remote sensing

MDPI

Article

Predicting Water Quality Distribution of Lakes through Linking Remote Sensing–Based Monitoring and Machine Learning Simulation

Mahdi Sedighkia [1,*], Bithin Datta [1], Parisa Saeedipour [2] and Asghar Abdoli [3]

[1] Civil and Environmental Engineering Department, College of Science and Engineering, James Cook University, Townsville, QLD 4811, Australia
[2] Computer Engineering Department, Engineering Faculty, Shahid Chamran University, Ahvaz 61357, Iran
[3] Environmental Science Research Institute, Tehran 19839, Iran
* Correspondence: mahdi.sedighkia@my.jcu.edu.au

Abstract: The present study links monitoring and simulation models to predict water quality distribution in lakes using an optimized neural network and remote sensing data processing. Two data driven models were developed. First, a monitoring model was established that is able to convert spectral images to TDS distribution. Moreover, a simulation model was developed to generate a TDS distribution map for unseen scenarios for which no spectral images are available. Outputs of the monitoring model were applied as the observations for training the simulation model. The Nash–Sutcliffe model efficiency coefficient (NSE) was utilized in the system performance measurement of the models. Based on the results in the case study, the monitoring model was sufficiently robust to convert the operational land imager spectral bands of Landsat 8 to the TDS distribution map. The NSE was more than 0.6 for the monitoring model, which confirms the predictive skills of the model. Furthermore, the simulation model was highly reliable in generating the TDS distribution map of the lakes. Three tests were carried out to demonstrate the reliability of the model. When comparing the results of the monitoring model and simulation model, an NSE of more than 0.6 was found for all the tests. It is recommendable to apply the proposed method instead of conventional hydrodynamic models that might be highly time consuming for simulating water quality parameters distribution in lakes. Low computational complexity is the main advantage of the proposed method.

Keywords: remote sensing; water quality modelling; neural networks; optimal number of hidden layers; Landsat 8

check for updates

Citation: Sedighkia, M.; Datta, B.; Saeedipour, P.; Abdoli, A. Predicting Water Quality Distribution of Lakes through Linking Remote Sensing–Based Monitoring and Machine Learning Simulation. *Remote Sens.* **2023**, *15*, 3302. https://doi.org/10.3390/rs15133302

Academic Editors: Venkatesh Merwade, John Kalogiros, Pierfranco Costabile and Jochen E. Schubert

Received: 27 May 2023
Revised: 15 June 2023
Accepted: 24 June 2023
Published: 27 June 2023

1. Introduction

The development of agriculture is vital for food security. However, the environmental impacts of agriculture on water bodies has been highlighted in recent decades [1]. On the one hand, using fertilizers is important to maximize yield of the production for the farmers. On the other hand, this might increase water pollutants in the agricultural runoff that could be drained to inland water bodies such as lakes [2]. Lakes are important habitats and require suitable water quality to maintain biological activities such as reproduction and ensure food supplies for various species. Draining agricultural runoff to lakes is a serious concern for environmental managers who want to mitigate the environmental impact of agricultural activities. Hence, monitoring water quality distribution in lakes is one of the main requirements for better management. Moreover, simulating the impact of unseen scenarios is important to improving environmental management. In other words, it is essential to simulate or predict the impact of agricultural development scenarios on the distribution of water quality parameters for managing lake habitats.

Due to the importance of water quality assessment in the lakes, field measurement is critical to assess water quality. However, traditional methods for measuring water quality parameters have some disadvantages that might reduce the applicability of these methods in practice. First, continuous monitoring of water quality in many points of the lake can be expensive. Secondly, advanced measurement tools that are able to record the water quality parameters online might not be available in all countries, especially in developing countries. Hence, using remote sensing data is a useful option, with minimal effort required for field measurements [3]. It is demonstrated that the spectral behavior of the water might be altered due to interactions of solar radiation and optically active components (OACs) [4]. Hence, detection and monitoring of a wide range of water quality parameters using remote sensing techniques have been highlighted in the literature [5,6]. For example, total suspended solids (TSS), total dissolved solids (TDS) and turbidity have been detected through the satellite images. It should be noted that several missions using satellites in remote sensing have been performed in recent decades. In other words, currently, several satellites capture the spectral images from earth through different installed sensors [7].

Data acquisition is one of the critical issues when using spectral images for the assessment of water quality in water bodies. Another important issue that might have a significant effect on the outputs of the remote sensing analysis is the choice of model for efficient conversion of the spectral images to concentrations of water quality parameters. According to the literature, regression models are widely used to process remote sensing data [8]. For instance, some previous studies have recommended using multiple linear regression to detect water quality parameters using remote sensing data [9]. However, many statistical studies have highlighted the drawbacks of regression methods to develop a robust predictive model. Hence, artificial intelligence (AI) methods have been developed to enhance the performance of the models in remote sensing analysis. The progress and challenges of using remote sensing in water quality analysis have been reviewed in the literature [10]. Different AI methods are utilized in remote sensing analysis. More details are provided in the literature [11].

It should be noted that AI methods are not only applicable in remote sensing analysis, they can be directly applied water quality modelling. Hence, it is necessary to review the principles of the most applicable AI methods. Artificial neural networks (ANNs) have been broadly applied in different branches of civil engineering [12]. In fact, proper response of these models has encouraged engineers to utilize neural networks as a robust method to predict or simulate different phenomena in environmental and civil engineering. For example, ANNs have been used to predict stream flow or inflow of the reservoirs in previous studies [13,14]. Furthermore, previous studies corroborate the robustness and applicability of the neural networks for simulating water quality in rivers and lakes [15–17]. According to the literature, ANNs have been recommended in remote sensing data analysis. ANNs have been used in other forms that might improve the results in some cases. For example, neuro fuzzy inference systems are one of the improved neural networks in this regard. A neural network conventionally contains an input layer, hidden layers (one or more hidden layers can be used) and an output layer. The input layer includes the effective inputs for the simulated outputs. It should be noted that the selection of inputs is critical for the output of the model. Moreover, the output layer includes output of the model as the purpose of the simulation. Hidden layers are the most important part of the ANNs; one or more layers can be used for connecting inputs to outputs. Development of the ANN generally has two steps, including the training process and testing process. In the training process, coefficients and constants for connecting the input layer, the hidden layers and the output layers should be optimized based on observations. In other words, the training process is the most important step to develop a successful model, in which availability of the observed or recorded data in the environment is highly effective. Different methods can be used to train the neural network; these have been reviewed in the literature [18]. Moreover, additional details regarding the theory and application of ANNs have been provided in previous studies [19].

The main motivation of the present study is the lack of robust models to simulate water quality distribution maps for lakes in which AI methods and remote sensing data analysis could be applied. Traditionally, hydrodynamic models such as MIKE could be used to simulate water quality distribution in the lakes [20]. However, their drawbacks, such as arduous verification processes, might reduce the applicability of the hydrodynamic models in practice. Moreover, conventional application of AI methods for limited points in lakes cannot simulate water quality distribution, because a wide range of points in the lake cannot be simulated. Hence, improving AI models for simulating water quality distribution in lakes is necessary. Combining AI methods with remote sensing data processing can provide a robust framework to predict water quality distribution maps; this is the main contribution and novelty of the present study. The proposed method might open new windows regarding the application of combined frameworks in which remote sensing analysis could be effectively applied for assessing water quality in lake environments.

2. Application and Methodology

2.1. Case Study and Data Acquisition

The proposed method was applied in Zrebar lake, one of the large freshwater lakes located in the Kurdistan province in Iran. This beautiful lake is one of the valuable habitats in the region; many native fish species and other aquatic organisms can be found here. The main economic activity for the people who live in the basin of the lake is agriculture. Cultivated lands have been developed in this basin in recent decades. Maximizing yield of production is the main purpose for the farmers, who would like to increase their net revenue. However, environmentalists have serious concerns regarding the water quality due to draining agricultural runoff, in which high concentrations of water pollutants have been reported. Hence, two tasks have been defined for the environmental engineers who are responsible for water quality control in this lake, including monitoring water quality parameter distribution at the surface of the lake and simulating water quality distribution for unseen scenarios. Hence, it is required to develop a robust system to monitor and predict distribution of water quality parameters in the lake. Figure 1 displays the location of the lake and the sampling points that were used in field studies for developing the data-driven models.

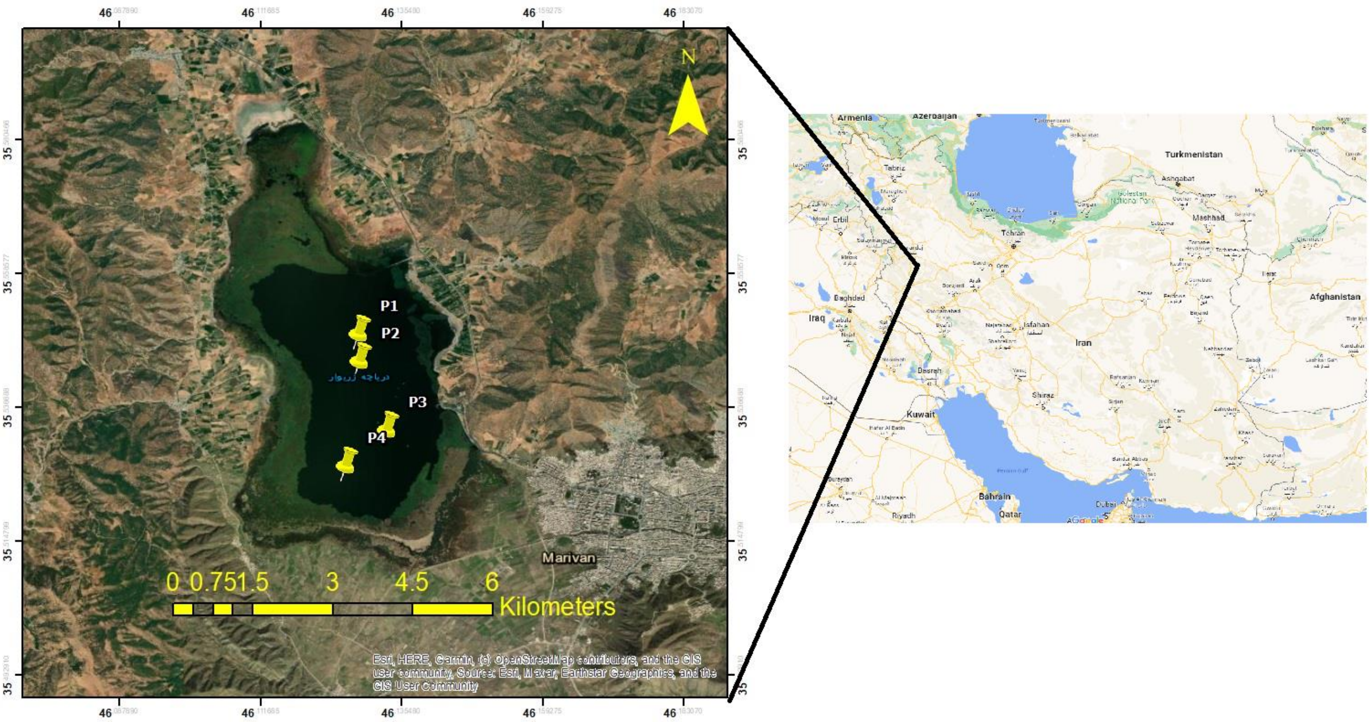

Figure 1. Location of Zrebar lake and sampling points including P1 to P4.

Data acquisition is one of the main steps for remote sensing analysis; data should be provided based on the needs and purposes of the study. Landsat 8 is one of the known satellites for capturing spectral images from the Earth through different sensors. Landsat 8 (formally the Landsat Data Continuity Mission, LDCM) was launched on an Atlas-V rocket from Vandenberg Air Force Base, California, on 11 February 2013. This satellite orbits the earth at an altitude of 705 km and completes one Earth orbit every 99 min. The captured images are available in 16-day intervals; this means several spectral images from the study area for a long-term assessment might be available. Hence, spectral images from the Landsat 8 are highly applicable for remote sensing analysis. The satellite has two types of sensors, including the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) instruments. Previous studies have demonstrated the applicability of OLI images for analyzing and monitoring water quality in water bodies. Thus, images generated by the OLI were applied in the present study. Eighteen images were downloaded from the USGS databank from 2013 to 2020, for which recorded data in the sampling points was available and images were proper in terms of requirements for remote sensing analysis, such as absence of clouds. Based on the initial ecological studies, TDS was selected as the most important water quality parameter for habitat suitability of the lake. Based on the technical definition, Total dissolved solids (TDS) can be defined based on concentration of all organic and inorganic materials dissolved in the water. This parameter is a measure of anything dissolved in the water. In the case study, TDS was selected as the representative parameter regarding water quality challenges. In other words, previous ecological sampling demonstrated that there is a strong relationship between the aquatic population and TDS. In fact, changing TDS affects all water quality parameters in the case study, which means it could be a good environmental index. Hence, we focused on the assessment of the TDS in the present study. TDS was measured in the flagged sampling points by the portable device. Mean monthly TDS as the inflow of the lake was simulated in previous studies and is displayed in Figure 2. It should be noted that the simulation of monthly TDS inflow to the lake is not in the scope of the present study, and we applied the generated data proposed by the regional water authority in the development of one of the data driven models. Pre-processing of remote sensing data was carried out based on the recommendations of previous studies. More details are provided in the literature (Maliki et al., 2020). Figure 3 displays one of the downloaded images (Landsat 8, OLI C2L2) showing seven bands. Additional details regarding the data-driven models are presented in the next sections.

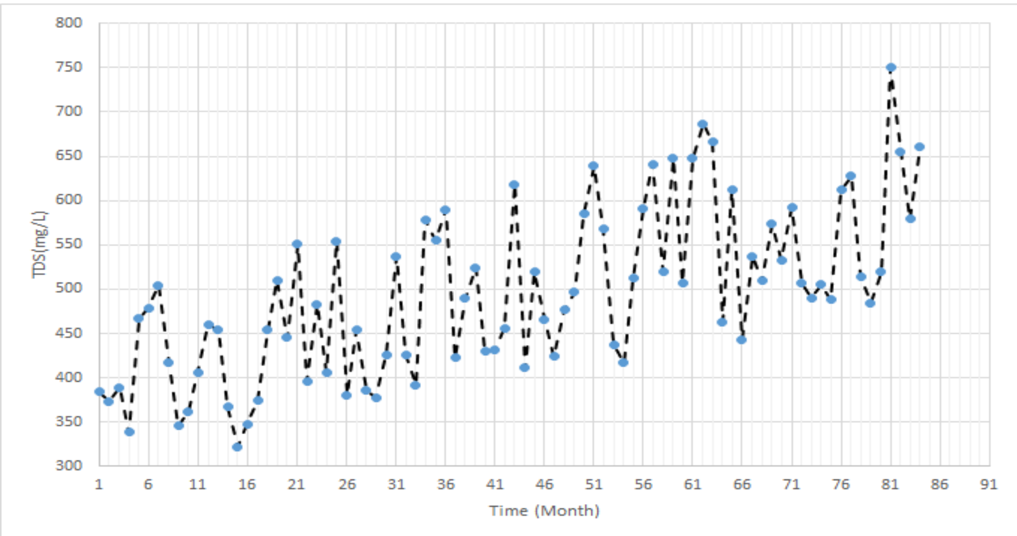

Figure 2. Simulated TDS inflow to the lake over long-term period (blue dots: simulated points).

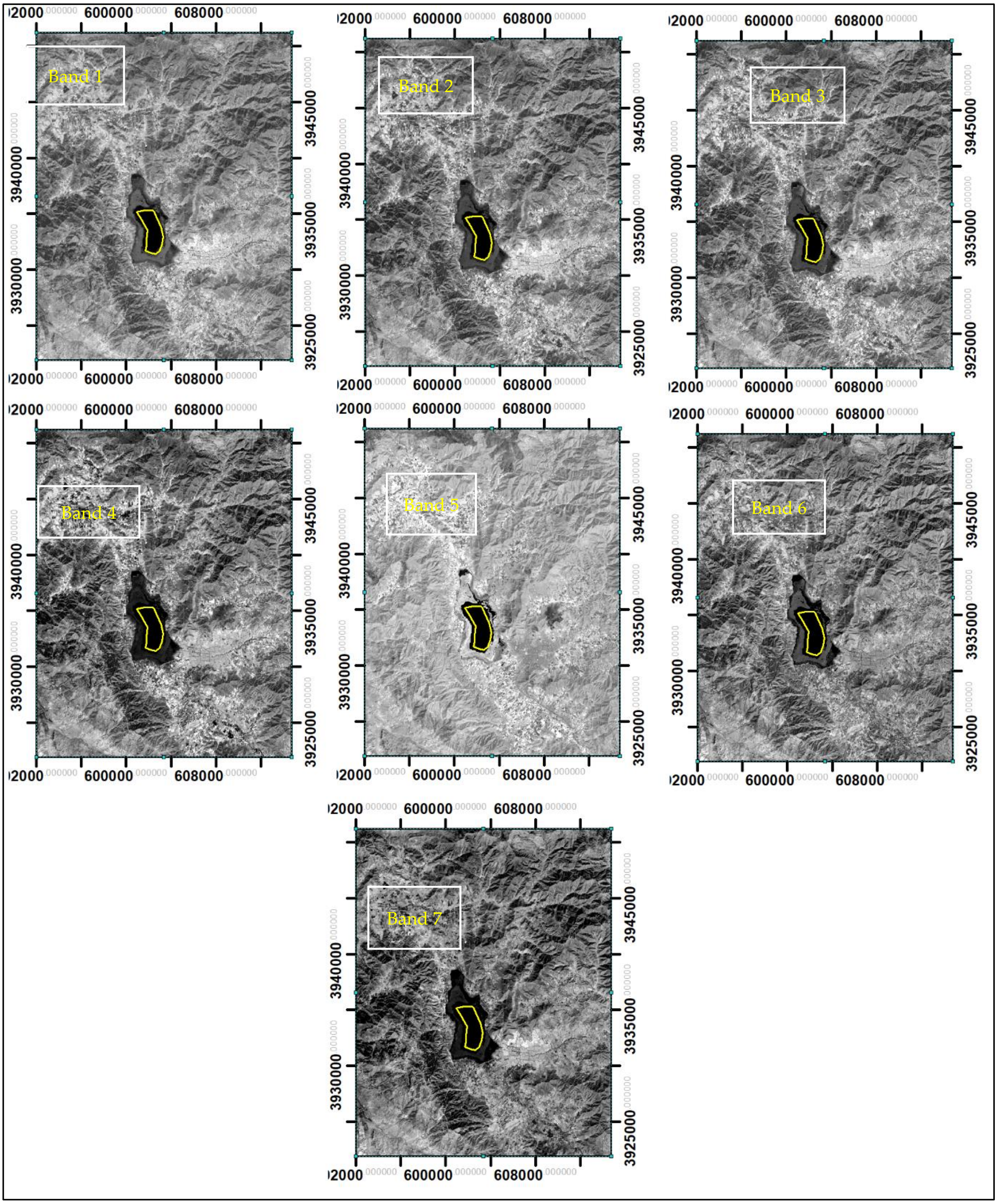

Figure 3. A sample of used spectral images (Landsat 8, OLI C2L2) in remote sensing data processing (selected region is displayed by yellow line).

2.2. Optimized Neural Network for Monitoring TDS Distribution in the Lake

The first data-driven model developed in the present study is an optimized neural network for converting the spectral images to the TDS distribution in the lake. Many previous studies applied ANNs for engineering problems. They commonly considered one or two hidden layers in the development of the model. However, it should be noted that the number of hidden layers highly impacts the output of the model. Hence, an optimized neural network was developed in the present study in which a number of hidden layers was optimized in the training process of the model. Figure 4 displays the general flowchart of ANNs, in which the number hidden layers is variable and the optimization process can find the best number of hidden layers. Table 1 displays the main characteristics of the neural network for converting the spectral images to the TDS concentration in the lake. Moreover, Figure 5 displays the workflow of the present study for developing the optimized neural network, in which the Marquardt–Levenberg method is used to train the network [21].

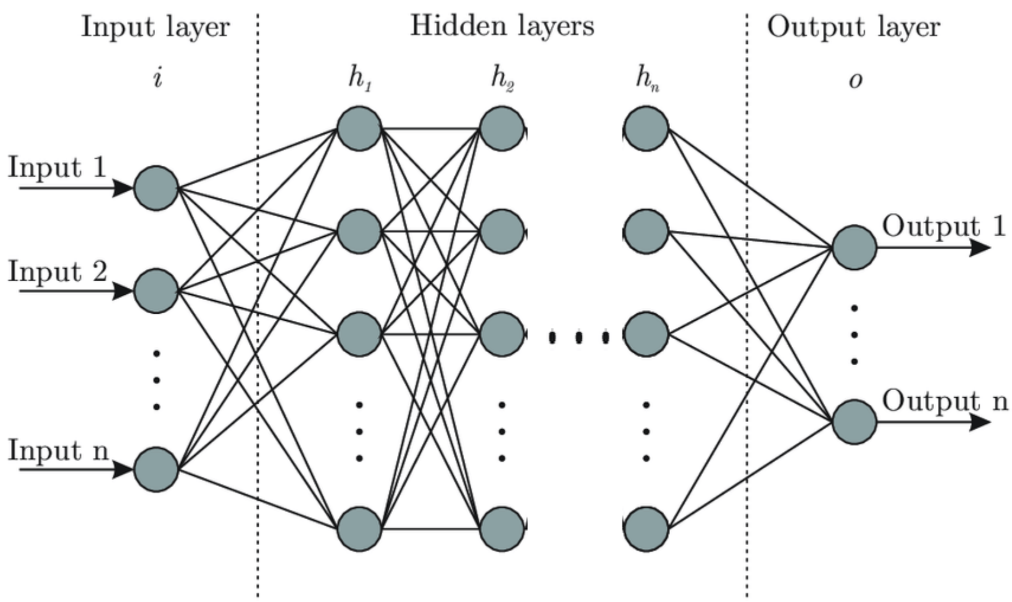

Figure 4. General architecture of ANNs.

Table 1. Main features of monitoring data-driven model.

Inputs	Output	Number of Hidden Layers	Training Method
Band 1 Visible (0.43–0.45 µm) 30 m	Total dissolved solids concentration (TDS)	Between 0 and 100	Subtractive Clustering
Band 2 Visible (0.450–0.51 µm) 30 m			
Band 3 Visible (0.53–0.59 µm) 30 m			
Band 4 Red (0.64–0.67 µm) 30 m			
Band 5 Near-Infrared (0.85–0.88 µm) 30 m			
Band 6 SWIR 1(1.57–1.65 µm) 30 m			
Band 7 SWIR 2 (2.11–2.29 µm) 30 m			

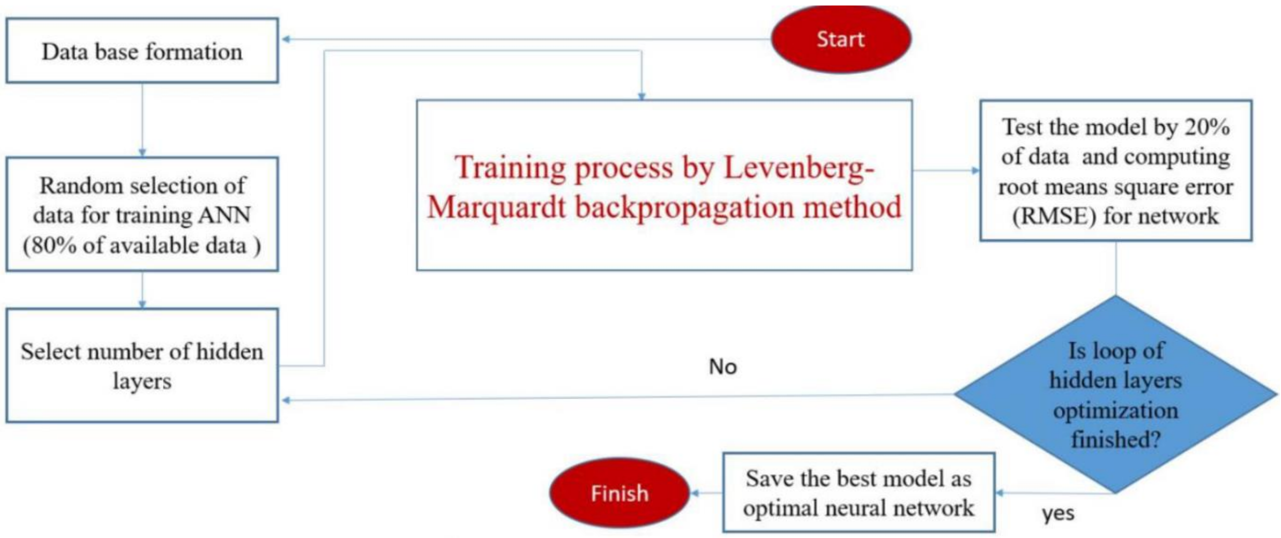

Figure 5. Workflow of developing optimal neural network.

2.3. Optimized Neural Network for Simulating TDS Distribution in the Lake

The second developed neural network in the present study is a data-driven model that might be able to simulate TDS concentration distribution for unseen scenarios. The architecture of the model for developing the optimized neural network is the same as the first model utilized for converting spectral images to TDS concentration. However, the inputs of the model are different, including TDS inflow in each time step (mean monthly TDS), longitude and latitude. The output of the model is distribution of TDS concentration. In fact, the output of model is a map, as will be presented in the next sections. In fact, this model considered the mean monthly TDS as the inflow of the model, which could be considered for generating the mean monthly TDS distribution map in the lake. It should be noted that most of the downloaded spectral images by the OLI were generated in the middle of the month, and this was adopted as the mean monthly TDS distribution. This assumption might not be perfectly applied in the real world. However, it is acceptable for generating the mean monthly TDS distribution map in the lake. The cell size of the OLI images is 30×30 m; this was considered the resolution of the simulation data-driven model. In other words, the latitude and longitude for each cell was considered in the data-driven model.

2.4. Measurement Indices

Each model requires some indices to measure the performance of the model. In fact, these indices measure how the predictive skills of the model are reliable. The Nash–Sutcliffe model efficiency coefficient (NSE) was applied to measure the robustness of the model [22]. Equation (1) displays the mathematical form of this index.

$$\mathrm{NSE} = 1 - \frac{\sum_{t=1}^{T}(M_t - O_t)^2}{\sum_{t=1}^{T}(O_t - O_m)^2} \tag{1}$$

where M_t is simulated data for sample t, O_t is observed or recorded data for sample t, O_m is mean observed or recorded data and T is the total number of simulated points. It should be noted that this index is one of the known indices for estimating the robustness of the data-driven models. Moreover, RMSE was applied as another index in the present study, as displayed in Equation (2).

$$\mathrm{RMSE} = \sqrt{\sum_{t=1}^{T}\frac{(O_t - M_t)^2}{T}} \tag{2}$$

3. Results

In the first step, it is essential to present the results of the development of the optimized neural network (ONN) in the present study. As presented, two ONNs were developed in the present study, including the ONN for converting spectral images to TDS concentration distribution in the simulated environment (DM1) and the ONN for simulating TDS distribution for unseen scenarios (DM2). Figure 6 displays the RMSE for DM1 in different numbers of hidden layers. It seems that the number of hidden layers has a great impact on the accuracy of the model; development of an ANN without optimization of the number of hidden layers might not reliable. Hence, it is recommendable to optimize the number of hidden layers in similar models for simulations. DM1 does not indicate a specific relationship between the number of hidden layers and the RMSE that supports the optimized process in each model. According to the results, eight hidden layers were selected as the optimal number of hidden layers for converting spectral images to the TDS concentration. Figure 7 shows the training and testing process of the DM1 (with optimal number of hidden layers) in which RMSE and NSE are displayed for the total period. As can be observed, NSE is close to 0.6, which demonstrates that the model is acceptable for converting the spectral images to the TDS concentration. According to the literature, if the NSE is more than 0.6, the developed model will be robust in terms of simulation of the outputs. Moreover, the RMSE is 63.6 mg/L, which indicates the mean error of the model is limited.

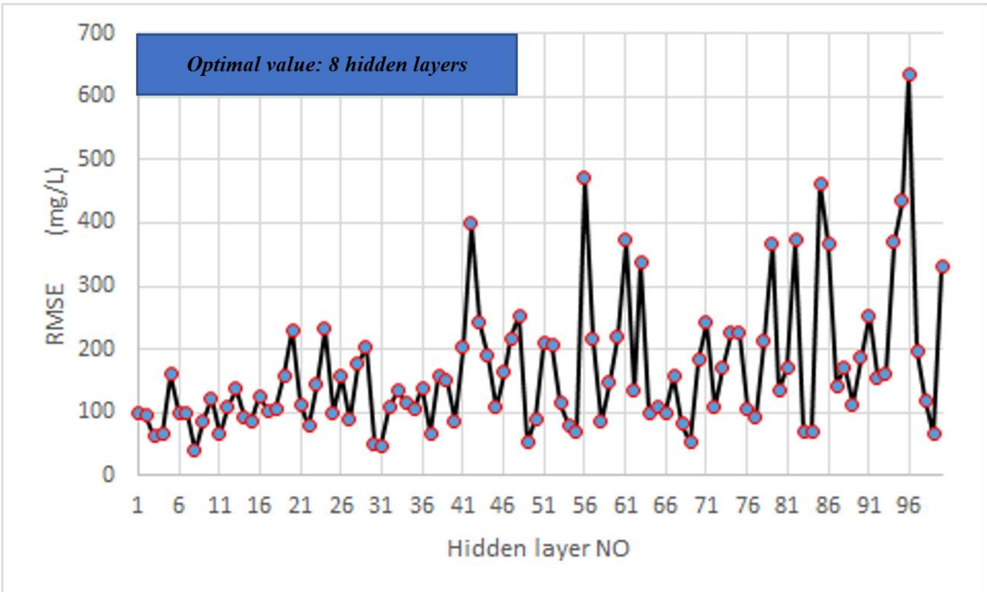

Figure 6. RMSE for determining optimal number of hidden layers (DM1).

In the next step, the output of DM2 should be presented. Figure 8 displays the RMSE for different numbers of hidden layers for DM2, which shows a different performance compared with DM1. In fact, it seems that there is a direct relationship between number of hidden layers and accuracy of the model. In fact, increasing the number of hidden layers will positively affect the accuracy of the model. Interestingly, DM1 and DM2 show different performance in the optimization of hidden layers. In the DM2, 96 hidden layers is the optimal number of hidden layers. Based on the simulated TDS inflow to the lake, three scenarios were selected to test the performance of the model for predicting TDS concentration. TDS inflows for test 1, test 2 and test 3 were 383.9, 373.4 and 389.3 mg/L, respectively. It should be noted that TDS inflows were selected close together to investigate the ability of the model to accurately simulate TDS distribution in the lake of the case study. In fact, the testing process can demonstrate how the machine learning model is able to predict or simulate TDS distribution in the lake. Figures 9–11 display the outputs by DM1 as the monitoring model of the TDS concentration and DM2 as the simulation model of the TDS concentration for the unseen scenarios or three simulated tests. The model is highly reliable and robust in simulating and predicting TDS concertation in the lake. However, the lower and upper limits of the simulations are slightly different from the data generated by DM1. This difference seems logical because the prediction of peak points is one of the conventional problems in using data-driven models in engineering. The RMSE and NSE for each generated map are displayed in the captions of the figures. The computed RMSE and NSE corroborate the robustness of the model to simulate TDS concentration distribution.

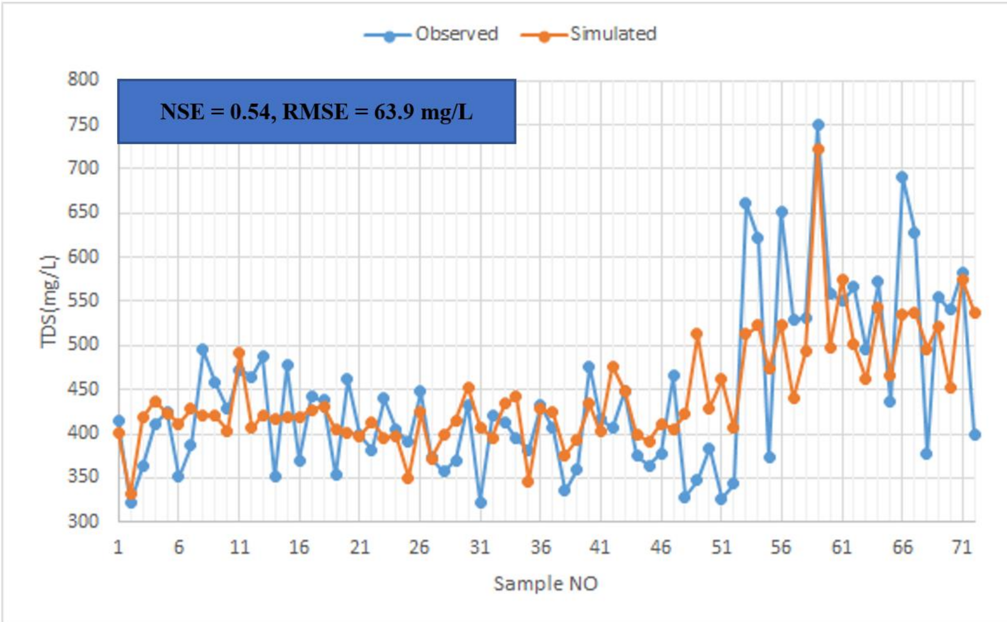

Figure 7. Training and testing process for DM1 (data-driven model for converting spectral images to TDS distribution map).

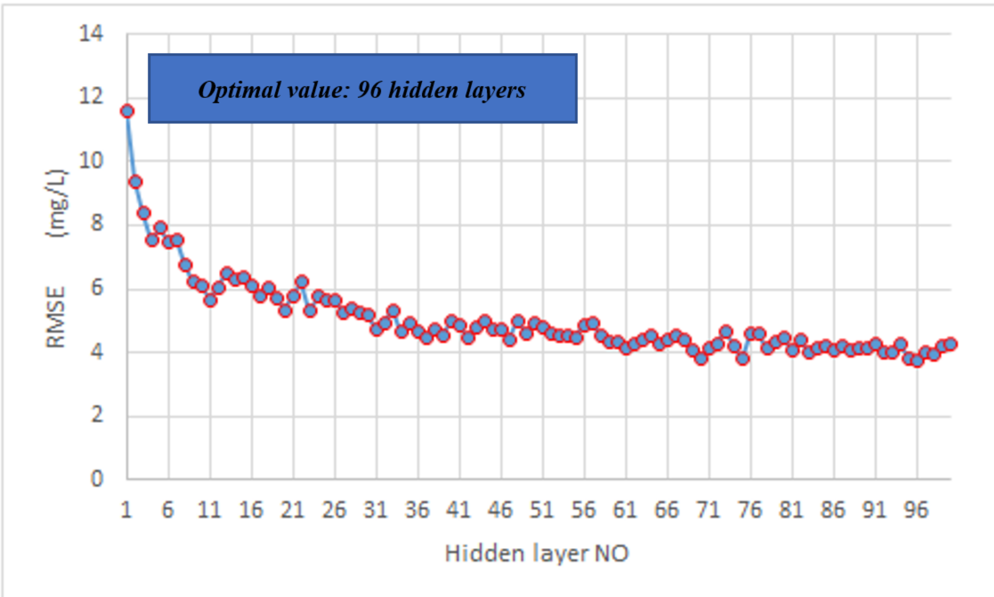

Figure 8. RMSE for determining optimal number of hidden layers (DM2).

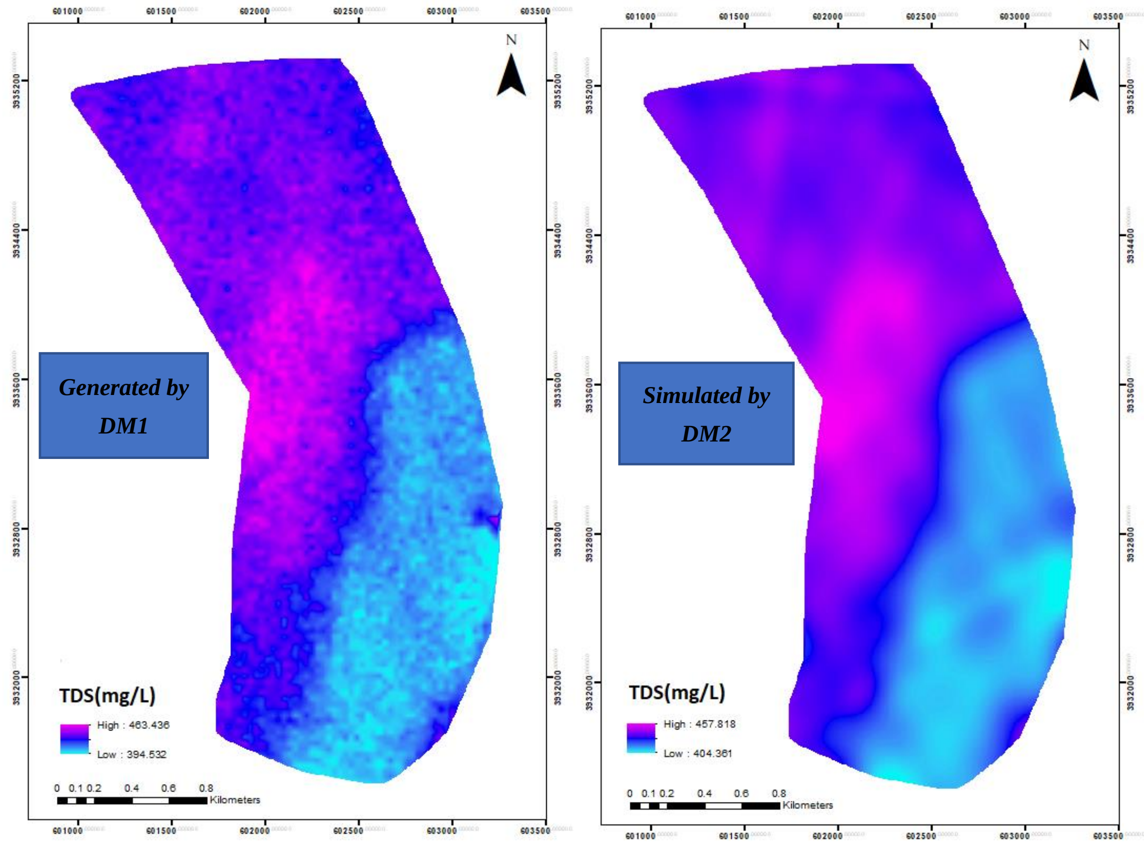

Figure 9. Outputs of test 1 (NSE = 0.93, RMSE = 3.47 mg/L).

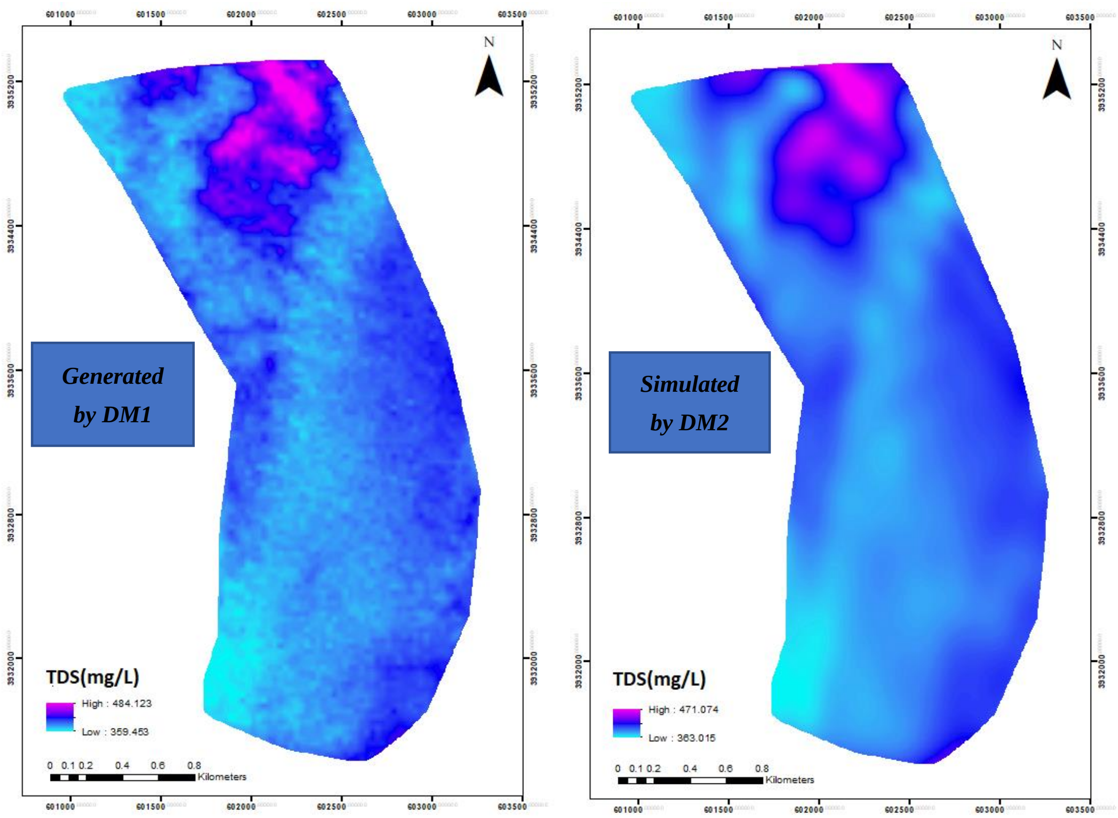

Figure 10. Outputs of test 2 (NSE = 0.90, RMSE = 5.28 mg/L).

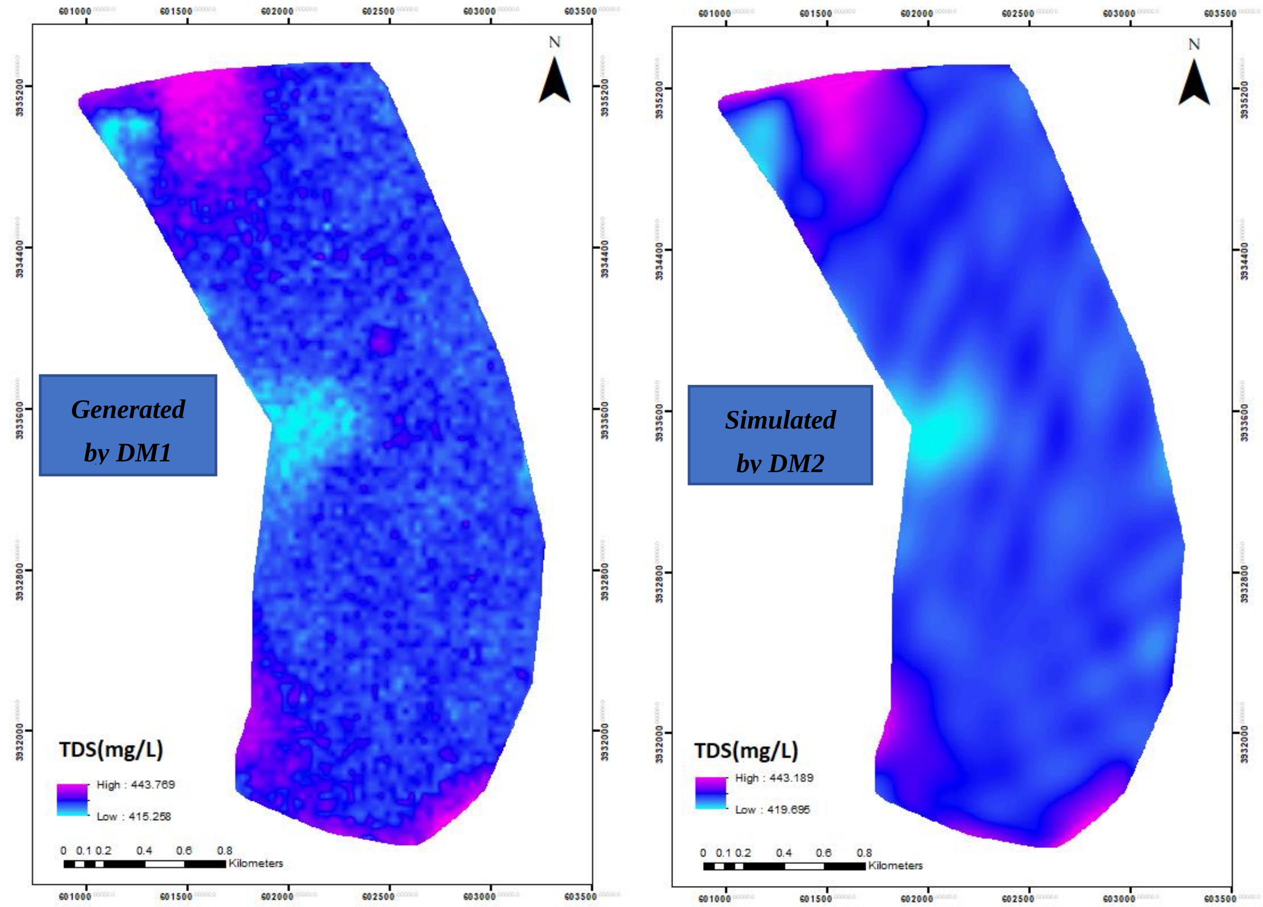

Figure 11. Outputs of test 3 (NSE = 0.72, RMSE = 1.62 mg/L).

In the final step, TDS suitability maps were generated to complete the testing process of the model. According to the initial ecological studies, 420 mg/L was determined as the threshold of TDS for classifying the habitats in the lakes. In other words, if TDS is higher than 420 mg/L, the aquatic habitat will be unsuitable. Conversely, if TDS is lower than the defined threshold, the habitat might be suitable for biological activities such as reproduction or searching for food. Figure 12 shows the habitat classification map generated based on DM1 and DM2 for each test. This figure corroborates that the developed data-driven model is highly reliable to classify the habitats of the lakes.

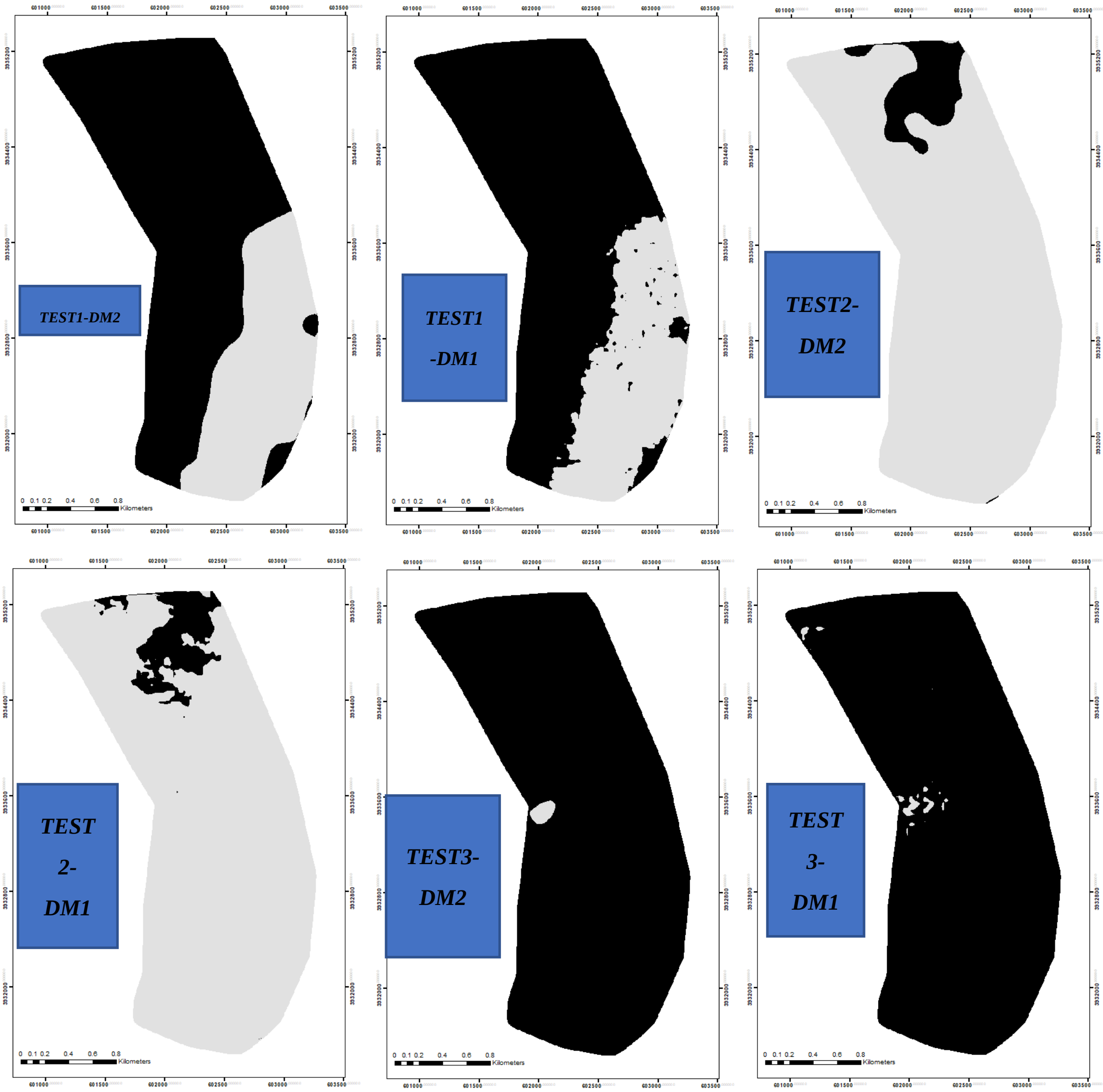

Figure 12. TDS suitability map for the lake in the three tests.

4. Discussion

A full discussion on different aspects of the model is essential. We applied an optimized neural network in the present study as it is highly advantageous compared with conventional neural networks in which the number of hidden layers is not optimized. In other words, the number of hidden layers is usually one or two in the conventional form of ANNs. However, other types of the data-driven models might be applicable as well. Many options might be available in this regard. For example, neuro fuzzy inference systems are another option for simulating water quality parameter distribution in the lakes. It is recommendable to apply these types of models, such as adaptive neuro fuzzy inference systems, in future studies in the aim to generate accurate results. The main advantage of the ANFIS-based model is higher interpretability. In fact, the ANNs work as a black box, which might be a disadvantage for these models. Furthermore, support vector machine (SVM) is another applicable data-driven model that could be applied in future studies. Deep learning methods are another option to simulate distribution of water quality parameters. For instance, convolutional neural networks (CNNs) might be an option in this regard. It seems that engineers face a wide range of models that could be applied in the process of prediction or simulation of water quality parameters associated with the remote sensing data processing.

The main advantage of the present study is development of a model that is able to simulate water quality distribution in the lakes. Conventionally, hydrodynamic models such as MIKE 21 could be utilized to generate distribution maps of the water quality parameters. However, using hydrodynamic models might be arduous. In fact, a significant drawback of these models might confine their applications in practice. For instance, considerable computational time is the most significant problem when using hydrodynamic models. In fact, the engineers need to cover a long-term period or to have numerous simulations for projects, which demonstrates that utilizing hydrodynamic models is not a proper solution in this regard. Moreover, the verification process of 2D hydrodynamic models might be very complex. In other words, changing calibration parameters is required to generate reliable results using the hydrodynamic models. Many trial and error attempts might be needed, which is a significant drawback for these models. The system developed by the present study is a robust alternative for 2D hydrodynamic models for simulating water quality parameter distribution in lakes. Similarly, the developed system could be used for coastal regions as well. However, some significant changes might be needed to develop a data-driven model for coastal regions.

Computational complexities are one of the important limitations in the application of AI methods. According to the official definition, computational complexities might be defined as the given time and memory needed by AI methods for optimization of the training process and simulation of the testing process. High computational complexities might reduce the applicability of the model considerably. Many are not willing to utilize models in which remarkable computational time for training or testing is required. Furthermore, complex models might need significant memory in the training process. The data-driven system developed in the present study is advantageous in terms of computational complexities. In other words, the developed data-driven model is efficient for the projects. It should be noted that other types of data-driven models might not be efficient in this regard. For example, ANFIS-based models might need more time for training processes.

The training method is another aspect in the application of the proposed method. We applied a known and familiar method for training the model. However, other advanced models might be useable as well. For example, evolutionary algorithms can be used in the training process of the models. The previous studies corroborated the applicability of the classic and new-generation algorithms in the training process of the data-driven models. Thus, it is recommendable to apply these algorithms in future studies for training the data-driven model.

The output of the present study demonstrated the strong relationship between changing the TDS concentration and spectral images. The previous studies corroborated the applicability of remote sensing analysis for other water quality parameters. Hence, the proposed model is not only able to simulate TDS concentration for unseen scenarios but might also be efficient for other water quality parameters, such as nitrogen or phosphate levels. The proposed method can be utilized for assessing the impact of agricultural development in the lake basins. It should be noted that using pesticides in the agricultural lands is a serious threat to the lakes due to draining agricultural runoff. An increasing population increases the need for food, which means that an increase in the cultivated areas is inevitable in future years. The developed method is helpful to investigate the environmental impact of agricultural development scenarios in which the area and use of biocides and fertilizers might be changed. In fact, the outputs of simulation tools such as the soil and water assessment tool (SWAT) can be utilized for the input of the proposed method to simulate distribution of water quality parameters such as TDS in lakes, as one of the important inland water ecosystems.

Another application of the proposed method is to investigate the potential impact of climate change in terms of increasing cultivated area in the future years. Linking climate change models, assessment tools such as SWAT and the proposed method in the present study is valuable for the projection of future environmental impacts on the lake ecosystems. It seems that the proposed methodology and other similar methods can improve the application of remote sensing analysis in water resource management. Hence, it is recommendable to link the proposed system to the optimization system of the water resources in the future studies. For example, linking the proposed method with reservoir operation models might be one of the options for simulating artificial lakes.

5. Conclusions

The present study developed an AI method for remote sensing data processing for simulating TDS concentration distribution in lakes using two optimal neural networks. The first data-driven model (monitoring model or DM1) was applied to convert the spectral images to the TDS distribution and is useable for monitoring water quality in the lakes. Inputs of the first model were the spectral bands of the OLI images by Landsat 8 (band 1 to band 7). The outputs demonstrated that an ANN in which the number of hidden layers has been optimized is highly efficient for water quality assessment using remote sensing analysis. The NSE was more than 0.6, which corroborates the robustness of the model in terms of monitoring water quality in the lakes. The second data-driven model, with the same structure as the first model, was developed to simulate TDS distribution in the lake for unseen scenarios for which no spectral image is available. The inputs of the second model were the simulated TDS inflow to the lake (mean monthly) as well as the longitude and latitude of each cell of the spectral image. In fact, the output of the second model is the TDS distribution map. The system performance analysis for the three tests demonstrated that the model is robust for simulation of the TDS distribution map for unseen scenarios. It is recommended to use the proposed model in other case studies and for other water quality parameters to corroborate its accuracy and efficiency.

Author Contributions: Methodology, M.S.; Software, M.S.; Resources, A.A.; Data curation, P.S. and A.A.; Writing—original draft, M.S.; Writing—review & editing, B.D.; Supervision, B.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All spectral images are freely available at https://earthexplorer.usgs.gov/ (accessed on 25 May 2023). Other data are available in the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ritchie, H.; Roser, M.; Environmental Impacts of Food Production. Our World Data. 2020. Available online: https://ourworldindata.org/ (accessed on 25 May 2023).
2. Mateo-Sagasta, J.; Zadeh, S.M.; Turral, H.; Burke, J. Water pollution from agriculture: A global review. In *Executive Summary*; FAO: Rome, Italy; IWMI: Colombo, Sri Lanka, 2017.
3. Wang, X.; Yang, W. Water quality monitoring and evaluation using remote sensing techniques in China: A systematic review. *Ecosyst. Health Sustain.* **2019**, *5*, 47–56. [CrossRef]
4. Topp, S.N.; Pavelsky, T.M.; Jensen, D.; Simard, M.; Ross, M.R. Research trends in the use of remote sensing for inland water quality science: Moving towards multidisciplinary applications. *Water* **2020**, *12*, 169. [CrossRef]
5. Maliki, A.A.; Chabuk, A.; Sultan, M.A.; Hashim, B.M.; Hussain, H.M.; Al-Ansari, N. Estimation of Total Dissolved Solids in Water Bodies by Spectral Indices Case Study: Shatt al-Arab River. *Water Air Soil Pollut.* **2020**, *231*, 482. [CrossRef]
6. Pizani, F.M.; Maillard, P.; Ferreira, A.F.; de Amorim, C.C. Estimation of water quality in a reservoir from sentinel-2 msi and landsat-8 oli sensors. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *3*, 401–408. [CrossRef]
7. Zhu, L.; Suomalainen, J.; Liu, J.; Hyyppä, J.; Kaartinen, H.; Haggren, H. A review: Remote sensing sensors. In *Multi-Purposeful Application of Geospatial Data*; IntechOpen: London, UK, 2018; pp. 19–42.
8. Forkuor, G.; Hounkpatin, O.K.; Welp, G.; Thiel, M. High resolution mapping of soil properties using remote sensing variables in south-western Burkina Faso: A comparison of machine learning and multiple linear regression models. *PLoS ONE* **2017**, *12*, e0170478. [CrossRef] [PubMed]
9. Singh, S.; Bhardwaj, A.; Verma, V.K. Remote sensing and GIS based analysis of temporal land use/land cover and water quality changes in Harike wetland ecosystem, Punjab, India. *J. Environ. Manag.* **2020**, *262*, 110355. [CrossRef] [PubMed]

10. Yang, H.; Kong, J.; Hu, H.; Du, Y.; Gao, M.; Chen, F. A review of remote sensing for Water Quality Retrieval: Progress and challenges. *Remote Sens.* **2022**, *14*, 1770. [CrossRef]
11. Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of machine-learning classification in remote sensing: An applied review. *Int. J. Remote Sens.* **2018**, *39*, 2784–2817. [CrossRef]
12. Huang, Y.; Fu, J. Review on application of artificial intelligence in civil engineering. *Comput. Model. Eng. Sci.* **2019**, *121*, 845–875. [CrossRef]
13. Poul, A.K.; Shourian, M.; Ebrahimi, H. A comparative study of MLR, KNN, ANN and ANFIS models with wavelet transform in monthly stream flow prediction. *Water Resour. Manag.* **2019**, *33*, 2907–2923. [CrossRef]
14. Sedighkia, M.; Datta, B. Utilizing Evolutionary algorithms for continuous simulation of long-term reservoir inflows. In *Proceedings of the Institution of Civil Engineers-Water Management*; Thomas Telford Ltd.: London, UK, 2021; pp. 1–35.
15. Wu, Y.; Ling, R.; Zhou, J.; Zhang, M.; Gao, W. Prediction of Water Quality based on artificial neural network. *J. Phys. Conf. Ser.* **2021**, *1738*, 012066. [CrossRef]
16. Chen, Y.; Song, L.; Liu, Y.; Yang, L.; Li, D. A review of the artificial neural network models for water quality prediction. *Appl. Sci.* **2020**, *10*, 5776. [CrossRef]
17. Lu, F.; Zhang, H.; Liu, W. Development and application of a GIS-based artificial neural network system for water quality prediction: A case study at the Lake Champlain area. *J. Oceanol. Limnol.* **2020**, *38*, 1835–1845. [CrossRef]
18. Stefenon, S.F.; Branco, N.W.; Nied, A.; Bertol, D.W.; Finardi, E.C.; Sartori, A.; Meyer, L.H.; Grebogi, R.B. Analysis of training techniques of ANN for classification of insulators in electrical power systems. *IET Gener. Transm. Distrib.* **2020**, *14*, 1591–1597. [CrossRef]
19. Cao, W.; Wang, X.; Ming, Z.; Gao, J. A review on neural networks with random weights. *Neurocomputing* **2018**, *275*, 278–287. [CrossRef]
20. Wang, Q.; Li, S.; Jia, P.; Qi, C.; Ding, F. A review of surface water quality models. *Sci. World J.* **2013**, *2013*, 231768. [CrossRef] [PubMed]
21. Yu, H.; Wilamowski, B.M. Levenberg–marquardt training. In *Intelligent Systems*; CRC Press: Boca Raton, FL, USA, 2018; pp. 12–21.
22. Gupta, H.V.; Kling, H. On typical range, sensitivity, and normalization of Mean Squared Error and Nash-Sutcliffe Efficiency type metrics. *Water Resour. Res.* **2011**, *47*, 10962. [CrossRef]