**Biospectra : ISSN: 0973-7057, Vol. 19(2), September, 2024, pp. 169-178**
*An International Biannual Refereed Journal of Life Sciences*

**Plant Science**

ISSN : 0973-7057

# Detection of water quality degradation of Punpun River, Patna using remote sensing and google earth engine

**Ashok Kumar Roy[a], Shelu Kumari[b]*, Ravi Ranjan Kumar[c], Ishita[c] & Abhinav Chauhan[d]**

[a]Department of Botany, P. C. Vigyaan Mahavidyalaya, J. P. University, Chapra, Bihar, India

[b]Department of Botany, J. D Women's College, Patliputra University, Patna, Bihar, India

[c]Department of Botany, T.P.S. College, Patliputra University, Patna, Bihar, India

[d]University Department of Botany, Patliputra University, Patna, Bihar, India

*Received : 25th May, 2024 ; Revised : 25th June, 2024*

*DOI:-https://doi.org/10.5281/zenodo.15106344*

**Abstract-** Water quality parameters are the most important indicators of water quality in inland water systems. Maintaining systems for monitoring physicochemical parameters is time-consuming and cost-intensive, as developing appropriate river management plans requires in-situ water quality data with high spatial and temporal resolution. In this study, we used 10-m Sentinel-2 pictures to map the spatial changes in the Punpun River's water quality. We used spectral predictors obtained from the satellite pictures to train one machine learning algorithm, Random Forest (RF), to predict concentrations of pH, DO, chlorophyll, BOD, COD, TSS, and turbidity. In addition, we computed a number of metrics to evaluate the accuracy of the water quality maps and the performance of the models, such as Mean Squared Error (MSE), Coefficient of determination (R2), and Root Mean Squared Error (RMSE). The modelled and measured concentrations of pH, DO, chlorophyll, BOD, COD, TSS, and turbidity exhibited good agreement with minor residual errors ranging between 0.201 mg/L and 0.241 mg/L, according to our results. Additionally, bands 5 (B5, vegetation red edge) and 8 (B8, NIR) were found to be significant predictors of parameter concentrations, and RF was found to be a dependable and efficient algorithm for doing so. The Punpun River had good to exceptional concentrations of pH, DO, chlorophyll, BOD, COD, TSS, and turbidity. The Punpun River's current state of water quality and the effectiveness of the management measures implemented to control and prevent eutrophic issues have been spatially illuminated by our findings.

**Key words: Punpun river, Water quality, Random Forest, Remote sensing**

## INTRODUCTION

Water is one of the most valuable resources on which all life depends. Water pollution degrades water quality and affects the health of marine life and therefore the people who use it. Therefore, it is crucial to monitor water quality and ensure the survival of marine life.[1] Understanding water quality concerns and issues is also critical to curbing and controlling water pollution. To assess the state of the nautical ecosystem, several governments around the world have begun to develop ecological water management programs. Approximately one billion people do not have access to clean drinking water, and two million people die every year as a result of contaminated water and poor sanitation and cleanliness. Therefore, maintaining freshwater quality is crucial.[2] Water quality is critical to the long-term viability of a

*Corresponding author :
Phone : 9308185460
E-mail : shelubhagat1981@gmail.com

diversion plan. Poor water quality can also be costly as resources must be directed toward repairing water supply infrastructure as soon as a problem occurs. To ensure safe drinking water at a reasonable cost, the demand for improved water management and water quality control has increased. To address these issues, systematic assessments of freshwater, disposal systems, and organizational monitoring issues are required.[3]

It is common knowledge that clean water is essential for a healthy life. Across the world, freshwater resources are threatened not only by over-exploitation and poor management, but also by ecological degradation. The main cause of freshwater pollution is the discharge of untreated waste, dumping of industrial wastewater and runoff from agricultural fields. Industrial growth, urbanization and the increasing use of synthetic organic substances are having serious and negative impacts on freshwater bodies. Freshwater resources are under severe and increasing environmental stress. The problem of freshwater pollution in India came to the fore in the early 1960s, with domestic sewage, sewage, sewage and industrial effluent discharges being the major sources of pollution in various cities. In recent years due to continuous population growth, rapid industrialization and various waste management technologies. The human factor is the main cause of river pollution. In Patna, Punpun River is the main source of freshwater water. In recent decades there has been great concern about the deterioration of water quality. The river has been polluted during its course, particularly between Madhopur and Patna urban areas, by indiscriminate dumping of domestic sewage, immersion of idols, dumping of plastics into drains, bathing of cattle and agricultural effluents as well as partially treated and untreated sewage in large quantities recorded. The pollutants flowing into the river come from the waste of the villages located on its banks.

The effects of severe hypoxia (e.g. oxygen levels below 2 mg/L) on mortality and population of aquatic species are devastating.[4] Decades of scientific research have highlighted the long-term impacts of reduced water quality on the operation of healthy water systems.[2,5,6] For example, empirical modelling results suggest that fish biases and anomalies, reduced population sizes and structure of various species (e.g. insects, macro invertebrates) may be related to the Punpun River.[7] As river characteristics change, it becomes increasingly necessary to maintain continuous spatiotemporal assessment of water quality variability to rapidly manage changing inland water quality. However, effective management requires comprehensive observations, integrated analysis, improved monitoring and prediction of water quality, and a synoptic view of numerous river locations simultaneously. To minimize the need for costly and time-consuming field monitoring, better technologies are required that are beneficial to resource managers.[8] Fortunately, remote sensing (RS) has been recognized as an ideal solution for monitoring water quality in freshwater systems and has demonstrated successful applications.[9] Remote sensing provides consistent observations, a synoptic perspective, and instant assessment of temporal land use patterns at various spatial resolutions. For example, recent studies have used medium spatial resolution images such as the 30 m Landsat[10] and the 10 m Sentinel-2[11] to obtain explicit spatial distributions create water quality parameters (e.g. pH, DO, chlorophyll, BOD, COD, TSS and turbidity). However, estimating water quality parameters using RS images still remains a challenge. Spectrally speaking, water quality is a non-optically active parameter as it does not change the spectral properties of water through absorption.[12] The lack of wavelength absorption for parameters means that it cannot be directly quantified by spectral analysis. Therefore, estimation of water quality from RS images could be achieved indirectly by introducing additional indicator variables or spectral indices into the models used. Predictive models that use machine learning techniques on RS images to assess water quality parameters have gained popularity in recent years. These machine learning methods have been shown in several studies to satisfactorily approximate water quality parameters in aquatic environments.[5,10,15] There are a number of these machine learning models, but they do not include key predictors of water quality parameters. Furthermore, there are only a few publications that used a satellite image with a resolution of 10 m to describe fluctuations in water quality over different time points.

In the current study, we proposed to extend the use of Sentinel-2 images to extract and map the concentration of physiochemical parameters in the Punpun River section using a machine learning technique, Random Forest (RF). Our specific objectives include: (1) the development of models to estimate the concentrations of physicochemical parameters of water from Sentinel-2 images, (2) the identification of important predictive variables for physiochemical parameters of water that are essential, readily available and are easy to measure, (3) create spatial distribution maps for physiochemical parameters of water at different locations in the selected study area. The results obtained in this study would have significant implications for selecting the most appropriate machine learning algorithm and key predictors under RF. Finally, mapping water quality parameters in the Punpun River system, let alone using a 10 m resolution satellite image. This research is required to obtain a comprehensive overview of the water quality parameters at different river locations and report their current status.

## MATERIALS & METHODS

### Study area

The study area is a 10 km long stretch of the Punpun River in the southwest of Patna city (Fig. 1). The urban drains of Badshahi flow into the Punpun River. It originates from Palamu district of Jharkhand and flows through all or parts of four districts of Bihar including Greene, Chatra, Aurangabad, Gaya and Patna and joins the Ganga River. The study area is subtropical-continental with an average annual temperature of 59.6 F to 89.1 F. The average annual precipitation ranges from 14 to 45.1 inches and increases toward the south and about one-third of the precipitation becomes surface runoff.[16] Commercial fertilizers N (nitrogen) and P (P in the form of $P_2O_5$-phosphate and $K_2O$-potash) are widely used in the basin as agriculture is the dominant land cover. These commercial fertilizers from row crops, along with livestock manure, are the main sources of nutrients in surface and groundwater.

### *In-situ* field monitoring sites

We deployed three in situ water quality monitoring parameters at strategic locations along the Punpun River (Fig. 1). The sites were also near major wastewater discharges into rivers. The sampling sites were divided into three locations: upstream, middle stream and downstream. Upstream, the upper part of the river was in Madhopur village, under block Patna Sadar of Patna District, Bihar and the middle river - before the discharge of Badshahi urban sewage - and downstream - after the discharge of sewage into the river. We collected the water sample in a plastic tap water bottle and transferred it to the laboratory for further analysis. We analysed the water quality parameters according to standard protocols[17] in the T.P.S. College, Patna (Patliputra University) laboratory.


**Fig. 1. The Punpun River study area covers one sampling site and one Google Earth Pro image.**

### Satellite remote sensing dataset

Remote sensing is capable of covering large areas of the river and allows for faster temporal analysis of pH, DO levels, chlorophyll, BOD, COD, TSS and turbidity. We used the Sentinel-2 multispectral images available in the European Space Agency (ESA) Scientific Data Hub (Sentinel-1 Scientific Data Hub, 2021). Sentinel-2 has 13 spectral bands: four bands with a resolution of 10 m, six bands with a resolution of 20 m and three bands with a spatial resolution of 60 m. The orbit width is 290 km. In this work, we downloaded the associated Sentinel-2 Level 2A scenes captured on October 1 to 20, 2023, cloud-free and available for all sampling locations using Google Earth Engine. Distributed Level 2 products have been atmospherically corrected by the Sen2Cor package. Both images matched the field campaign days. We resampled all bands at a resolution of 20 m to 10 m to maintain consistency with the four native bands (band 2 in blue, band 3 in green, band 4 in red, and band 8 in NIR). Using Sentinel-2 imagery, we delineated the concentrations of pH, dissolved oxygen, chlorophyll, BOD, COD, TSS, and turbidity in the Punpun River using spectral water indices such as the Automated Water Extraction Index (no shade) (AWEInsh)[18] and the Sentinel Water Mask (SWM)[19]. The AWEInsh maximizes the separability of water and non-water pixels through band differentiation and addition and application of various coefficients. SWM increases the ability to detect water by enhancing visual contrast and value separability between water and non-water pixels. The AWEInsh for Sentinel is expressed by equation. (1). where the corresponding bands for Sentinel are Band 3 for Green, Band 8 for NIR, Band 11 for SWIR1 and Band 12 for SWIR2. The SWM for Sentinel is expressed by Equation. (2).

To estimate turbidity in water bodies, the Normalize Difference Turbidity Index (NDTI) is used, which is estimated based on the spectral reflectance values of the water pixels. Therefore, as the turbidity increases, the reflectivity of the red spectrum also increases. The NDTI for Sentinel is expressed by Equation. (3).

We created a map of water indices for the October 2023 images. For parts of the river that are narrow (less than one pixel) that the algorithm was unable to extract, we manually intervened by adjusting the river width using high spatial resolution images from Google Earth as base maps. Although manual feature extraction is subjective, we believe that the intervention has clear and methodological advantages over automated classification methods.[20] Using Google Earth images as auxiliary datasets improves post-processing, fills missing patches, and improves the accuracy of the final classification product.[21]

**Classification algorithms**

We used random forest classification algorithms to map the pH, DO, chlorophyll, BOD, COD, TSS and turbidity concentrations by running a set of codes in Google Earth Engine.[22] RF are popular algorithms for mapping rivers and other surface waters[9,23] including water quality parameters using remote sensing data.[13,24-26] RF provides a way to select important covariates based on changes in prediction accuracy when variables are added to or deleted from models. RF is a non-parametric supervised classifier that uses the classification and regression tree (CART) by bagging, randomly selecting a set of features and building a classifier with a bootstrapping sample of the training data to create a tree.[27] When selecting RF training data, it is possible that the same sample will be selected multiple times, while others may not be selected at all. Besides RF being quite robust for highly collinear variables, the random selection process at each tree node results in low correlation between trees and avoids overfitting.[28]

**Covariates, training, and test datasets**

To build the RF models for mapping the Punpun River, we used a set of covariates for image month (October) and a set of nine features. The set of covariates included the following features: Band 2 (492.1 nm), Band 3 (559.0 nm), Band 4 (665.0 nm), Band 5 (703.8 nm), Band 6 (740.2 nm), band 7 (782.5 nm), Band 8 (835.1 nm), Band 8A (864.8 nm), Band 11 (1613.7 nm), Band 12 (2202.4 nm), AWEI and SWM. We made the selection of spectral bands and indices from previous studies in which machine learning algorithms were applied to remote sensing data to classify water bodies and water quality parameters.[1,29-31] We used a specific range of pH, DO, chlorophyll, BOD, COD, TSS and turbidity datasets within days in October 2023 to correspond to the time the satellite images were acquired. We manually checked the pH, DO, chlorophyll, BOD, COD, TSS, and turbidity values for abnormal fluctuations within the time period. Before running the models, we split the final dataset into training and testing sets. We used the Google Earth Engine code from Brus *et al.* (2011)[32] applied to achieve this step a split criterion of 70-30, with 70% of the sample data used for calibration and 30% for validation.

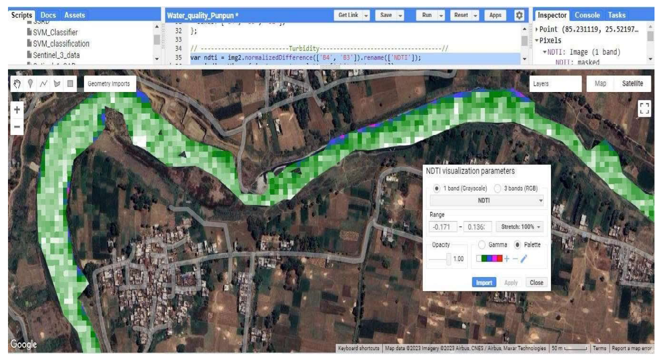



**Fig.2. Run the classification algorithms in google earth engine.**

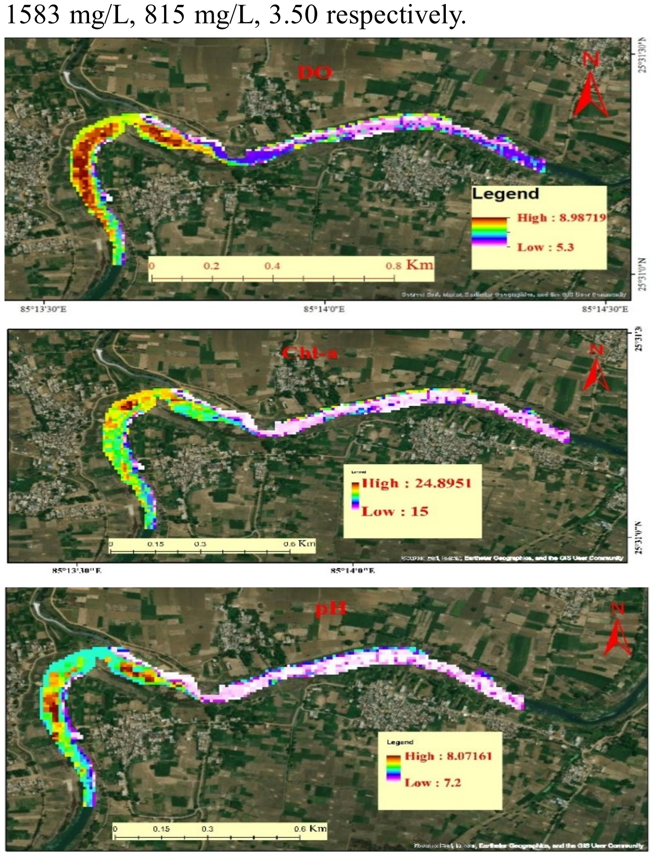
**Fig.3. Sampling site photographs and sampling collection.**

## RESULT

**Posterior probability water quality maps**

We created pH, DO, chlorophyll, BOD, COD, TSS and turbidity maps for October data. The average physiochemical parameters for RF mapping results in the form of posterior probability maps for subsections of the Punpun river basin are shown in Fig. 4. The average values of water quality parameters were all above concentration value of 7.52, 7.35 mg/L, 24.27 mg/L, 146.10 mg/L, 118.05 mg/L, 19.16 ug/L for the entire section of the Punpun River in October. We made comparisons between upstream and downstream flow by creating difference maps for the entire river section and enlarging the sections of the three sampling sites for viewing. These sampling sites in the upper section are surrounded by farmland, while the sampling sites in the lower section are generally located in urban areas with minimal agricultural land use. Comparing both sections, it was found that the lower section had lower modelled DO values than the upper section, especially in October. The summary of prediction errors for RF models in predicting physiochemical parameters is shown in Fig. 4. For both RFs, the difference between the predicted and actual values was comparable. Error values were also similarly consistent between parameters, with RF for October having a range value of 0.98 to 0.86. Using the prediction error values, we evaluated the map accuracy using the two evaluators in Table 2. RMSE and MSE for RF were calculated.

Magnitudes of errors shown in MSE, and RMSE were considerably similar. No unexplained distribution skewness from very large values were observed in the RMSE. We also presented the validation results in graphical forms - scatter and spatial bubble plots -We further evaluated our results based on a predictive variable. Our results showed that the predictive capabilities of RF were reduced from MSE of 0.02, 0.56 mg/L, 20.83 mg/L, 1583 mg/L, 815 mg/L, 3.50 respectively.



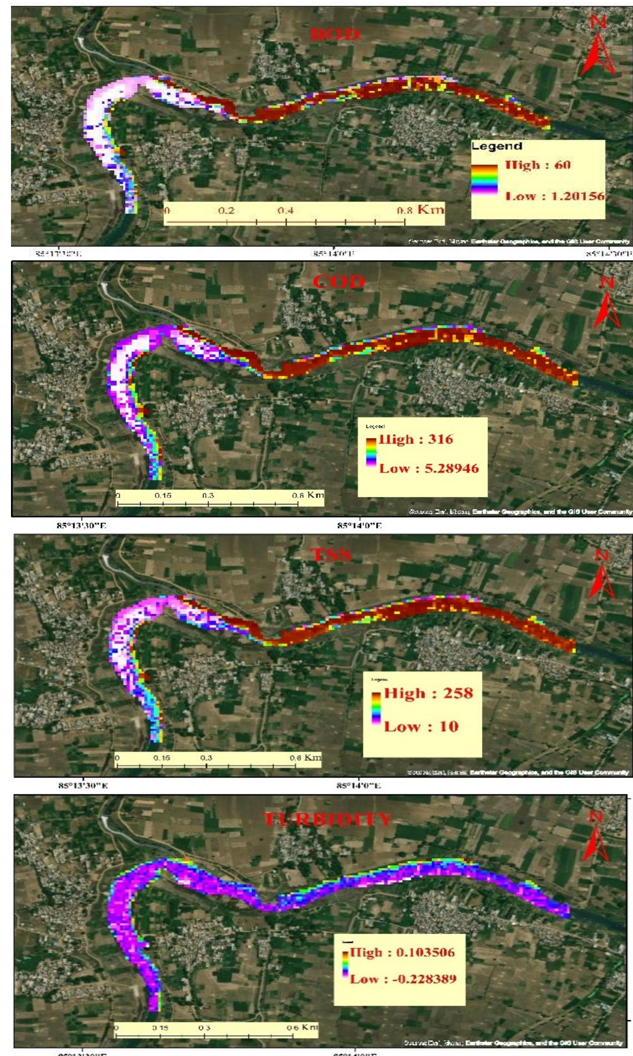




**Fig. 4. Averaged pH, DO, Chlorophyll, BOD, COD, TSS, and Turbidity (mg/L) for RF results in forms of posterior probability maps for the study reach.**

**Table.1. Descriptive statistics of all parameters along the stretch of the Punpun River.**

| Parameter | Min | Max | Mean | STD | SE |
|---|---|---|---|---|---|
| PH | 7.2 | 8.08 | 7.63 | 0.44 | 0.25 |
| DO | 5.3 | 9 | 7.63 | 2.03 | 1.17 |
| BOD | 1.2 | 60 | 21.16 | 33.63 | 19.41 |
| COD | 5 | 316 | 108.76 | 179.46 | 103.61 |
| TSS | 10 | 258 | 93.06 | 142.83 | 82.46 |
| CHL | 15 | 25 | 20 | 5 | 2.88 |

**Table.2. Summary of map quality measures for the two mapping methods, RF.**

| Parameter | $R^2$ | RMSE | MSE |
|---|---|---|---|
| PH | 0.86 | 0.048 | 0.022 |
| DO | 0.91 | 0.147 | 0.565 |
| BOD | 0.98 | 1.832 | 20.83 |
| COD | 0.97 | 9.342 | 1583.6 |
| TSS | 0.97 | 11.89 | 815.09 |
| CHL | 0.91 | 0.016 | 3.5 |

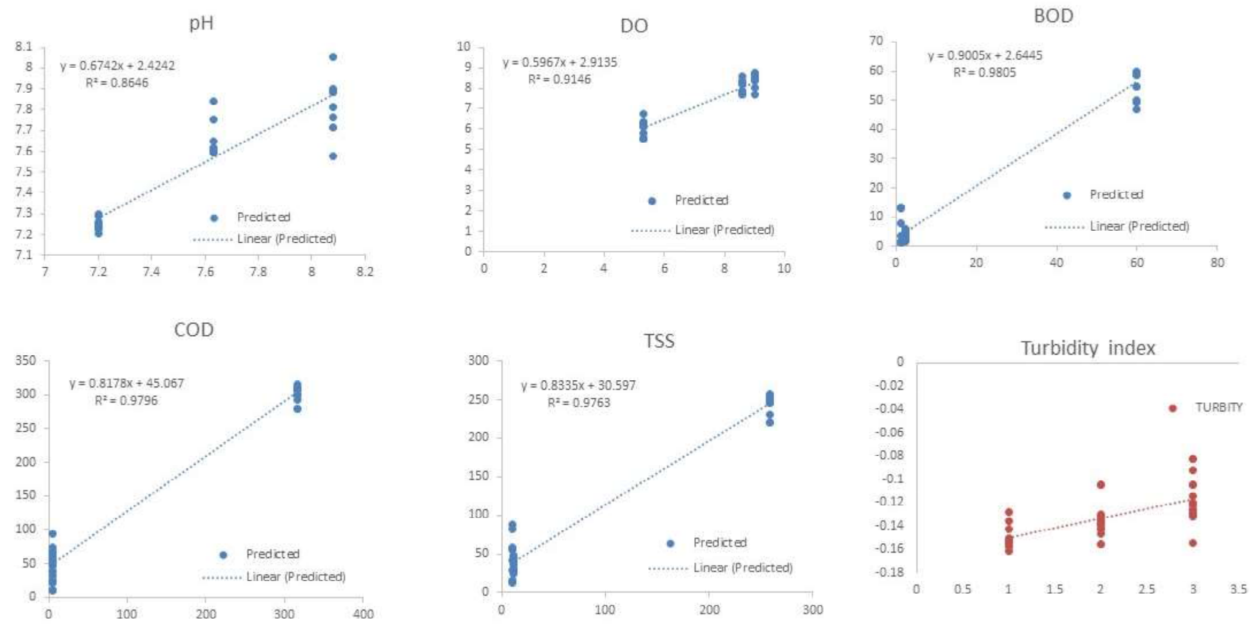



**Fig.5. Graphical representation of Model validation with coefficient of determination.**

## DISCUSSION

### Model performance and Sentinel images

The results of our investigation showed that Sentinel-2 satellite data with a resolution of 10 m was efficient in retrieving and predicting pH, DO content, chlorophyll, BOD, COD, TSS and turbidity in rivers for specific observation months. Furthermore, the results showed that parameters could be successfully predicted through calibration and validation using RF using Sentinel-2. The RF algorithms were shown to be effective in estimating parameter concentrations along the Punpun River and provide highly acceptable results. The promising ability of RF to spatially map pH, DO, chlorophyll, BOD, COD, TSS and turbidity is not surprising as these models have already demonstrated remarkable predictive accuracy.[2,5,6] The RMSE and MSE values for checking various parameters showed that the RF is extremely accurate; even the r values were equal to or greater than 0.86 in all cases. The good agreement between modelled and measured pH, DO, chlorophyll, BOD, COD, TSS and turbidity concentrations, further reinforced by the minimal residual errors ranging from 0.016 mg/L to 11.89 mg/L, highlights the robustness of the model predictions for both dates. These inconsistent results should not be interpreted as a limitation on the potential applicability of the algorithms in predicting DO. Instead, they served as a reminder that indirectly determining precise pH, DO, chlorophyll, BOD, COD, TSS and turbidity concentrations from temporal satellite imagery could only be done with the right training data set. If a prediction is made across the same modelling space using data from a different time period, the results may be affected.[33] However, when we applied our model with pooled covariates from sentinel images of different dates, it had minimal impact on the predictions of pH, DO, chlorophyll, BOD, COD, TSS, and turbidity for self-trained single images.

### Water quality mapping and concentration

Over the length of the studied river, pH, DO, chlorophyll, BOD, COD, TSS and turbidity concentration values in the Punpun River were in the good to excellent range based on the minimum wastewater habitat criterion of DO of 5.3 mg/L if wastewater quality is lacking. Although there is no consistent trend data for individual sampling sites, higher average BOD, COD, TSS and turbidity concentrations were predicted for the lower portion of the Punpun River, where sites were generally located in urban areas with minimal agricultural land use. These sites were predicted to have higher DO and Chl concentrations in the upper reaches of the river. Although we lacked an independent data set to confirm the sources of larger DO concentration gradients in the lower section compared to the upper section, it is possible that this may be due to lower biological oxygen demand (BOD).[10] In a biological and water quality study conducted in the Punpun River basin, lower DO concentrations were found in the headwater streams with the highest five-day BOD.

## CONCLUSIONS

This study demonstrated the applicability of Sentinel-2 imagery to map the spatial distribution of pH, DO, chlorophyll, BOD, COD, TSS and turbidity concentrations in the Punpun River using machine learning techniques. The resulting overall accuracy of the various parameter extraction from our models demonstrated that it is suitable to use the Sentinel-2 dataset to examine the temporal changes in water quality parameter concentrations in the Punpun River section. With the right training data, the remote sensing methods and algorithms used here could be used to indirectly determine pH, DO, chlorophyll, BOD, COD, TSS and turbidity concentrations in other river systems. However, the physical, chemical and biological processes within watersheds are complicated, particularly when processes such as nutrient conversion in rivers are not well understood. Using predictive models to represent the spatiotemporal variability of parameters would always lead to uncertainties in research. The machine learning method has its own advantages and disadvantages, and no single algorithm is suitable for all applications. Because our models are data-driven, it is important to apply an adaptive management approach to future model changes. As new input variables become available, we recommend making changes to the choices and re-running the models. Finally, we provided the Google Earth Engine script used to run our models so that future research can extend the conclusions of this work to other river systems.

## ACKNOWLEDGEMENT

We thank TPS College, Patna for their assistance in the experimental setup and data collection. The experimental design was greatly improved by discussions with Assistant Professor Dr Vinay Bhushan Kumar and Prof. A. K. Nag. The authors declare no competing financial interest.

## AUTHOR CONTRIBUTION

Data collection, analysis of the data, and conceived of the idea, developed the methods, and edited the manuscript. Earth observations conducted the experiment and ran analyses. All authors approved the submitted version of the manuscript and agreed to be listed.

## REFERENCES

1. **Abobakr Yahya, A.S., Ahmed, A.N., Binti Othman, F., Ibrahim, R.K., Afan, H.A., El-Shafie, A., Fai, C.M., Hossain, M.S., Ehteram, M., Elshafie, A., 2019.** Water quality prediction model-based support vector machine model for ungauged river catchment under dual scenarios. *Water* **11:**1231. https://doi.org/ 10.3390/ w11061231.

2. **Ahmed, M.H., Lin, L.-S., 2021.** Dissolved oxygen concentration predictions for running waters with different land use land cover using a quantile regression forest machine learning technique. *J. Hydrol*. ISSN: 0022-1694 **597:**126213 https://doi.org/10.1016/ j.jhydrol.2021.126213.

3. **Akshaya, D., 2012.** Impacts of Land Use/land Cover and Soil on Water Quality in the Upper Little Miami River Sub-basin, 553. https://corescholar.libraries. wright.edu/cgi/viewcontent.cgi?article=1692& context=etd_all/. (Accessed 23 February 2022).

4. **Brett, J., Blackburn, J., 1981.** Oxygen requirements for growth of your coho salmon (Orconhynchus kisutch) and sockey (*O. nerka*) salmon at 15 degrees Celsius. *Can. J. Fish. Aquat. Sci*. **38:**399–404.

5. **Valera, M., Walter, R.K., Bailey, B.A., Castillo, J.E., 2020.** Machine learning based predictions of dissolved oxygen in a small coastal embayment. *J. Mar. Sci. Eng*. **8:**1007. https://doi.org/10.3390/jmse8121007.

6. **Heddam, S., 2021.** Intelligent data analytics approaches for predicting dissolved oxygen concentration in river: extremely randomized tree versus Random Forest, MLPNN and MLR. In: Deo, R., Samui, P., Kisi, O., Yaseen, Z. (Eds.), Intelligent Data Analytics for Decision-Support Systems in Hazard Mitigation. Springer Transactions in Civil and Environmental Engineering. Springer, Singapore. https://doi.org/10.1007/978-981-15-5772-9_5.

7. **Evans, R.L., Miller, M.C., 2006.** Nutrients, eutrophic response, and fish anomalies in the Little Miami River, Ohio. *Ohio J. Sci*. **106:**146–155.

8. **Weiqi, H.E., Chen, S., Xuehua, L., Chen, J., 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*. **2:** 163–171. https://doi.org/10.1007/s11783-008-0027-7.

9. **Liu, Q., Huang, C., Shi, Z., Zhang, S., 2020.** Probabilistic river water mapping from Landsat-8 using the Support Vector Machine method. Rem. Sens. **12:**1374. https:// doi.org/10.3390/rs12091374.

10. **Al-Shaibah, B., Liu, X., Zhang, J., Tong, Z., Zhang, M., El-Zeiny, A., Faichia, C., Hussain, M., Tayyab, M., 2021.** Modeling water quality parameters using Landsat multispectral images: a case study of Erlong Lake, Northeast China. *Rem. Sens.* **13:**1603. https://doi.org/10.3390/rs13091603.

11. **Peterson, K.T., Sagan, V., Sloan, J., 2020.** Deep learning-based water quality estimation and anomaly detection using Landsat-8/Sentinel-2 virtual constellation and cloud computing. *GIsci. Remote Sens.* **57:** 510–525. https://doi.org/10.1080/15481603.2020.1738061.

12. **Dornh ̈ ofer, ̈ K., Oppelt, N., 2016.** Remote sensing for lake research and monitoring – recent advances. Ecol. Indicat. **64:**105–122. https://doi.org/10.1016/j.ecolind.2015.12.009.

13. **Wagle, M., Acharya, T.D., Lee, D.H., 2020.** Estimating chlorophyll-a and dissolved oxygen based on Landsat 8 bands using support vector machine and recursive partitioning tree regressions. *Proceedings* **42:** 25. https://doi.org/10.3390/ecsa-6-06573.

14. **Anderson, A., Nagar, R., Sarkar, D., 2007.** Spatial and temporal trends in surface water quality in a segment of the San Antonio river, Texas. In: Sarkar, D., Datta, R., Hannigan, R. (Eds.), Concepts and Applications in Environmental Geochemistry. Elsevier, Amsterdam, pp. 591–606.

15. **Boyle, S.A., Kennedy, C.M., Torres, J., Colman, K., P ́erez-Estigarribia, P.E., de la Sancha, N.U., 2014.** High-resolution satellite imagery is an important yet underutilized resource in conservation biology. *PLoS One.* **9(1):** e86908. https://doi.org/10.1371/journal.pone.0086908.

16. **Ranjan, S., & Singh, V., 2022.** HEC-HMS based rainfall-runoff model for Punpun river basin. *Water Practice & Technology*, **17(5):**986-1001.

17. **APHA. 2017.** Standard methods for the examination of water and wastewater, 23rd edition edited by R. B. Baird, A. D. Eaton and E. W. Rice. American Public Health Association (APHA). American Water Works Association (AWWA) and Water Environment Federation (WEF), Washington, D.C., USA

18. **Feyisa, G.L., Meilby, H., Fensholt, R., Proud, S.R., 2014.** Automated water extraction index: a new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ*. **140:** 23–35. https://doi.org/10.1016/j.rse.2013.08.029.

19. **Robak, A., Gadawska, A., Milczarek, M., Lewinski, ́ S., 2017.** The detection of water on Sentinel-2 optical imagery based on water indices. *Teledetekcja Srodowiska*. ́ **56:** 59–72.

20. **Castilla, G., Hay, G.J., Ruiz-Gallardo, J.R., 2008**. Size-constrained region merging (SCRM): an automated delineation tool for assisted photo interpretation. *Photogramm. Eng. Rem. Sens*. **74 (4):** 409–419.

21. **Zhao, C.P., Qin, C.Z., 2021.** A detailed mangrove map of China for 2019 derived from Sentinel-1 and -2 images and Google Earth images, 00 *Geosci. Data J.* 1–15. https://doi.org/10.1002/gdj3.119.

22. **FAO, 2018.** Soil organic carbon mapping cookbook. Yigini, Y., Olmedo, G.F., Reiter, S., Baritz, R., Viatkin, K., Vargas, R.R. (Eds.) 2nd Edition, Rome.

23. **Wang, Z., Liu, J., Li, J., Zhang, D.D., 2018.** Multi-Spectral Water Index (MuWI): a native 10-m multi-spectral water index for accurate water mapping on Sentinel-2. *Rem. Sens*. **10:** 1643. https://doi.org/10.3390/rs10101643.

24. **Urbanski, J.A., Wochna, A., Bubak, I., Grzybowski, W., Lukawska-Matuszewska, K., £¹cka, M., Sliwi ́ nska, ́ S., Wojtasiewicz, B., Zaj¹czkowski, M., 2016.** Application of Landsat 8 imagery to regional-scale assessment of lake water quality. *Int. J. Appl. Earth Obs. Geoinf*. **51:** 28–36. https://doi.org/10.1016/j.jag.2016.04.004.

25. **Li, X., Sha, J., Wang, Z.L., 2017.** A comparative study of multiple linear regression, artificial neural network and support vector machine for the prediction of dissolved oxygen. *Hydrol. Res.* **48:** 1214–1225. https://doi.org/10.2166/nh.2016.149.

26. **Wei, L., Zhang, Y., Huang, C., Wang, Z., Huang, Q., Yin, F., Guo, Y., Cao, L., 2020.** Inland lakes mapping for monitoring water quality using a detail/smoothingbalanced conditional random field based on Landsat-8/levels data. *Sensors* **20:**1345. https://doi.org/10.3390/s20051345.

27. **Breiman, L., 2001.** Random forests. Mach. *Learn*. **45:** 5–32.

28. **Salas, E.A.L., Subburayalu, S.K., 2019.** Modified shape index for object-based random forest image classification of agricultural systems using airborne hyperspectral datasets. *PLoS One* **14:** e0213356. https://doi.org/10.1371/journal.pone.0213356.

29. **Sarp, G., Ozcelik, M., 2017.** Water body extraction and change detection using time series: a case study of Lake Burdur, Turkey. *J. Taibah Univ. Sci*. ISSN: 1658-3655 **11:** 381–391. https://doi.org/10.1016/j.jtusci.2016.04.005

30. **Possa, ˆ E.M., Maillard, P., 2018.** Precise delineation of small water bodies from Sentinel-1 data using Support Vector Machine Classification. *Can. J. Rem. Sens*. **44:** 179–190. https://doi.org/10.1080/07038992.2018.1478723.

31. **Rana, V.K., Suryanarayana, T.M.V., 2020.** Performance evaluation of MLE, RF and SVM classification algorithms for watershed scale land use/land cover mapping using sentinel 2 bands. *Remote Sens. Appl.: Soc. Environ*. ISSN: 2352-9385 **19:** 100351 https://doi.org/10.1016/j.rsase.2020.100351.

32. **Brus, D.J., Kempen, B., Heuvelink, G.B.M., 2011.** Sampling for validation of digital soil maps. *Eur. J. Soil Sci.* **62 (3):** 394–407. https://doi.org/10.1111/j.1365- 2389.2011.01364.x.

33. **Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carr´e, G., Marqu´ez, J.R.G., Gruber, B., Lafourcade, B., Leit˜ ao, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schroder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013.** Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36:**27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x.


***