Nep J Environ Sci (2023), 11(1), 27-36
https://doi.org/10.3126/njes.v11i1.53406
ISSN 2350-8647 (Print) 2542-2901 (Online)

Research Article

Forecasting Seasonal TDS using a Normalized Difference Water Index in the Ganges River in Bangladesh

Marzana Rahman Khuku, Golam Shabbir Sattar, Md. Golam Mostafa*
Institute of Environmental Science, University of Rajshahi, Rajshahi-6203, Bangladesh
(Received: 20 March 2023; Revised: 06 August 2023; Accepted: 14 August 2023)

Abstract
The Ganges is one of the most remarkable rivers in Bangladesh. The river contributes to the surface water supply, which is significantly high in the northwest region of Bangladesh. The water quality of this river changed based on many environmental events, and rainfall was one of them. The research focused on finding the mathematical relation between the field observations of water quality parameters and the observed data from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensors. This study treated the Total Dissolved Solids (TDS) as one of the crucial parameters in the water quality index, which was monitored and forecasted using multiple linear regression from satellite-derived Normalized Difference Water Index (NDWI) from the Landsat imageries and Electrical Conductivity (EC) from the field observations for the pre-monsoon, monsoon, and post-monsoon periods. The predicted values returned significant correlation coefficient values compared to the field TDS values of 0.64, 0.71, and 0.52 for the three periods, respectively. The regression equations found from this study may be further utilized to predict the TDS values in the future years.

***Keywords*:** Moderate resolution imaging spectroradiometer (MODIS), multiple regression, TDS, the Ganges

Introduction
The rivers in Bangladesh are mostly from the Himalayas in Nepal. The Ganges is one of the most remarkable transboundary rivers in Bangladesh. The Ganges is the main hydrodynamic system that created the greatest delta in the world, which is in huge portions of West Bengal in India and Bangladesh (BWDB, 2011). The river migrated southeast during the extensive construction of the Ganges Delta, arriving at its current location in the Bengal lowlands. The mighty Ganges and the fluvio-hydrological environment of the Bengal Basin are closely tied to the hydrology and drainage systems of the Ganges Delta in southwestern Bangladesh. The Ganges and its numerous distributaries and tributaries have had a significant role in the formation of the delta for a very long time. The Ganges-Padma-Meghna system's deltaic estuaries drain the combined flows of these rivers, which average 35,000 cumec. However, the Ganges' discharge swells to 76,000 cumec during the monsoon, increasing its silt burden in the process (Banglapedia, 2012). The Ganges and its numerous distributaries and tributaries have had a significant role in the formation of the delta for a very long time. The water quality of this river may change based on many environmental events such as river erosion, large-scale sediment transport, rainfall, and so forth. The river's low-level discharge is on the order of 15,000 cumec, and during this time the river naturally carries very little silt (Rahman et al., 2019). The river's width in the deltaic region fluctuates from 1.6 to 8 km, and while it is a meandering channel, it occasionally exhibits a braided nature (Ji & Gong, 2018).

The river's water quality has gradually declined as a result of pollution from numerous point and non-point sources brought on by rising living standards, unplanned urbanization, and rapid industry (Islam & Mostafa, 2022a, 2022b, 2022c). Untreated water is nevertheless used for a variety of reasons despite the worrisome amount of contamination, which could influence human health (Rafiqul & Mostafa, 2022). In order to raise public awareness and assist planners and government authorities in managing and conserving water bodies, water quality evaluation is crucial as the first step. In its simplest form, measuring water quality by quantifying numerous physicochemical parameters is exceedingly difficult to understand and only loosely approximates the situation at hand. In order to properly integrate all the elements and reflect water quality through a single numerical value, an evaluation tool is needed. The water quality index (WQI) (Smith, 1990), one of many evaluation techniques used for water quality study, is the most widely used indexing tool that incorporates all the criteria and compares them to the requirements advised by government agencies to protect human health(Harkins, 1974; Sutadian et al., 2015; Islam & Mostafa, 2021a, 2021b, 2021c). Total Dissolved Solids (TDS) is one of the vital parameters used to evaluate the water quality index of any stream water. TDS are calculated as the mass of residue left over after evaporating a specified amount of filtered water (Abbasi & Abbasi, 2012).

Corresponding author: mgmostafa@ru.ac.bd

Total solids in water are made up of dissolved solids along with suspended and settleable solids Calcium, chlorides, nitrate, phosphorus, iron, sulfur, and other ions are among the dissolved solids in stream water that would flow through a filter with holes as small as 2 microns (0.002 cm). Plankton, algae, fine organic waste, silt and clay particles, and other particulate materials are examples of suspended solids. These are particles that a 2-micron filter will not be able to capture. The water balance in the cells of aquatic organisms is influenced by the concentration of total dissolved solids (Uddin et al., 2015).

A water body can "stick out" against the ground and vegetation thanks to the Normalized Difference Water Index (NDWI) (Gao, 1996), which is used to emphasize open water features in satellite images.

Materials and Methods

Study area and data collection

In this study, water quality parameters of the river Ganges were utilized. The river Ganges enters Bangladesh through the Chapai Nawabganj district and then continued through Rajshahi and Pabna district. In Figure 1, the initial and ending location of the study area were illustrated. Two types of data were collected to forecast the TDS of the river Ganges from Godagari, Rajshahi (24.45, 88.31) to the Pakshey, Pabna (24.07, 89.03). Field data was collected from 80 points in the river from Godagari to Pakshey. The field data was collected for three different seasons for Bangladesh in the pre-monsoon (March-May), monsoon (June- September), and post-monsoon (October-February).

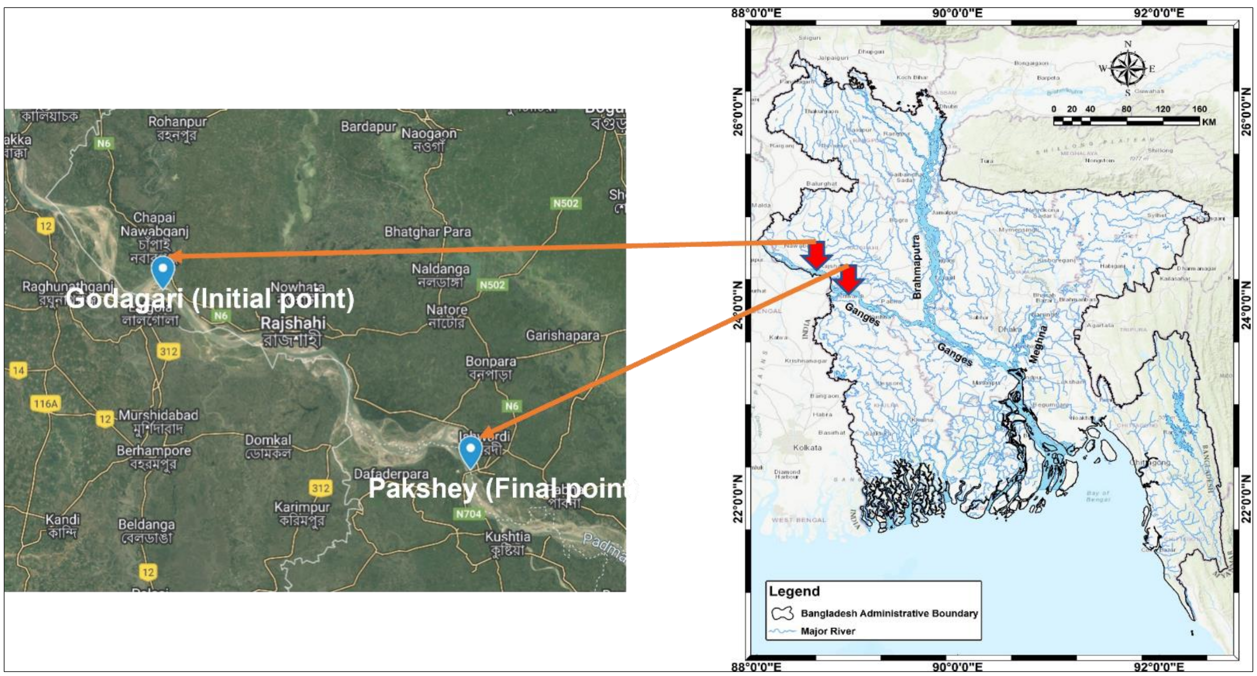

Figure 1 Initial and Final point of the study area

McFeeters first introduced the NDWI index in 1996. Its main function today is to identify and keep track of minute changes in the water content of the bodies of water. The near-infrared (NIR) and visible green (GREEN) spectral bands are used by the NDWI to enhance water bodies in satellite images (Acharya et al., 2018; Chen et al., 2020; Masocha et al., 2018; Mirchooli et al., 2020). The index has a drawback in that it can overestimate water bodies because it is sensitive to constructed buildings.

The NDWI makes effective use of this characteristic to identify water bodies on a map and track turbidity. Maps are used to depict how values change over time as a curve, and graphs are used to show how data taken from a satellite picture using the NDWI index is represented. On a map, the higher values that are getting close to +1 typically display blue and represent either a high-water content or a water surface, whereas the lower values that go all the way to -1 represent drought conditions, unless the area of interest is a non-aqueous surface (Raman et al., 2022). Table 1 includes the NDWI value interpretation.

Table 1 Interpretation of NDWI value range

NDWI Value	Interpretation
0.2 to 1.0	Water surface
0.0 to 0.2	Flooding, Humidity
-0.3 to 0.0	Moderate drought, non-aqueous surfaces
-1 to -0.3	Drought, non-aqueous surfaces

The Normalized Difference Water Index (NDWI) was calculated from the Landsat 8 data, downloaded from the United States Geological Survey (USGS) Earth Explorer website. The NDWI value was determined using the following formula for the Landsat 8 the OLI and TIR sensor data,

$$NDWI = \frac{Band\ 3 - band\ 5}{Band\ 3 + Band\ 5}$$

The OLI sensor offers data with a spatial resolution of 30 m in eight bands covering the visible, near-infrared, and shortwave infrared spectrums, as well as an additional panchromatic band with a resolution of 15 m. Using two bands situated in the atmospheric window between 10 and 12 m, the TIRS detects Thermal Infrared (TIR) light with a spatial resolution of 100 m. Table 2 lists the band designations for Landsat 8 (USGS, 2019).

Table 2 Landsat 8-9 operational land imager (OLI) and thermal infrared sensor (TIRS) descriptions (Barsi et al., 2014)

Bands	Wavelength (micrometers)	Resolution (meters)
Band 1 - Coastal aerosol	0.43-0.45	30
Band 2 – Blue	0.45-0.51	30
Band 3 – Green	0.53-0.59	30
Band 4 – Red	0.64-0.67	30
Band 5 - Near Infrared (NIR)	0.85-0.88	30
Band 6 - SWIR 1	1.57-1.65	30
Band 7 - SWIR 2	2.11-2.29	30
Band 8 – Panchromatic	0.50-0.68	15
Band 9 – Cirrus	1.36-1.38	30
Band 10 - Thermal Infrared (TIRS) 1	10.6-11.19	100
Band 11 - Thermal Infrared (TIRS) 2	11.50-12.51	100

Multiple regression

Multiple regression is a widely used technique for different kinds of data prediction. By employing dummy variables with values of 0 or 1, depending on whether a given characteristic is true, the multiple regression model can handle categorical variables. The least squares method, like simple regression, identifies the coefficient estimates that reduce the total squared prediction errors (Freund et al., 2010).

The methods for estimating the coefficients, determining the error variance, and drawing conclusions about the model parameter are all discussed. There is also a brief explanation of correlations, which expresses the strength of linear relationships between various variables. Regression models include more independent variables than are required to get a complete picture of the data. By choosing a subset of independent variables to be included in the model, the issue of an excessive number of independent variables can be resolved (Smith, 2015).

Mathematically,

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_nX_n \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where

$$\hat{Y} = Predicted\ Variable$$
$$b_0 = Intercept\ at\ Y\ axis$$
$$b_1, b_2, b_3, \ldots\ldots\ldots, b_n = Regression\ Coefficients$$
$$X_1, X_2, X_3, \ldots\ldots\ldots\ldots\ldots\ldots, X_n = Independent\ variables$$

The multiple regression was carried out in R programming software which is a renowned statistical environment for data analysis and statistical interpretation (R Core Team, 2016).

Results and Discussion

The values of TDS exposed an upward trend from upstream to downstream in three specified periods in this study (Fig. 2). In the pre-monsoon period, based on the observed values of TDS, the trend was observed as increasing as the river went to the downstream region. The minimum value of TDS was recorded as 100 mg/L, and the highest value was 490 mg/L. The mean TDS was 325.25 mg/L. In the predicted model, the minimum TDS was forecasted as 160 mg/L and the highest TDS was found 530 mg/L, while the mean TDS was 342 mg/L. Moreover, 41 stations showed higher values than the average for the observed TDS and 39 stations showed values below the average. Additionally, in the case of the predicted values of TDS, 37 stations showed more than the average values and 43 stations showed values below the average.

During the monsoon period, the tendency was seen to be rising as the river moved into the downstream zone based on the observed TDS readings. TDS values ranged from 250 mg/L at the lowest end to 590 mg/L at the highest. TDS on average was 432.25 mg/L. The mean TDS was 432.25 mg/L, while the anticipated lowest and maximum TDS values were 285 mg/L and 585 mg/L, respectively. Additionally, 40 stations displayed values below the average and 40 stations had values above the average for the recorded TDS. Additionally, 41 stations displayed values below the average and 39 stations had values above the average for the projected TDS values.

In the post-monsoon period, the trend was seen as growing as the river moved into the downstream zone based on the recorded TDS values. TDS measurements ranged from 10 mg/L at the lowest end to 220 mg/L at the maximum. The average TDS was 142 mg/L. In the anticipated model, the lowest TDS was predicted to be 55 mg/L and the highest TDS was discovered to be 205 mg/L, with a mean TDS of 142 mg/L. In addition, 32 stations had values below the average for the studied TDS, whereas 48 stations displayed higher values. Additionally, 36 stations exhibited readings below the average and 44 stations displayed more than the average TDS values. While predicting, the R-squared values were found 0.98, 0.96, and 0.90 in the pre-monsoon, during monsoon and the post-monsoon period respectively. In the prediction, the upward trend was also observed for all the three seasons illustrated in the Figure 3.

6

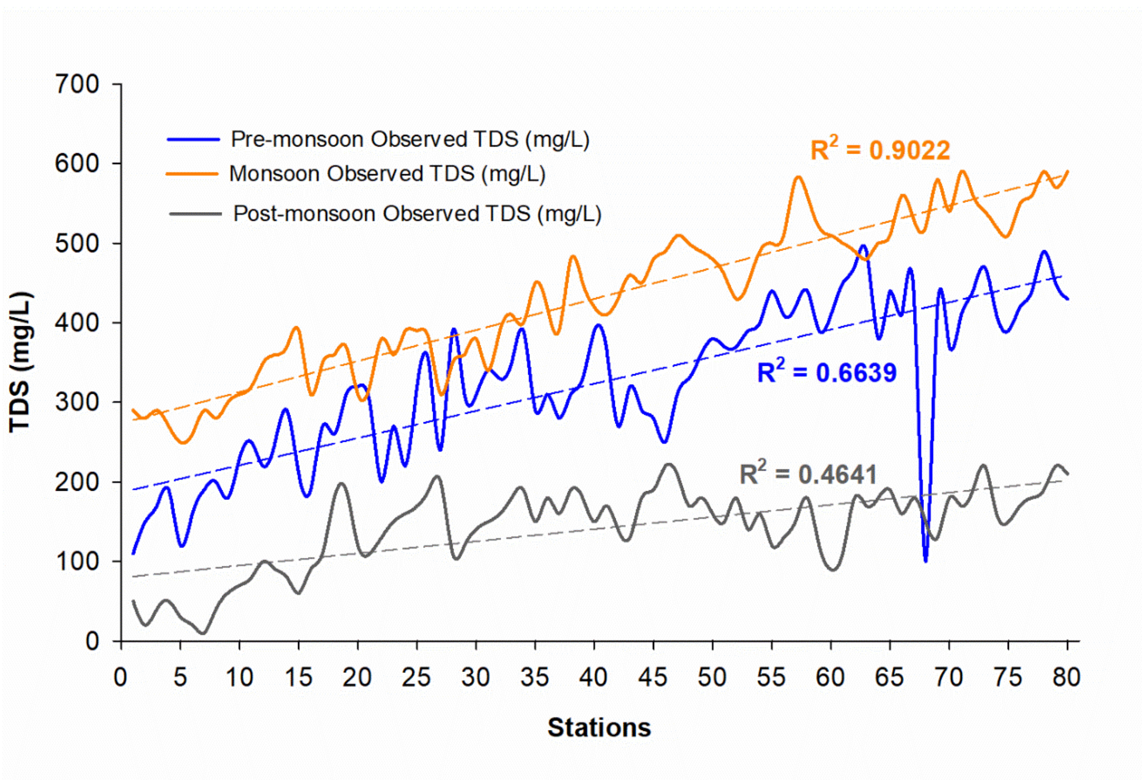

Figure 2 Trend of observed TDS values (mg/L) in different seasons

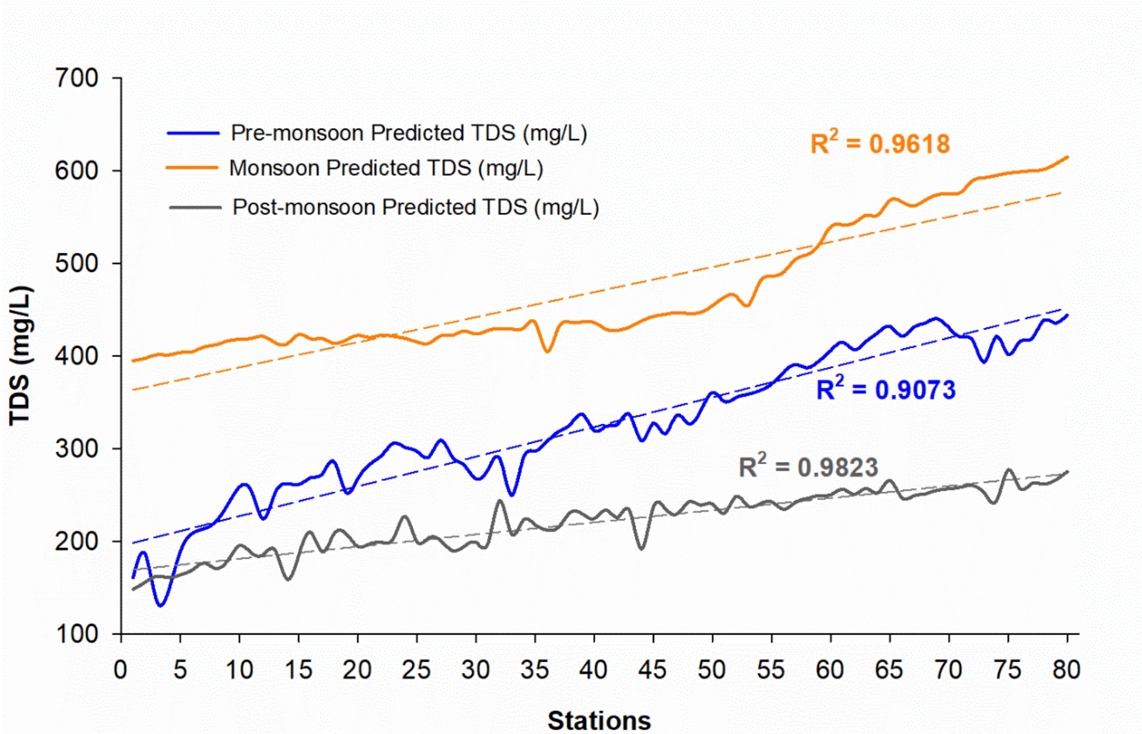

Figure 3 Trend of predicted TDS values (mg/L) during the pre-monsoon, monsoon, and post-monsoon periods

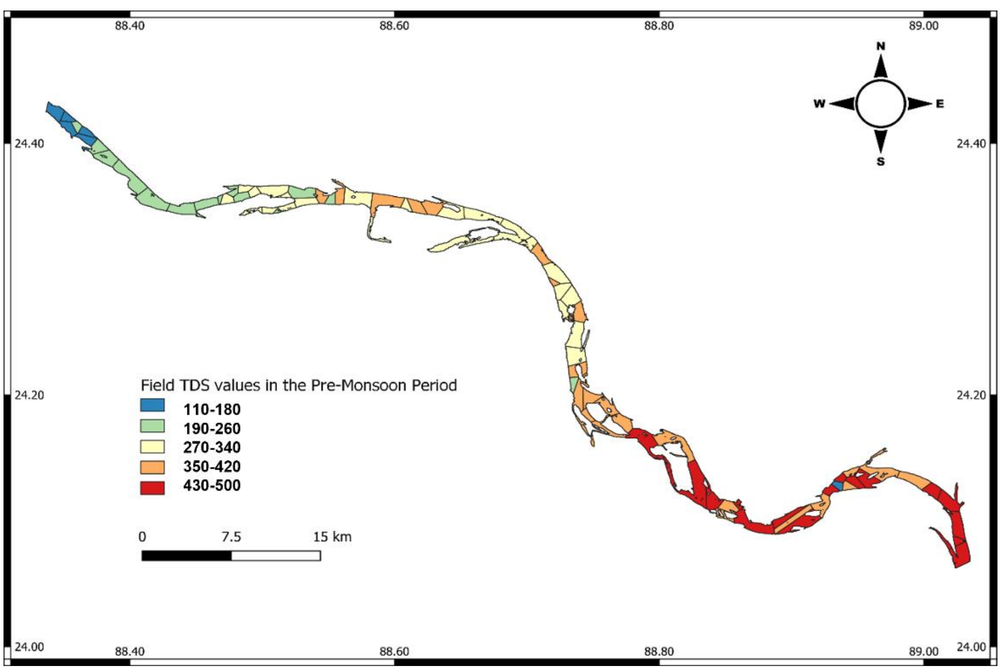

Figure 4 Observed TDS in the pre-monsoon period (in mg/L)

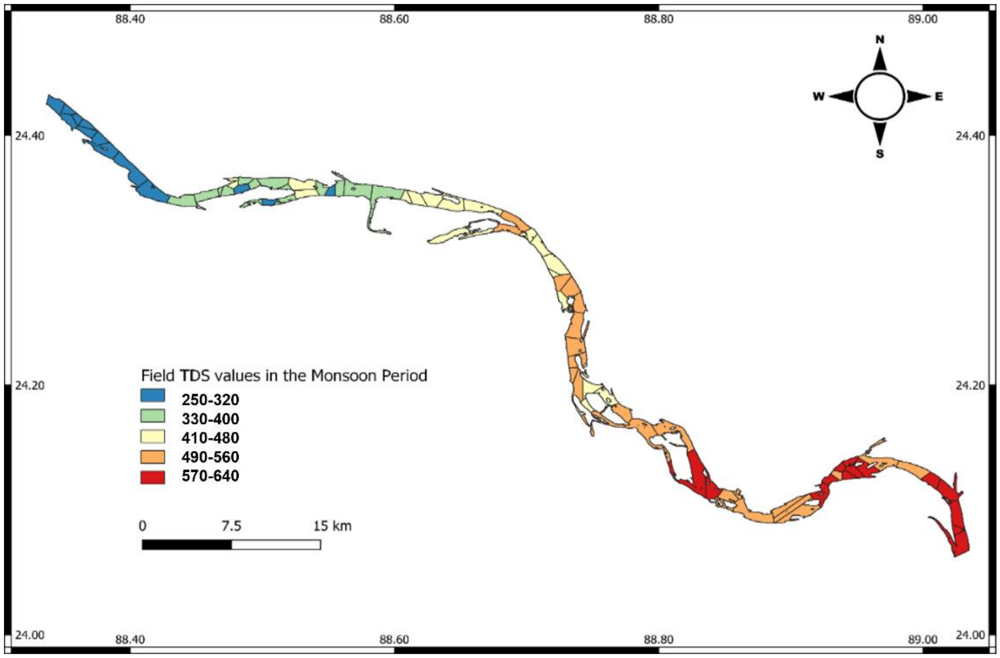

Figure 5 Observed TDS in the monsoon period (in mg/L)

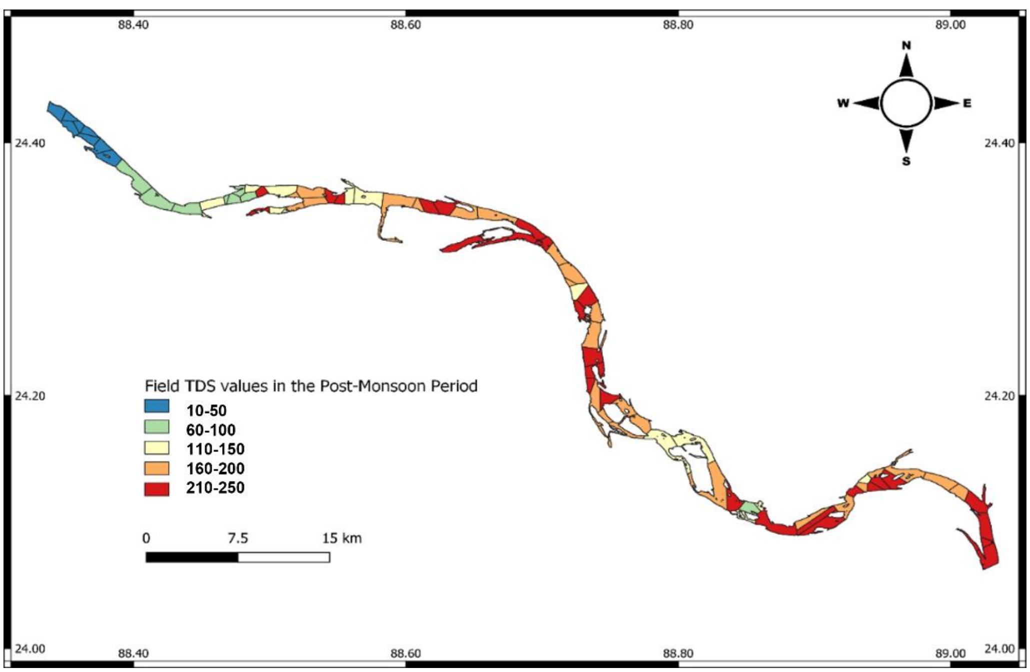

Figure 6 Observed TDS in the post-monsoon period (in mg/L)

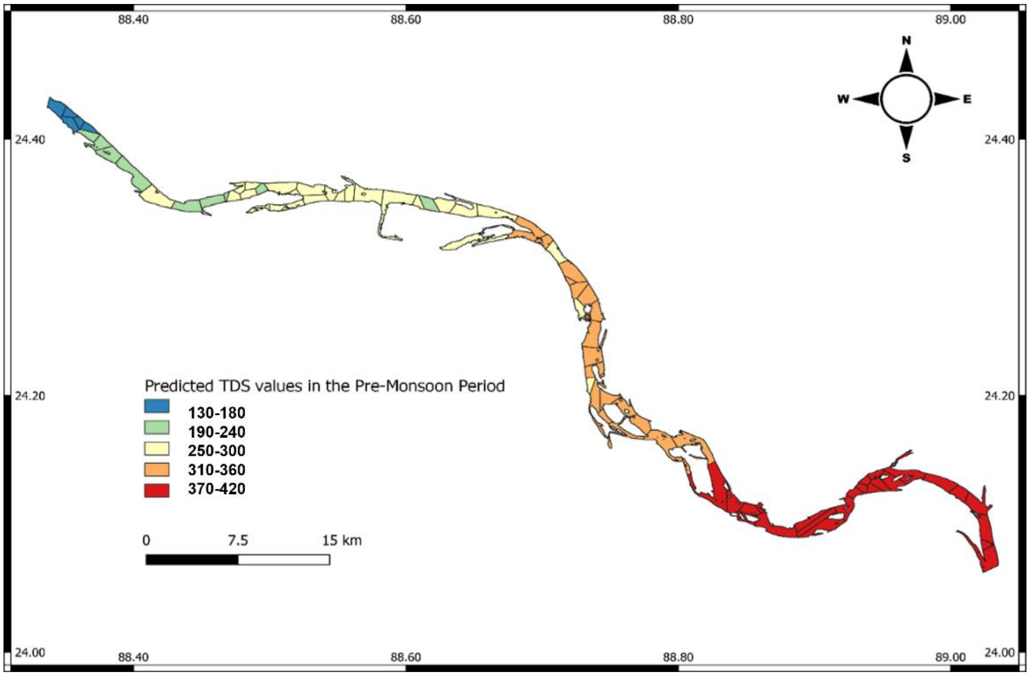

Figure 7 Predicted TDS in the pre-monsoon period (in mg/L)

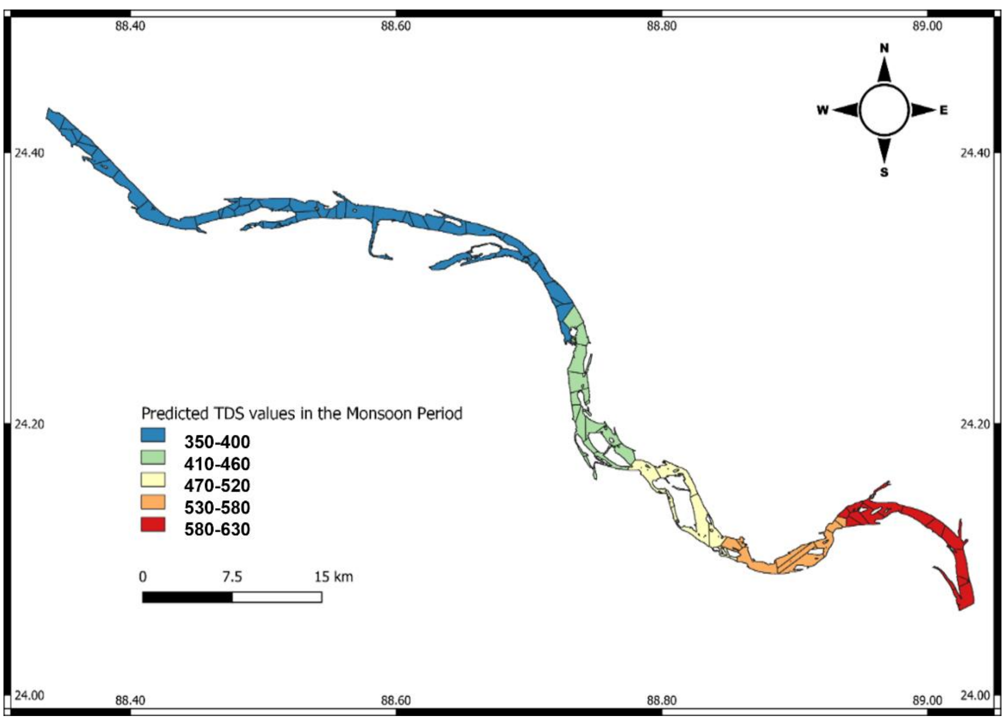

Figure 8 Predicted TDS in the monsoon period (in mg/L)

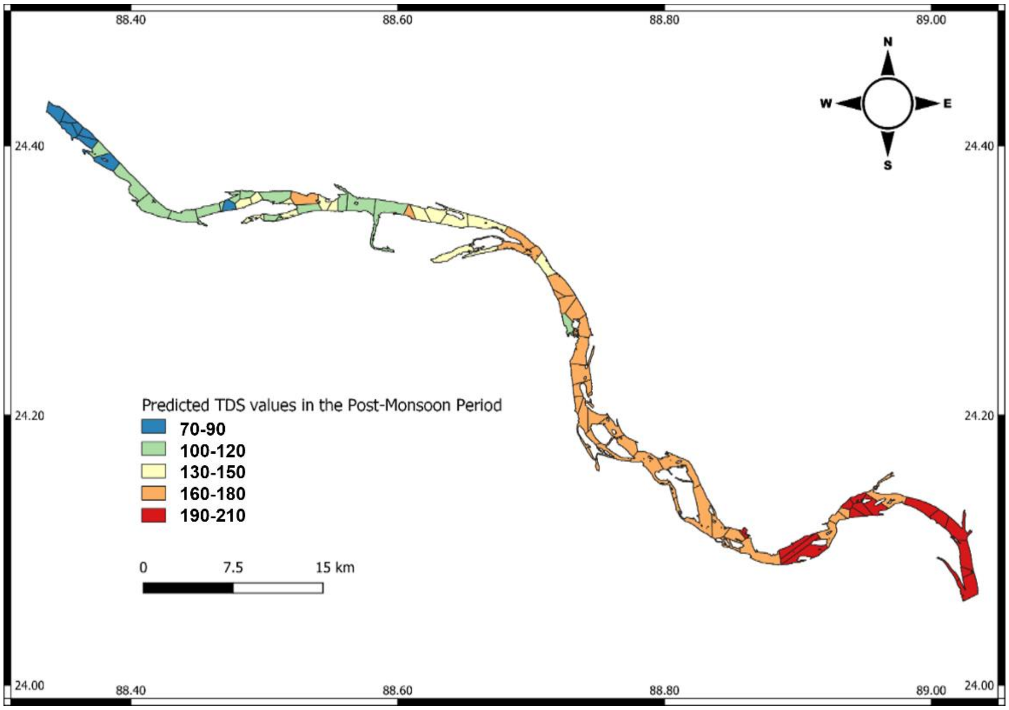

Figure 9 Predicted TDS in the post-monsoon period (mg/L)

The values of multiple R-squared and adjusted R-squared indicate the regression accuracy and the model accuracy respectively in Table 3. The multiple regression equations found from the models are:

$$TDS = -0.409 + 2.6 * EC - 0.074 * NDWI$$
$$TDS = 0.055 + 1.012 * EC - 0.060 * NDWI$$
$$TDS = -0.323 + 1.919 * EC + 0.135 * NDWI$$

Table 3 Model accuracy for three different seasons

Season	Multiple R-squared	Adjusted R-squared
Pre-Monsoon	0.64	0.63
Monsoon	0.71	0.70
Post-Monsoon	0.52	0.51

In this study, a remotely sensed water index (NDWI) and a physical parameter of water (Electrical Conductivity) were used to predict another useful parameter, Total Dissolved Solids (TDS) in the three different seasons based on the rainfall events. A linear multiple regression model was applied in the three phases for this study.

Among the three regression models, the monsoon period showed consistent accuracy both in R-squared value and adjusted R-squared value. Moreover, the highest accuracy was found in the monsoon period for the observed and predicted data (71%). In the pre-monsoon period, the predicted data showed 64% correlation with the observed values. Additionally, in the post-monsoon period, the association between the observed and predicted TDS values was found 52%. In the three periods, the adjusted R-squared values indicated the fitting accuracy of the respective models that were statistically significant.

The adjusted R-squared values were found 0.63, 0.70, and 0.51 for the pre-monsoon, monsoon, and post-monsoon period. By dividing the residual mean square error by the overall mean square error, adjusted R-squared is determined. Also, it shows how much of the variance in the target field is explained by the inputs. From this analysis, the model found in the pre-monsoon period, explained 63% variance of the predicted data. Similarly, 70% and 51% of variance of the modeled data were explained by the models from monsoon and post-monsoon period.

Table 4 Highest and Lowest values of TDS in the study area

	Pre-Monsoon		Monsoon		Post-Monsoon	
	Observed	**Predicted**	**Observed**	**Predicted**	**Observed**	**Predicted**
TDS (Maximum)	490 (Taltola)	530 (Pakshey)	590 (Saraghat)	585 (Pakshey)	220 (Khoyerhat)	205 (Pakshey)
TDS (Minimum)	100 (Godagari)	160 (Godagari)	250 (Ujanpara)	285 (Godagari)	10 (Premtali)	55 (Godagari)

Based on the data from Table 4, the locations containing highest and lowest values of the observed and predicted TDS values were illustrated. From the table, the downstream locations showed an upward trend in both parameters whereas the upstream contained lower values of TDS compared to the downstream region. Godagari, which was the initial location of the study area, showed minimum TDS value mostly over the year whether the monsoon period may exist or not. Premtali and Ujanpara also showed minimum TDS values along with the Godagari point. On the contrary, Pakshey showed the maximum TDS values in the three temporal phases described in this study. But the spatial variety based on the maximum TDS values was significant than the minimum TDS values. These locations were also illustrated in Figures 4 to 9 using qGIS software which is an open-source Geographic Information Software (GIS).

From this study, it is clear that the TDS values are going upwards from the upstream to the downstream side of the river Ganges in pre-monsoon, monsoon, and post-monsoon period. In the pre-monsoon period, the value of TDS is mostly average of the pre-monsoon and the post-monsoon period from Figure 2. Thus, it can be said that the downstream of the Ganges carries more TDS from the downstream than the upstream side. Moreover, the p-value found from the regression models were 0.002 (<0.05), 0.005 (<0.05), and 0.003 (<0.05) for the pre-monsoon, monsoon, and post-monsoon period respectively which make this study as a significant one.

Conclusions

The water quality parameters are highly variable concerning time and location. The remote sensing index, NDWI, was included along with the field data, so the surface water monitoring program would encompass both the field and space data. Thus, the empirical model would be used for specific points where the surface water stations are absent. In developing countries like Bangladesh, the construction of large-scale water gauge stations is highly recommended, despite their expensive maintenance costs. To resolve this problem, this study may be useful to a great extent. The existing water quality monitoring parameters can be merged to find out the TDS values in the nearest locations between adjacent water stations using the remote sensing data from MODIS or Landsat eight (8). The higher the number of sampling points, the higher the accuracy of the results. Moreover, the water quality parameters may change along the banksides and in the middle of the river. Thus, the data can be categorized into two types: the water quality near the bankside and in the middle of the river. The maximum number of quality parameters should be considered for a more effective NDWI.

Acknowledgements

The authors would like to thank NASA for providing the Landsat 8 imageries from the United States Geological Survey (USGS) website. The first author would like to acknowledge the Institute of Environmental Science, University of Rajshahi, Bangladesh, for providing a fellowship.

Author Contributions: Marzana Rahman Khuku: Participated in the design and performance of the experiments, data analysis, and writing of the primary manuscript draft and revised draft preparation.
G.S. Sattar: Drafted the revision and finalized the manuscript.
M.G. Mostafa: conceptualized the design, analyzed the data, drafted the revision, and finalized the manuscript for submission.

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

Data Availability Statement: The data that support the finding of this study are available from the corresponding author, upon reasonable request.

References

Abbasi, T., & Abbasi, S.A. (2012). Water-Quality Indices: Looking Back, Looking Ahead. Elsevier. doi 10.1016/B978-0-444-54304-2.00016-6.

Acharya, T.D., Subedi, A., & Lee, D.H. (2018). Evaluation of Water Indices for Surface Water Extraction in a Landsat 8 Scene of Nepal. *Sensors (Basel, Switzerland)*, 18(8), 2580. doi 10.3390/s18082580.

Banglapedia (2012). National encyclopedia of Bangladesh. Retrieved March 8, 2023 from https://en.banglapedia.org/index.php/Ganges_River.

Barsi, J.A., Lee, K., Kvaran, G., Markham, B.L., & Pedelty, J.A. (2014). The spectral response of the Landsat-8 operational land imager. *Remote Sensing*, 6(10), 10232–10251. doi 10.3390/rs61010232.

BWDB. (2011). *Analysis of water level along the Brahmaputra-Jamuna, Ganga-Padma & Surma-Meghna River System*. Bangladesh Water Development Board.

Chen, F., Chen, X., Van de Voorde, T., Roberts, D., Jiang, H., & Xu, W. (2020). Open water detection in urban environments using high spatial resolution remote sensing imagery. *Remote Sensing of Environment*, 242, 111706. doi 10.1016/j.rse.2020.111706.

Freund, R.J., Wilson, W.J., & Mohr, D.L. (2010). Statistical Methods,(Third Edition), Academic Press. doi 10.1016/B978-0-12-374970-3.00008-1.

Gao, B. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3), 257–266. doi 10.1016/S0034-4257(96)00067-3.

Harkins, R.D. (1974). An objective water quality index. *Journal (Water Pollution Control Federation)*, 46(3), 588-591.

Islam, M.S., & Mostafa, M.G. (2021a). Exposure of Trace Metals and Their Effects on Human Health: A Case Study for Groundwater in Part of the Ganges Basin Areas, Intl. J. Wat. Resource. and Env. Sci., 10(1): 01-13.

Islam, M.S., & Mostafa, M.G. (2021b). Influence of chemical fertilizers on arsenic mobilization in the alluvial Bengal delta plain: a critical review. *Journal of Water Supply: Research and Technology-Aqua*, 70(7), 948. doi 10.2166/aqua.2021.043.

Islam, M.S., & Mostafa, M.G. (2021c). Development of an integrated irrigation water quality index ($IIWQ_{Index}$) model. *Journal of Water Supply*, 22(2), 2322- 2337. doi 10.2166/ws.2021.378.

Islam, M.S., & Mostafa, M.G. (2022a). Comparison of classical and developed indexing methods for assessing the groundwater suitability for irrigation. *Journal of Sustainable Agriculture and Environment*, 1(9), 226-239.

Islam, M.S., & Mostafa, M.G. (2022b). Suitability of water quality index methods for assessing groundwater quality in the Ganges River basin area. *H_2 Open Journal*, 5(2), 198–220. doi 10.2166/h2oj.2022.145.

Islam, M.S., & Mostafa, M.G. (2022c). Evaluation of hydrogeochemical processes in groundwater using geochemical and geostatistical approaches in the Upper Bengal Basin. *Geofluids*, 9591717. doi 10.1155/2022/9591717.

Ji, L., & Gong, P. (2018). Surface water. In S. Liang (Ed.), *Comprehensive Remote Sensing*. Elsevier, pp. 258–294. doi 10.1016/B978-0-12-409548-9.10360-4

Masocha, M., Dube, T., Makore, M., Shekede, M.D., & Funani, J. (2018). Surface water bodies mapping in Zimbabwe using landsat 8 OLI multispectral imagery: A comparison of multiple water indices. *Physics and Chemistry of the Earth, Parts A/B/C*, 106, 63–67. doi 10.1016/j.pce.2018.05.005.

Mirchooli, F., Sadeghi, S.H., & Khaledi Darvishan, A. (2020). Analyzing spatial variations of relationships between Land Surface Temperature and some remotely sensed indices in different land uses. *Remote Sensing Applications: Society and Environment*, 19, 100359. doi 10.1016/j.rsase.2020.100359.

Rafiqul Islam, M. & Mostafa, M.G. (2022). Adsorption kinetics, isotherms and thermodynamic studies of methyl blue in textile dye effluent on natural clay adsorbent. *Sustainable Water Resources Management*, 8, 52. doi 10.1007/s40899-022-00640-1.

Rafiqul Islam, M. & Mostafa, M.G. (2020). Characterization of textile dyeing effluent and its treatment using polyaluminum chloride. *Applied Water Science,* 10, 119. doi 10.1007/s13201-020-01204-4.

Rahman, K.S., Islam, Z., Navera, U.K., & Ludwig, F. (2019). A critical review of the Ganges water sharing arrangement. *Water Policy*, 21(2), 259–276. doi 10.2166/wp.2019.164.

Raman, R.K., Bhor, M., Manna, R.K., Samanta, S., & Das, B.K. (2022). Statistical and geostatistical modelling approach for spatio-temporal assessment of river water quality: a case study from lower stretch of River Ganga. *Environment, Development and Sustainability*, 25, 9963-9989. doi 10.1007/s10668-022-02472-7.

R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. Retrieved from https://www.R-project.org/

Smith, D.G. (1990). A better water quality indexing system for rivers and streams. *Water Research*, 24(10), 1237–1244. doi 10.1016/0043-1354(90)90047-A.

Smith, G. (2015). Multiple Regression. In S. Smith (Ed.), *Essential Statistics, Regression, and Econometrics* (Second Edition); Academic Press, pp. 301–337. doi 10.1016/B978-0-12-803459-0.00010-8.

Sutadian, A.D., Muttil, N., Yilmaz, A.G., & Perera, B.J.C. (2015). Development of river water quality indices—a review. *Environmental Monitoring and Assessment*, 188(1), 58. doi 10.1007/s10661-015-5050-0.

USGS. (2019). *Landsat 8 Data Users Handbook. The United States Geological Survey*. Retrieved November 8, 2022 from https://landsat.usgs.gov/documents/Landsat8DataUsersHandbook.pdf.

Uddin, M., Alam, M., Mobin, M., & Miah, M. (2015). An Assessment of the River Water Quality Parameters: A case of Jamuna River. *Journal of Environmental Science and Natural Resources*, *7*(1), 249–256. doi 10.3329/jesnr.v7i1.22179.