mg/L)	0.027	0.2	5	0.034441764	13.5	0.464963815

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Vol. 8, Issue 6, June 2021
DOI: 10.17148/IARJSET.2021.8612

CADMIUM (mg/L)	0.009	2	0.5	0.003444176	0.45	0.001549879
LEAD (mg/L)	0.016	0.1	10	0.068883528	16	1.10213645
						2.381376378

WQI=£Wnqn/£(Wn) 2.381376378

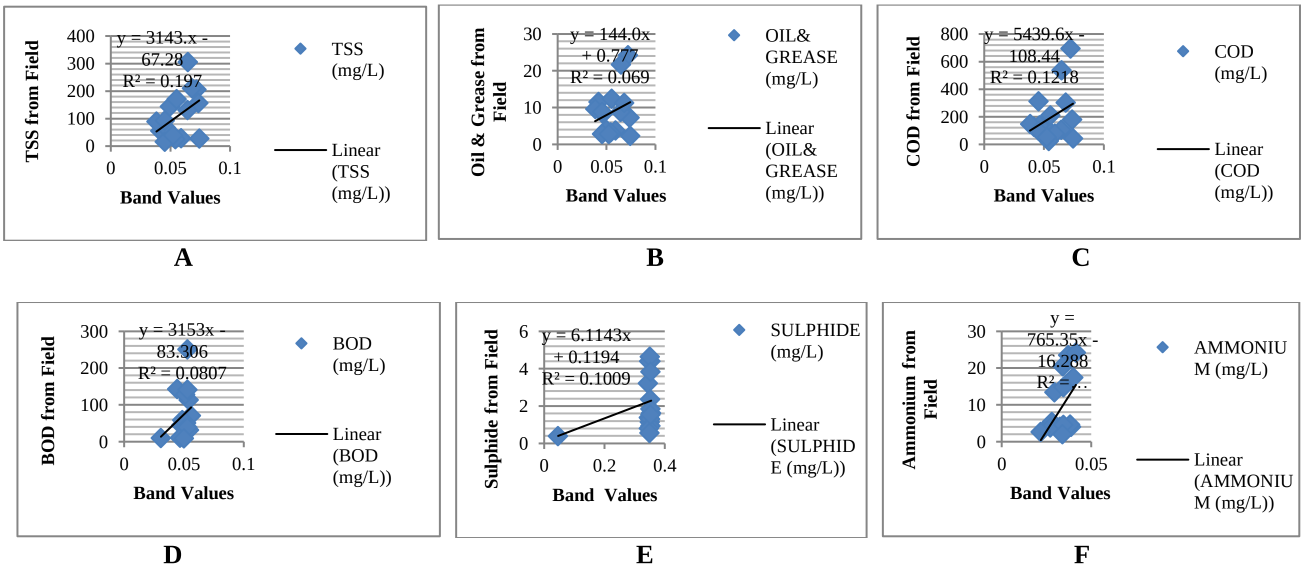
Based on the index value the water quality status of the area was compared with the standard and accordingly categorised . From the output it is evident that the water quality of the area is falling under very bad to moderate quality. Present day Geo-spatial techniques are widely used to assess water quality . Particularly to measure the concentration of dissolved and suspended particles in water bodies and to provide their spatial-temporal variations. The more suspended particles, the more difficult for light to travel through the water and therefore, the higher the water's turbidity . Apart from this with increasing amount of dissolved inorganic matters in water bodies causes the peak of visible reflectance to shift from the green region toward the red region of the electromagnetic spectrum. The presence of complex substances in water also change the reflectance of the water body and cause variation in colors, so it becomes necessary to calculate a number of bands ratios to identify the concentrated materials because available literature confirmed that no single band can be used with high confidence to perform an appropriate model to measure the reflectance of water resulting from DO, COD, and BOD. Keeping this in view in the present study the conversion of image radiance values into reflectance values and calculations of different band ratios along with assessment of correlation Water quality index were made (Table-7).

Table VII **:** Correlation between Spectral reflectance values and WQI

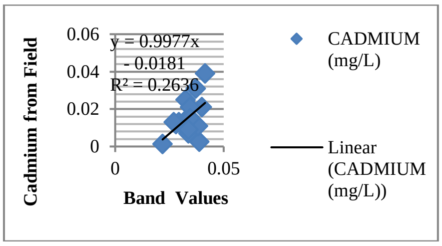
Sampling Points	**B1**	**B2**	**B3**	**B4**	**B5**	**B6**	**B7**	**WQI**
1	0.174	0.151	0.141	0.133	0.183	0.113	0.081	1.643051
2	0.181	0.159	0.133	0.131	0.175	0.169	0.144	4.2461
3	0.177	0.155	0.126	0.108	0.216	0.165	0.107	4.41736
4	0.181	0.161	0.137	0.107	0.059	0.018	0.01	2.88768
5	0.186	0.165	0.142	0.112	0.062	0.02	0.01	4.09581
6	0.181	0.16	0.134	0.109	0.079	0.043	0.024	1.78361
7	0.179	0.156	0.128	0.106	0.136	0.087	0.05	5.50885
R Square	**0.06598**	**0.021934**	**0.328519**	**0.103301**	**0.044133**	**0.084982**	**0.059049**	

Subsequently correlation- regression analysis have been done to find the relationship between the independent variables or the image data with the dependent variables or the sample tested data (Figure-2, A-H).

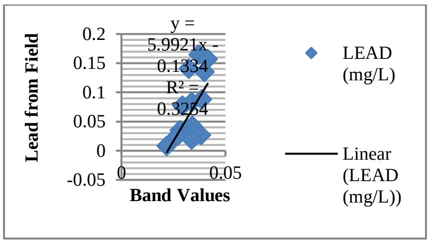
Figure 2 : Relation between sample parameters and Landsat 8 image reflectance



IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Vol. 8, Issue 6, June 2021
DOI: 10.17148/IARJSET.2021.8612



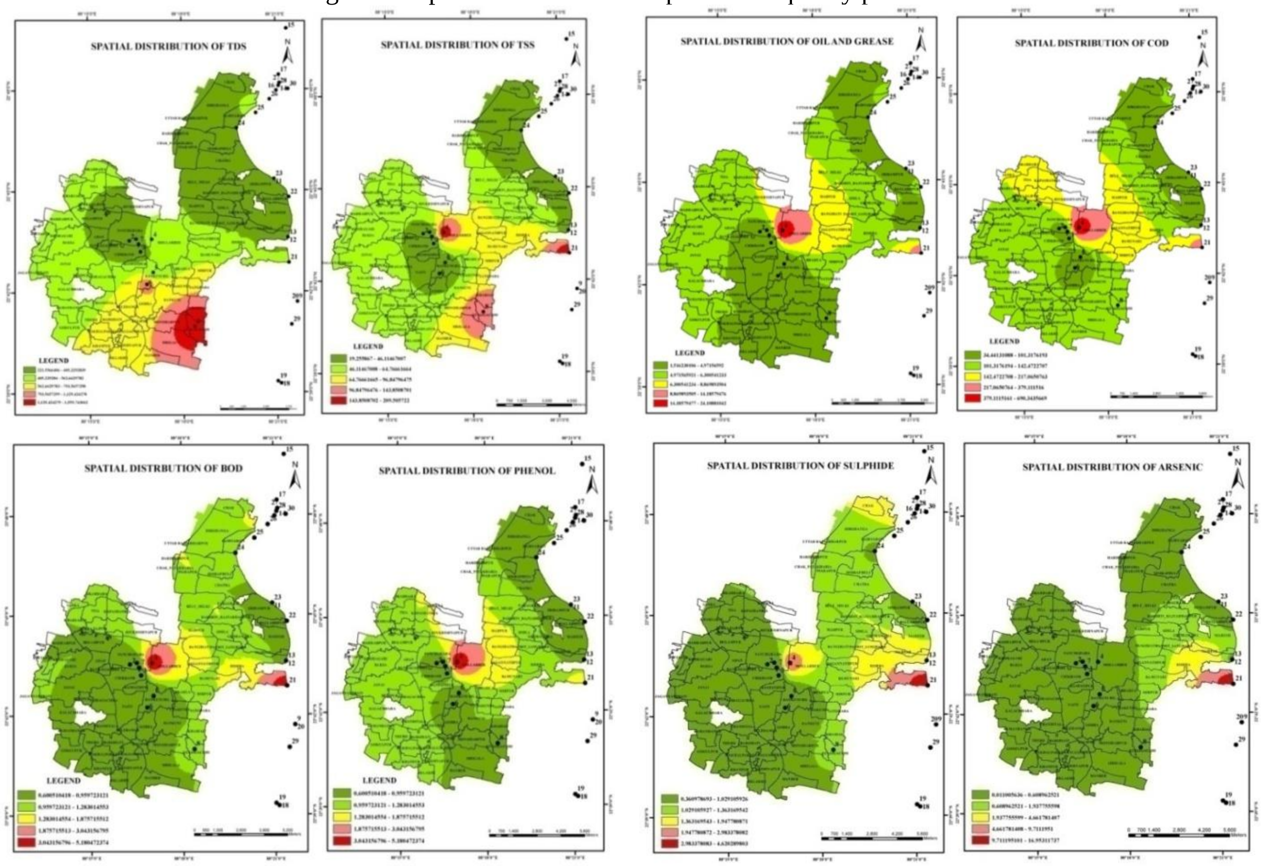
G



H

The analysis shows there is a positive correlation between the parameters like TSS, Oil & Grease, COD, BOD, Sulphide, Ammonium, Cadmium and Lead with the first four bands of Landsat 8 data whereas for the rest of the parameters it was negative. Finally, based on this relationship spatial distribution of the parameters were measured and subsequent mapping (Figure- 3) were done using IDW method at Arc GIS 10.3 platform for future environmental management and planning purposes.

Figure 3: Spatial Distribution Map of water quality parameter



V. CONCLUSIONS

The objective of this study was to assess the potential application of Geo-spatial techniques for measuring water quality status in an industrial belt along with statistical and mathematical methods. Analysis of Variance (ANOVA) and P-value analysis were also good predictors for finding the significance of the parameters. The analysis of One-way ANOVA test for significant difference shows that the samples don't belong to the same population (F=23.5133>$Fcrit$ = 1.523774). Therefore, null hypothesis is rejected, all the parameters have significant differences. The obtained p-values also ascertained that there is significant differences among the physico-chemical characteristics of the parameters at all the sampling points. Calculation of Water Quality Index became very useful to assess the overall water quality of the area. Subsequently, the creation of band ratios coupled with correlation- regression models proven the way in establishing the correlation between image spectral value (independent variables) and the sample tested data (dependant variables). As there was little difference between satellite-estimated values and measured concentrations this research demonstrates that Landsat 8 data may prove useful for identifying and demarcating the spatio-temporal variability of WQPs in surface waters. In the present study, band 2- 4 reflectance values displayed the strongest

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 6, June 2021

DOI: 10.17148/IARJSET.2021.8612

relationship with WQPs. Finally the Spatial Distribution Map of each and every parameter suggests that water quality treatments are required to keep the area clean and healthy. The algorithm and methods applied in this study could be applied to other industrial areas also for better environmental management and planning purposes.

ACKNOWLEDGEMENT

The Authors are thankful to the Additional Chief Secretary , Department of Science and Technology and Biotechnology, Government of West Bengal for providing all kinds of facilities to carry out projects related to this subject matter. Authors acknowledge the immense help received from the scholars and authorities who were directly or indirectly involved with this work. The authors are really grateful to authors/ editors/ publishers/ of all those articles, journals and books from where various information regarding this article has been taken and also included in references of this manuscript.

REFERENCE

1. Zhou F et al. (2007a) . Spatio-temporal patterns and source apportionment of coastal water pollution in eastern Hong Kong. Water Research, 41(15): 3429–3439.

2. Jonathan et al. (2008). A re-assessment of indicators of national water scarcity. Water International 33(1):5-18.

3. Mohammad Haji Gholizadeh et al.(2016). A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques , www.mdpi.com/journal/sensors, 16 (8):1298.

4. S. Chander et al., (2019).Water quality assessment of River Ganga and Chilika lagoon using AVIRIS-NG hyperspectral data. Current Science 116(7):1172-1181.

5. Crosa et al. (2006) .Spatial and seasonal variations in the water quality of the Amu Darya River (Central Asia). Water Research, 40(11):2237-45.

6. Sarkar et al. (2007). Water quality management in the lower stretch of the river Ganges, east coast of India: an approach through environmental education . Journal of Cleaner Production ,15 ,1559-1567.

7. Zhou F et al. (2007a) . Spatio-temporal patterns and source apportionment of coastal water pollution in eastern Hong Kong. Water Resour, 41(15): 3429–3439.

8. Verdin, J., 1985. Monitoring Water Quality Conditions in a Large Western Reservoir with Landsat imagery. Photogrammetric Engineering And Remote Sensing, Vol. 51, No. 3, 343-353.

9. S. Tassan (1993). An improved in-water algorithm for the determination of chlorophyll and suspended sediment concentration from Thematic Mapper data in coastal waters. International Journal of Remote Sensing Volume 14, Issue 6.

10. Arenz et al. (1996). Determination of chlorophyll and dissolved organic carbon from reflectance data for Colorado reservoirs. International Journal of Remote Sensing ,17, 1547–1566.

11. Dewidar et al.(2005). Remote Sensing of Water Quality for Burullus Lake, Egypt. Geocarto International , Vol. 20, Issue 3.

12. He, W. et al.(2008). Water quality monitoring in a slightly-polluted inland water body through remote sensing—Case study of the Guanting Reservoir in Beijing, China. Front. Environ. Sci. Eng. China ,2, 163–171.

13. Kallio, K et al. (2008). Landsat ETM+ images in the estimation of seasonal lake water quality in boreal river basins. Environmental Management. 42, 511–522.

14. APHA (2005). Standard Methods for the Examination of Water and Wastewater. 21st Edition, American Public Health Association/American Water Works Association/Water Environment Federation, Washington DC.

15. N. Das. (1976): Statistical Method, Volume – II, Tata McGraw – Hill Publishing Company Ltd. New Delhi.

16. Spiegel, M.R et al. (2010): Statistics, Tata McGraw – Hill Publishing Company Ltd. New Delhi.

17. Bharti, N et al. (2011). Water quality indices used for surface water vulnerability assessment. International Journal of Environmental Sciences, 2, 154.

18. Deepshikha Sharma et al. (2011). Water quality analysis of River Yamuna using water quality index in the national capital territory, India (2000–2009). Appl Water Science , 1:147–157.

19. Sahoo, M.M. et al.(2015). Inference of Water Quality Index Using ANFIA and PCA. Aquat. Procedia, 4, 1099–1106.

20. Liu et al (2003). Quantification of shallow water quality parameters by means of remote sensing. Progress in Physical Geography, 27,1 , 24–43.