Estimating the Concentration of Physico-Chemical Parameters in Hydroelectric Power Plant Reservoir

by

Hieda Adriana Nascimento Silva

Ph.D. in Information and Communication Technology - XXX Cycle

University of Rome "La Sapienza"

September 2018

Author ..... Hieda Adriana Nascimento Silva

Department of Information Engineering, Electronics and Telecommunications

Certified by ..... Massimo Panella and Rita Asquini

Thesis Supervisor

This thesis was evaluated by the two following external referees:

Prof. Francesco Riganti Fulginei, Roma Tre University, Rome, Italy.

Prof. Giovanni Costantini, University Tor Vergata, Rome, Italy.

The time and effort of the external referees in evaluating this thesis, as well as their valuable and constructive suggestions, are very much appreciated and greatly acknowledged.

I would like to dedicate this thesis to my Mum

Abstract

The United Nations Educational, Scientific and Cultural Organization (UNESCO) defines the amazon region and adjacent areas, such as the Pantanal, as world heritage territories, since they possess unique flora and fauna and great biodiversity. Unfortunately, these regions have increasingly been suffering from anthropogenic impacts. One of the main anthropogenic impacts in the last decades has been the construction of hydroelectric power plants.

As a result, dramatic altering of these ecosystems has been observed, including changes in water levels, decreased oxygenation and loss of downstream organic matter, with consequent intense land use and population influxes after the filling and operation of these reservoirs. This, in turn, leads to extreme loss of biodiversity in these areas, due to the large-scale deforestation. The fishing industry in place before construction of dams and reservoirs, for example, has become much more intense, attracting large populations in search of work, employment and income.

Environmental monitoring is fundamental for reservoir management, and several studies around the world have been performed in order to evaluate the water quality of these ecosystems. The Brazilian Amazon, in particular, goes through well defined annual hydrological cycles, which are very importante since their study aids in monitoring anthropogenic environmental impacts and can lead to policy and decision making with regard to environmental management of this area. The water quality of amazon reservoirs is greatly influenced by this defined hydrological cycle, which, in turn, causes variations of microbiological, physical and chemical characteristics.

Eutrophication, one of the main processes leading to water deterioration in lentic environments, is mostly caused by anthropogenic activities, such as the releases of industrial and domestic effluents into water bodies.

Physico-chemical water parameters typically related to eutrophication are, among others, chlorophyll-a levels, transparency and total suspended solids, which can, thus, be used to assess the eutrophic state of water bodies.

Usually, these parameters must be investigated by going out to the field and manually measuring water transparency with the use of a Secchi disk, and taking water samples to the laboratory in order to obtain chlorophyll-a and total suspended solid concentrations. These processes are time- consuming and require trained personnel. However, we have proposed other techniques to environmental monitoring studies which do not require fieldwork, such as remote sensing and computational intelligence.

Simulations in different reservoirs were performed to determine a relationship between these physico-chemical parameters and the spectral response. Based on the in situ measurements, empirical models were established to relate the reflectance of the reservoir measured by the satellites. The images were calibrated and corrected atmospherically.

Statistical analysis using error estimation was used to evaluate the most accurate methodology. The Neural Networks were trained by hydrological cycle, and were useful to estimate the physical-chemical parameters of the water from the reflectance of visible bands and NIR of satellite images, with better results for the period with few clouds in the regions analyzed.

The present study shows the application of wavelet neural network to estimate water quality parameters using concentration of the water samples collected in the Amazon reservoir and Cefni reservoir, UK. Sattelite imagens from Landsats and Sentinel-2 were used to train the ANN by hydrological cycle.

The trained ANNs demonstrated good results between observed and estimated after Atmospheric corrections in satellites images. The ANNs showed in the results are useful to estimate these concentrations using remote sensing and wavelet transform for image processing.

Therefore, the techniques proposed and applied in the present study are noteworthy since they can aid in evaluating important physico-chemical parameters, which, in turn, allows for

identification of possible anthropogenic impacts, being relevant in environmental management and policy decision-making processes.

The tests results showed that the predicted values have good accurate. Improving efficiency to monitor water quality parameters and confirm the reliability and accuracy of the approaches proposed for monitoring water reservoirs.

This thesis contributes to the evaluation of the accuracy of different methods in the estimation of physical-chemical parameters, from satellite images and artificial neural networks. For future work, the accuracy of the results can be improved by adding more satellite images and testing new neural networks with applications in new water reservoirs.

Acknowledgments

Firstly, I am grateful to The God for enlightening my mind, giving me the skill and perseverance necessary to complete those studies, and the opportunity to enjoy such a wonderful experience.

I would like to express my sincere gratitude to my supervisor Prof. Massimo Panella and my coadvisor Prof. Rita Asquini for his support and guidance, for their advice, your continuous support, passion for research, patience, motivation, and immense knowledge for letting me try, fail, re-try and and finally get success in the results. Thank you!

I would also like to thank the University of Rome, La Sapienza, my lectures and coordination, Maria-Gabriella Di Benedett, always very dedicated and available to help us. Also, Thank you the Federal University of Para - Brazil, for allowing me to be here to do this doctorate.

My sincere thanks also goes to Prof. Giovanni Laneve (La Sapienza), for their contribution and for constantly reviewing my work, thank you for believing in me. I would like to extend my thanks to Prof. Antonio Morais da Silveira (UFPA) and Profa. Terezinha Ferreira de Oliveira (UFPA) for their contribution in the initial step of this project.

My sincere thanks to my friend Paola Gabriela Vinueza, For all the days we studied together and research collaboration.

My sincere gratitude to my laboratory colleagues Antonello Rosato, Francesca Ortolani and Rosa atilio, your support, research and exchange of experience, passion for research and motivation helped me to develop more and more.

I would also like to thank the European Union for its financially through the Euro Braziliam Windows (EBW+) Program, with financial support granted by the European Commission through the Erasmus Mundus Programme, grants scholarships to Brazilian students and staff members to undertake a mobility period, funding my research activities in Italy - Rome, "La Sapienza". Also, Thank you to the Brazilian Government for funding my work.

Thank you also the companies that gave the databases for the researches to be developed: Eletrobras - Eletronorte - Brazil and Welsh Water and Hamdden Ltd - UK;

I would like to thank my husband Robert Burnage, for he has always supported me on this long journey, I am encouraged with love and patience and often correcting my "English". So also can collaborate to improve my writing. He was always helping at any time, my partner in life, on travel, good times and difficult times.

Of course, infinite thanks to my family, my classmates, Mônica Yonashiro, Reginaldo Guiraldelli e Bruna Brogin, You contributed to my story in Rome. I'd like to say thank you to all my friends who contributed, even at a distance, to make this dream a reality.

I could only conclude that: "If I have seen further it is by standing on the shoulders of Giants." ... I miss you Roma. Ciao! Grazie e Arrivederci.

List of Figures

List of Tables

Nomenclature

($\mu s/cm$) MicroSiemens per centimeter

($Chla$) Chlorophyll-a Levels

(mg/lit) Milligram per liters

$ANFIS$ Adaptive Neuro Fuzzy Inference System

ANN Artificial Neural Network

BOD Biochemical Oxygen Demand

BP Back Propagation

BPA Back Propagation Algorithm

COD Chemical Oxygen Demand

$COSQC$ Central Organization for Standardization and Quality Control

Cr Chromium element

CWT Continuous Wavelet Transform

DNs Digital Numbers

DO Dissolved oxygen

DOS Dark Object Subtraction

EC Electrical conductivity

$FBNN$ Feedforward Neural Network

Fe Iron element

FIS Fuzzy Inference System

MLP Multi layer perceptron

$MLPANN$ Multi-Layer Rerceptron Articial Neural Network

MSE Mean Squared Error

$NERC$ Natural Environment Research Council

Rrs Remote Sensing Reflectances

SCP Semi-automatic Classification Plugin

SD Secchi Disk Depths

SVM Support Vector Machine

T Turbidity

TDS Total Dissolved Solids

$Temp$ Temperature

TN Total Nitrogen

TP Total Phosphorus

TSS Total Suspended Solids

$USGS$ United States Geological Survey

$WANN$ Wavelet Artificial Neural Network

WB Waterbody

WT Wavelet

Zn Zinc element

PCPC Physico-Chemical Parameters Concentrations

Contents

Chapter 1

Introduction

Contents

1.1 Introduction

This chapter is for introducing the basic concepts that concern to Artificial Neural Networks (ANNs), Wavelet Artificial Neural Networks (WANN) and Remote Sensing (RS), in order to estimate water quality variables, to understand the purpose of our work, justifying and identifying applications.

Giving a short description of the structures and technological paradigms, and describe some challenges and the proposed objectives to address them.

ANNs are one of the most commonly used artificial intelligence methods and are very applicable to meteorology and hydrology studies especially for modelling and forecasting [1] [2]. The main structure of the ANNs is very similar to the human brain which can recognize patterns and learn from examples. The WANN denotes the conjunction of wavelet decomposition and ANN. The wavelet decomposition is employed to decompose an input time series into approximations and details components. In this research, the decomposed time series are used as inputs to ANNs.

A wavelet pre-processing of input data (wavelet neural networks WANN) was used to improve the accuracy of ANN models for physico-chemical parameters estimation. A number of different ANN and WANN models for all parameters have been tested. Moreover, the performance of

WANN models was investigated using several mother wavelets including Haar wavelet (db1). The forecast results of all models were compared using five performance measures (NMSE, MSE, SNR, MAPE, MARE) .

A comparison has been done between observed and estimated data, the results of this study indicate that the coupled wavelet neural network (WANN) models obtained better results for the season with few clouds in the satellite images.

In previous works, the prediction of physical chemical parameters using WANN models had not considered the hydrological cycle of the region. This analysis was made in the present study, during the whole hydrological cycle of the region, from 2007 to 2014, constituted by a full, emptying, empty and filling stages.

The satellite images analyzed were obtained from the U.S. Geological Survey (USGS) (available at https://espa.cr.usgs.gov/), they include spectral band 2 (Blue), band 3 (Green), band 4 (Red) and NIR. All satellite images used were calibrated for radiance values and, subsequently, for reflectance values, as described in many papers [3, 4, 5, 6].

The results of this approach showed that WANN using satellite imagery is fast, precise, very efficient, and can effectively predict physico-chemical parameters in water reservoirs.

The proposed methodology is, therefore, relevant in environmental management and policy decision-making processes, in order to mitigate these impacts and attempt the recovery of degraded water bodies. They can be applied to other ecosystems that also suffer the effects of a well-defined hydrological cycle.

The present study investigates artificial intelligence techniques for estimating physico-chemical parameters in water reservoirs (in this case the Tucurui in Brazil and Cefni in UK) and to compare their performance with a traditional time series modelling technique.

The structure of the proposed model was simplified in WANN models, being very convenient with good accuracy. The final forecast models can be utilized for other water quality parameters and other water reservoirs.

1.2 Importance of Research and motivation

Water is one of the most essential components of human life. This natural resource generates socioeconomic development for society in general, more specifically for industry, agriculture and public use. Water quality involves physico-chemical and biological processes. These processes are necessary for the existence of life as it is an important factor for health. Therefore, the monitoring of water quality variables is of fundamental importance for society in general.

The values of the physico-chemical parameters present in the water should be within the standards allowed by the legislation, since the presence of components in high concentrations can be harmful to human health and the ecosystem. However, the presence of these components in the Water is important for the geochemical cycle and environmental health, so it is necessary to analyze and periodically monitor these components[7].

Due to the development and economic and population growth, Brazil is increasingly constructing water reservoirs, thus increasing the demand for drinking water and hydroelectric potential. [8]

Water reservoirs have different seasonal characteristics and several studies have been developed to identify and characterize seasonal differences as well as physico-chemical parameters in aquatic environments.[9] [10]

Some researches were developed with the objective of identifying physico-chemical parameters in aquatic environments using Computational intelligence [11] [12] [13] [14]. However, these studies do not analyze the hydrological cycles of the regions.

In this context, this research proposes a method, using the techniques of computational intelligence to infer levels of the physico-chemical parameters in bodies of water in reservoirs, using historical data, collecting water several years, analyzing and predicting the physico-chemical parameters per hydrological cycle.

This thesis aims to investigate and propose a computational solution, using computational intelligence and remote sensing techniques to infer the levels of physico-chemical parameters present in the body of water.

Satellite images were used, captured in the same period of the water collection, verified the reflectance of the substances in the images, so we developed a neural network to predict these parameters.

We investigated the reflectance condition in water bodies of two reservoirs, one in the Brazilian Amazon, considered a deep reservoir (reservoir of the hydroelectric plant of Tucuruí) and another located in the United Kingdom, Cefni, in order to validate the proposed model.

1.3 Context of research

The construction of hydroelectric power plants in the Amazon region has generated social and environmental impacts in the last decades, and statistical data indicate that renewable energy, including hydroelectric power, is expected to increase by almost 85% between now and 2030 [15].

While Russia and China have natural gas and coal reserves, Brazil relies on hydro power for 85% of its electricity needs [16].

The changes in the aquatic environment due to the construction of reservoirs have been studied by several authors[17][18]. These studies have shown that the changes in physico-chemical parameters can directly affect water quality and local biota because freshwater ecosystems are an important natural resource, essential for multiple purposes such as drinking, domestic use, industrial cooling, power generation, agriculture, waste disposal, and transportation routes.[19].

The presence of physico-chemical parameters in waterbody is also an important component for the geochemical cycle and biodiversity, but high concentrations can negatively influence ecosystems when they are in quantities not allowed by legislation[11] [20] [21]. For these reasons, many scientists have studied the influence of physico-chemical parameters in freshwater and how this affects ecosystems.

Studies involving Remote Sensing (RS), Artificial Neural Networks (ANN), Wavelet Transform (WT), Adaptive neuro-fuzzy inference systems (ANFIS) with cross-validation and Statistical Analysis have recently gained attention in the literature for the monitoring of water quality. [22] [23] [24] [25] [26] [27][28] [29][30][31][32][33] [34] Among the monitoring techniques used are those using remote sensing [35]and computational intelligence [36].

The artificial neural network (ANN) technique is a tool for modeling real-world problems and has been used to evaluate the physico-chemical parameters of water.

Another technique that has also been gaining attention in prediction models for water quality monitoring is the adaptive neuro-diffuse inference system (ANFIS) [37].

In some studies, the performance of the ANFIS model was compared with an artificial neural network model. The ANFIS model was able to provide greater accuracy, particularly in the case of extreme events [24]. Considered an option with greater precision and reliability for the treatment of forecasting problems involving training and prediction of concentrations of various parameters.

Among the several important parameters, we can mention Levels of Chlorophyll (C), Total Suspended Solids (TSS) and Transparency (T), as parameters that can influence ecosystems and were considered important factors for monitoring water quality in reservoirs[38][39] [40] [41].

All these parameters are currently evaluated by companies, however, the methods currently used for water analysis are very time-consuming, extremely expensive because they require sample collection, trained personnel and specialized laboratories.

We propose a less expensive and more dynamic method to monitor these parameters, using techniques of artificial intelligence and remote sensing.

The proposed method uses satellite images and the "in situ" measurements made by the responsible companies. The water samples were collected in 7 sampling stations called: Caraipé 1 (C1), Caraipé 2 (C2), Breu Branco (MBB), Jacunda (MJV), Upstrem 1 (M1) Upstrem 3 (M3), Ipixuna (MIP).

The Satellite Sensor chosen was Landsat 7 ETM +, Landsat 8 OLI (Operational Land Imager), and Sentinel-2. Landsat 7 ETM + has a spatial resolution of 30 m for the six reflective bands, 60 m for the thermal band, and includes a panchromatic band (pan) with a resolution of 15 m.

Landsat 7 has a 378 gigabit(Gb) Solid State Recorder (SSR) that can hold 42 min (approximately 100 scenes) of sensor data and 29 h of housekeeping telemetry concurrently (L7 Science Data User's Handbook[1]).

We downloaded the satellite images from 2007 to 2014 from the Earth Explorer USGS [2]. These images were classified and converted into vector format, which served as inputs for the neural model. The output of the neural network was validated with the samples analyzed in a chemical laboratory performed by the companies responsible.

The WANN techniques was developed to monitor the physico-chemical parameters in two reservoirs: Tucurui-Brazil and Cefni-UK.

[1] http://landsathandbook.gsfc.nasa.gov/handbook.html
[2] https://espa.cr.usgs.gov/

1.4 Objectives

The degradation of surface water quality occurs due to the presence of various types of pollutants generated from human, agricultural, and industrial activities. Thus, mapping concentrations of different surface water quality parameters (SWQPs), such as turbidity, total suspended solids (TSS), chemical oxygen demand (COD), biological oxygen demand (BOD), and dissolved oxygen (DO), is indeed critical for providing the appropriate treatment to the affected waterbodies.

Traditionally, concentrations of SWQPs have been measured through intensive field work. Additionally, quite a lot of studies have attempted to retrieve concentrations of SWQPs from satellite images using regression-based methods.

The remote-sensing method based on the back-propagation neural network is developed to quantify concentrations of different SWQPs from satellite images of Landsat7, Landsat8 and Sentinal-2.

In this sense, this research has the following general objective and specific objectives:

1.4.1 General objective

The general objective of this research is based on computer intelligence paradigms to aid in the monitoring of water bodies in reservoirs. In this sense, this proposal aims at the creation of computational solutions with remote sensing and wavelet neural networks to analyze the reflectances of the satellite images and to estimate the physico-chemical parameters concentrations present in the waterbody. The goal is to create an artificial neural model to predict the parameters and infer the future level of these substances.

1.4.2 Specific objectives

- Investigate the reflectance in the satellite images, Landsat7, Landsat8 and Sentinel-2.
- Investigate physico-chemical parameters in water reservoirs, Tucurui-Brazil and Cefni-UK.
- Analyze the available historical data base of the physico-chemical parameters from reservoirs.
- Analyze the satellite images corresponding to the dates of the water samples collected by the responsible company.
- Develop a neural model using Computational Intelligence techniques and satellite imagery to monitor water quality parameters.
- Pre-processing of satellite images.
- Use data mining techniques to identify patterns of behavior in historical data series.
- Propose a solution, low cost, using Computational Intelligence techniques to infer levels of physico-chemical parameters present in the reservoirs water bodies, where the reflectance conditions manifest favorable to infer the future level of these substances, from reservoir images.
- To evaluate and propose the repositioning of the water sampling points of the Tucurui hydroelectric reservoir through the interpretation of satellite images of these reservoirs using paradigms of Computational Intelligence techniques, using regression, classification, Clustering and Statistical Analysis.
- To contribute to the debate on the issues investigated, presenting proposals for contemporary solutions.

1.5 Problems investigated

- The construction of several hydroelectric power plants in the last decades, due to the increasing demand of energy.

- Monitoring of water quality variables in reservoirs.

- Prediction and analysis of physico-chemical parameters, ANNs and WANN.
- Eutrophication and physico-chemical parameters.

Eutrophication, one of the main processes leading to water deterioration in lentic environments, is mostly caused by anthropogenic activities, such as the releases of industrial and domestic effluents into water bodies. Evidence of eutrophication, includes the increase in turbidity, suspended solids and algae blooms, and the change in water colour toward green due to an increase in phytoplankton abundance.

The increased probability of algal blooms occurring is of major concern, especially where these blooms are due to (toxic) cyanobacteria species. Under natural conditions in aquatic ecosystems, a balance exists between cyanobacteria and other phytoplankton groups [42].

Microcystis aeruginosa in particular is considered a cyanobacterial harmful algal bloom (CyanoHAB) organism because it can impede rec- reational use of waterbodies, reduce esthetics, lower dissolved oxygen concentration, and cause taste and odor problems in drinking water, as well as produce microcystins, powerful hepatotoxins associated with liver cancer and tumors in humans and wildlife [43]. For these reasons, monitoring of algae blooms in drinking water and reservoirs is very importante and necessary.

Cyanobacteria have adapted to survive in a variety of environments and have been found globally. Toxin-producing cyanobacterial harmful algal blooms (CHABs) have been increasing in frequency worldwide and pose a threat to drinking and recreational water [44].

Cyanobacterial harmful algal blooms (CHABs) are excessive growths of cyanobacteria (also known as blue-green algae) in water. Not all cyanobacteria produce toxins, but it has been hypothesized that the aquatic environment with increasing water temperatures will favor the bloom of toxin-produc- ing harmful cyanobacteria, such as Microcystis sp [45]. These blooms can produce a variety of cyanotoxins as well as taste and odor compounds that pose a nuisance to water supplies worldwide [44].

Toxic cyanobacteria poisonings (CTPs) are responsible for almost all known cases of fresh and brackish water intoxication involving phycotoxins (1, 2). Toxic water blooms can be found in many eutrophic to hypereutrophic lakes, ponds and rivers throughout the world and are responsible for sporadic but recurrent episodes of wild and domestic animal illness and death. They are also implicated in human poisonings from certain municipal and recreational water supplies. The predominance of cyanobacteria over other species occurs under specific environmental conditions, including optimal light intensity and water temperature, nutrient availability and stability in the water column [18].

1.6 Physico-Chemical parameters analyzed for water reservoirs.

The parameters mentioned in the list below were analyzed in the reservoirs. Models using Wavelet Neural Network was implemented to infer future levels of these concentrations. Then, statistical models were applied to validate the results obtained, considering the hydrological cycle of the study area as: Full, Emptying, Empty, Filling.

The following parameters were monitored:

- Transparency (cm)
- Total Suspended Solids (mg/L)
- Chlorophyll (mg/L)
- Temperature (oC)
- Conductivity (µS/cm)
- pH
- Ammonia (mg/L)
- NO3 (mg/L)

- Turbidity (NTU)
- dissolved oxygen (mg/L)
- PO4 (mg/L)
- P Total (mg / L).

The methodology used involves water collecting, satellite images, remote sensing and Artifical Intelligence.

1.7 Water collection

The water data were provided by the Eletronorte/Eletrobras Company for Tucurui reservoir and the points chosen are considered the most important by this company for water analysis. These data were collected from January 2007 to December 2014.

For cefni in the centre of Anglesey, we obtained the data by Welsh Water and Hamdden Ltd. In the Cefni Reservoir samples will have been collected, where practical, from a boat, and from the same location. If boat sampling is not practical, sampling will have taken place from a fixed location, often the outlet tower. In severe weather, or in times of staffing pressure, samples may be missed.

Sample point names may have confusing annotations, e.g. the name "Queen Elizabeth 1m depth" does NOT refer to the depth at which a sample was collected, merely the location.

Results generated from an individual point in the water column are linked by a unique identifier in the table column "Tag Number".

The depth in the water column from which a sample is collected is stated against the determinand "sample/pump depth below ground". If only one sample was collected on a particular visit, the result value may be null, in which case a value of 1 may be assumed.

Chlorophyll , temperature, and dissolved oxygen values are measured on site with a Xylem YSI 6600 V2 Sonde.

Discrete water samples at depth are collected using a Patalas sampler, similar to a Schindler-Patalas sampler, but without an attached net or side port.

The relationship between Chlorophyll-a Levels, Total Suspended Solids, Transparency and spectral response of the riverwater was determined using the physico-chemical water samples collected. These data have been extracted from the samples and analyzed. We compared it to the proposed parameters level extracted from the remote sensing images, analyzed with an ANN method. This was done for the entire hydrological cycle of the area.

1.8 Introduction to Artificial Neural Networks (ANN)

Neural networks are a class of flexible nonlinear models that can discover patterns adaptively from the data. Theoretically, it has been shown that given an appropriate number of nonlinear processing units, neural networks can learn from experience and estimate any complex functional relationship with high accuracy. Although many types of neural network models have been proposed, the most popular one for time series forecasting is the feed-forward model.

A feed-forward ANN comprises a system of units, analogous to neurons that are arranged in layers [46]. A multilayer perceptron (MLP) is the most popular neural network architecture. MLPs have often been used in hydrology for their simplicity. MLP is typically composed of several layers of nodes (neurons). An input layer, which is the first layer, receives external information. The problem solution is obtained in an output layer which is the last layer. One or more intermediate layers, which are called hidden layers, separate the input and output layers. The nodes in adjacent layers are usually fully connected by acyclic arcs, which are called synapses,

from the input layer to the output layer [47]. The proposed model structure is presented in the next section

An ANN is a parallel-distributed processor that resembles the human brain by acquiring knowledge through a learning process and, then, stores the knowledge in the connection strength between computational units called neurons[48]. We used for this purpose a simple feedforward neural network in which the information moves in only one direction, from the input to the output nodes. The preprocessed data is given as the input of the ANN, and the outputs are: C, S and T.

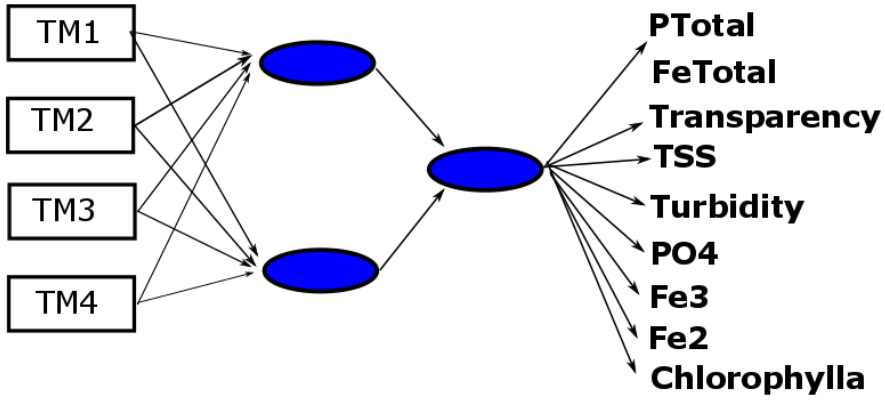


Figure 1.1. ANN Model by TM

The Fig 1.1 shows Inputs for the neural network from image landsat satellite, sensor ETM, band 1, band 2, band 3 and band 4. The process was repeated for each year from 2007 to 2014. As a result, estimates for PTotal, FeTotal, Transparency, TSS, Turbidity, PO4, Fe3, Fe2 and Chlorophyll-a.

1.9 Introduction to Wavelet Artificial Neural Networks (WANN)

The wavelet transform is an integral transform whose kernel is a class of special functions, called wavelets [49]. The main advantage of this method compared to other methods is its spectral location capability in space and frequency, which allows for the analysis of non-stationary signals in their various scales [50].

The wavelet transform used in the present study was the discreet transformed, with allows for the multi-resolution analysis of a signal, decomposing said signal into approximations and details. The approaches are high ranges, i.e., low-frequency signal components. The details are the low ranges, i.e., high frequency components[51].

The Haar family with a degree of decomposition in the Matlab software Mathworks 2016b package was used to implement the proposed model.

One sampling station was initially chosen for analysis and an image of 32x32 pixels was cropped, corresponding to an array containing 1024 pixels. Subsequently, the wavelet transform was applied, with only one level of decomposition, resulting in a matrix array of 16x16 pixels for each of the following three components: Wavelet - Horizontal (H), vertical (V) and diagonal (D).

The conversion of the arrays with the Wavelet H, V and D components to their respective column-matrices was performed, and subsequently a concatenation of the three arrays (each containing 256 pixels) was executed, generating a vector with 768 column size (256 x 3). This data was used as the ANN input.

Neural networks are parallel distributed systems consisting of two basic types of components: the processing units, arranged in one or more layers, interconnected, called neurons; and the synapses, which are the connections between the processing units.

In the present study, the ANN paradigm applied was the Direct Multilayer Perceptron, developed using the Matlab software package [52]. Post-processed images of the ten sampling sites per water cycle were used for the ANN input and the output variables were the the physico-chemical parameters of water.

1.10 Contributions

This thesis contributes to the development of a method to predict physico-chemical parameters such as chlorophyll-a levels, transparency and total suspended solids in the waters reservoirs by Wavelet Transform of satellite images, which demonstrate sensitivity to the presence of suspended particulate matter in water bodies. WANNs were developed and analyzed, using the pixels of the satellite images as input, considering the seasonal characteristics of the regions .

The method resulted in good approximations of laboratory results regarding the same water samples. It is clear that the proposed method in this study is specific to the Tucuruí hydroelectric power plant in Brazil and to Cefni reservoir in UK, although nothing prevents this method to be used in monitoring waters from other reservoirs, which in turn should result in less expenditures on environmental monitoring processes.

The techniques proposed and applied in the present study can aid in evaluating important physico-chemical parameters which, in turn, allows for identification of possible anthropogenic impacts. The proposed methodology is, therefore, relevant in environmental management and policy decision-making processes, in order to mitigate these impacts and attempt the recovery of degraded water bodies. They can be applied to other ecosystems for water quality analysis, thus demonstrating worldwide relevance.

1.11 Thesis Structure

The remainder of the thesis is organized as follows:

- Chapter 2: Presents the highlights of the research issues related to satellite remote sensing, with proposed methods, field survey methodology and study area.
- Chapter 3: Presents the Wavelet Artificial Neural Networks (WANN), methodology proposed with the wavelet transform research.
- Chapter 4: Presents the survey of the state of the art on the subject and water quality research.
- Chapter 5: In this chapter we present the data processing, with Dark Object Subtraction (DOS) method and factor analysis results.
- Chapter 6: In this chapter we present Inference Models implemented.
- Chapter 7: Case Study: Tucuruí Hydroelectric Reservoir (THR) with detailing the study area, characterization of sampling stations, the process of analysis of physicochemical parameters and sampling stations.
- Chapter 8: Case Study:Cefni Reservoir, Anglesey, UK with the process of acquisition of spatial data, statistical methods and used as well as remote sensing and image processing.

- Chapter 9: Results for Tucurui and Cefni Reservoirs.
- Chapter 10: Conclusion and Future Research.

Publications:

- Lobato, T. C., R. A. Hauser-Davis, T. F. Oliveira, A. M. Silveira, **H. A. N. Silva**, M. R. M. Tavares, and A. C. F. Saraiva. "Construction of a novel water quality index and quality indicator for reservoir water quality evaluation: A case study in the Amazon region." *Journal of hydrology* 522 (2015): 674-683
- **H.A. Nascimento Silva**, G. Laneve, A. Rosato, and M. Panella, "Retrieving Chlorophyll-a Levels, Transparency and TSS Concentration from Multispectral Satellite Data by Using Artificial Neural Networks", *Proc. of Progress in Electromagnetics Research Symposium - Fall* (PIERS - FALL 2017), pp. 2876-2883, ISBN: 978-1-5386-1211-8, DOI: 10.1109/PIERS-FALL.2017.8293624, IEEE, Singapore, November19-22, 2017.
- **H.A. Nascimento Silva**, A. Rosato, R. Altilio , and M. Panella, "Water Quality Prediction Based on Wavelet Neural Networks and Remote Sensing", *Proc. of International Joint Conference on Neural Networks* (IJCNN 2018), pp. 1-6, IEEE, Rio de Janeiro, Brasil, July 8-13, 2018.
- **H.A. Nascimento Silva** and M. Panella, "Eutrophication Analysis of Water Reservoirs by Remote Sensing and Neural Networks", *Proc. of Progress in Electromagnetics Research Symposium* (PIERS 2018), pp. 1-7, IEEE, Toyama, Japan, August 1-4, 2018.
- **H.A. Nascimento Silva** and M. Panella, "Neural Networks for Multispectral Analysis of Water Reservoirs", submitted to *Expert Systems with Applications*, 2018.

Chapter 2

Satellite Remote Sensing

Contents

2.1 Introduction

Monitor water quality variables is a very challenging task, parameters such as algae Blooms, Suspended Solids and Turbidity, Transparency, among others, which includes dead, inert, and degraded particles are parameters analyzed to assess the water quality monitoring.

Remote sensing techniques can be used to calculate suspended sediment concentrations. The discrimination of Suspended Sediments or Total Suspended Solids from water reflectance is based on the relationship between the scattering and absorption properties of water and its constituents. Most of the scattering is caused by suspended sediments and the absorption is controlled by chlorophyll-a and coloured dissolved organic matter. These absorptives in-water components have been shown to lower the reflectance in a substantial way. However, these absorptive effects are generally found at wavelengths less than 500 nm, and can be monitored by remote sensing techniques [53, 54, 55].

Recent studies carried out show successful application of freely available remotely sensed data in extracting critical water quality parameters, such as Algae Blooms, chlorophyl-a, suspended solids, turbidity [56, 57, 58], and used remote sensing technology to assess water quality levels[53, 55, 59, 54, 60]. Remote sensing techniques have substantially improved our ability to observe the environment and its processes. However, remote sensing technologies are underutilized in environmental management [61, 62].

Other research claims that eutrophication and climate change are two processes that may promote the proliferation and expansion of cyanobacterial harmful algal blooms, and in man-made reservoirs has received considerable attention over time due to its harmful effects on the aquatic environment and on human and animal health [63].

The satellite images approach recently showed good experiences in Mapping Lake Water Quality Parameters. The suitability of Sentinel-2 Multispectral Imager's (MSI) and data from

another satellites for mapping different lake water quality parameters was tested with good results in some recent studies. [55, 59].

Sentinel-2 MSI images were used to estimate cyanobacteria density by means of empirical models and significant progress has been done in the development of methods for retrieving information on phytoplankton biomass and the detection and monitoring of cyanobacterial forests in lakes, reservoirs and large rivers [64] emphasized the value of MSI images based on the blue to green spectral region for assessing waters with a low to medium amount of biomass of blue-green algae.

The study of Algal, Suspendes Solids and Turbidity in the Cefni reservoir, Anglesey, UK, is of great importance, due to the multiple uses of its water. However, the reservoir has only been mentioned in a few studies [65, 66, 67].

Monitoring algal with satellite images and neural networks has not performed in previous studies. Here we propose a multidisciplinary approach that aims to evaluate the extension of algal blooms using satellites images and artificial neural networks to predict algae, suspended solids and turbidity.

The approach was based on the water limnology of the Cefni and forecasting, consisting use of three inference systems for data regression based on neural and fuzzy neural networks.

The algal samples are used to feed three different models of function approximation and techniques based on neural and diffuse neural networks, which are adequate to predict data sequences that usually show a chaotic behavior: radial base function (RBF); the adaptive system of neuro-diffuse inference (ANFIS) [68].

As approached in [69], a predictor can be based on function approximation models, whose parameters are estimated by data-driven learning procedures.

In this study, the estimation of these important parameters of water quality and mitigation efforts were reviewed, focusing in the Cefni Reservoir. The results contribute to analysis in other reservoirs in the other parts of the world facing similar problems due to algae, suspended solids and turbidity in their freshwater environments. The approach was developed and tested from data collected in the reservoir.

The main goal of this study was to investigate if the combination between water limnology, remote sensing and neural networks is a suitable approach to identify algae extent in the Cefni reservoir.

This approach was based on the characteristics of the water conditions collected in the stations. The application of satellite remote sensing techniques may help to compensate for the limited spatial dimension of traditional in situ methods. It permits acquisition of necessary information at different spatial and temporal scales, allowing a more complete analysis of aquatic ecosystems, and it is a functional analysis in synoptic order.

The outcomes of this study aim to support the management of water reservoirs, with special attention to water quality variables. We introduce in this research the function approximation models that are used to predict time series. The application is verified by several benchmark results and extensive computational simulations, the results prove the validity of the proposed method.

2.2 Satellite Images

Thirty two satellite images from Landsat 7, sensor ETM+ were acquired for testing of the Tucurui reservoir. In order to obtain directly the TSS, C and T concentration from the reflectance of the satellite images, all satellite-image bands from visible and NIR were first calibrated for radiance values and, subsequently, for reflectance values. Image-based methods for atmospheric correction can estimate path radiance without using atmospheric properties, their accuracy is

highly dependent on what is captured in a scene, as described in many papers: [3], [4], [5], [6], [70]. The characteristics of the analyzed bands are reported in table 2.5.1

Table 2.1. Characteristics of Visible and NIR Bands of the Analyzed Sensors

Satellite	Bands	Spectral Resolution(nm)	Spatial Resolution(m)	Temporal Resolution
Landsat7(ETM+)	TM1	450-520	30	16 days
Landsat7(ETM+)	TM2	520-600	30	16 days
Landsat7(ETM+)	TM3	630-690	30	16 days
Landsat7(ETM+)	TM4	760-900	30	16 days

The proposed method corrects the atmospheric effect by estimating the path radiance spectrum based on the dark object subtraction (DOS) method so that the spectrum meets general spectral characteristics of path radiance. The atmospheric effects that influence the signal registered by remote sensors might be minimized in order to provide reliable spectral information.

In aquatic systems, the application of atmospheric correction avoids the under or overestimation of remote sensing reflectance (Rrs). Accurately Rrs provides better information about the state of aquatic system establishing the concentration of aquatic compounds more precisely [71].

The DOS method with semi-automatic classification plugin was used, as described elsewhere [72][73]. Afterwards, a relative scattering model was chosen based on the atmospheric conditions of the image at the acquisition time and the initial haze value for the other spectral bands were then calculated.

Equations and parameters to convert calibrated Digital Numbers (DNs) to physical units, such as at-sensor radiance and reflectance, have been presented in a "sensor-specific" manner elsewhere [74].

DN to Radiance: There are two formulas that can be used to convert DNs to radiance; the method you use depends on the scene calibration data available in the header file(s). One method uses the Gain and Bias (or Offset) values from the header file. The longer method uses the LMin and LMax spectral radiance scaling factors.

Conversion to spectral radiance is a substantial improvement over use of DNs in analysis. When transformed, all individual sensor measurements are in comparable physical units. This is generally accomplished through information supplied by the instrument developer in the form [75],[76] :

$$rad_\lambda = LMIN_\lambda + (Q_{cal} - Q_{calmin}) \left(\frac{LMAX_\lambda - LMIN_\lambda}{Q_{calmax} - Q_{cmin}} \right), \tag{2.1}$$

$$ref_{i,j} = CH_{i,j} \times GAIN\_BAND + OFFSET\_BAND. \tag{2.2}$$

where:
Lλ: Spectral radiance at the sensor's aperture $[W/(m^2 \times sr \times \mu m)]$
Qcal: Quantized calibrated pixel value [DN]
Qcalmin: Minimum quantized calibrated pixel value corresponding to LMINλ[DN]
Qcalmax: Maximum quantized calibrated pixel value corresponding to LMAXλ[DN]
LMINλ: Spectral at-sensor radiance that is scaled to Qcalmin $[W/(m^2 \times sr \times \mu m)]$
LMAXλ: Spectral at-sensor radiance that is scaled to Qcalmax $[W/(m^2 \times sr \times \mu m)]$
GAIN_BAND: Band-specific rescaling gain factor $[(W/(m^2 \times sr \times \mu m))/DN]$
Brescale: Band-specific rescaling bias factor $[W/(m^2 \times sr \times \mu m)]$
ESUN: Mean exoatmospheric solea irradiance

SEA: Sun elevation
SZA: 90.0 - SEA
OFFSET_BAND: The DN value where zero radiance is detected 2.2

Table 2.2. ETM+ spectral radiance range ($W/m^2 - sr - \mu m$)

Band Number	Low gain		High gain	
	LMIN	LMAX	LMIN	LMAX
1	- 6.2	293.7	-6.2	191.6
2	-6.4	300.9	-6.4	196.5
3	-5.0	234.4	-5.0	152.9
4	-5.1	241.1	-5.1	157.4

Images from 2007 to 2013 were used to train the system while the images from 2014 was used to validate the methodology.The images acquisition date are presented in table 2.3

Table 2.3. Images used for Training and Validation

Training							Validation
2007	2008	2009	2010	2011	2012	2013	2014
07-MAY	09-MAY	12-MAY	15-MAY	02-MAY	04-MAY	23-MAY	26-MAY
11-AUG	13-AUG	16-AUG	19-AUG	22-AUG	24-AUG	11-AUG	14-AUG
15-NOV	01-NOV	20-NOV	07-NOV	10-NOV	12-NOV	15-NOV	02-NOV
15-JAN	18-JAN	20-JAN	23-JAN	10-JAN	13-JAN	31-JAN	18-JAN

2.3 Proposed Methods

Cefni Reservoir on the Isle of Anglesey was overflown as part of the UK Natural Environment Research Council (NERC) Airborne Remote Sensing Campaign 1992. The lake is shallow, with a maximum depth of approximately 4 m and contains beds of submersed, floating-leaved and emergent aquatic macrophyte species. It is also known to support dense growths of toxic blue-green algae during summer. The reservoir is surrounded by an approximately 100-m wide plantation of coniferous trees with agricultural fields beyond [65].

Cefni Reservoir is managed by Welsh Water and Hamdden Ltd, while the fishery is managed by the Cefni Angling Association, area: 86 ha and length: 2.3 km. The collection of water were done periodically in Cefni reservoir.

2.4 Field Survey Methodology

Field data were collected during the summer, winter, autumn, and spring of 2015, 2016, 2017. The different cycles are a consequence of the differences in rainfall during the year and they influence a lot the water monitoring in the reservoir, Fig. 2.1. The area is of highly interest and water monitoring is one of the important things that have to be in place to ensure the sustainability of the reservoir.

Water samples have been collected, where practical, from a boat, and from the same location. If boat sampling is not practical, sampling will have taken place from a fixed location.

Sample point name, called “Queen Elizabeth 1m depth”, does not refer to the depth at which a sample was collected, merely the location. Results generated from an individual point in the water column are linked by a unique identifier in the table column “Tag Number”.

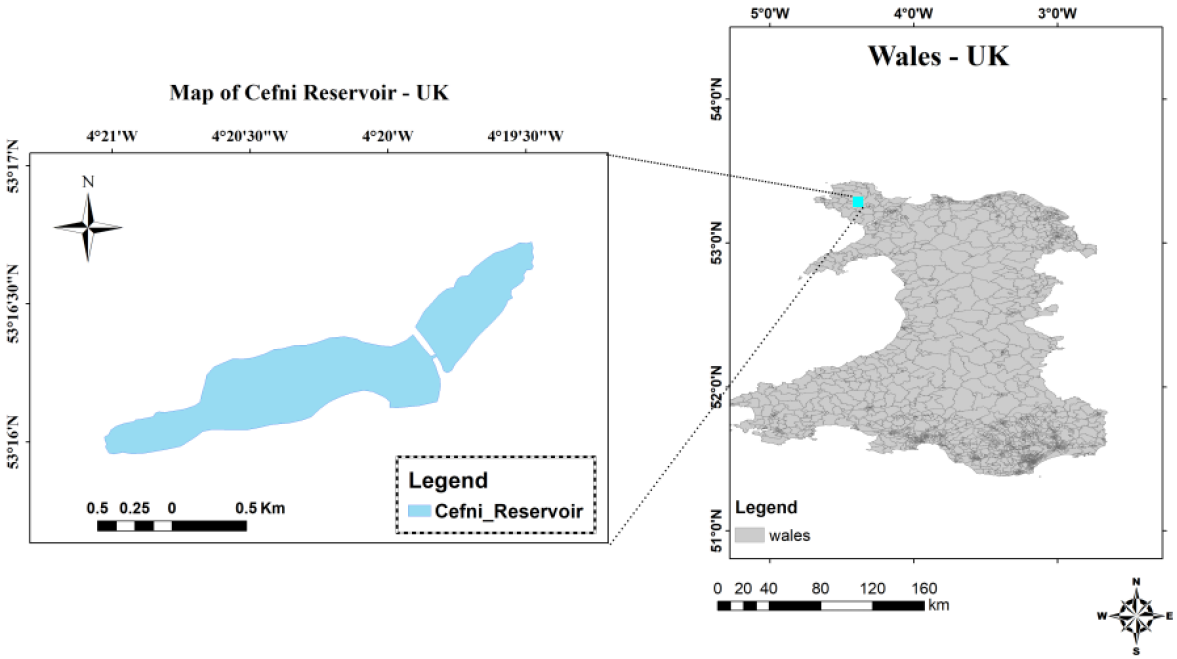


Figure 2.1. Satellite image of Cefni reservoir, Anglesey, U.K.

Algal samples for identification are collected by a full column vertical net haul using a 57µm mesh net in the case of boat samples, and are therefore concentrated, and "dipped" by a Patalas sampler for fixed location samples and are therefore not concentrated. All the results including the field readings are entered into the laboratory system, and are then archived. If the result value is null, a dip sample may be assumed.

In the laboratory, algal samples are observed using a microscope, and results coded. The data storage system does not store alpha results, only numeric ones, so each algal genus has been assigned a number, corresponding to type and genus with the abundance as defined by the ACFOR scale (see below) represented by the decimal fraction. The alga table has a reference column called "result number" which matches to the value for each algal result. This can be used to match up the results and discover which alga was in which sample. The ACFOR scale :

- A = Abundant (greater than/equal to 30 %)
- C = Common (20 to 29%)
- F = Frequent (10 to 19%)
- O = Occasional (5 to 9%)
- R = Rare (1 to 4%)

Water from the Cefni Reservoir is currently abstracted and treated for use as drinking water by Welsh Water. The Cefni reservoir is fed by an inflow from a river draining the Cors Erddreiniog fen, which is notably high in coloured Dissolved organic carbon (DOC) [77], fed by base rich waters from an escarpment of limestone bedrock. The site water chosen as this location has been identified as the main source of DOC within the catchment [77]. The presence of natural DOC in drinking water supplies is undesirable due to its reaction with chlorine during water treatment and the resulting formation of disinfection by-products (DBPs).

Rapid growth of algae during the summer months contributes up to 45 of all the DOC inputs, although the lower molecular weight of the carbon compounds ensures that algae do not generate trihalomethanes (THMs) as readily as terrestrially-derived DOC during water treatment.

2.4.1 Water Sampling and Analysis

The limnological measurements were made available by Welsh Water and Hamdden Ltd, they seem to the current owners of this particular reservoir. It covers the period from February 2015, which is the earliest period for which we have reservoir level available, to 31 December 2017, used for neural network training. The data is for Cefni Reservoir which is located on the Island of Anglesey, North Wales latitude 53.271131, longitude -4.3353295.

Water samples were measured periodically by Welsh Water company, 4 samples of water are collected per month, from 2015 to 2017, hich were used to train the neural network, as showed one example of collection in Tab. 2.5.1.

The satellite images were collected during the same period in the winter, spring, summer, autumn from the Sentinel 2 satellite, Tab.2.5.1, using band 8, band 4, band 3 and band 2. These images were used as input to the neural network after band composite processing, and then applying atmospheric correction,using the dark object subtraction (DOS) method, as explained in [67].

Table 2.4. Algae Identification

Tag Number	Date	Value	ACFOR
000006114930	02/02/2015	62.4	Common
000006114930	09/02/2015	753.3	Frequent
000006114930	17/02/2015	792.3	Frequent
000006114930	24/02/2015	563.2	Occasional
000006114930	03/03/2015	112.2	Occasional

2.5 Sentinel-2 Satellite

The Sentinel-2 mission is a land monitoring constellation of two satellites that provide high resolution optical imagery and provide continuity for the current SPOT and Landsat missions.The mission provides a global coverage of the Earth's land surface every 10 days with one satellite and 5 days with 2 satellites, making the data of great use in on-going studies [78].

Of the selected scenes, only the data of those presenting clear skies and cloud-free conditions over the Algae sample sites were considered. The free availability of data from the United States Geological Survey (USGS) (http://earthexplorer.usgs.gov/), as well as the temporal resolution of 5 days of the sensor, provide certain advantages over other sensors. Orthorectified and terrain corrected MSI imagery was obtained. Imagery of the USGS website is provided in GeoTIFF.

The fieldwork data were collected the same days before and/or after the overpass of the sensor for all scenes. The main characteristics of the Sentinel - 2 sensor are summarized in Tab. 2.5.1.

Processing of the Sentinel 2 images was undertaken using the pixel extraction and map layouts using the latest version of SeaDAS (version 7.4, NASA, Washington, DC, USA) available on https://seadas.gsfc.nasa.gov/ and ArcMap (version 10.5.1, ESRI, Redlands, CA, USA), respectively. Image data for optical band 8, band 4, band 3 and band 2 were extracted for further processing.

Then the data were radiometrically calibrated from relative to absolute radiance values using a linear model. The imagery was then geometrically corrected and atmospheric correction was applied as described in [67].

The atmospheric correction of satellite measurements in aquatic ecosystems is very important for the reason that a large part of radiation detected by the sensor is backscatter from the atmosphere. Thus, to properly identify the pixel content in an image in terms of water quality,

the atmospheric correction presents a critical step and very important in data processing of satellite images [79] [80] [58].

The Sentinel 2 satellites have a Multi-Spectral Instrument (MSI) with 13 spectral bands that range from the visible range to the shortwave infrared (SWIR).

2.5.1 Band Composition

Bands come in variable resolutions from 10 to 60 meter and their wavelenght is determined based on specific purposes. The spatial resolution of SENTINEL-2 is dependent on the particular spectral band. The list of bands used in this approach with their central wavelengtns and resolutions are shown below in the tab 2.5.1:

Spacial Resolution (m)	Bands	Central Wavelength (nm)
10	B02 -Blue	490
	B03 - Green	560
	B04 - Red	665
	B08 - NIR	842

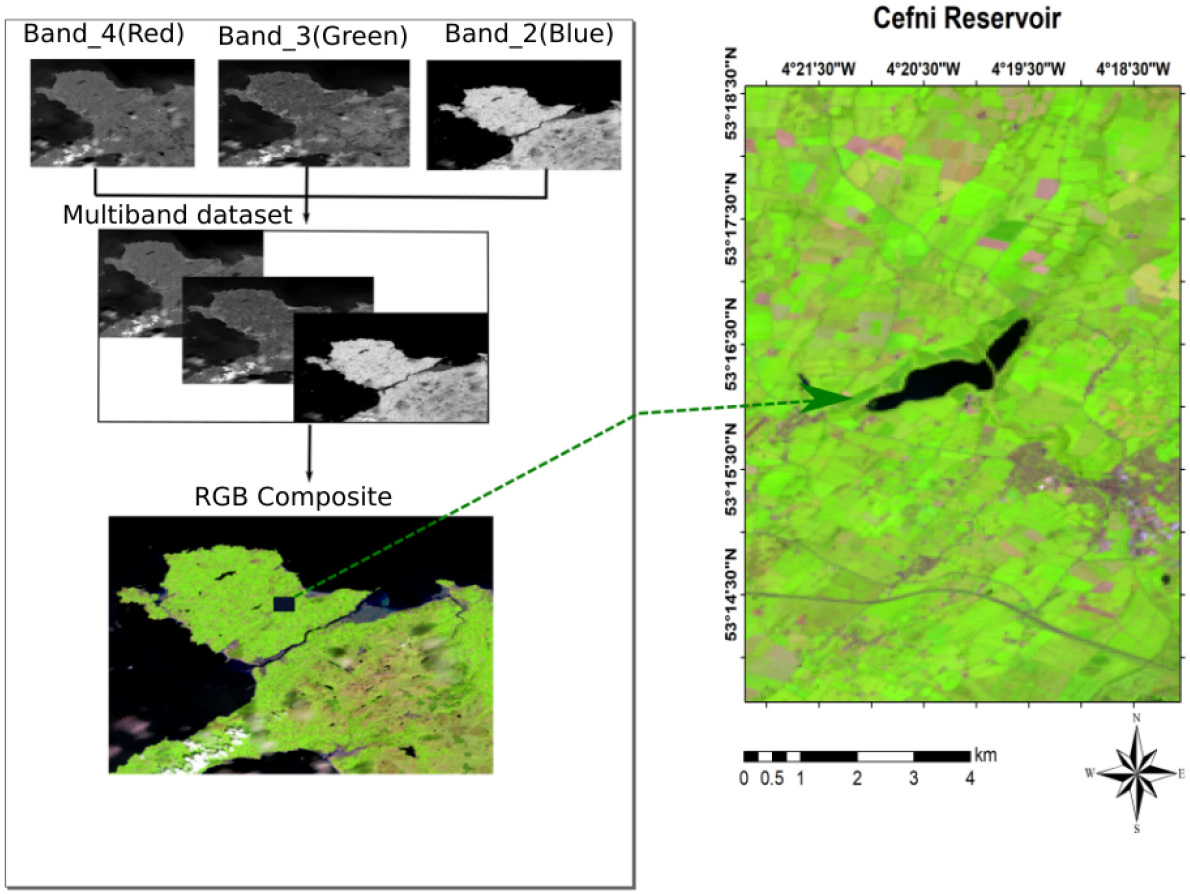


Figure 2.2. Satellite image of Cefni reservoir, Anglesey, U.K.

Chapter 3

Wavelet Artificial Neural Networks (WANN)

Contents

3.1 Introduction to WANN

The goal of this section is to review the concepts of wavelet decomposition, which is employed to decompose an input time series into approximations and details components. The decomposed time series are used as inputs to ANN model. Wavelet transform (WT) of remote sensing images and analysis by an ANN method, described in this section, was used for the implementation of the proposed model.

Wavelets are rapidly finding application as a tool for the analysis of nonstationary signals e is particularly suitable for signal processing. The multiresolution approach of Mallat and Meyer, originally used in image processing [81] [82] [83] [84],.

The Fourier transform was one of the first methods for studying fluctuations in signals, but it has limitations especially when the signal or time series has a non-uniform structure over time. The WT is an advanced modification of the short time Fourier transform in which the window function is completely flexible and can be changed over time based on the shape and compactness of the signal or time series [85].

The Discrete Wavelet Transform (DWT) was implemented in this research, as a preprocessing step for the ANN model and therefore these models were named WANN. For the ANN models the original time series are inputs and for the WANN models the time series are decomposed via the DWT and the newly created time series are used for the ANN model 3.1, 3.2. Based on these expressions, the WANN models are exactly the same as ANN models; therefore, three groups of ANN models and three groups of WANN models were developed, based on the input variables.

3.1.1 Wavelet Decomposition

Wavelets are mathematical functions that give a time-scale representation of the time series and their relationships to analyse time series that contain nonstationarities. Wavelet analysis allows the use of long time intervals for low frequency information and shorter intervals for high frequency

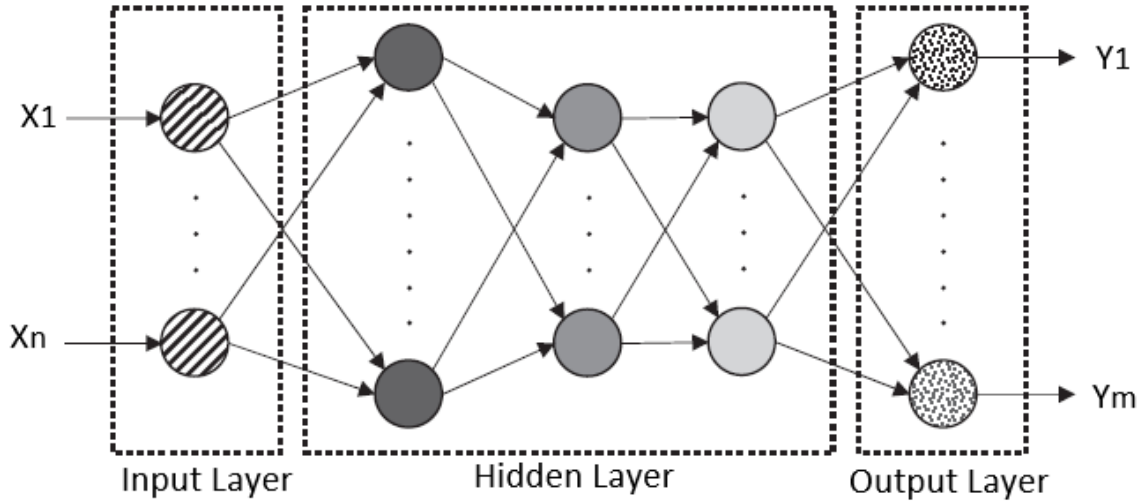


Figure 3.1. WANN

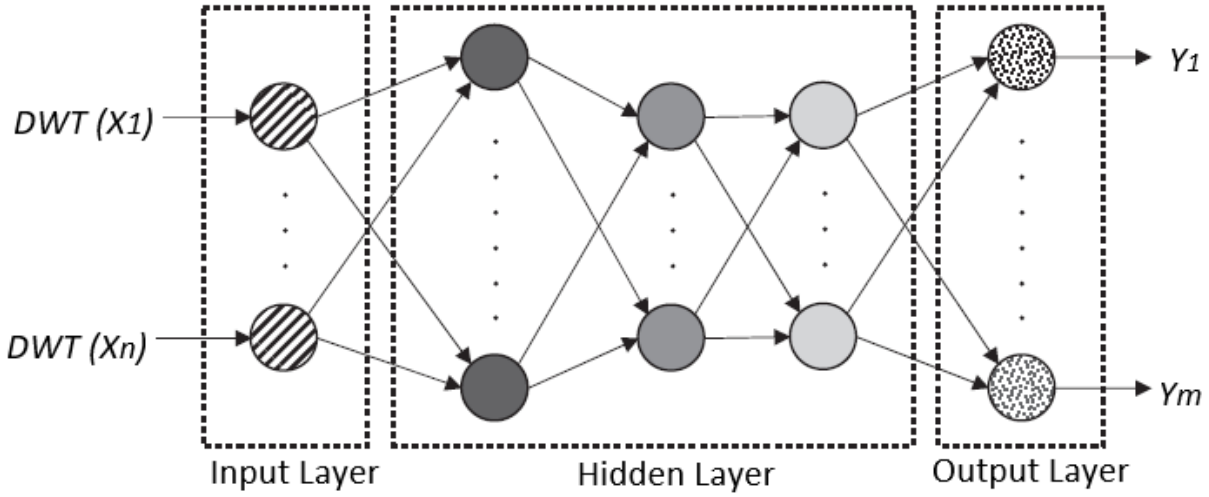


Figure 3.2. Wavelet DWT

information. Wavelet analysis is capable of revealing aspects of data like trends, breakdown points, and discontinuities that other signal analysis techniques might miss. Furthermore, it can often compress or de-noise a signal [86] [87]

The continuous wavelet transform (CWT) at time t for a time series f(t) is defined as follows [88]:

$$W_{a,b}(t) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{a}} \psi^* \left(\frac{t-b}{a} \right) \mathrm{d}x \tag{3.1}$$

where * refers to the complex conjugate of the function. The entire range of the signal is analysed by the wavelet function by using two parameters, namely, 'a' and 'b'. The parameters 'a' and 'b' are known as the dilation (scale) parameter and translation (position) parameter, respectively [89].

A disadvantage of these non-orthogonal wavelets is that the CWT of a given signal is

characterized by redundancy of information among the wavelet coefficients. This redundancy, on account of correlation between coefficients, is intrinsic to the wavelet-kernel and not a characteristic of the analysed signal. As an alternative, for practical applications (as in the study of noise reduction models for communication systems and image and signal compression), DiscreteWavelet Transform (DWT) is usually preferred [90]. The DWT scale and position are based on the power of two (dyadic scales and positions) and can be defined as [88]:

$$W(a,b)_D = 2^{-j/2} \int_{j=1}^{j=J} \psi^*(2^{-j/2} - k) f(t)\, \mathrm{d}x \tag{3.2}$$

where the real numbers j and k are the integers that control the wavelet dilation and translation, respectively. In the DWT, a signal S passes through the low pass filter Lo and the high pass filter Ho followed by down sampling (keeping only one data point out of two) with both filters. At decomposition level n, the high pass filter followed by down sampling produces the detail information D(n) of the signal while the low pass filter associated with the scaling function produces the approximation A(n). This decomposition (filtering and down sampling) process continues until the desired level is achieved [89].

The DWT is generally a non-redundant transform and, accordingly, only a minimally required number of wavelet coefficients are preserved at each level of decomposition which, as a further consequence, enables reconstruction of the original signal from a reduced number of wavelet coefficients. While this property is useful in applications such as data and image compression, this type of non-redundant DWT, however, is prone to shift sensitivity and is therefore an undesirable feature when applied to problems related to singularity detection, forecasting and nonparametric regression [90].

3.2 Performance Measures

The forecast performance of all developed models was evaluated using the following measures of goodness of fit:

HAAR

In this section, we present the discrete wavelets like Haar Wavelet and Daubechies Wavelet for implementation in a still image compression system. This is implemented in software using MATLAB Wavelet Toolbox. The experiments and results is carried out on .jpg format images. These results provide a good reference for application developers to choose a good wavelet compression system for their application.

Haar functions have been used from 1910 when they were introduced by the Hungarian mathematician Alfred Haar [91]. Haar wavelet is discontinuous, and resembles a step function. It represents the same wavelet as Daubechies db1. Haar used these functions to give an example of an orthonormal system for the space of square integrable function on the unit interval [0,1] 3.3.

Haar Transform

$$(b,a)(t) = \frac{1}{\sqrt{a}} f\left(\frac{x-b}{a}\right). \tag{3.3}$$

In Haar's system, a particular family of piecewise linear functions can be used as a basis for representing any given function of finite energy. He used the pattern $\chi_n^{(k)}(s)$ with n and k representing what are now referred to as scale and time. $\chi_0(s)$ is simply 1 on the interval $[0,1]$, and $\chi_1(s)$ is 1 from $[0,\frac{1}{2})$ and -1 from $(\frac{1}{2},1]$. In a more modern phrasing, $\chi_0(s)$ is $\phi(t)$, the scaling function, and $\chi_n^{(k)}(s)$ for $n > 0$ is $\psi_{j,k}(t)$, the wavelet function at each scale and time.

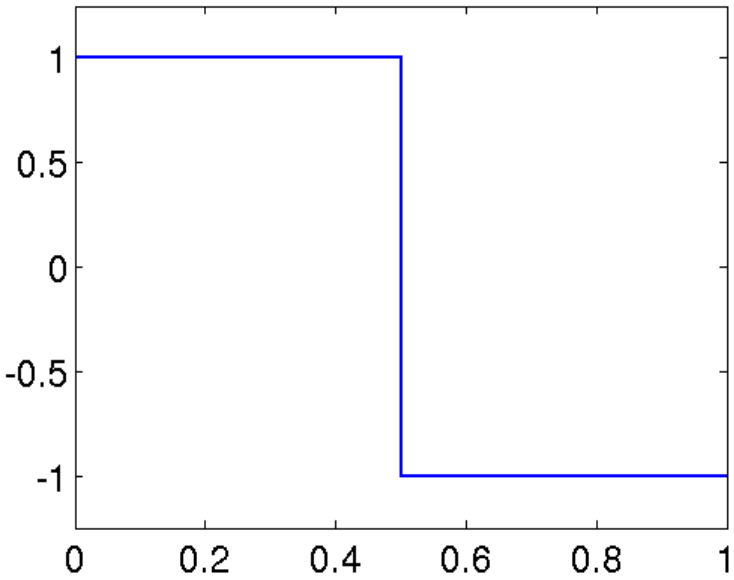


Figure 3.3. HAAR Wavelet

As scale increases by integers, the wavelet functions at that scale become shorter in time by a factor of 2, and amplitude higher by a factor of $\sqrt{2}$ (preserving norm). The k parameter delays these shortened functions in time. Taken together, an expression relates the wavelet functions at higher scale to the lowest-scaled version, or *mother wavelet*:

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \tag{3.4}$$

Haar's insight was that, as this set of functions are orthogonal to one another and have unit norm, we can project a function onto these scaled and shifted χ functions and obtain a complete description of the function in this new domain. Here instead of t (or s), we have j, k (or n, k). A synthesis expression for a time-domain function could be written as

$$f(t) = \sum_{j,k} a_{j,k}\psi_{j,k}(t)$$

And thanks to the orthonormality of the basis functions, we can determine the scaling coefficients $a_{j,k}$ with inner products, or projections, onto those basis functions:

$$a_{j,k} = \langle \psi_{j,k}(t), f(t) \rangle$$

And finally, those Hilbert space projections have a simple integral form:

$$\langle \psi_{j,k}, f(t) \rangle = \int_{-\infty}^{\infty} \psi^*_{j,k}(t) f(t)\, \mathrm{d}x$$

(In this classic orthonormal wavelet basis, the same basic idea applies as in the Fourier example earlier, but with a different change of variables instead of switching between pure time and pure frequency.)

The ANN and WANN models were trained, validated and tested. As an example, the performance of ANN and WANN models for physico-chemical parameters are presented in the results section.

Chapter 4

State-of-the-Art

Contents

The importance of lakes and reservoirs leads to the high need for monitoring lake water quality both at local and global scales. The aim of the study was to test suitability of Sentinel-2 Multispectral Imager's (MSI) and Landsat data for mapping different lake water quality parameters. Data of chlorophyll-a (Chl a), Total Suspended Solids and Transparency from two small and large lakes were compared with band ratio algorithms derived from Sentinel-2 Level-1C and atmospherically corrected.

4.1 Emerging applications areas satellite image requirements

we conduct an in-depth study of the existing work in principle, O conjunto de dados in situ foi limitado em número, mas cobriu um gama de propriedades ópticas da água. The results allow us to assume that Sentinel-2 and Landsat will be a valuable tool for lake monitoring and research, especially taking into account that the data will be available routinely for many years, the imagery will be frequent, and free of charge.

4.1.1 Remote Sensing Data Processing

The free availability of data from the United States Geological Survey (USGS), as well as the temporal resolution of 16 days of the OLI sensor, provide certain advantages over other sensors. Orthorectified and terrain corrected Level1 OLI imagery was obtained. Imagery of the USGS website is processed by the Level 1 Product generation System (LPGS) and is provided in GeoTIFF format with UTM projection and WGS84 datum. The fieldwork data were collected approximately three days before and/or after the overpass of the OLI sensor for all scenes. The main characteristics of the OLI sensor are summarized in Table 4.1 and the main characteristics of the Sentinel-2, MSI (Multispectral Imager) sensor are summarized in Table 4.2

The atmospheric correction of satellite measurements in aquatic ecosystems is very important for the reason that a large part of radiation detected by the sensor is backscatter from the atmosphere. Thus, to properly identify the pixel content in an image in terms of water quality, the atmospheric correction presents a critical step in data processing of satellite images [80].

The atmospheric correction was performed using Dark Object Subtraction (DOS) method using the Semi-Automatic Classification Plugin (SACP). SACP is a free plugin for QGIS (open

Table 4.1. L8/Operational Land Imager (OLI) bands with wavelength and ground sampling distance (GSD).

Band	Wavelength (nm) Range	GSD (m)
1 (Coastal/aerosol)	433–453	30
2 (Blue)	450–515	30
3 (Green)	525–600	30
4 (Red)	630–680	30
5 (NIR)	845–885	30
6 (SWIR 1)	1560–1660	30
7 (SWIR 2)	2100–2300	30
8 (PAN)	500–680	15
9 (CIRRUS)	1360–1390	30

Table 4.2. Sentinel-2 Satellite Sensor Specifications

Band	Wavelength (nm) Range	GSD (m)
Band 1 – Coastal aerosol	0.433	60
Band 2 – Blue	0.490	10
Band 3 – Green	0.560	10
Band 4 – Red	0.665	10
Band 5 – Vegetation Red Edge	0.705	20
Band 6 – Vegetation Red Edge	0.740	20
Band 7 – Vegetation Red Edge	0.783	20
Band 8 – NIR	0.842	10
Band 8A – Narrow NIR	0.865	20
Band 9 – Water vapour	0.945	60
Band 10 – SWIR – Cirrus	1.375	60
Band 11 – SWIR	1.610	20
Band 12 – SWIR	2.190	20

source) thatallows for the semi-automatic supervised classification of remote sensing images, providing tools to expedite the creation of ROIs (training areas) through region growing or multiple ROI creation. The spectral signatures of training areas can be automatically calculated and displayed in a spectral signature plot. It is possible to import spectral signatures from external sources. Also, a tool allows for the selection and download of spectral signatures from the USGS Spectral Library. Several tools are available for the pre processing phase (image clipping, Landsat conversion to reflectance), the classification process (Minimum Distance, Maximum Likelihood, Spectral Angle Mapping algorithms, and classification previews), and the post processing phase (conversion to vector, accuracy assessment, land cover change, classification report) [92] .

4.2 Related Works

Several researches have been done to identify and characterize seasonal variations of physico-chemical parameters in aquatic environments. Among the most used techniques are: ANFIS, multivariate analysis and artificial intelligence techniques.

Recently, the artificial neural networks have been examined for similar prediction applications and showed great potential to tackle and detect its nonlinearity behavior. According to the author [93]Heavy metal toxicity is a matter of considerable concern for environmental researchers. A highly cause of heavy metal toxicity in the aquatic environments is considered a serious issue that required full attention to understand in order to solve it. Heavy metal accumulation is a vital parameter for studying the water quality. Therefore, there is a need to develop an accurate prediction model for heavy metal accumulation. The author developed a model radial basis function neural network algorithm to investigate and mimic the relationship of heavy metals with the climatic and pollution conditions in lake water bodies.

The model was implemented in different climatic conditions as well as polluted and non-polluted lakes. Weekly records of physico-chemical data parameters (e.g., PH, EC, WT, DO, TDS, TSS, CL, NO3, PO4 and SO4) and Climatic parameters (e.g., air temperature, humidity And rainfall) were used as input data for the modeling, whereas the heavy metal concentration was the output of the model. Three different scenarios for modeling the input architecture considering the climate, pollution or both Have been investigated and the results obtained from all the scenarios are positively encouraging with high-performance Accuracy. Furthermore, the results showed that Isolated model for each condition achieves a better prediction Accuracy level rather than developing one general Model for all conditions.

The author [94] determined ecological stream health (ESH) and analyze trophic relations of nutrients (N, P) – chlorophyll and macroinvertebrate – fish, which is associated with stream morphology, land-use patterns, and water chemistry. The neural network modeling of a self-organizing map (SOM) suggested that clustering of trained SOM units reflected stream morphology, land-use patterns, and water chemistry, which influenced community structures and tolerances of top trophic level fish species in the ecosystem. Lotic ecosystem health, based on a multi-metric approach (MF-IBI model), was clearly demonstrated by a multivariate analysis (PCA); important factors were watershed characteristics (land-use patterns), nutrient levels (N, P), organic matter (BOD, COD) regimes, and biological components (trophic and tolerance guilds). In the research [95] twenty-five water quality parameters, including eight heavy metals, were studied at four sampling sites over a stretch of 63 km between Beas and Harike towns for pre-monsoon, post-monsoon and winter seasons. Artificial neural network models were fitted to the data. Correlations between the target values from ANN for turbidity, Biochemical Oxygen Demand (BOD)and bands 2 (green), 3 (red) and 4 (near infra-red) were highly significant.

The degradation of water quality is a major problem worldwide and often leads to serious environmental impacts and concerns about public health [96]. The water quality monitoring and

assessmen of the Lakes has been done by many authors as following:

In the Koumoundourou Lake, a brackish urban shallow lake located in the northeastern part of Elefsis Bay (Greece), were evaluated. A number of water quality parameters (pH, temperature, dissolved oxygen concentration, electrical conductivity, turbidity, nutrients, and chlorophyll-a concentration) were analyzed in water samples collected bimonthly over a 1-year period from five stations throughout the lake. Statistical analysis was performed in order to evaluate the water quality of the lake and distinguish sources of variation measured in the samples. Satellite images of Landsat 5 Thematic Mapper were used in order for algorithms to be developed and calculate the concentration of chlorophyll-a (Chl-a). The trophic status of the lake was characterized as oligotrophic based on phosphorus and as mesotrophic–eutrophic based on Chl-a concentrations. The results of the remote sensing application indicated a relatively high coefficient of determination (R2) among point sampling results and the remotely sensed data, which implies that the selected algorithm is reliable and could be used for the monitoring of Chl-a concentration in the particular water body when no field data are available[96].

In binh dai ben tre, Vietnam, Monitoring surface water quality was also one of the essential missions especially in the context of increasing freshwater demands and loads of wastewater fluxes. The method of Fault Movement Potential (FMP) was used to assess the Surface water resources played a fundamental role in sustainable development of agriculture and aquaculture. Recently, remote sensing technology has been widely applied in monitoring and mapping water quality at a regional scale replacing traditional field-based approaches. This study assessed the application of the Landsat 8 (OLI) images for estimating Chemical Oxygen Demand (COD) as well as detecting spatial changes of the COD concentration in river reaches of the Binh Dai district, Ben Tre province, a downstream area of the delta. The results applied the Artificial Neuron Network (ANN) approach. [97]

A model that predicts the monthly water quality for a reservoir was constructed based on a newly developed programming system, the genetic algorithm operation tree (GAOT) was recently proposed. GAOT, which consists of genetic algorithm (GA) and operation tree (OT), is to find the best function, and to explore complex relationships between inputs and outputs when physical models cannot be defined in advance. In this study, was applied GAOT to estimate the total phosphorous (TP) in Feitsui Reservoir of Taiwan. From GAOT, the three significant input variables was extracted from 15 input variables, including the TP concentration of the Diyu Creek tributary, the TP concentration of main inflow Peishih Creek, the maximum rainfall in the watershed and TP concentration in reservoir, and expressed them appropriately in a sophisticated mathematical manner with accepted complexity. The sensitivity analyses reconfirm the effectiveness of the selected variables in the nonlinear mathematical equations [98].

Another research worked issues inherent to the design of navigation planning and control systems required for adaptive monitoring of pollutants in inland waters. Proposed a new system for estimating water quality, in particular the chlorophyll-a concentration, by using satellite remote sensing data. The aim was to develop an intelligent model based on supervised learning, with the goal of improving the precision of the evaluation of chlorophyll-a concentration. To achieve this, an intelligent system based on statistical learning was used to Classify the waters a priori, before estimating the chlorophyll-a concentration with neural network models. therefore, was developed several models for the same surface of water, based on the spectral signature of the samples acquired in-situ. A control architecture was proposed to guide the trajectory of an aquatic platform to collect in-situ measurements It uses a multi-model classification/regression system to determine and forecast the spatial distribution of chlorophyll-a. Experimental results were presented to validate the approach using data collected on Lake Winnipeg in Canada[99] .

Remote-sensing framework based on the back-propagation neural network (BPNN) also wass developed to quantify concentrations of different surface water quality parameters (SWQPs) from

the Landsat8 satellite imagery. Estimating turbidity, total suspended solids (TSS), chemical oxygen demand (COD), biological oxygen demand (BOD), and dissolved oxygen (DO), Considering the mapping concentrations of different SWQPs critical for providing the appropriate treatment to the affected waterbodies. [100]

In the Albufera de Valencia, Spain, also was develop an integrated algorithm for data fusion and mining of satellite remote sensing images to generate daily estimates of some water quality parameters of interest, such as chlorophyll-a concentrations and water transparency, showed that the spatiotemporal variations of water transparency and chlorophyll-a concentrations may be assessed simultaneously on a daily basis throughout the lake for environmental management using a genetic programming (GP) models[101].

Although these studies demonstrate the application of artificial intelligence and remote sensing to water quality monitoring, none of these studies addresses the application of this technique to water quality monitoring considering the regional hydrological cycle, seasonally, mainly in Amazon reservoirs where Cycle changes four times a year. Thus, this research becomes a pioneer in the application of artificial intelligence and the monitoring of water quality in reservoirs through remote sensing images, training the Wavelet Neural Network by hydrological cycle, that is, a network for each cycle, demonstrating that it is possible to achieve the objective of the research using these techniques.

Chlorophyll-a, TSS and Transparency have evaluated in this research.Chlorophyll-a is an important constitutent reflecting both water quality status and ecosystem state because it is required for phytoplankton existence and can be considered an indicator of algal growth or an indirect indicator of nutrients [102]. Excessive growth of algae blooms in oceans and coastal areas decreases the amount of dissolved oxygen and causes eutrophication in rivers and streams. Because measuring chlorophyll-a is relatively simpler than algae biomass, chlorophyll-a is more often used as a trophic indicator[103].

Chlorophyll-a measurement is costly and time consuming, but remote sensing can provide a spatial view and long term trend of this parameter. Reflectance of chlorophyll-a concentration varies between blue and green sections. In other words, a higher concentration of chlorophyll-a increases reflectance in blue wavelengths and increases the reflectance in green wavelengths[104].

Dissolved organic matter (except phytoplankton) in the water as well as chlorophyll-a concentration can affect this ratio, however; radiation is highly absorbed by chlorophyll-a at about 450 and 670 nm[104][105] concluded that the peak reflectance of different concentrations of chlorophyll-a in a lake is about 700 nm wavelength. Most empirical ocean color algorithms for determining chlorophyll-a concentration are based on the correlations between chlorophyll-a concentration and spectral blue-to-green upward spectral radiance.

Maximum absorption of chlorophyll-a occurs in the blue waveband located in the maximum phytoplankton absorption (440 nm); however, the minimum phytoplankton absorption (550–555 nm) occurs in the green waveband [106]. Uncertainties may occur in determining chlorophyll-a concentration during cyanobacteria blooms[107] or in monitoring surface water with high suspended sediment concentration. In eithor case above, the conventional blue-green ratio is less applicable because the blue light signal decreases with increasing chlorophyll-a concentration. The fluorescence signal will be more efficient in eutrophic waters for monitoring chlorophyll-a concentration, however, because (a) chlorophyll-a has a dominant spectral signature, (b) simple atmospheric correction is not required, and (c) fluorescence increases with intensifying chlorophyll-a concentration [108].

Hyperspectral remote sensing is more reliable for monitoring chlorophyll-a concentration than multispectral remote sensing because it can measuring the reflectance of the extremely narrow wavebands[105], and therefore has high potential to monitor chlorophyll-a concentration in water bodies [107].

Suspended sediments play an important role in transporting nutrients and contaminants because a considerable amount derives from soil and bedrock erosion. The presence of suspended sediments in surface waters has negative effects on aquatic life. In addition, a high concentration of suspended sediments shortens the beneficial and efficient life of lakes and reservoirs [109]

Turbidity and SSSC are related to the suspended sediment fluxes in rivers lakes, and reservoirs, and can help monitoring the sediment discharge, and more generally the sediment budget within catchments, seasonal variability and evolution over time. In turn, the sediment budget is controlling the silting of the dams, which impacts the sustainability of hydroelectric structures and the supply of water for treatment plants. SSSC in inland waters also contributes to pollution and public health issues. Indeed, a signifi- cant correlation exists between the concentration of parasites and bacteria and several water quality parameters including SSSC and turbidity[110].

Suspended particles can carry viruses and bacteria pathogenic to humans [111] and foster their development [112].

High SSSC and turbidity can therefore be considered as a vector of microbiological contaminants which cause diarrheal diseases[109].

Water turbidity and SCCC in lakes or reservoirs may evolve through time, for instance in response to land use changes, modifi- cation of soil erosion, transport and deposition over the watershed, as well as exceptional rainfall events [109].

The quality ofin-situ monitoring networks depends on the number of sampling stations, their spatial representativeness and the frequency of the measurements. In many regions of the world, monitoring networks are decreasing[113], and in some regions, such as West Africa, they are very poor or non-existent.

The Surface Suspended Sediments (SSS) absorb and scatter light, thereby affecting the spectral response of surface waters. Turbidity refers to optical properties of water and has been shown to impact water reflectance in the visible and near-infrared domain. In that context, remote sensing may be a solution in mitigating the data gaps or lack of in-situ network in many areas worldwide[109].

Satellite data and field data were integrated for monitoring tributaries of the Amazon River in Peru, and found that MODIS images could be used to study the SSSC and, combined to river discharge data, to assess the sediment discharge[109].

The objectives of our study are Estimate turbidity, TSS and chlorophyll-a and analysis of its spatio-temporal variability in two reservoirs (Tucurui and Cefni) using LandSat satellite data. So we can do comparisons between the results obtained from two reservoirs and improve the already existing techniques, thus contributing to the state of the art.

Chapter 5

Data processing

Contents

5.1 The Dark Object Subtraction (DOS)

Digital analysis of remotely sensed data has become an important component of a wide variety of earth science studies. However, regardless of the analysis to be done, most remotely sensed data are processed through a set of preprocessing or clean-up routines that remove or correct unwanted artifacts (both geometric and radiometric). A routine that corrects for atmospheric scattering effects is often included in the radiometric preprocessing [114]. The atmospheric correction methods were applied to bands 1–4 of the scene using methods based in the Dark Object Subtraction (DOS) of satellite images, [115].

The DOS method assumes that atmospheric path radiance is the radiance value measured by the satellite for the darkest object within the image, usually clear water bodies or areas in complete shadow. This model corrects only the atmospheric additive scattering component. The Improved DOS method uses a relative atmospheric scattering model to consider the multiplicative efects. Therefore, the atmospheric correction procedure was based on an Improved DOS technique. The spectral band histograms were analysed and selected a start band DOS haze value.

After a relative scattering model had been chosen based on the atmospheric conditions of the image at the acquisition time the initial haze value for the others spectral bands was calculated. The DN value that will be subtracted for each band was calculated after normalization [115], using equation:

$$DN(HAZE_{\mathrm{i}}) = NORM_{\mathrm{i}} * HAZE_{\mathrm{i}} + OFFSET_{\mathrm{i}} \tag{5.1}$$

where:

- DN ($HAZE_{\mathrm{i}}$): is the DN value that predicted $HAZE_{\mathrm{i}}$ value will be mapped in band $_{\mathrm{i}}$,
- $NORM_{\mathrm{i}}$: is the normalized gain value,
- HAZEi: is the predicted haze value for band i using a relative scatterring model and the offset corrected starting haze value, and

- $OFFSET_i$: is the offset value used for band i.

The basic assumption is that within the image some pixels are in complete shadow and their radiances received at the satellite are due to atmospheric scattering (path radiance). This assumption is combined with the fact that very few targets on the Earth's surface are absolute black, so an assumed one-percent minimum reflectance is better than zero percent.

It is worth pointing out that the accuracy of image-based techniques is generally lower than physically-based corrections, but they are very useful when no atmospheric measurements are available as they can improve the estimation of land surface reflectance. The path radiance is given by [116]:

$$L_{\mathrm{p}} = L_{\mathrm{min}} - L_{\mathrm{DO1\%}} \tag{5.2}$$

Where:

- L_{min}: "radiance that corresponds to a digital count value for which the sum of all the pixels with digital counts lower or equal to this value is equal to the 0.01% of all the pixels from the image considered" [116], therefore the radiance obtained with that digital count value DN_{min}.

- $L_{DO1\%}$: radiance of Dark Object, assumed to have a reflectance value of 0.01.

Reflectance results after correction were compared with the in-situ reflectance data. Results have shown that the DOS-based methods have the best performance and therefore the smallest errors, corroborating with other comparative studies on water bodies, which show that these simple methods provide reasonable correction in the bands [117].

Different algorithms to determine [Chl-a], [TSS], and Transparency were obtained from ground-based data and TM reflectance values through simple regression techniques. The entire dataset was split into two groups, one for the algorithmdevelopment (2/3 data) and the second for model testing (1/3 data) as as described in [41]. We focused on bands TM1- TM4, since these bands correspond to the reflectance of the analyzed parameters. The sensors used have a spectral resolution suitable for monitoring water quality variables.

The absolute maximum reflectance of the [Chl-a] falls within the TM2 (528– 609 nm) spectral range, the [TSS] spectral response shows a relative maximum at 810 nm, within the spectral range of band 4 (776–904 nm). For Transparency, the author [118] obtained good results estimation of different combinations of TM bands (1-4).

Regressions showing the best performance based on statistic parameters were selected. Statistical analyses for the comparison between algorithm predictions and field measurements were done as described in [41], including the rootmean- square error (RMSE).

For these image corrections we used the Semi-Automatic Classification Plugin is a free plugin for QGIS (open source) that allows for the semi-automatic supervised classification of remote sensing images, providing tools to expedite the creation of ROIs (training areas) through region growing or multiple ROI creation [92], see Fig. 5.1 . The empirical models established based on in situ measurements were applied for real satellite data.

5.2 Factor Analysis

In factor analysis (FA), we assume that there is a set of unobservable, latent factors Z_j, j = 1,...,k, which when acting in combination generate x. The goal is to characterize the dependency among the observed variables by means of a smaller number of factors [119].

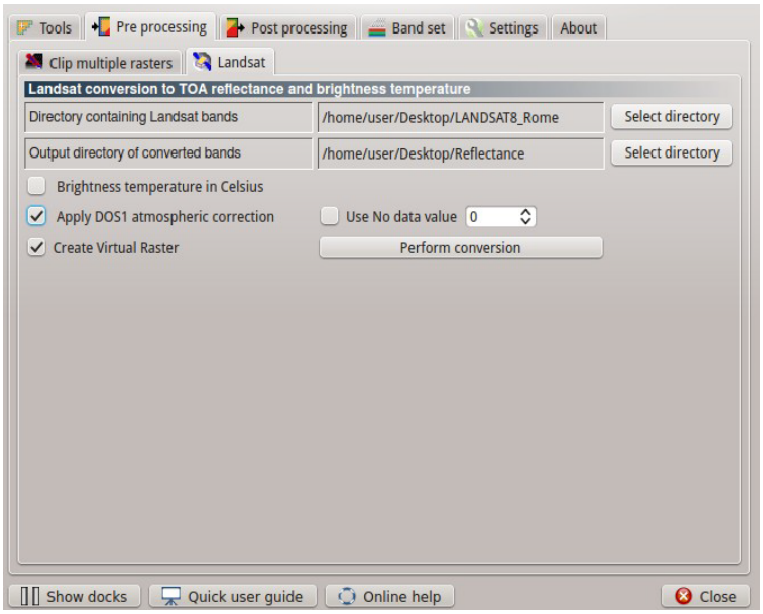


Figure 5.1. Pre processing: Conversion of raster bands from DN to Reflectance

5.2.1 Introduction to Factor Analysis

Descriptive statistics were used to analyze the physico-chemical Parameters of water. Factor analysis on the database for validity was performed using tests contained in SPSS. Exploratory Factor Analysis (EFA) was applied to the all parameters. Tests Used In Factorial Analysis: Bartlett's sphericity test and Kaiser-Meyer-Olkin (KMO). Kaiser-Meyer-Olkin (KMO): This measure is represented by an index (KMO) that assesses the adequacy of the factorial analysis, being calculated by:

$$KMO = \frac{\sum \sum_{j \neq k} r_{jk}^2}{\sum \sum_{j \neq k} r_{jk}^2 + \sum \sum_{j \neq k} q_{jk}^2} \tag{5.3}$$

where the correlation matrix is $R = [r_{ij}]$ and the partial covariance matrix is $Q = [q_{ij}]$. The overall KMO measure of sample adequacy is given by the above formula taken over all combinations and $j \neq k$.

If the partial correlation is near to zero, the PCA can perform efficiently the factorization because the variables are highly related: $KMO \cong 1$.

First, the Keiser-Meyer-Olkin (KMO) test for sampling adequacy and Bartlett's test for sphericity was done. The KMO value was 0.78, in this case, KMO is over then 0.5, indicate that Factor Analyses Method is appropriate for this analyses.

The Bartlett's test checks if the observed correlation matrix R=(rij)(p x p) diverges significantly from the identity matrix (theoretical matrix under H0: the variables are orthogonal). Principal Component Analysis (PCA) is the most widely used unsupervised dimensionality reduction approach. In recent research, several robust PCA algorithms were presented to enhance the robustness of PCA model[120]. The PCA can perform a compression of the available information only if we reject the null hypothesis [121].

In order to measure the overall relation between the variables, we compute the determinant of the correlation matrix |R|. Under H0, |R| = 1; if the variables are highly correlated, we have

|R|= 0. The Bartlett's test statistic indicates to what extent we deviate from the reference situation |R| = 1. It uses the following formula.

$$X^2 = -\left[(n-1) - \left(\frac{2P+5}{6}\right)\right] \ln |R| \tag{5.4}$$

Where:
n: sample size
p: number of variables
|R|: determinant of the correlation matrix

Under H0, it follows a distribution with a [p x (p-1) / 2] degree of freedom.

The Bartlett test presented significant value. Six factors were obtained with total variance explained of 68.68. Fig. 5.3 shows the factors, total and cumulative variances. The method of factor extraction was Principal Components and the orthogonal varimax rotation.

In Fig. 5.5 Component Matrix, the parameters PTotal, FeTotal, Turbidity, Transparency, Fe2, Total Suspended Solids, PO4 and Fe3 accounted for the greatest amount of common variance compared to the rest of components.

This is again reflected in Fig. 5.4 the scree plot for physico-chemical Parameters. It had six values above the eigenvalue of 1. Even though the seven score (0.92), and eight score (0.86) and nine score (0.71) were below eigenvalue of 1 and did not contribute sufficiently to the model, its presence, nevertheless, was indicative that with sufficient power, its score could increase to above eigenvalue of 1. This could result in the formation of more three components.

The Principal Component Analysis(PCA) extraction method component matrix clearly demonstrated that PTotal, FeTotal, Turbidity, Transparency, Fe2, Total Suspended Solids, PO4, Fe2 and Fe3 parameters were related to Factor 1, observed in Fig. 5.6 , this means a strong correlation between these parameters.

The figure 5.2 refers to commonalities, which are quantities of variance, that is, correlations of each variable explained by the factors. High value of commonality means that the variable has great power of explanation. The minimum acceptable value is 0.5 and the maximum is 0.8.

The figure 5.5 in this case showed the relation of the elements PTotal, FeTotal, Turbidity, Transparency, Fe2, Total Suspended Solids, PO4 and Fe3, grouped in Factor 1 (F1), with a strong correlation between these parameters. We can conclude that these elements present great reflectance in bodies of water, being possible the analysis through satellite images.

Table 5.1. KMO and Bartlett's Test

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	0.78	
Bartlett's Test of Sphericity	Approx. Chi-Square	1414.6
	df	190.00
	Sig.	0.00

Figure 5.2Extraction Method: Principal Component Analysis.

Figure 5.3 Extraction Method: Principal Component Analysis.

Figure 5.5Extraction Method: Principal Component Analysis. 6 components extracted.

Figure 5.6 Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 8 iterations.

Figure 5.7 Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Figure 5.8 Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Component Scores.

	Initial	Extraction
Turbidity	1.000	.75
Total Suspended Solids	1.000	.69
PO4	1.000	.61
NO3	1.000	.52
NH4	1.000	.69
Mg	1.000	.62
Na	1.000	.67
K	1.000	.68
Fe2	1.000	.71
Fe3	1.000	.55
Ca	1.000	.59
Chloro	1.000	.74
FeTotal	1.000	.78
Conductivity	1.000	.69
pH	1.000	.75
Dissolved Oxygen	1.000	.62
Transparency	1.000	.79
Temperature	1.000	.74
Chlorophyll	1.000	.71
PTOTAL	1.000	.83

Figure 5.2. Communalities

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.977	29.884	29.884	5.977	29.884	29.884	5.424	27.118	27.118
2	2.469	12.347	42.231	2.469	12.347	42.231	2.008	10.040	37.159
3	1.528	7.638	49.869	1.528	7.638	49.869	1.653	8.263	45.421
4	1.456	7.282	57.151	1.456	7.282	57.151	1.637	8.183	53.605
5	1.295	6.474	63.624	1.295	6.474	63.624	1.623	8.115	61.719
6	1.012	5.061	68.685	1.012	5.061	68.685	1.393	6.966	68.685
7	.924	4.621	73.306						
8	.869	4.347	77.654						
9	.717	3.587	81.241						
10	.709	3.544	84.785						
11	.562	2.812	87.597						
12	.461	2.307	89.904						
13	.442	2.209	92.112						
14	.376	1.878	93.990						
15	.275	1.375	95.365						
16	.261	1.304	96.669						
17	.240	1.198	97.867						
18	.189	.946	98.813						
19	.135	.674	99.487						
20	.103	.513	100.000						

Figure 5.3. Total Variance Explained

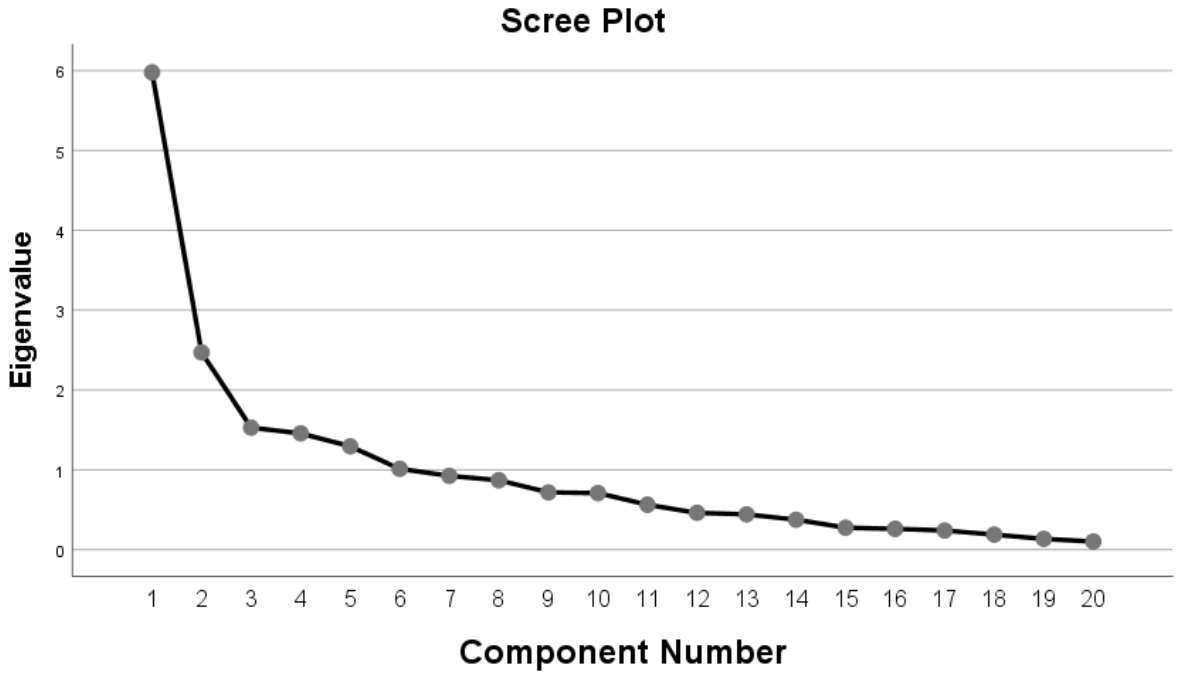


Figure 5.4. Scree Plot

	Component					
	1	2	3	4	5	6
PTOTAL	.893	.086	-.019	-.025	-.130	-.095
FeTotal	.869	.081	-.020	-.038	-.102	-.088
Turbidity	.812	-.279	-.083	.043	-.035	.031
Transparency	-.799	.033	-.038	-.014	-.278	.263
Fe2	.763	-.095	-.202	-.055	-.202	.189
Total Suspended Solids	.760	-.136	.176	.036	-.006	-.242
PO4	.733	.011	-.147	-.155	.035	-.155
Fe3	.711	.119	.001	-.086	-.070	-.116
Temperature	-.533	.374	.291	-.011	-.044	-.483
K	.378	.675	-.112	.127	-.184	.141
Na	.165	.630	-.206	.430	-.053	-.140
Cloro	.055	.521	-.445	.371	.058	.360
pH	.361	-.516	.012	.070	.413	.418
Mg	.165	.235	.693	.133	.157	.114
NH4	.322	.117	.664	.198	-.032	.309
Ca	-.174	-.471	-.061	.571	.081	-.079
Dissolved Oxygen	.090	-.463	-.156	.463	.255	-.303
Chlorophyll	.092	.445	.047	.319	.626	-.071
Conductivity	.258	.284	.045	-.510	.528	.052
NO3	.304	-.143	.364	.337	-.401	.041

Figure 5.5. Component Matrix

	Component					
	1	2	3	4	5	6
PTotal	.882	.145	.114	.089	-.094	-.052
FeTotal	.857	.133	.108	.097	-.094	-.025
Transparency	-.805	.001	-.131	-.056	-.177	-.295
Total Suspended Solids	.780	-.107	.219	.014	.136	-.003
Turbidity	.767	-.032	.045	.362	.124	-.097
PO4	.757	.036	-.080	.097	-.065	.119
Fe3	.715	.099	.086	.021	-.124	.026
Fe2	.699	.124	-.043	.388	-.133	-.193
Chloro	-.088	.826	-.096	.199	-.051	.059
Na	.155	.743	.035	-.303	.059	.029
K	.303	.667	.134	-.091	-.341	-.024
NH4	.151	.038	.799	.125	-.095	-.086
Mg	.047	.006	.761	-.087	-.069	.155
Temperature	-.379	-.031	.093	-.762	-.038	.101
pH	.185	-.202	.153	.737	.268	.181
Dissolved Oxygen	.134	-.078	-.113	.067	.759	.022
Ca	-.183	-.040	-.002	.126	.709	-.203
Conductivity	.217	-.065	.033	.090	-.376	.700
Chlorophyll	.006	.443	.255	-.096	.236	.618
NO3	.261	-.005	.434	.002	.135	-.497

Figure 5.6. Rotated Component Matrix

Component	1	2	3	4	5	6
1	.942	.122	.179	.254	-.032	.030
2	-.031	.700	.158	-.425	-.479	.271
3	-.059	-.392	.886	-.226	-.083	-.011
4	-.093	.538	.299	.004	.734	-.271
5	-.088	-.028	.079	.218	.314	.916
6	-.304	.225	.252	.810	-.353	-.116

Figure 5.7. Component Transformation Matrix

	Component					
	1	2	3	4	5	6
Turbidity	.125	-.017	-.027	.114	.057	-.062
Total Suspended Solids	.186	-.109	.063	-.165	.114	.004
PO4	.175	-.036	-.131	-.067	-.013	.078
NO3	.029	.014	.266	-.051	.065	-.368
NH4	-.079	.012	.518	.137	-.076	-.086
Mg	-.057	-.037	.487	-.018	-.019	.095
Na	.044	.364	-.024	-.191	.141	-.030
K	.018	.310	.039	-.002	-.157	-.093
Fe2	.090	.066	-.077	.195	-.117	-.160
Fe3	.156	-.008	-.021	-.095	-.047	-.003
Ca	-.038	.070	.032	.035	.431	-.092
Chloro	-.117	.479	-.054	.278	-.002	-.008
FeTotal	.173	.015	-.017	-.062	-.032	-.042
Conductivity	.017	-.114	-.008	.085	-.206	.495
pH	-.095	-.032	.128	.506	.087	.176
Dissolved Oxygen	.070	.009	-.082	-.091	.499	.080
Transparency	-.184	.062	.002	.129	-.166	-.224
Temperature	.049	-.090	.052	-.525	.067	.063
Chlorophyll	-.035	.205	.144	-.030	.248	.441
PTotal	.179	.020	-.016	-.072	-.031	-.063

Figure 5.8. Component Score Coefficient Matrix

Chapter 6

Neural network models

Contents

6.1 Inference Models

Work on artificial neural network has been motivated right from its inception by the recognition that the human brain computes in an entirely different way from the conventional digital computer.

The brain is a highly complex, nonlinear and parallel information processing system. It has the capability to organize its structural constituents, known as neurons, so as to perform certain computations many times faster than the fastest digital computer in existence today.

The brain routinely accomplishes perceptual recognition tasks, e.g. recognizing a familiar face embedded in an unfamiliar scene, in approximately 100-200 ms, whereas tasks of much lesser complexity may take days on a conventional computer. [122].

A neural network is a machine that is designed to model the way in which the brain performs a particular task. The network is implemented by using electronic components or is simulated in software on a digital computer. A neural network is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experimental knowledge and making it available for use. [123]

According to [48] An ANN is a parallel-distributed processor that resembles the human brain by acquiring knowledge through a learning process and, then, stores the knowledge in the connection strength between computational units called neurons.

Neural networks and fuzzy logic are very suited to solving a prediction problem dealing with real-world sequences. In this research, neural models of this type are proposed for the study of Chlorophyll [Chl-a], considering a time series of the reservoirs.

They are introduced in the following, considering as a baseline for benchmarking in the successive tests: the linear predictor reported in Equation 9.7, where S(t) can be predicted by considering a function approximation $Y = f(\underline{x})$, $f : \mathbb{R}^N \rightarrow \mathbb{R}$, by using linear models each

input vector $\underline{x}_t$ is made of N subsequent samples of $S(t)$, and the output y_t is the sample to be predicted.

In this sense, the [Chl-a] was estimated through a common least-squares estimator (LSE) [124]; radial basis function (RBF) [125][69]; the well-known ARIMA model, which handles non-stationarity and seasonality and whose parameters are determined by a maximum likelihood approach [126] and ANFIS [127] models. All models were estimated using samples of the water samples as input.

Chlorophyll values, previously collected and analyzed in the laboratory, were used to train the neural models. The application and results of the models are shown as follows:

6.1.1 LSE

The basic Least Square Estimator is used to estimate the Algae values of the decomposed images at each sampling station. The three wavelet output can be viewed as a sequence S of values. The real value of the parameter in analysis (e.i. algal blooms) is attached at the end of the sequence, resulting as the last sample of the sequence itself. This is the sample to be estimated. The function estimator has this form:

$$\begin{aligned} l\underline{x}_t = \left[S(t)\ S(t-1)\cdots S(t-N+1)\right] , \\ y_t = S(t+m) , \\ f_{\text{lin}}(\underline{x}_t) = \sum_{j=1}^{N} \lambda_j x_{tj} , \end{aligned} \tag{6.1}$$

We obtain:

$$\widetilde{S}(t+m) = \sum_{j=1}^{N} \lambda_j S(t-j+1) , \tag{6.2}$$

considering the statistical properties of $S(t)$, as the autocorrelation function, it is possible to determinate the parameters λ_j, $j = 1 \ldots N$, of function f_{lin}.

Table 6.1. LSE Results

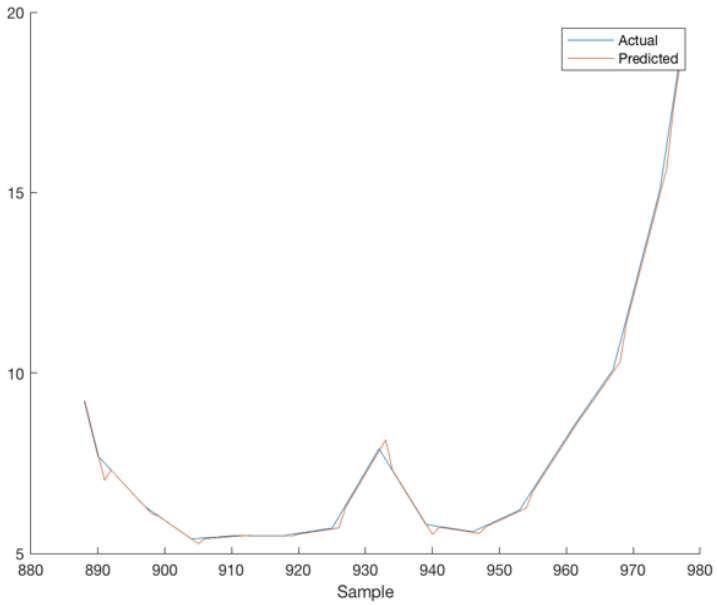
Error on Test set			
Prediction Model		1 year	2 years
lSE	NMSE	0.00251	0.00270
	MSE	0.02005	0.02153
	MARE	0.00569	0.00621

6.1.2 RBF

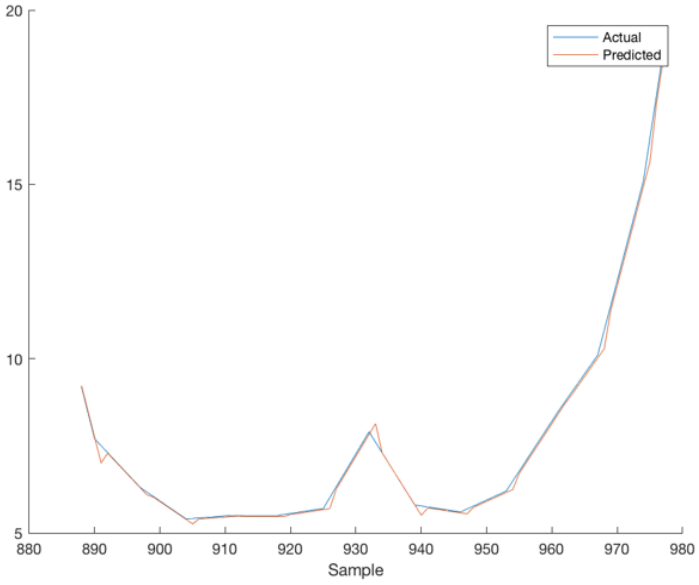
An RBF neural network is used to build up a function approximation model having the following structure:

$$lf(\underline{x}) = \sum_{i=1}^{M} \lambda_i \phi(\|\underline{x} - \underline{c}_i\|),$$

where $\underline{x} \in \mathbb{R}^N$ is the input vector, $\phi(\cdot)$ is a radial basis function centered in $\underline{c}_i$ and weighted by an appropriate coefficient λ_i. The choice of $\phi(\cdot)$ and $\underline{c}_i$ must be considered for the ability of the network in its approximation capability. Commonly used types of radial basis functions include:



(a) Prediction behavior for one-year training



(b) Prediction behavior for two-years training

Figure 6.1. Prediction using LSE

- *Gaussian*

$$l\phi(r) = e^{-(\epsilon r)^2};$$

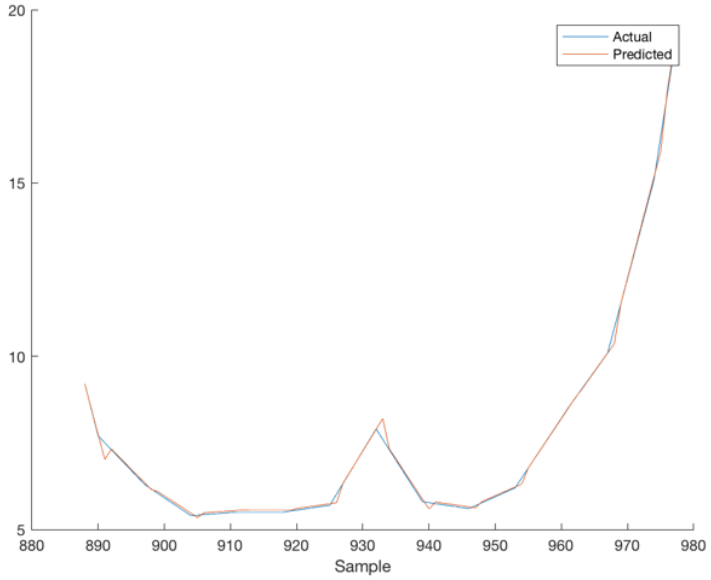
- *Multiquadric*

$$l\phi(r) = \sqrt{1 + (\epsilon r)^2};$$

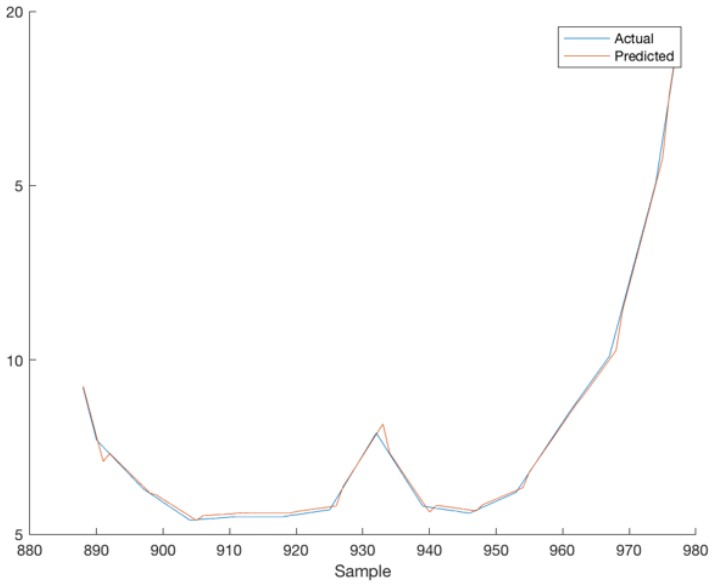
- *Inverse Quadratic*

$$l\phi(r) = \frac{1}{1 + (\epsilon r)^2}.$$

Several methods can be used to minimize the error between desired output and model output and hence, to identify the parameters $\underline{c}_i$ and λ_i [128]. The NN built in this way can be used to estimate a single parameter.



(a) Prediction behavior for one-year training



(b) Prediction behavior for two-years training

Figure 6.2. Prediction using RBF

Table 6.2. RBF Results

Error on Test set			
Prediction Model		1 year	2 years
RBF	NMSE	0.00176	0.00226
	MSE	0.01403	0.01805
	MARE	0.00469	0.00682

6.1.3 ANFIS

This approach is based on neural networks and fuzzy logic, and the inference system is treated as a function approximation problem. An ANFIS neural network implements a fuzzy inference system to approximate the function $y = f(\underline{x})$, $f : \mathbb{R}^N \to \mathbb{R}$. It is composed of M rules of Sugeno first-order type, where the kth rule, $k = 1 \dots M$, is:

$$\begin{aligned} &\text{If } x_1 \text{ is } B_1^{(k)}, \dots, \text{ and } x_N \text{ is } B_N^{(k)} \text{ then} \\ &\qquad y^{(k)} = \sum_{j=1}^{N} a_j^{(k)} x_j + a_0^{(k)} \ , \end{aligned} \tag{6.3}$$

where $\underline{x} = [x_1\ x_2 \cdots x_N]$ is the input to the network and $y^{(k)}$ is the output of the rule. The antecedent part of the rule depends on the membership functions (MFs) $\mu_{B_j^{(k)}}(x_j)$ of the fuzzy input variables $B_j^{(k)}$, $j = 1 \dots N$; the consequent part is determined by the coefficients $a_j^{(k)}$, $j = 0 \dots N$, of the crisp output $y^{(k)}$. By using standard options for composing the input MFs and combining the rule outputs [68, 124], the output of the ANFIS network is represented by:

$$\widetilde{y} = \frac{\sum_{k=1}^{M} \mu_{\underline{B}^{(k)}}(\underline{x})\ y^{(k)}}{\sum_{k=1}^{M} \mu_{\underline{B}^{(k)}}(\underline{x})} \ , \tag{6.4}$$

where $\widetilde{y}$ is the estimation of y and $\mu_{\underline{B}^{(k)}}(\underline{x})$ is the composed MF of the kth rule.

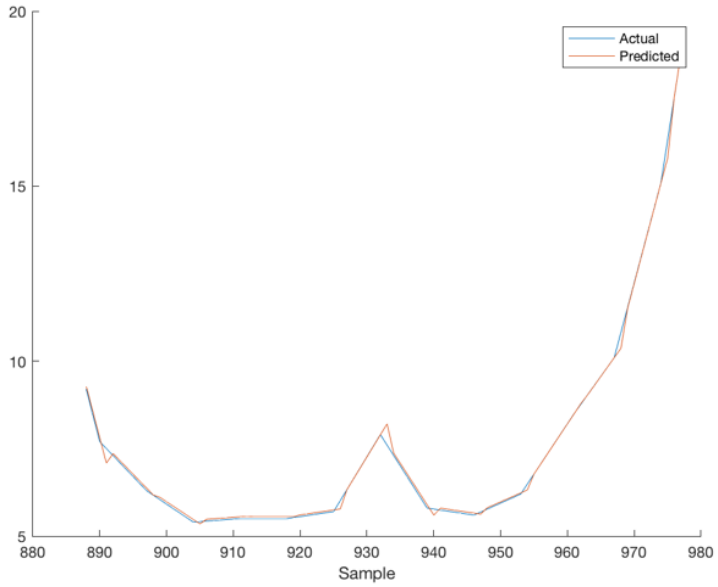
Several clustering algorithms, followed by a suited classification procedure, can be applied to associate each ANFIS rule with the right input pattern. Considering a single sample output we used ANFIS to estimate C, S and T values.

Table 6.3. ANFIS Results

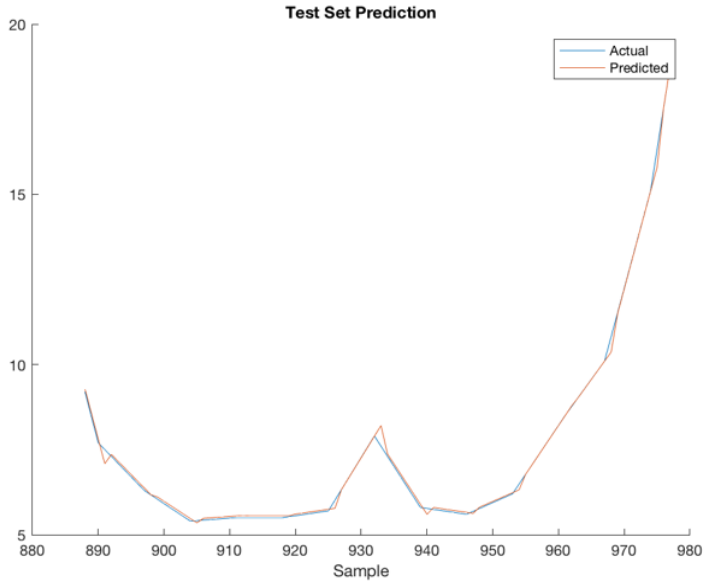
Error on Test set			
Prediction Model		1 year	2 years
ANFIS	NMSE	0.00188	0.00216
	MSE	0.01500	0.01725
	MARE	0.00502	0.00567

6.1.4 ARIMA

Autoregressive integrated moving average (ARIMA) is one of the popular linear models in time series forecasting during the past three decades. One of the most important and widely used time series models is the autoregressive integrated moving average (ARIMA) model.



(a) Prediction behavior for one-year training



(b) Prediction behavior for two-years training

Figure 6.3. Prediction using ANFIS

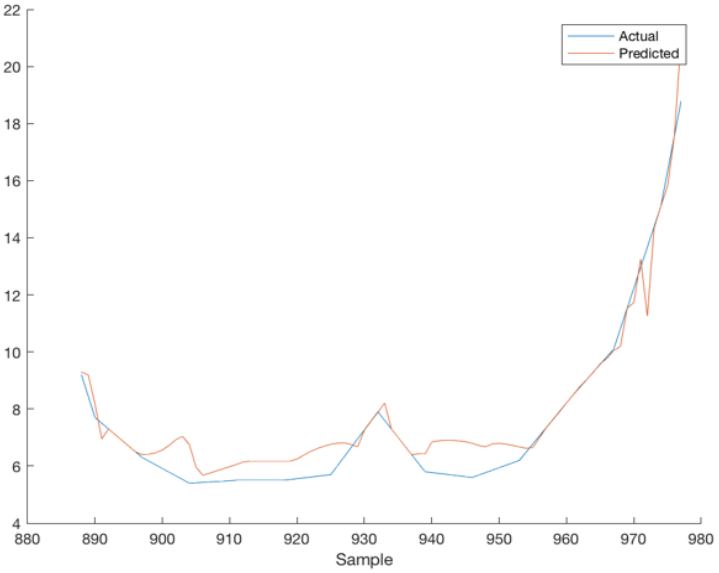
The ARIMA model

In an autoregressive integrated moving average model, the future value of a variable is assumed to be a linear function of several past observations and random errors. That is, the underlying process that generate the time series has the form [129]:

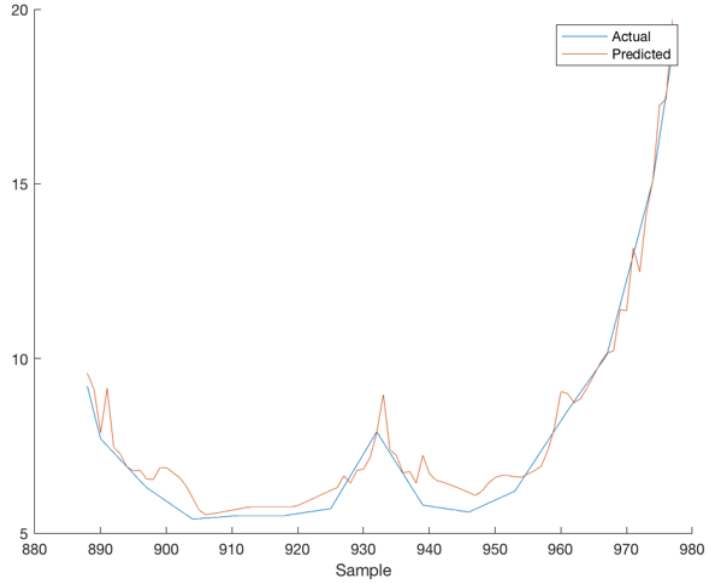
$$\begin{aligned} y_t = \theta_0 + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \ldots + \theta_p y_{t-p} \\ + \epsilon_t \text{ - } \theta_1 \epsilon_{t-1} \text{ - } \theta_2 \epsilon_{t-2} \text{ - } \ldots \text{ - } \theta_q \epsilon_{t-q} \end{aligned} \tag{6.5}$$

where y_t and ϵ_t are the actual value and random error at time period t, respectively; θ_i

$(i = 1, 2, ..., p)$ and θ_j $(j = 0, 1, 2, ..., q)$ are model parameters. p and q are integers and often referred to as orders of the model.



(a) Prediction behavior for one-year training



(b) Prediction behavior for two-years training

Figure 6.4. Prediction using ARIMA

6.1.5 MLP

The classical feedforward multilayer perceptron (MLP) when environmental data sequences are to be predicted this models that can easilyfail, since the mappings to be approximated are often noisy, nonconvex and multi-valued. The principal limitation is due to the conventional least-squares estimation of the model parameters and, more generally, to a lack of an accurate

Table 6.4. ARIMA Results

Error on Test set			
Prediction Model		1 year	2 years
ARIMA	NMSE	0.06413	0.03590
	MSE	0.51185	0.28649
	MARE	0.03839	0.03079

and complete description of data. The latter can be provided in a statistical sense in terms of the probabilitydensity $p(x; y)$ in the joint input–output data space [130].

MLP neural networks learn to model a relationship during a supervised training procedure, when they are repeatedly presented with series of input and associated output data.

Learning of MLP is accomplished by adjusting the weights of the connections between neurons. A typical feed-for- ward neural network is shown in 6.5 as was well explained in [131]. The back-propagation algorithmThis model is represented by the following definition:

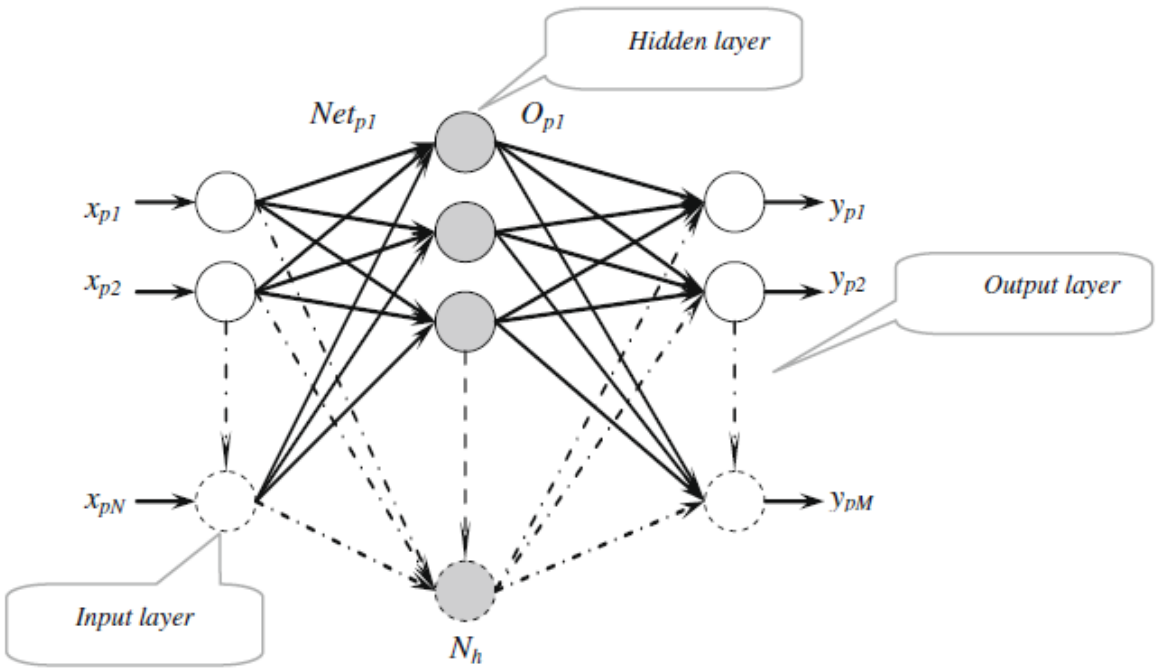


Figure 6.5. Feed-forward neural network

Table 6.5. MLP Results

Results by hydrological cycle					
Prediction Model		Filling	Full	Emptying	Emptying
MLP	MSE Training	$3.36x10^{-22}$	$1.05x10^{-21}$	$3.87x10^{-23}$	$1.34x10^{-21}$
	MSE Test	0.0044	0.0048	0.0183	$2.39x0.0180$
	Relative Error	0.0240	0.0246	0.0767	0.1642

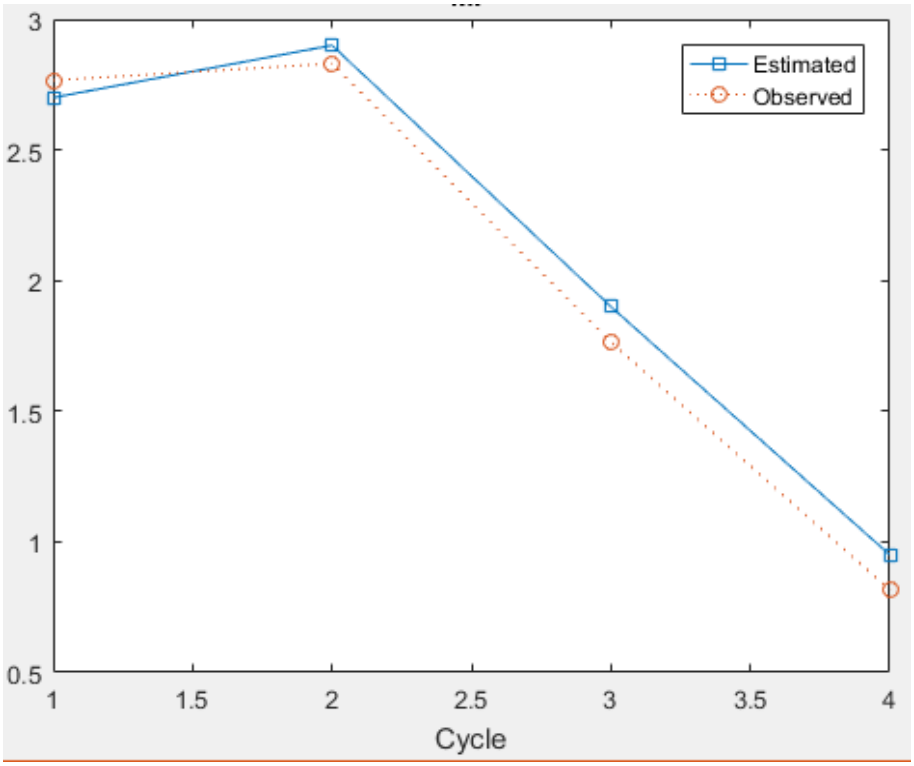


Figure 6.6. Feed-forward neural network

6.2 Conclusions to Inference Models

The performance of the forecast is highly affected by the seasonality of the region, intrinsic to the time of water collection. In any case, the LSE, RBF and ANFIS achieve the best results for almost all training sets with in relation to the other proposed neural models, and all of them surpass both the ARIMA approach benchmark, which is the least feasible solution in this case, especially when we consider the periods with many clouds in the regions.

The results reinforce the fact that the analysis of the physical chemical parameters of water is a promising field for the application of neural and diffuse neural approaches along with remote sensing techniques.

Chapter 7

Case Study: Tucuruí Hydroelectric Reservoir(THR)

Contents

7.1 The Tucurui Reservoir

Studies for the construction of a hydroelectric power plant on the Tocantins River to make use of the area's exploitation potential began in 1957, with the Tucuruí hydroelectric power plant inaugurated in 1984. This plant was built to supply energy for aluminum production, stimulate the regional industry, articulate the links and produce energy to power the country on a national scale [132]. The first stage of deployment occurred between 1975 and 1989, with twelve main units with a total capacity of 3960 MW. Subsequently, two auxiliary units increased capacity to 4000 MW. The second stage opened in late 2008 increased the installed capacity to 7960 MW. The damming led to the formation of a large lake of about 200 km in extension and an area of approximately 2875 km2.

The construction of the Tucuruí hydroelectric power plant caused a large increase in the surrounding population and displacement of the rural population due to flooding of the area, development of mining projects in adjacent regions and agricultural colonization in the vicinity of the Trans-Amazon Highway. The wide availability of fishing resources generated by the reservoir also attracted a large number of people looking for work, employment and income, causing extreme anthropogenic impacts in this area[132].

For development of this project were chosen two reservoirs: the reservoir of the hydroelectric plant of Tucurui[133], [134], viewed as a deep reservoir with a maximum depth of 77 m average depth of 198m and Cefni reservoir, viewed as a shallow reservoir [65].

7.1.1 Location

The Tucuruí hydroelectric power plant is located in the state of Pará, Brazil. The reservoir is located at coordinates: latitude 03^o 45' 03"S, longitude 49^o 40' 03"W. The plant was constructed in Tocantins river, about 7 km from the town of Tucuruí and 300 km from the city of Belem, the state capital. The reservoir has a total flooded area of approximately 2,850 m2, with approximately 50.8 million m3 of water. It is the first large-scale (25 units) hydroelectric project in the Brazilian Amazon rainforest, with an installed capacity of 8370 MW. The main purpose of the dam is hydroelectric power production to the Brazilian states of Maranhão and Pará and navigation between the upper and lower Tocantins river[133].

Tucurui is considered as a deep reservoir, with a maximum depth of 77 m and an average depth of 19-8 m. The power plant reservoir was built in Tocantins river, about 7 km from the city of Tucuruí. The reservoir has a total flooded area of approximately 2850km^2, with approximately 50.8 million m^3 of water.

In the Brazilian amazon region there are 5 reservoirs in operation: Couracy Nunes, Curua Una, Tucurui, Balbina, and Samuel. The UHE Tucuruí plant is a large-scale hydroelectric power plant that is located in the state of Para on the Tocantins River 7.1.

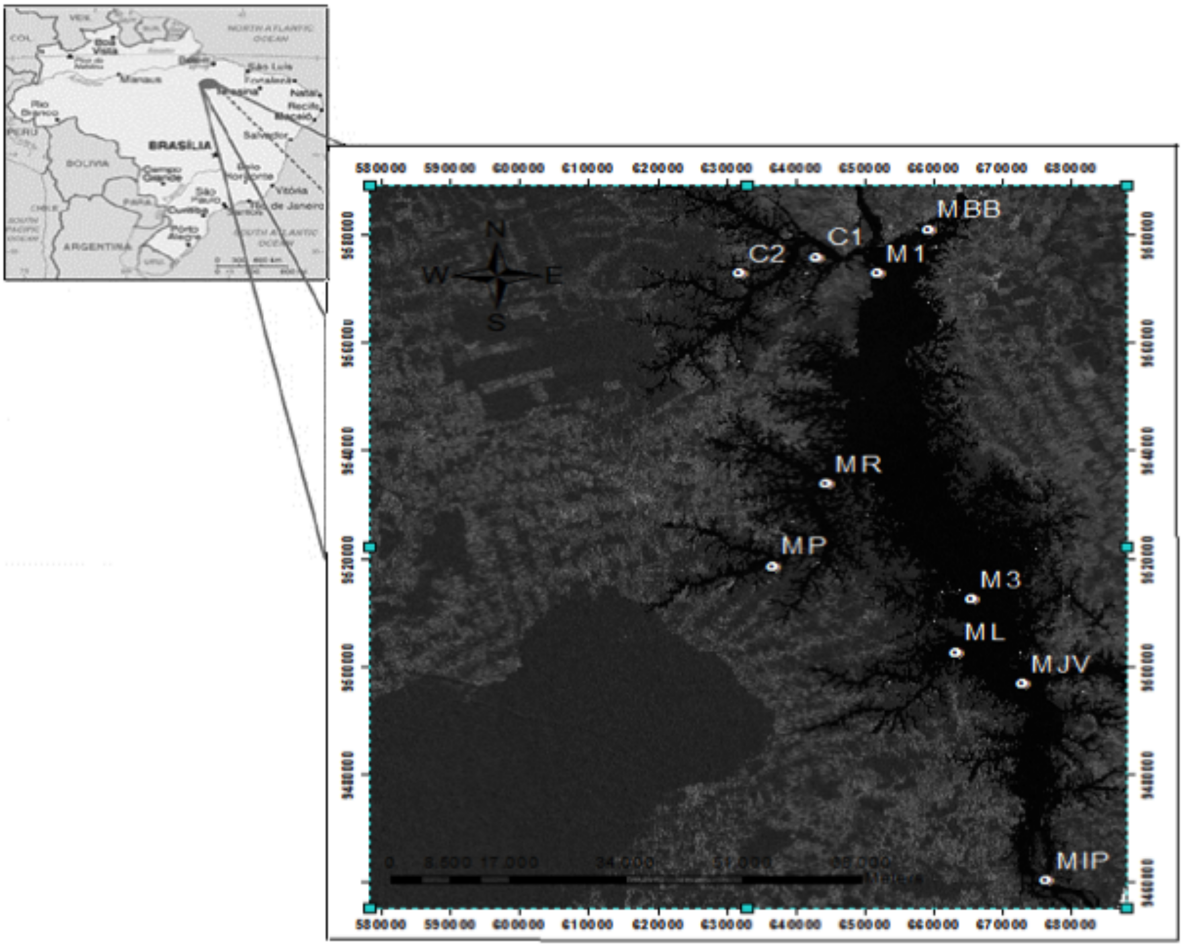


Figure 7.1. Sampling Stations

Table 7.1. Geographical location of sampling points located upstream and downstream of the Tucuruí Hydroelectric Power Plant dam, Brazil- Tucuruí

Sampling Stations	Name	Latitude	Longitude
C1	Caraipé 1	04°32.91'9"S	49°26.44'2"W
C2	Caraipé 2	04°29.46'2"S	49°31.54'2"W
M1	Montante 1	03°45.84'5"S	49°39.54'4"W
M3	Montante 3	04°25'21.7" S	49°30'29.4"W
MR	Montantante Novo Repartimento	04°13.05'7"S	49°41.96'3"W
MBB	Montante Breu Branco	03°49.75'0"S	49°38.89'5"W
ML	Montante Lontra	04°29'20.2"S	49°31'17.5"W
MBL	Montante Belauto	04°14'04.3"S	49°27'55.0"W
MP	Montante Pucuruí	04°21'22.8"S	49°46'05.7"W
MJV	Montante Jacundá Velho	04°32'58.5"S	49°26'24.5"W

7.1.2 Physico-chemical Parametres analyzed in Tucurui Reservoir

One of the main impacts in the Brazilian Amazon in the last decades has been the construction of hydroelectric power plants, with their accompanying dams and reservoirs, resulting in dramatic alterations to these ecosystems, such as loss of diversity and large-scale deforestation. Monitoring is fundamental for reservoir management and the evaluation of anthropogenic environmental impacts. The water quality of reservoirs in the Brazilian Amazon is greatly influenced by hydrological cycles, that in turn cause variations of microbiological, physico-chemical characteristics. There are, however, scarce reports in areas that suffer well-defined hydrological cycles, such as the Brazilian Amazon. In this context, this study presents an alternative method for predicting PTotal, FeTotal, Turbidity, Transparency, Fe2, Total Suspended Solids, PO4, Fe3, Temperature and Chorophylla in the Tucuruí Hydroelectric Power Plant reservoir, in the Brazilian Amazon, by applying Wavelet transformation of data obtained from remote sensing images, taking into account the hydrological cycles of the area, from 2007 to 2014, which were then analyzed by Artificial Neural Networks and compared to laboratory results.

7.1.3 Identification and Location of Collection Stations - Tucuruí

For this study, the following sampling stations were chosen: C1, C2, M1, M3, MR, MBB, ML, MBL, MP, MJV 7.2.

The water collections occurred at the ten points indicated above[135]. these water collection points have established important differences in morphometry of the system, which directly influenced the circulation of the flow of water in the region. The table 7.3 shows some of the relevant descriptions of each point, which will make it possible to understand of the results found:

7.1.4 Sampling Stations: Tucurui reservoir

Description of the collection stations in Tucurui Reservoir:

Sampling Station: Montante 1(M1)

Located 2 km upstream from the dam on the original Tocantins river channel. This is important in the monitoring, because it represents the water to be captured by the generating units and also the water to be sent downstream. When the reservoir is at the maximum operating level (74m in relation to the sea) depths up to 70m can be verified.

Sampling Stations: Caraipé 1 (C1) e Caraipé 2 (C2)

Table 7.2. Water collection stations along the reservoir of the Tucurui Hydroelectric Power Plant - BRAZIL - PA

MSE Validation by Cycle					
Parameters	Sampling Stations	Full	Emptying	Dry	Filling
CHLOROPHYLLa	C1	1.1593	17.1529	0.2679	5.4940
	C2	0.1555	4.3905	0.4366	0.0889
	MBB	0.3346	0.0070	0.8778	0.1564
TRANSPARENCY	MJV	0.1789	0.0736	0.0156	0.0828
	M1	0.0010	0.4436	0.4957	0.0272
	M3	0.0966	0.2470	0.2272	0.2318
TSS	M3	0.0444	1.1343	0.0006	1.6069
	MJV	0.0106	1.1881	0.0471	0.2483
	MIP	1.1363	0.0135	11.3284	0.7024

Table 7.3. Characteristics of the sampling points of the Hydroelectric Tucuruí - Pará - Brazil

Sampling Stations	Name	Characteristics
C1	Caraipé 1	Islands
C2	Caraipé 2	Islands
M1	Montante 1	Next to dam
M3	Montante 3	Lake - Central channel
MR	Montantante Novo Repartimento	Lake - left bank of the reservoir
MBB	Montante Breu Branco	Next to dam
ML	Montante Lontra	Lake - Central channel
MBL	Montante Belauto	Lake - right bank of the reservoir
MP	Montante Pucuruí	Lake - left bank of the reservoir
MJV	Montante Jacundá Velho	Lake

Located on the left bank of the reservoir in the region currently called "Caraipé Region" where since 2002 the Sustainable Development Reserve (RDS) . In this region there are many inhabitants, around 5.000 inhabitants, with a high level of anthropization characterized by the high level of deforestation in this region. The two stations are located in the old channel of the Caraipé region presenting depths of up to 28m and 22m (when the reservoir is full).

Sampling Stations: Montante Breu Branco (MBB)

Located on the right bank of the reservoir, in front of the city of Breu Branco, region where much of the surface drainage of the city is launched. The sampling station has a maximum depth of 32 m. This region is located near the urban nucleus of the municipality of Breu Branco, with few areas of primary forest.

Sampling Stations: Montante Belauto (MBL)

Located on the right bank of the reservoir, full protection area (ie, no residents allowed). In this region, there are large areas with primary forest, however this part of the reservoir is characterized by having a retention time superior to the average of the reservoir, contributing to the fact that the water in this region presents different physico-chemical characteristics. It has a maximum depth of 26 meters.

Sampling Stations: Montante Repartimento (MR) e Montante Pucuruí (MP)

Located on the left bank of the reservoir, in the region of the Sustainable Development Reserve (RDS). This region is the most dendritic of the reservoir, that is, it occupied large area,

but presents low depth. However, the two stations have maximum depths of 20m (MP) and 32m (MR) due to their location in the old pipeline of the Pucuruí and Pucuruizinho streams, respectively. In this region, there is a place of fish landing "in natura" called "Polo Pesqueiro" to follow the municipality of Novo Repartimento.

Sampling Stations: Montante Jacundá Velho (MJV)

Located on the right bank, where there was the old urban nucleus of the city of Jacundá. In this region, there are two landing sites for fresh fish, one of which is called "Porto Novo" and "Porto da Colônia". This region is quite anthropized with small areas of native vegetation. The sampling station is located in the old Jacundá river channel with a depth of up to 22 m (when the reservoir is full).

Sampling Stations: Montante Lontra (ML)

Located on the right bank, near the Indigenous Land of the Parakanã Indians (today the largest extension continues with native vegetation in the Lake of Tucuruí region), located in the old Bacuri channel, presenting a maximum depth of 22 m (when the reservoir is full).

Sampling Stations: Montante 3 (M3)

Located in the central part of the reservoir, distant approximately 60 km straight from the Tucurui dam, located in the old Tocantins river, it has a maximum depth of up to 52 meters (when the reservoir is full).

Sampling Stations: Montante Ipixuna (MIP)

Located in the central part of the reservoir, approximately 130 km straight from the Tucurui dam, located in the old trough of the river Tocantins. When in the filling period of the reservoir, it presents depths of up to 42 meters. This presents the characteristic of behaving like reservoir in the period of full and like river in the dry period.

7.1.5 Pre-processing satellite images: Tucurui Reservoir

The satellite images were obtained from the ESPA (https://espa.cr.usgs.gov), captured by the LandSat 7 of the reservoir from 2007 to 2014. Four images were collected per year, corresponding to each hydrological cycle for all sampling stations.

Initially a point is chosen for analysis, from this it is cut out an image A in its surroundings corresponding to 32x32 pixels. Obtaining an array containing 1024 pixels of information. For the next step, the wavelet transform with 1 level of decomposition is applied, resulting in a reduced image of 16x16 pixels, with 3 images (H, V, D) of 16x16 pixels representing the Horizontal, Vertical of reduced image.

For the input of the neural network, the matrices H, V and D were converted to their respective column-arrays and then merged with H, V and D images, generating a T column vector of size 768. This procedure was performed for the images of 2007,2008,2009, 2010, 2011, 2012, 2013, 2014, generating an input M matrix.

During the validation of the Neural Network, the images of 2014 were used, and it is necessary to perform the same image processing performed on the previous images. The process described from image A to generation of the M matrix and subsequent submission to the analysis and training in the neural network was performed by hydrological cycle for each point of water collection.

7.1.6 Wavelet Transform and ANN applied to the remote sensing images

The wavelet transform is an integral transform whose kernel is a class of special functions, called wavelets [49]. The main advantage of this method compared to other methods is its spectral location capability in space and frequency, which allows for the analysis of non-stationary signals in their various scales [50]. The wavelet transform used in the present study was the discrete

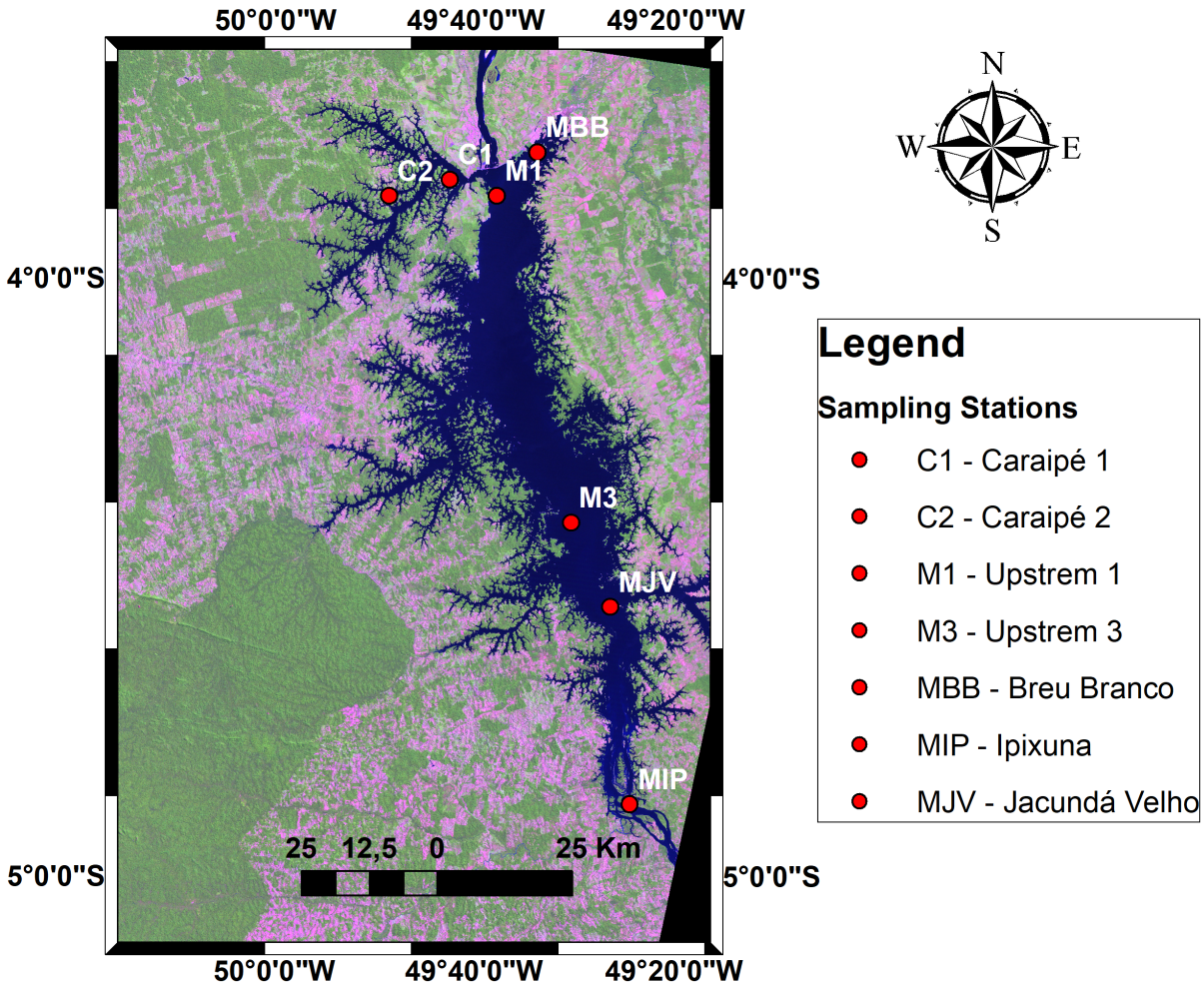


Figure 7.2. Sampling Stations Tucurui Reservoir - Amazon Region

transformed, with allows for the multi-resolution analysis of a signal, decomposing said signal into approximations and details. The approximations are high ranges, i.e., low-frequency signal components. The details are the low ranges, i.e., high frequency components[51]. The Haar family with a degree of decomposition in the Matlab software package was used.

One sampling station was initially chosen for analysis and a geographics image of the water sampling station, of 32x32 pixels, was cropped, corresponding to an array containing 1024 pixels. Each digital pixel value corresponds to an average of radiance values, emittance or backscatter of the different targets that can be contained in the pixel from the vicinity of the water sampling stations, as displayed in an example in Fig. 7.6

Subsequently, the wavelet transform was applied, with only one level of decomposition, resulting in a matrix array of 16x16 pixels for each of the following three components: Horizontal (H), vertical (V) and diagonal (D).

The conversion of the arrays to the H, V and D components to their respective column-matrices was performed, and subsequently a concatenation of the three arrays (each containing 256 pixels) was executed, generating a vector with column size (256 x 3).

This data, the image of the geographical area containing the water sampling collection point,

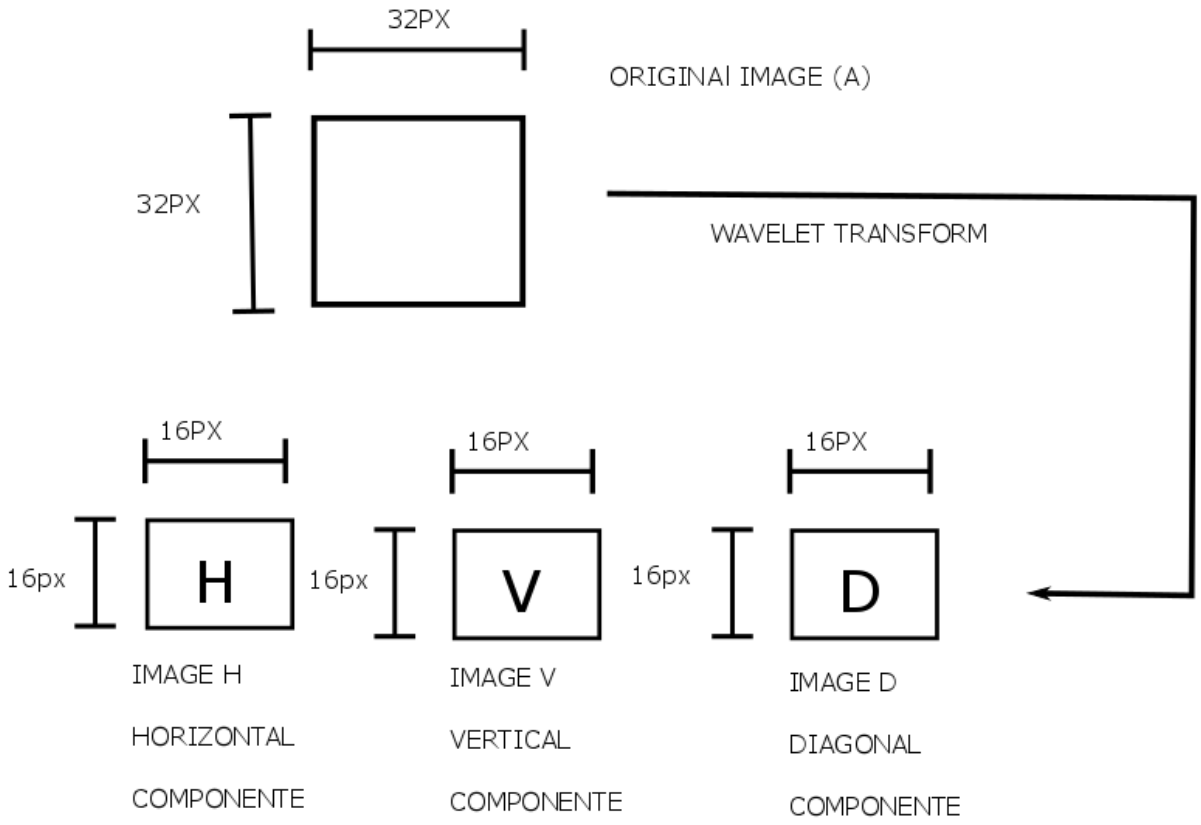


Figure 7.3. Pre processing: Satellite Image

decomposed via wavelet into its three wavelet components, was used as the ANN input. Tests were conducted considering the image representations isolated for each wavelet component, with satisfactory results.

However, when the input data of the three wavelet components was considered, the approximations were even better, which motivated the choice of this arrangement in the proposed solution.

The digital values of the pixels of the images cut in the vicinity of the collection stations of water samples were used as input for the ANN. The digital pixel value is an average of radiance values, emittance or backscatter of the different targets that can be contained in the pixel. Thus, the possible differences between the digital values of the images of the different hydrological cycles used in the study were related to the output data of chlorophyll-a levels, water transparency and total suspended solids, forming the input/output pairs for the ANN training. The figure 7.7 displays a pixel matrix and its corresponding digital values.

The data obtained in the laboratory (estimated) refer to the ANN execution, which are then compared to the data really observed in 2014. This validates the ANN output data.

The images from the satellites were obtained during the same timeframe as the water samplings. For example, if a water sampling was conducted in March 2008, a satellite image was retrieved in March 2008.

After processing, satellite images were used as inputs to the neural network.

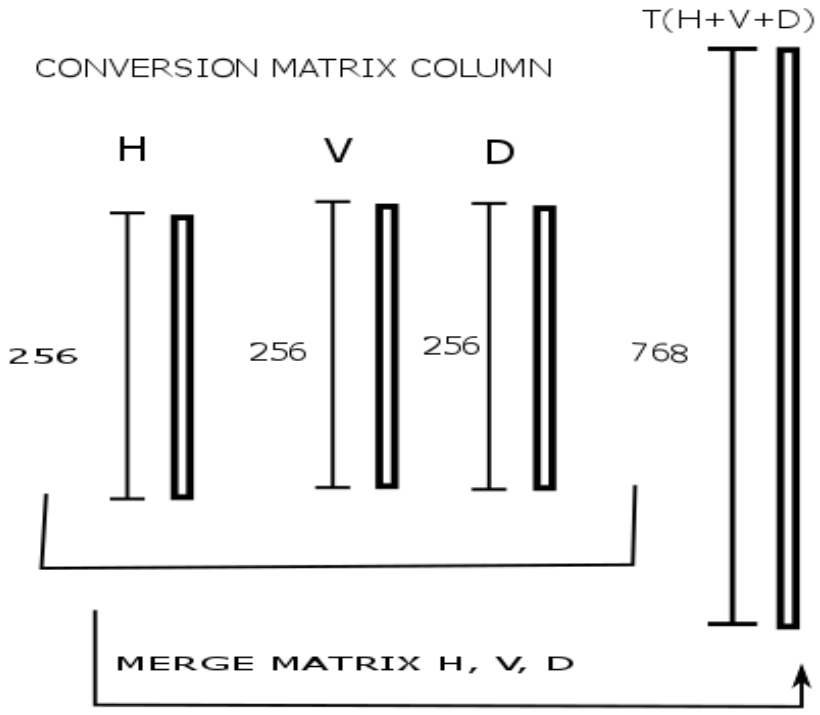


Figure 7.4. Pre Processing Satellite Image: Conversion Matrix Column

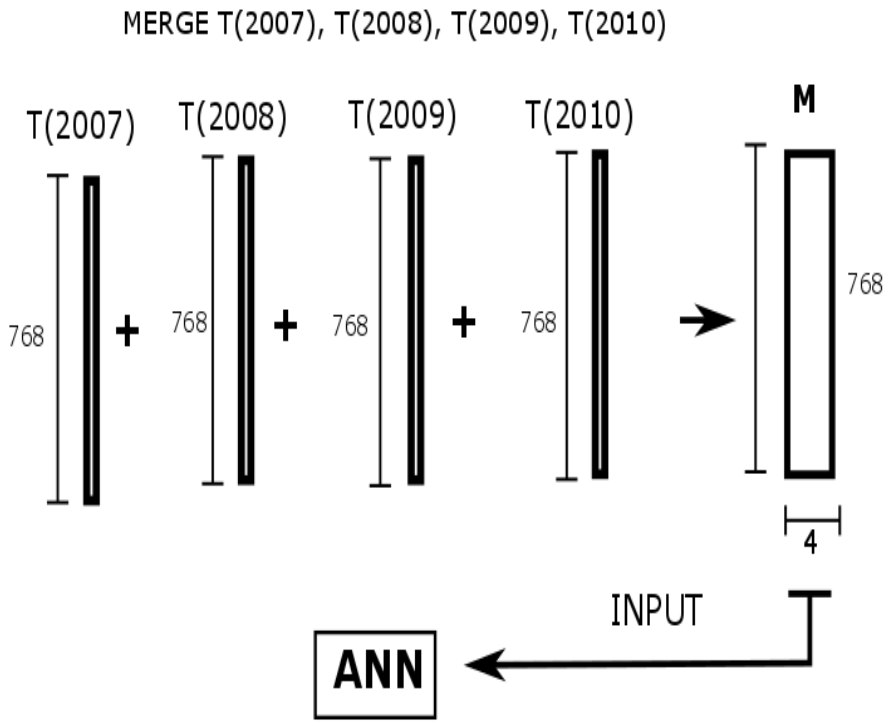
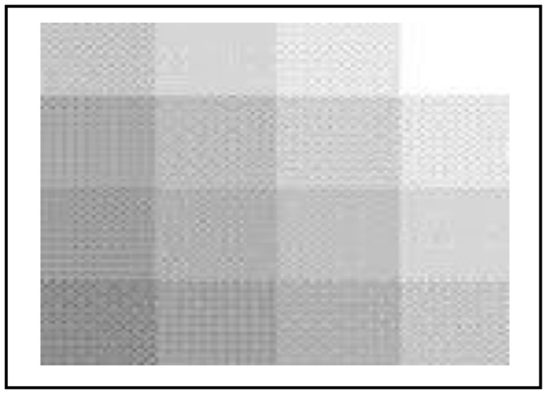
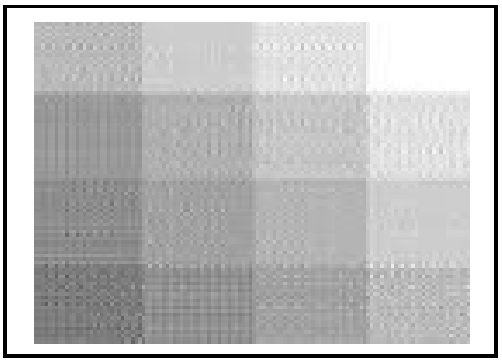


Figure 7.5. Pre processing3: Satellite Image



124	135	141	255
110	120	128	141
109	115	120	135
98	110	119	128

Figure 7.6. Example of a pixel matrix and it corresponding digital values



124	135	141	255
110	120	128	141
109	115	120	135
98	110	119	128

Figure 7.7. Digital values of the pixels of the images cut in the vicinity

7.1.7 Neural Networks Model proposed for Tucurui Reservoir

Neural networks are parallel distributed systems consisting of two basic types of components: the processing units, arranged in one or more layers, interconnected, called neurons; and the synapses, which are the connections between the processing units. In the present study, the ANN paradigm applied was the Direct Multilayer Perceptron, developed using the Matlab software package (Mathworks, 2016b). Post-processed images of the ten sampling sites per water cycle were used for the ANN input and the output variables were the variables chlorophyll-a, total

suspended solids and transparency. The architecture of the ANN consisted of three layers: the input layer, the hidden layer and the output layer. The validation process of the ANN was conducted with 2014 images processed according to the method described in section 2.4. Figure 7.8 displays the architecture ANN with the column vector Pi (i = 1,2, .., 728) as input. The ANN was trained with the following parameters:

- Learning rate: 0.01.
- Transference function: tansig in all the neurons of the hidden layer and purelin in the output layer.
- Network training function: Gradient descent with momentum and adaptive learning rate backpropagation.

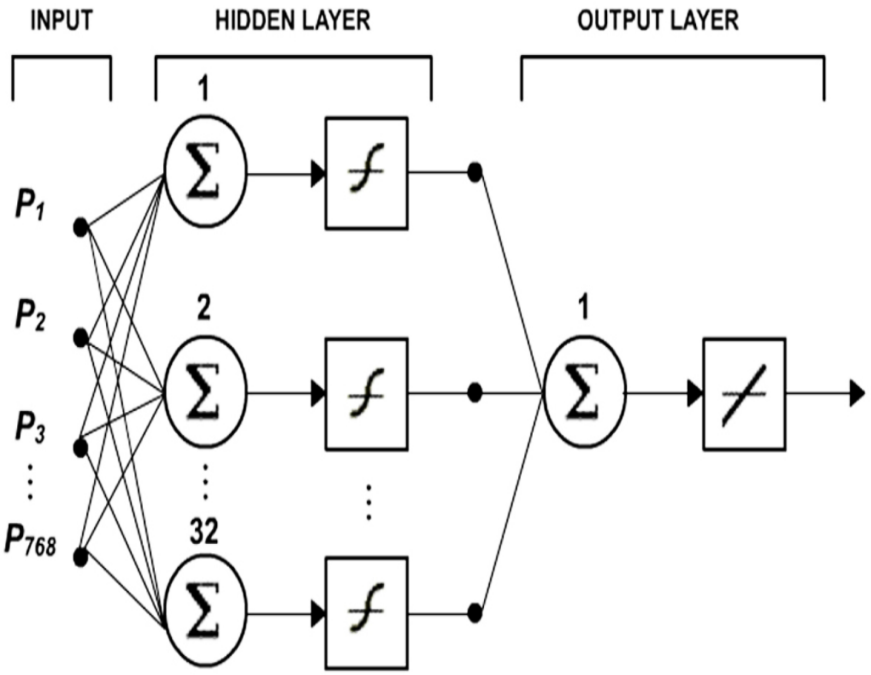


Figure 7.8. Schematic of the ANN architecture used in the present study

The digital values of the pixels of the images cut in the vicinity of the collection stations of water samples were used as input data for the ANN. Thus, the possible differences between the digital values of the images of the different hydrological cycles used in the study were related to the output data of chlorophyll-a levels, water transparency and total suspended solids, forming the input/output pairs for the ANN training.

The Figure 7.10 shows the procedure performed with the 2007, 2008, 2009, 2010, 2011, 2012, 2013 images, generating an ANN input matrix with 768x7 dimensions for each measurement of the hydrological cycle (full, emptying, dry and filling) of the investigated study years. Water samples were collected from the 12 sampling points, but for 2 points located south of the reservoir (downstream) there were missing images within the four analyzed years, which would undermine the analysis in the study. So only the following points sampling points were analyzed:C1, C2, M1, MBB, MR, MP, M3, ML, MJV , MIP.

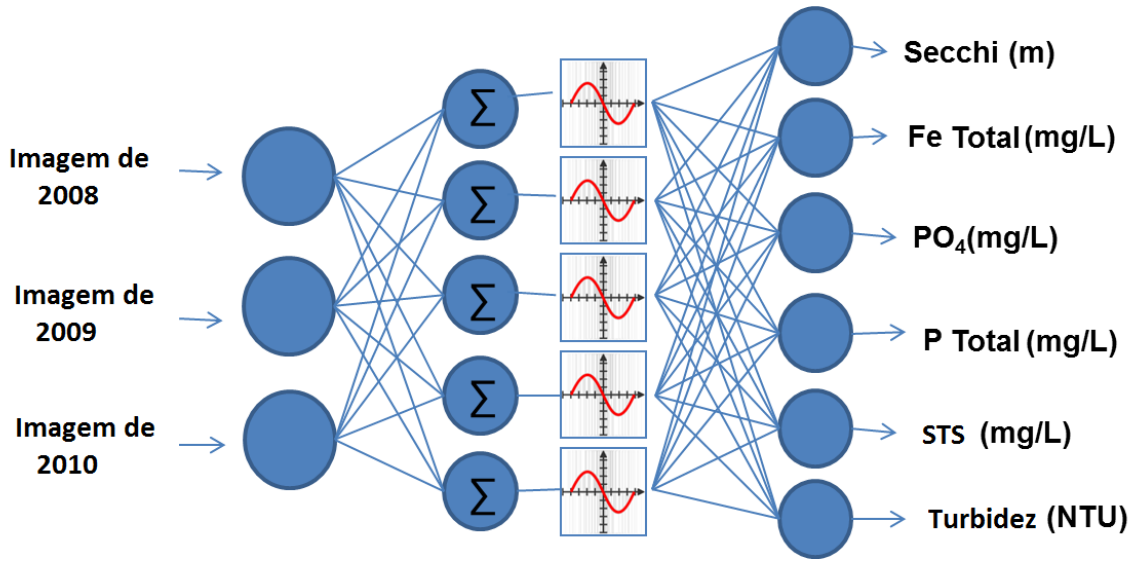


Figure 7.9. Schematic of the ANN architecture proposed in this research.

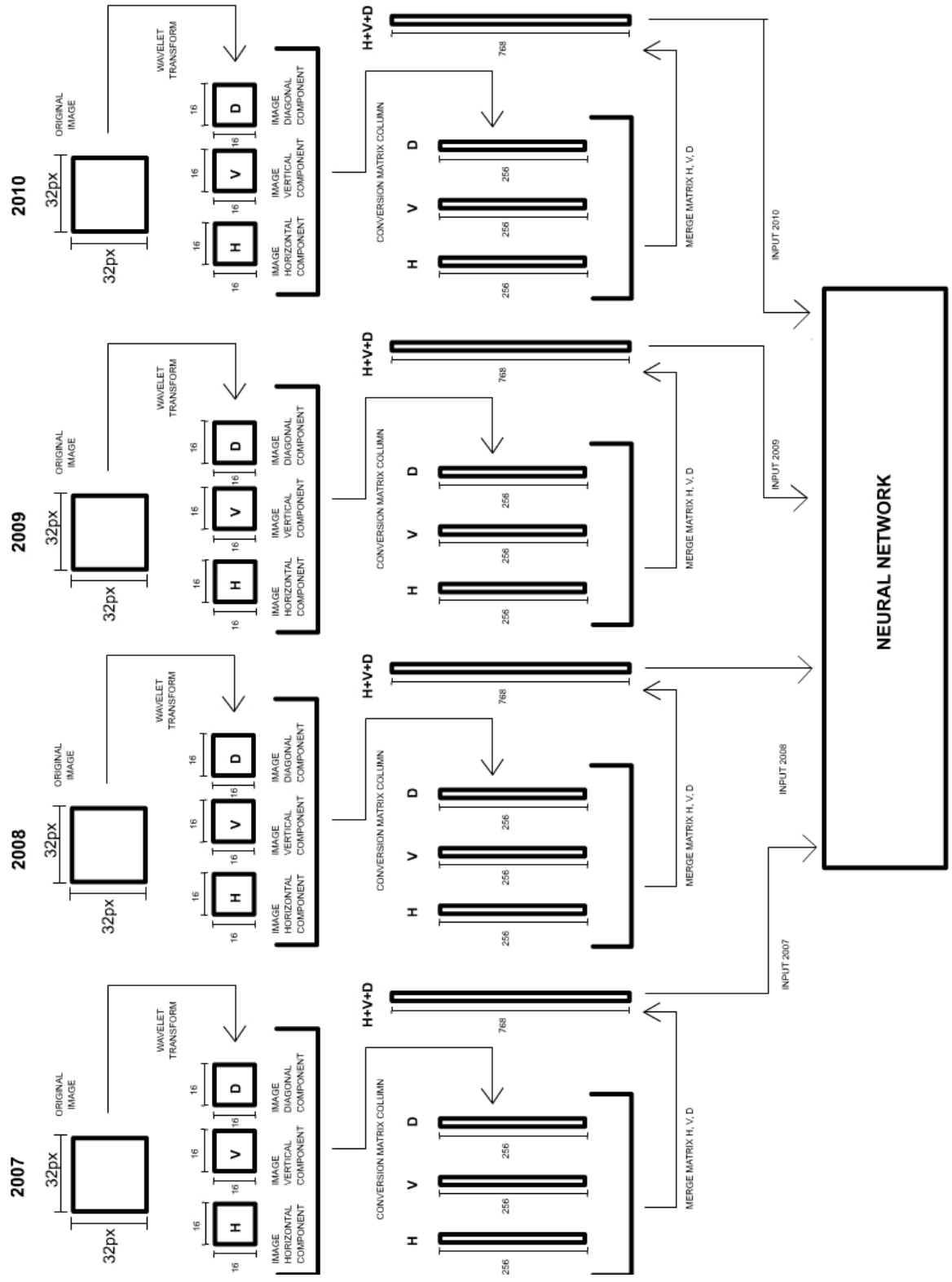


Figure 7.10. Conversion of the image of a water sampling station, by integrating the wavelet transform and artificial neural network techniques

Chapter 8

Case Study:Cefni Reservoir, Anglesey, UK.

Contents

8.1 The Cefni Reservoir

The Cefni reservoir is located in the center of Anglesey, Wales, UK, and it is managed by Welsh Water and Hamdden Ltd, while the associated fishery is managed by the Cefni Angling Association. The associated area is 860000 squared meters (86 ha), with a length of 2.3 km (Fig. 8.1).

Cefni Reservoir on the Isle of Anglesey was overflown as part of the UK Natural Environment Research Council (NERC). The lake is shallow, with a maximum depth of approximately 4 m and contains beds of submersed, floating-leaved and emergent aquatic macrophyte species. It is also known to support dense growths of toxic blue-green algae during summer.

The reservoir is surrounded by an approximately 100-m wide plantation of coniferous trees with agricultural fields beyond[65]. The collection of water was done periodically in this reservoir.

8.1.1 Water collection

Water samplings for physico-chemical water quality evaluations are performed periodically at Cefni and the associated data were provided by the Welsh Water and Hamdden Ltd Company. These data were collected from January 2013 to December 2017.

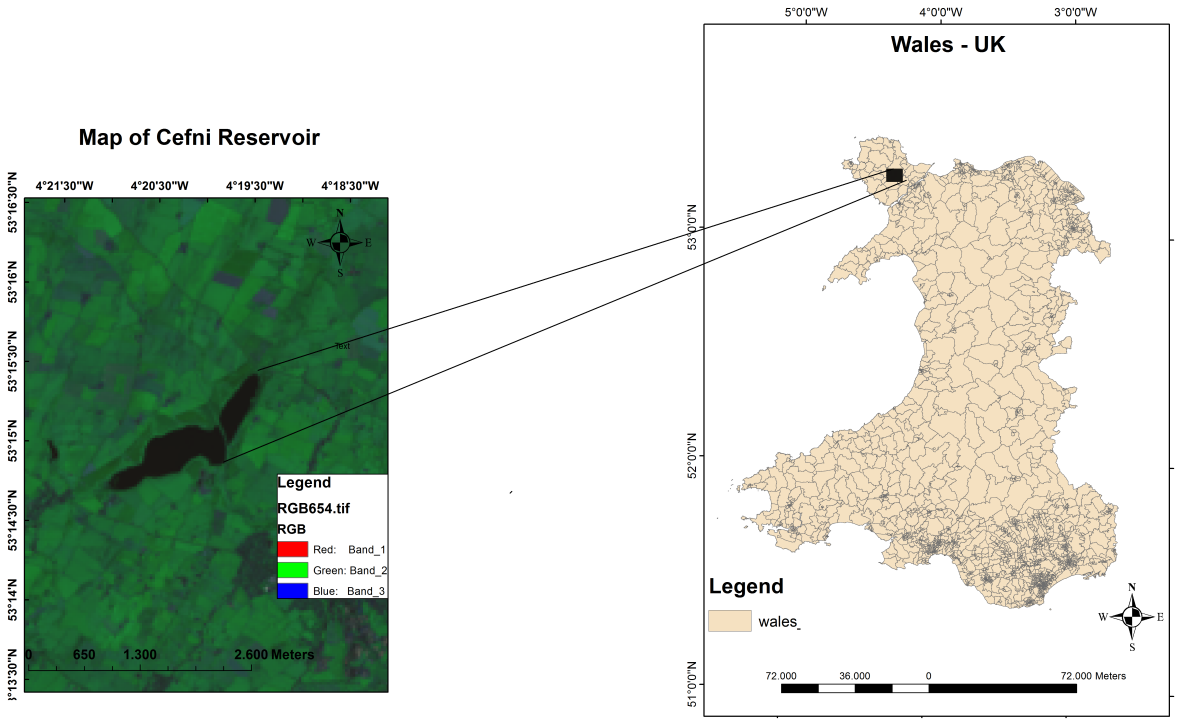


Figure 8.1. Cefni Reservoir, Anglesey, UK.

The relationship between Chlorophyll-a Levels, Suspended Solids, Turbidity and spectral response of the reservoir was determined using the physical water samples collected.

These data have been extracted from the samples and analyzed. We compared it to the proposed parameters level extracted from the remote sensing images, analyzed with a WANN method, described below.

8.1.2 Remote Sensing applied to the Cefni Reservoir

An analysis of the Landsat-8 images has been performed. Landsat-8 was launched on 11 February 2013 with two Earth imaging sensors on-board: the Operational Land Imager (OLI) (that will be discussed in this research) and the Thermal Infrared Sensor (TIRS). The OLI images are in 9 spectral bands as described in [136]. Using measurements "in situ" and a series of Landsat-8 images it is possible to monitor the variables that change the dynamics of ecosystems, such as water level, loss of organic matter or a loss of biodiversity that could be observed in some reservoirs.

Four images per year were collected, corresponding to each season cycle of sampling station, from 2013 to 2017 in order to obtain directly the C, T, and SS concentration from the reflectance of the satellite images. This research has considered the seasonal characteristics of the region, analyzing summer cycle (June to September), autumn cycle (September to December), winter cycle (December to March) and spring cycle (March to June). Landsat-8 data are obtained from the U.S. Geological Survey (USGS) (available at https://espa.cr.usgs.gov/), they include spectral band 2 (Blue), band 3 (Green) and band 4 (Red). All satellite images used were calibrated for radiance values and, subsequently, for reflectance values, as described in many papers [3, 4, 5, 6].

8.1.3 Atmospheric Correction methods proposed for satellite images in the Cefni Reservoir

Atmospheric interference calculation and correction in the satellite images were done using an improved Dark Object Subtraction (DOS) method for Landsat-8 multispectral satellite image [137]. DOS is perhaps the simplest yet most widely used image-based absolute atmospheric correction approach for classification and change detection applications[138, 139, 137]. DOS approach assumes the existence of dark objects (zero or small surface reflectance) throughout a Landsat scene and a horizontally homogeneous atmosphere. The minimum digital number (DN) value in the histogram from the entire scene is thus attributed to the effect of the atmosphere and is subtracted from all the pixels as described in[140].

The method was chosen so the spectrum meets general spectral characteristics of path radiance. These include bad weather conditions, such as clouds, that affect both the amount of incoming solar radiation reaching the water surface and the fraction of light leaving the water surface and reaching the satellite sensor [141]. DOS method with semi-automatic classification plugin was used, as described elsewhere[73].

The reflectance calibration of the OLI was performed using methods based on atmospheric modeling as showed in (8.1). Various models exist with different degrees of spectral resolution, number of atmospheric constituents, cloud management, etc. These models are generally based on the radiative transfer equation, which is given below in its generalized form.

Many atmospheric correction methods have been proposed for use with multi-spectral satellite imagery [142]. Such methods can be grouped in the following categories:

$$L_{sensor} = \underbrace{\frac{E_o * cos(\theta) * \tau * \rho}{\pi} + \frac{E_d * \tau * \rho}{\pi}}_{\text{Ground-reflected}} + L_{path}, \tag{8.1}$$

where:
L_{sensor}: Radiance received at the sensor.
E_o: Solar radiance at the top of the atmosphere.
θ: Incidence angle of solar radiance on the surface (0 for vertical, 90 for horizontal).
τ: Transmittance factor .
ρ: Reflectance factor.
E_d: Scattered background radiation (of the sky).
L_{path}: Path scattered radiation reaching the sensor.

This equation describes the core process of many models. Atmospheric models need to take into account the complex time and space varying composition of the atmosphere, as well as the wavelength dependent interactions caused by the radiance reaching the sensor. The primary characteristics of the OLI instrument relevant to this research are presented in table 8.1, reporting characteristics of the analyzed bands: band 2, band 3 and band 4. Throughout this document, the spectral bands will be referred to by their band number, name or center wavelength.

8.2 Physico-chemical Parametres

Physico-chemical parameters of the water such as PH , CONDUCTIVITY, COLOUR, CHLOROPHYLL-a are important variables for the analysis of Freshwater ecosystems, that are significant not only for human populations but also essential for plant and animal diversity.

Table 8.1. Landsat-8 Operational Land Imager (OLI).

Landsat-8			
Band	***Band***	***Wavelength(nm)***	***Bandwidth(nm)***
2	Blue	482	60
3	Green	561	57
4	Red	655	38

Freshwater lakes are significant not only for human populations but also essential for plant and animal diversity. These aquatic systems are unique and rich in biodiversity at the same time are under constant threat due to bludgeoning human populations and their demand for land[143].

The concentration of the hydrogen ion in water is usually measured in terms of pH. The pH or negative logarithm of the hydrogen ion concentration is a master variable in water quality because the hydrogen ion influences many reactions. The optimum pH range for most aquatic organisms is 6.5–8.5, and the acid and alkaline death points are around pH 4 and pH 11, respectively. Most living organisms do not tolerate large variations in pH and may die. [144][145].

The pH of natural waters is strongly influenced by the concentration of carbon dioxide, which is considered an acid gas[145]. Because dissolved carbon dioxide is acidic, rainwater that is saturated with this gas is naturally acidic—usually about pH 5.6.[144].

Water bodies with moderate to high alkalinity are well-buffered against wide daily swings in pH resulting from net removal of carbon dioxide by photosynthesis during daytime and return of carbon dioxide to the water by respiratory process at night when there is no photosynthesis. [144]

Conductivity defines the ability of water to conduct electricity.This parameter provides a good indication of changes in the composition of water as pollutant particles. This type of measurement accesses the concentration of ions in a solution. More the ions higher will be the conductivity. For water to be pure it's conductivity should be poor.

Generally there are two types of conductivity sensors: two electrodes and multiple electrodes from which two electrodes sensor is commonly used. It is made by using two platinum plates deposited on two parallel glass or inner wall of glass tube. Conductivity of water measures in μs/cm or mA [146].

Colour is water quality parameters that detract from the appearance of water, colour refers to Transparency condition of water,organic material that has dissolved into solution and important in determining Secchi disk depths among reservoir [147].

The Chlorophyll-a (Chl-a) concentration is commonly used as a proxy for phytoplankton biomass and as indicator for eutrophication and it can be retrieved from remote sensing data[148].Chl-a is a direct indicator used to evaluate the ecological state of a waterbody, such as algal blooms that degrade the water quality in lakes, reservoirs and estuaries[149].

Many scientists have studied on the influence of freshwater physico-chemical parameters to the changing ecosystems. Studies involving the statistical analysis, wavelet signals, neural network, remote sensing and water sampling have been used in the monitoring of ecosystems [30][31][32][33] [34].

Several operational monitoring systems based on remote sensing are in place to monitor the reservoir. However, evaluations of reservoir monitoring systems based on satellite data are scarce. The methods currently used for water analysis are time-consuming, extremely costly, because it requires sample collection, trained people and specialized laboratories. Predicting

these parameters helps decision-making in the present and planning in the future.

Artificial neural network algorithm can be used for simulates human learning processes through establishment and reinforcement of linkages between the input and output data and can make relationship of a dependent variable with independent variables.

Even with a correct model applied by a well-trained analyst, all predictions remain subject to fundamental uncertainties, especially with regards to variation in aspects such as actual weather conditions[150].In order to reduce this uncertainty, many investigations are conducted to find a more efficient model that allows the researchers to infer with greater precision the estimation of water quality parameters.

Predictions based on time series are efficient for treatment of uncertainty. Thus arises interest in prediction physico-chemical parameters, in which the difficulty involves to estimate these parameters through the reflectance of satellite images associating a set of data available in data collection.

In the context of water quality parameters, we propose an alternative for retrieval through prediction with ANFIS that are adjusted, or trained, so that a given input leads to a specific target output.

To this end, a case study with the data of the Cefni reservoir was applied, aiming to clarify the benefits of the ANFIS, emphasizing its efficiency and simplicity of implementation.

8.2.1 Pre-processing satellite images: Cefni Reservoir

For the Cefni Reservoir, the remote sensing images obtained from the Landsat 8 and Sentinel-2 satellites were analyzed for the years 2007 to 2016, using spectral band 1, band 2, band 3 and band 4, applying the band compositions technique for satellite image, and then applied to atmospheric correction, as described in the following sessions:

8.3 Atmospheric corrections

The value recorded at any pixel location on a remotely sensed image does not represent the true ground-leaving radiance at that point. Part of the signal is due to the properties of the target of interest and the remaining is generated by the properties of the atmosphere itself. The separation of contributions is not known a priori, so the objective of atmospheric correction is to quantify these two components. In this respect, the analysis can be based on the corrected target reflectance or radiance values. The process of atmospheric correction is the removal of atmospheric effects to retrieve physical parameters of the earth's surface, generally the surface spectral reflectance. Also emissivity and temperature, when applicable, can be retrieved after atmospheric corrections [151].

Therefore, atmospheric correction is an essential part of pre-processing and a prerequisite for the derivation of certain value adding products. The total signal at a remote sensing sensor consists of four components [151]; see Figure 8.2.

1. path radiance: radiation scattered by the atmosphere (photons without ground contact);

2. reflected radiation from the viewed pixel;

3. scattered radiation from the neighbourhood and scattered into the view direction;

4. terrain radiation reflected to the pixel (e.g. from opposite hills, according to the terrain view factor).

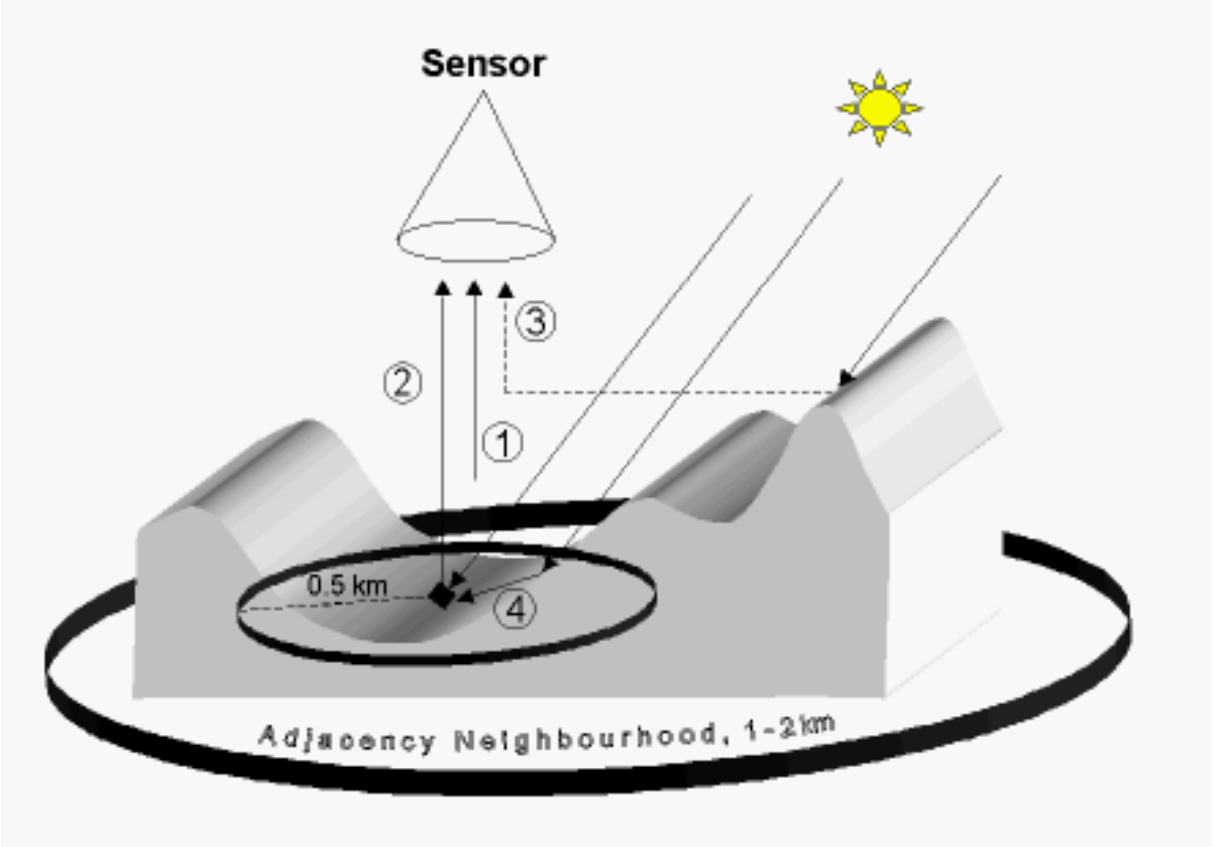


Figure 8.2. Representation of the signal at a remote sensing sensor

Only component 2 contains information from the viewed pixel. The task of the atmospheric corrections is the calculation and removal of components 1, 3 and 4, and the retrieval of the ground reflectance from component 2.

8.4 WANN applied to the remote sensing images in the Cefni Reservoir

The wavelet transform is an integral transform whose kernel is a class of special functions, called wavelets [49]. The main advantage of this method, compared to other methods, is in its spectral location capability in space and frequency, which allows the analysis of non-stationary signals in their various scales [152][153][50]. WANN was trained by season, four images are collected per year, corresponding to four different cycles of the reservoir.

Its essence is to express a function through the telescopic or translation of basic *mother* wavelet function $\psi(t)$.

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \tag{8.2}$$

$$y^{(k)} = \sum_{j=1}^{N} \omega_i \psi_j \left(\sum_{k=1}^{M} \frac{x_k * U_k j}{a_j}\right) , \tag{8.3}$$

where $\underline{x} = [x_1\ x_2 \cdots x_N]$ is the input, composed by the pixels of satellite images, and $y^{(k)}$ is the predicted output value (C, T, SS). U_{kj} is the connection weight from the kth node (input layer) to the jth node (hidden layer); ψ_j is the activation function of the jth neuron (hidden layer); ω_j

is the layer weight from the jth nodes (hidden layer) to the output; a_j is the expansion parameter of wavelet function; and b_j is the translation parameter of wavelet function.

WANN is based on the topology and structure of the back propagation neural network. The wavelet basis function is the incentive function of the neurons. By using the advantage of the wavelet signal analysis and combining the function of the neural network training and prediction, the signal is transmitted forward and the error is propagated backward, so as to achieve a more accurate predictive value signal.

This study adopted a three-layer network structure: the input with m nodes, the hidden layer with n nodes, and the output with a single node [154]. The WANN network structure is shown in Fig. 8.5.

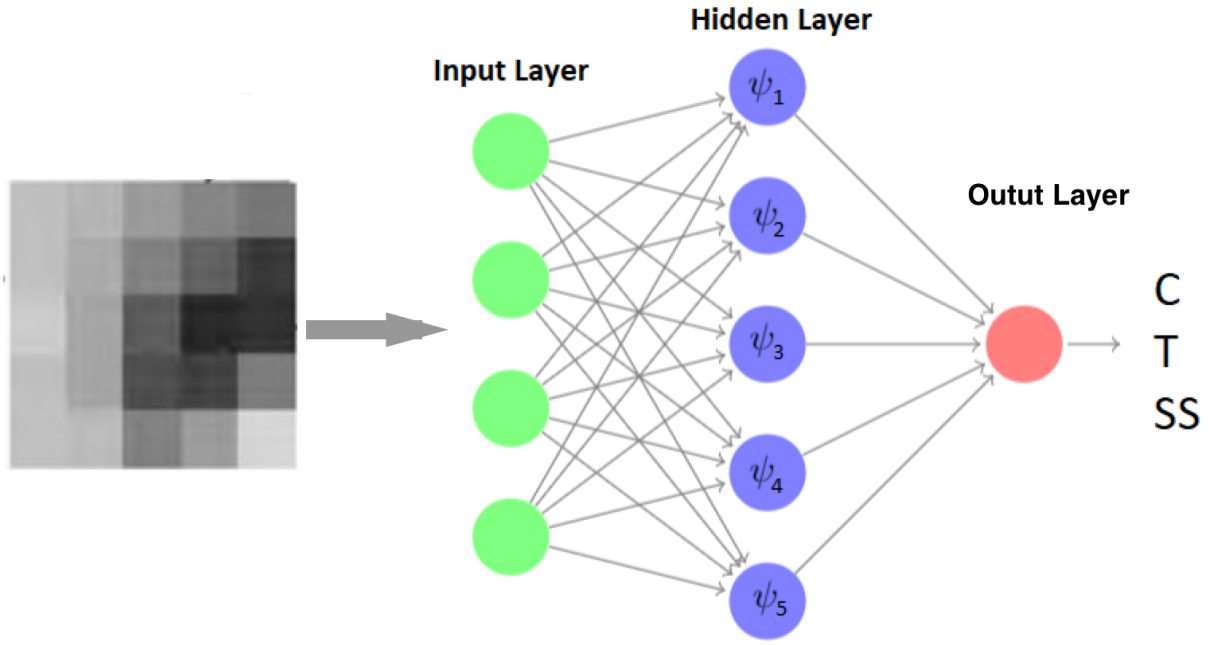


Figure 8.3. WANN

The validation process of the WANN was conducted with images from the year 2017. Fig. 8.5 shows the architecture of the ANN with the column vector $P_i, (i = 1, 2, .., 768)$ as input. The WANN was trained with the following parameters: Learning rate 0.01, *tansig* transference function in all the neurons of the hidden layer and *purelin* in the output layer. The network training function is a gradient descent with momentum and adaptive learning rate. The wavelet transform used in the present study is the discrete transform, which allows for the multi-resolution analysis of a signal, decomposing said signal into approximations and details.

The approximations are high ranges, i.e., low-frequency signal components. The details are the low ranges, i.e., high frequency components [155][156]. One sampling station was initially chosen for analysis and a geographic image of the water sampling station, of 32x32 pixels, was cropped, corresponding to an array containing 1024 pixels. Each digital pixel value corresponds to an average of radiance values, emittance or backscatter of the different targets that can be contained in the pixel from the vicinity of the water sampling station, as displayed in an example in Fig. 8.4.

Subsequently, the Haar Wavelet Transform was applied, with only one level of decomposition, resulting in a matrix array of 16x16 pixels for each of the following three components: Horizontal

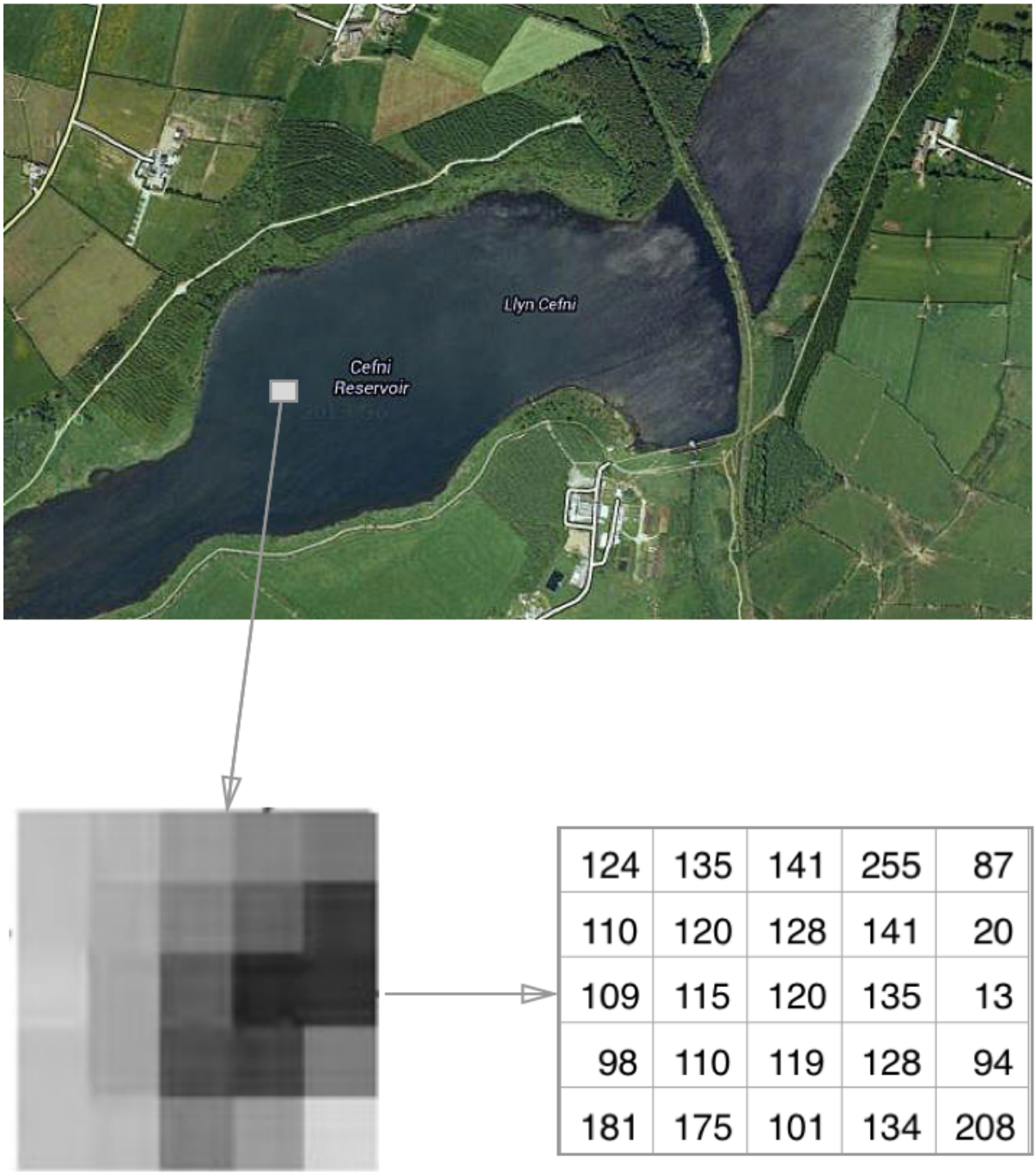


Figure 8.4. Pre Processing Images: Cefni reservoir with group of pixels area in gray values and corresponding digital numbers (DN)

(H), vertical (V) and diagonal (D). The conversion of the arrays of the H, V and D components to their respective column-matrices was performed, and subsequently, a concatenation of the three arrays (each containing 256 pixels) was executed, generating a vector with 768 column size (256 x 3), as shown in Fig. 8.5

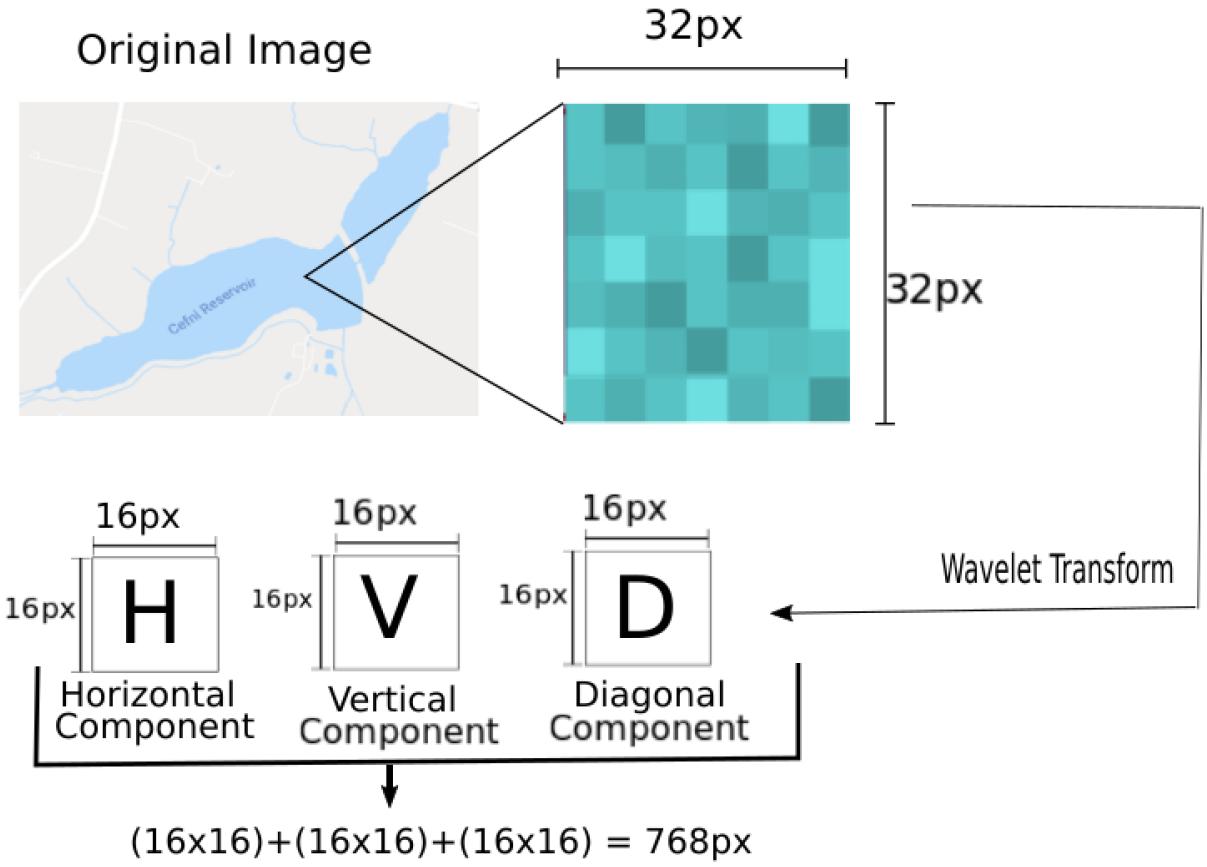


Figure 8.5. Pre Processing Images: Wavelet Transform

The images of the geographical area containing the water sampling collection point, decomposed via wavelet into its three wavelet components (Horizontal, Vertical, Diagonal) was used as the input for WANN as shown in Fig.8.6 .

The wavelet transform result in 768 input elements for the season cycle for each year, with each season composed of 4 stages, or 768X4. Thus, we have 4 Artificial Neural Network, each one responsible for validating each cycle with four stages for each of the study years. In Figure 8.6 , onlye the dimension 768 was specified, since the modeling is open to specify 768 elements x number of years. In this case, 768 x 4 years (2013, 2014, 2015 and 2016).

Preliminary tests were conducted considering the image representations isolated for each wavelet component, with satisfactory results. However, when the input data of the three wavelet components was considered, the approximations were even better, which motivated the choice of this arrangement in the proposed solution.

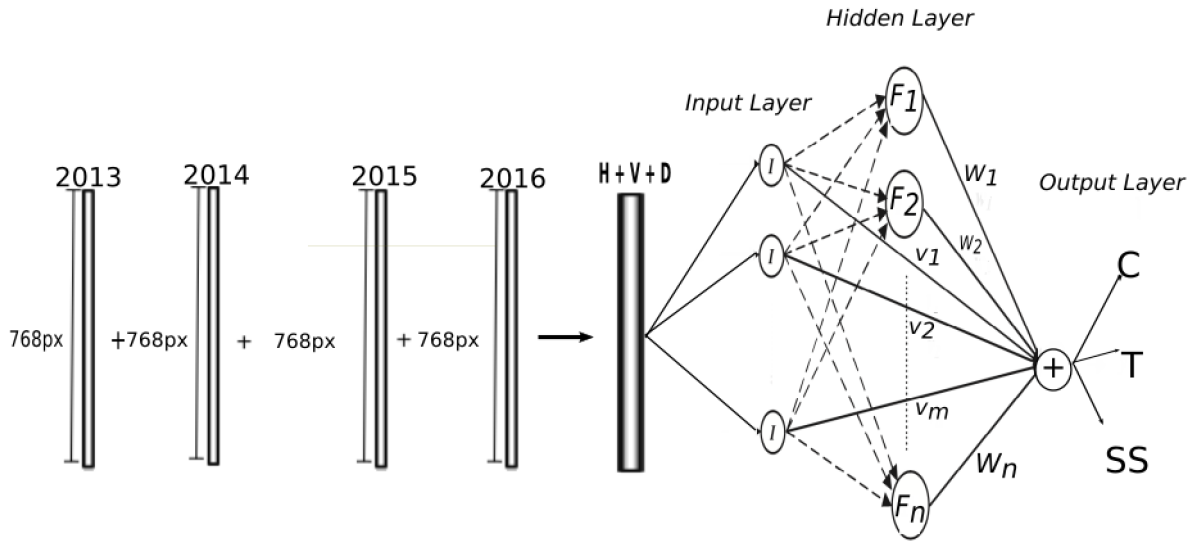


Figure 8.6. The input for WANN

8.5 Predicting Physico-Chemical Parameters in the Reservoirs by WANN and Satellite Imagery

In the following, the results of the prediction are shown. The table 8.2 shows the WANN training results for Cefni. The mean square errors (MSE) for neural network training are shown to be low for all the Chlorophyll-a, Turbidity and Suspended Solids.

Table 8.2. Mean square errors in the WANN training conducted in the present study

MSE Training				
Var	***Summer***	***Autumn***	***Winter***	***Spring***
C	9.1×10^{-22}	3.5×10^{-09}	4.1×10^{-22}	5.7×10^{-22}
T	1.6×10^{-06}	1.2×10^{-04}	3.6×10^{-09}	2.3×10^{-09}
SS	2.6×10^{-21}	2.5×10^{-07}	1.6×10^{-06}	1.0×10^{-21}

The relative errors were calculated by the following formula in (9.7) and shown in Table 8.3:

$$RelativeError(E_r) = \frac{|\chi_e - \chi_o|}{\chi_o} \tag{8.4}$$

where:
χ_e: Estimated Value
χ_o: Observed Value

Validation results for 2017 are shown in the Figures 8.7, 9.13, 9.14. The laboratory results correspond to the "observed values" and the analysis by ANN correspond to the "estimated values" (namely C, T and SS).

The X-axes of figures represent the season cycles (1, 2, 3 and 4), respectively, summer, autumn, winter, and spring. The Y-axis represents the quantitative value of the analyzed parameters.

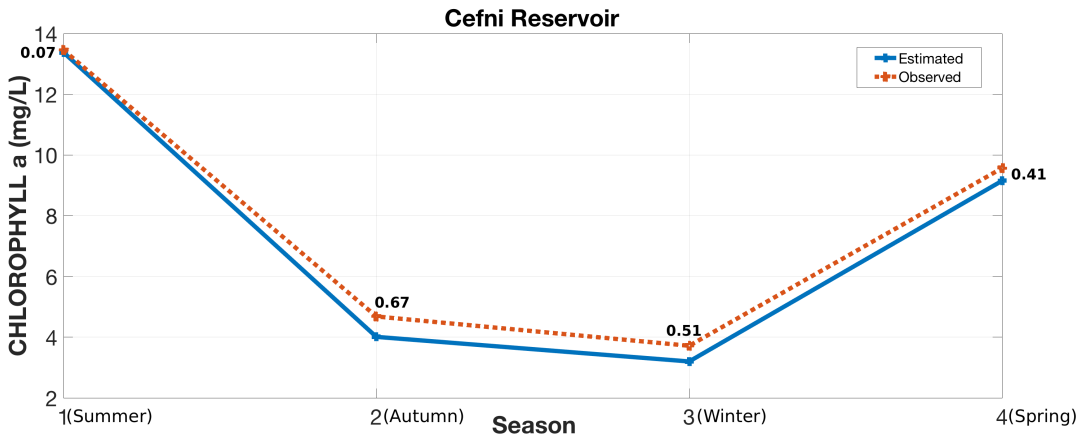


Figure 8.7. Predicting Chlorophyll-a Levels in the Cefni reservoir by WANN and satellite images.

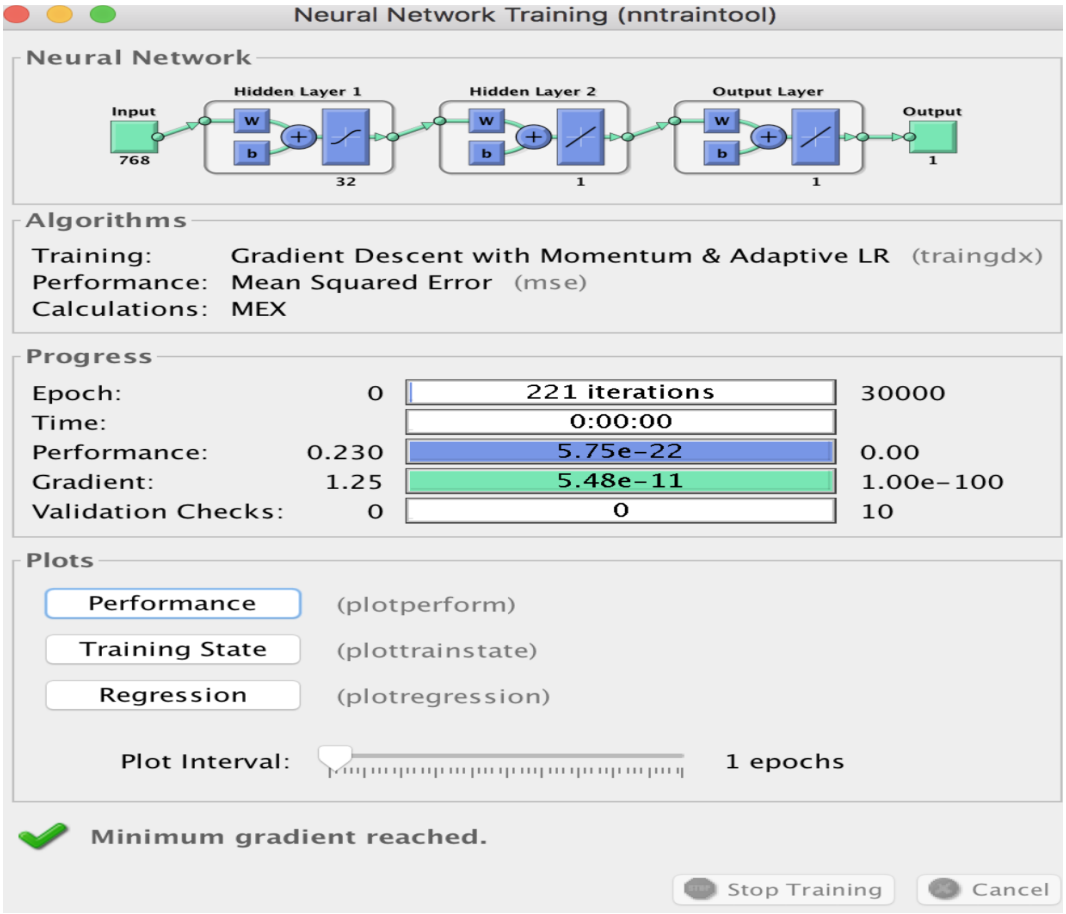


Figure 8.8. Neural Network Epoch Iterations: Predicting Chlorophyll-a Levels

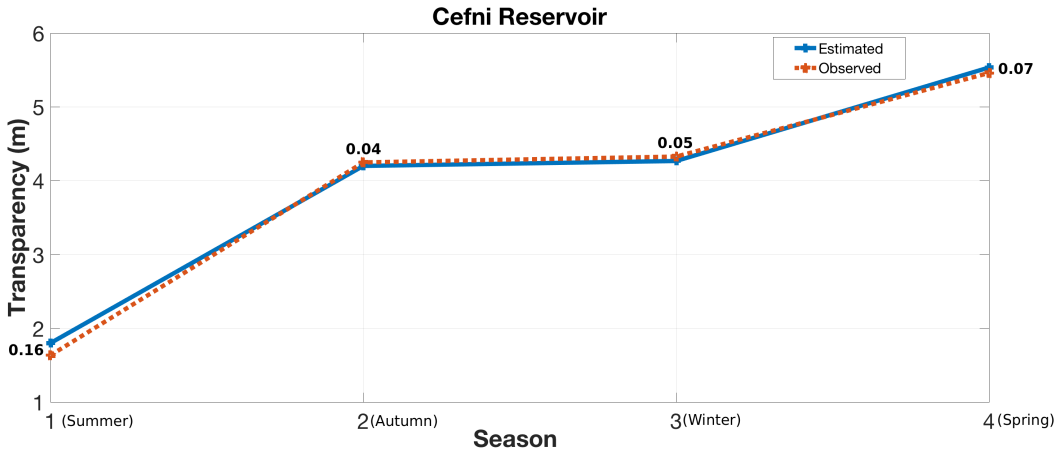


Figure 8.9. Predicting Turbidity in the Cefni reservoir by WANN and satellite images.

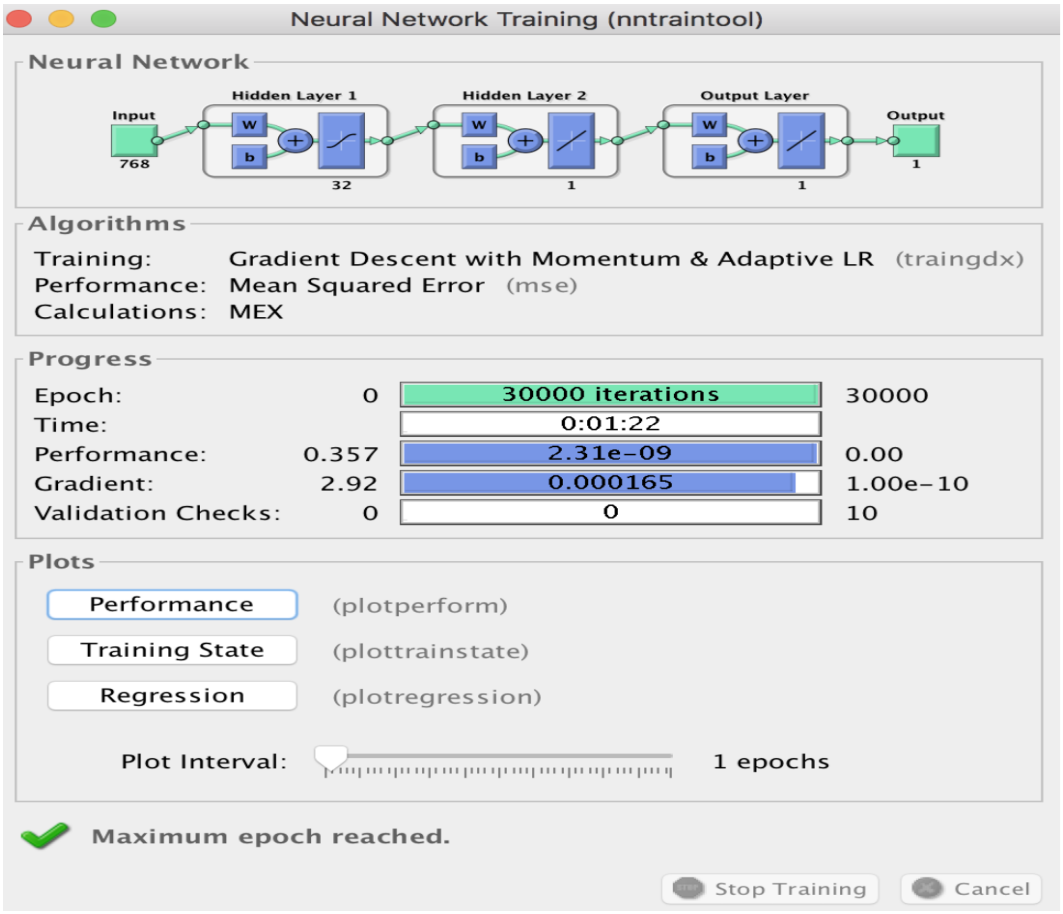


Figure 8.10. Neural Network Epoch Iterations: Predicting Turbidity

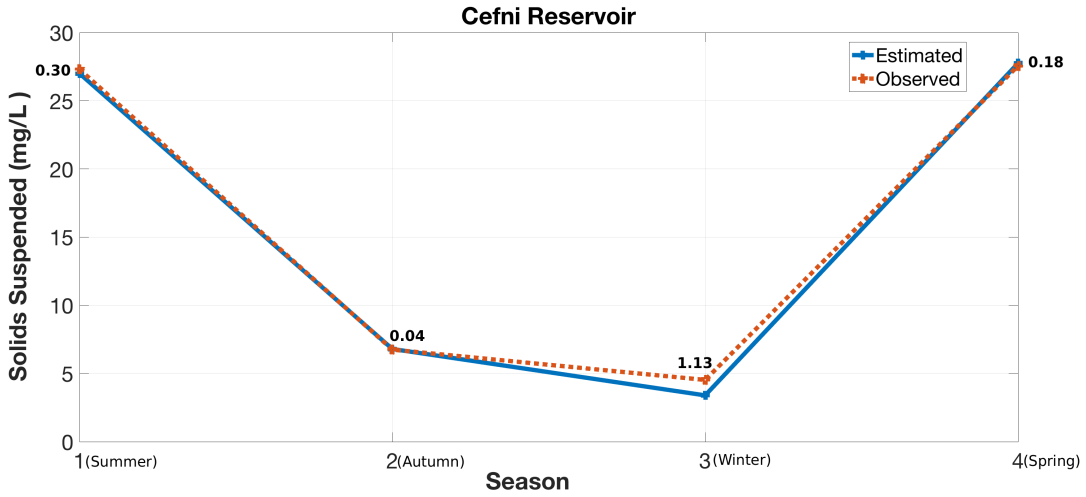


Figure 8.11. Predicting Solids Suspended in the Cefni reservoir by WANN and satellite images.

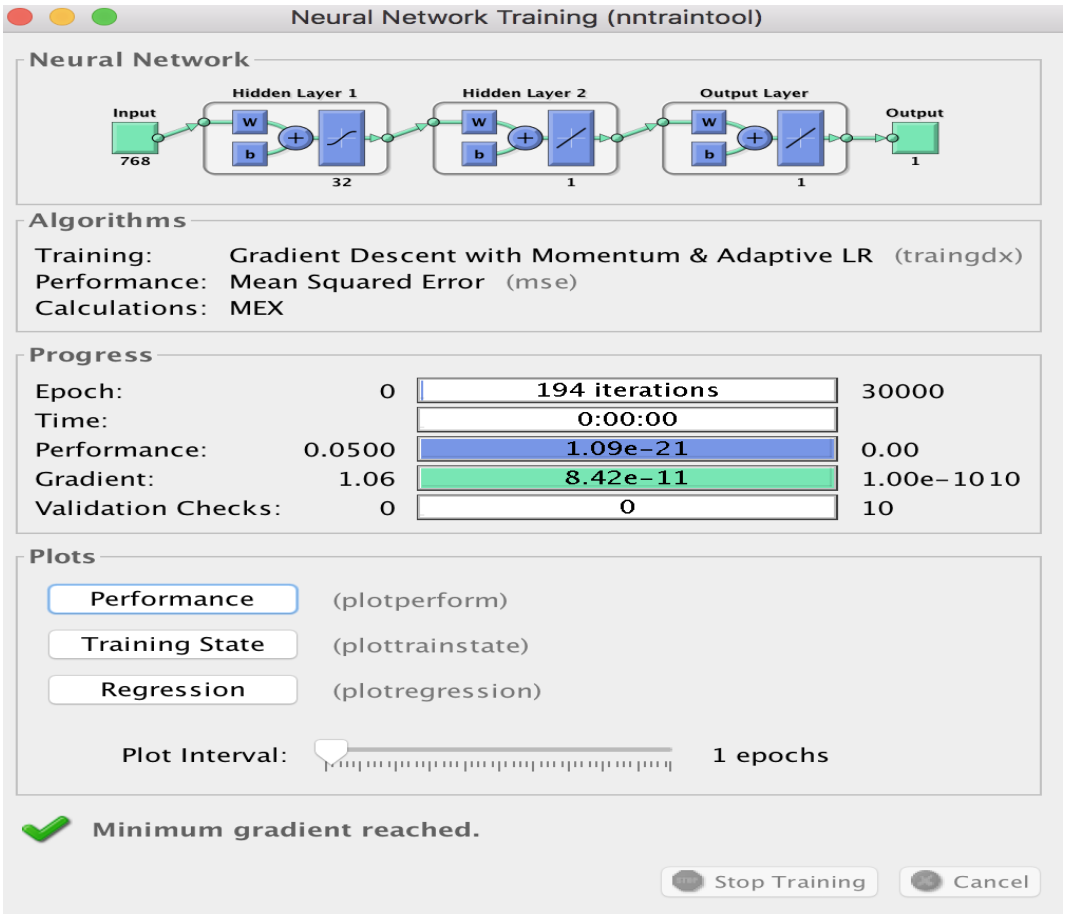


Figure 8.12. Neural Network Epoch Iterations: Predicting Solids Suspended

Table 8.3. Relative Error (Er) in the sampling station, evaluated parameter by season cycle

Relative Error				
Var	***Summer***	***Autumn***	***Winter***	***Spring***
C	0.0059	0.1435	0.1391	0.0435
T	0.1010	0.0112	0.0138	0.0136
SS	0.0112	0.0070	0.2499	0.0068

The Table 8.4 shows the errors between expected and observed values. The Best results were obtained in the Summer season for Chlorophyll-a Levels, cycle 1; in the Autumn for Turbidity and Solids Suspended, cycle 2. These periods corresponds to the less cloudy period in the region, these are associated with the best estimates by WANN and satellite images from Landsat-8. The results are satisfactory for the other cycles and the other parameters.

Table 8.4. Approximation errors of the proposed method for 2017 in the sampling station by seasonal cycle.

Absolute differences				
Var	***Summer***	***Autumn***	***Winter***	***Spring***
C	0.07	0.67	0.51	0.41
T	0.16	0.04	0.05	0.07
SS	0.30	0.04	1.13	0.18

However, in the present study, slightly different values between estimated and in-situ data were observed in some cases, attributed to either the presence of clouds or sun glint in the remote sensing images, corroborating previous studies [141].

This should, thus, be taken into account when applying this type of methodology to environmental monitoring of reservoirs, even though differences were very slight, and corrections to these questions have been done here by atmospheric correction method, the errors can still be improved in the future, providing focus for future studies.

The absolute differences between the observed and estimated values were calculated as showed in (8.5)

$$Err(i) = Y_{lab}(i) - Y(i) \tag{8.5}$$

Where:
$Y_{lab}(i)$ is the value obtained in the laboratory for this parameter and $y(i)$ is the output variable obtained by the ANN.

Parameters	Summer			Autumn			Winter			Spring		
	EV	OV	AD	EV	OV	AD	EV	OV	AD	EV	OV	AD
Chlorophylla	13.38	13.45	**0.07**	4.01	4.68	**0.67**	3.20	3.71	**0.51**	9.15	9.56	**0.41**
Turbidity	1.80	1.63	**0.16**	4.20	4.24	**0.04**	4.26	4.32	**0.05**	5.53	5.45	**0.07**
Suspended Solids	27.00	27.30	**0.30**	6.80	6.75	**0.04**	3.40	4.53	**1.13**	27.75	27.56	**0.18**

Table 8.5. Absoluted Difference (AD) between Estimated Value (EV) and Observed Value (OV)

The results estimated in the present study when compared with those observed in the

laboratory proved extremely close to each other, demonstrating the adequate efficiency of the proposed method, WANN showed good results in the evaluation of physico-chemical parameters.

8.6 Conclusions for estimations using WANN

The present study proposes a method to estimate Chlorophyll-a Levels, Turbidity and Suspended Solids using Wavelet Artificial Neural Network and satellite images, band 2, band 3, band 4 from Landsat-8, which demonstrated sensitivity to the presence of suspended particulate matter in water bodies. The summer, autumn, winter, and spring seasons were considered for this study from 2013 to 2017.

The images from 2013 to 2016 were used to train neural networks, and 2017 used for validation. The proposed method using neural networks and wavelet transform resulted in satisfactory approximations between the observed values and the estimated values in relation to the same water samples.

The wavelet artificial neural network demonstrates good results between observed and estimated values after atmospheric corrections in satellites images, better results were observed in the summer and Autumn season, when there are few clouds in the region.

The techniques proposed in the present study are noteworthy since they can aid in evaluating important physico-chemical parameters, allowing for identification of possible anthropogenic impacts, being relevant in environmental management and policy decision-making processes. This work contributes the evaluation of different methods accuracy in estimating of the water quality.

Chapter 9

Results

Contents

9.1 Results for Cefni Reservoir

The prediction performances of the considered models were investigated by means of several simulation tests carried out using data from Cefni Reservoir, Anglesey, UK. Water from the Cefni is currently abstracted and treated for use as drinking water by Welsh Water.

The samples were collected in 2015, 2016 and 2017 and were analyzed in the laboratory by specialists using a microscope and results coded.

In order to predict algae, the neural networks were trained using water collections and satellite imagery. All of the experiments had been performed using Matlab 2016b, running on a 2.9-GHz Intel Core i7 platform equipped with 16 GB of memory.

In addition to the three proposed neural and fuzzy neural models, the linear (LSE) predictor was adopted for benchmarking. LSE did not have parameters to be set in advance, while for RBF and ANFIS, we had adopted the default options provided by the software platform for training and model regularization (RBF and ANFIS models were trained by using the supported functions in the econometrics toolbox, neural network toolbox and fuzzy logic toolbox of MATLAB, respectively).

The prediction performances were measured by metrics commonly adopted for time series, which were independent of the said procedure for data normalization:

- *Normalized Mean Square Error (NMSE)*

$$NMSE = \frac{\sum_{k=1}^{N_T} (y_t - \widetilde{y}_t)^{(2)}}{\sum_{k=1}^{N_T} (y_t - \widetilde{y}_t)^{(2)}} , \tag{9.1}$$

Where $\widetilde{y}$ is the average value of samples y_t in the test set;

- *Mean Absolute Range Error (MARE)*

$$MARE = \frac{1}{N_T} \sum_{t=1}^{N_T} \frac{|y_t - \widetilde{y}_t|}{y_{max} - y_{min}} \ , \tag{9.2}$$

- *Mean Square Error (MSE)*

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (X_i^{estimated} - X_i^{measured})^2 \ , \tag{9.3}$$

- *Mean Absolute Percentage Error (MAPE)*

$$MAPE = \left\{ Mean \left(\frac{1}{n} \sum_{i=1}^{n} \left| \frac{X_i^{\text{estimated}} - X_i^{\text{measured}}}{X} \right| \right) \right\} \ , \tag{9.4}$$

- *signal-to-noise ratio (SNR)*

$$SNR = 10 \log_b \frac{L^2}{MSE} \ , \tag{9.5}$$

Where L is the dynamic range of allowable image pixel intensity. [157]. ;

The numerical results for Algae in each year tested, 2015, 2016, 2017 are reported in the table 9.1, where the performance of the best model is marked in bold font.

According to the results shown in this research, neural networks models ensure good performances. Algae, Turbidity and Suspended Solids Training Results are reported in the Fig. 9.1, 9.2, 9.3, where the graphic illustration of the current time series (blue line) and predicted by the neural network (red line), along the four training conditions, by seasonality of the time series considered, winter, spring, summer, autumn.

The Test Results are reported from Fig. 9.4, 9.5, 9.6, 9.7, 9.8, 9.9, 9.10, 9.11, 9.12. In the x-axis is reported the cumulative index of samples of the considered day, being shown the number of samples for two analyzed years.

Table 9.1. Algae: Prediction results for two years

	NMSE - ALGAE			
Prediction Model	Summer	Autumn	Winter	Spring
LSE training	0.00284	0.00272	0.00262	**0.00302**
LSE test	0.00270	**0.00176**	**0.03219**	**0.00190**
ANFIS training	0.00215	0.00204	0.00198	0.00220
ANFIS test	**0.00216**	0.00286	0.18846	0.00400
RBF training	0.00185	0.00166	0.00165	0.00184
RBF test	0.00226	0.00216	0.12106	0.00283

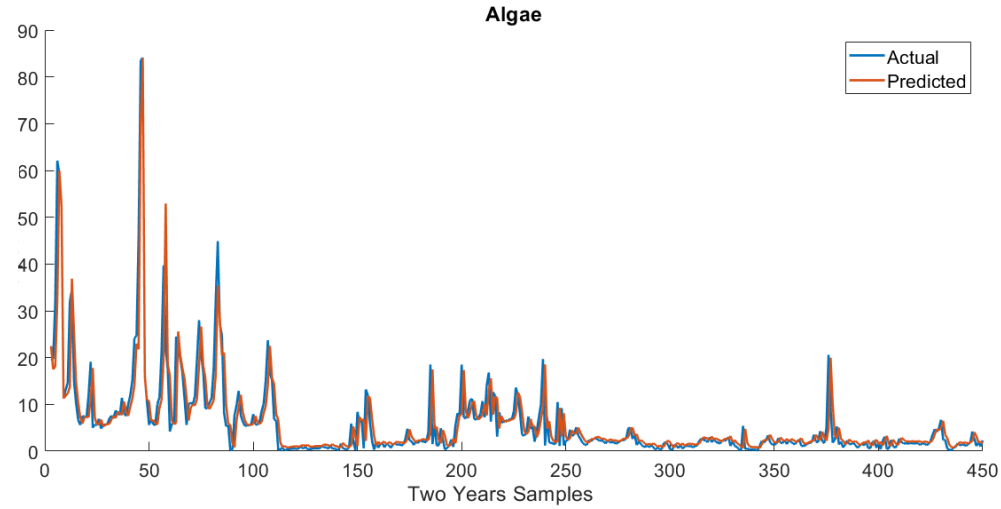


Figure 9.1. Algae Training Results

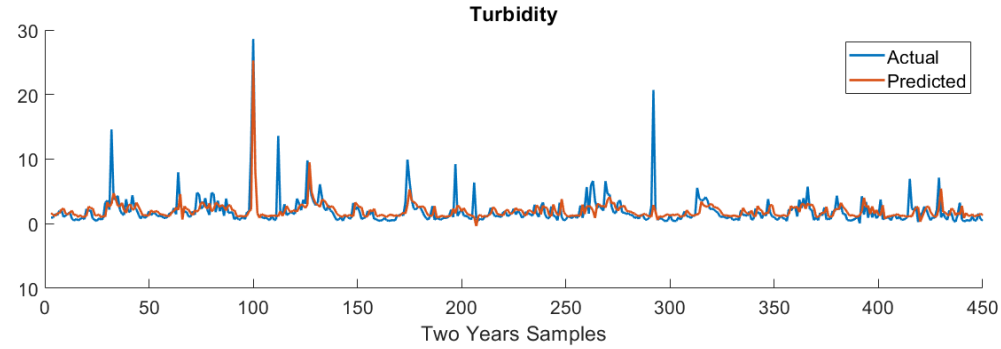


Figure 9.2. Turbidity Training Results

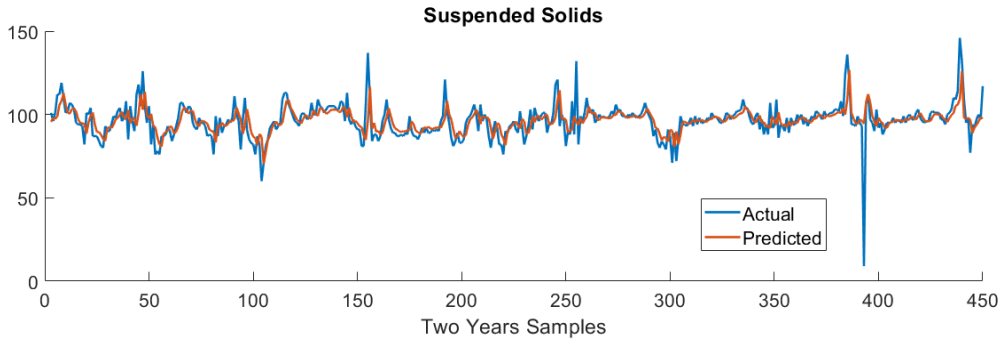


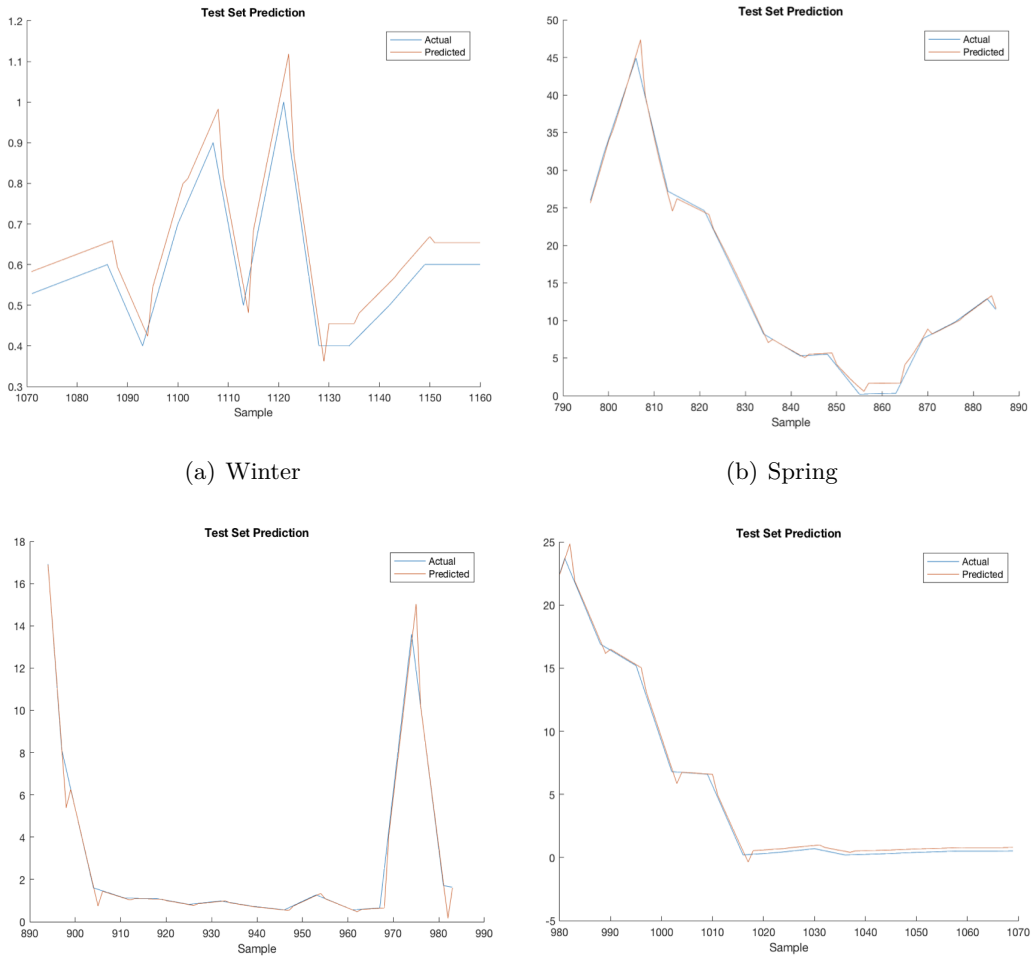
Figure 9.3. Suspended Solids Training Results

The numerical results for Turbidity in each year tested, 2015, 2016, 2017 are reported in the table 9.2, where the performance of the best model is marked in bold font.

The numerical results for Suspended Solids in each year tested, 2015, 2016, 2017 are reported in the table 9.3 where the performance of the best model is marked in bold font.

9.2 Conclusions for Cefni using ANFIS, RBF and LSE

This study proposes a multidisciplinary approach to asses the feasibility for the Sentinel-2 sensor to monitor Algae, Turbidity and Suspended Solids in the Cefni reservoir. An understanding of the ecological characteristics of reservoirs, including bio-physical and chemical features is



(a) Winter (b) Spring

(c) Summer (d) Autumn

Figure 9.4. ANFIS for Algae: Winter, Spring, Summer, Autumn from 2015 to 2017

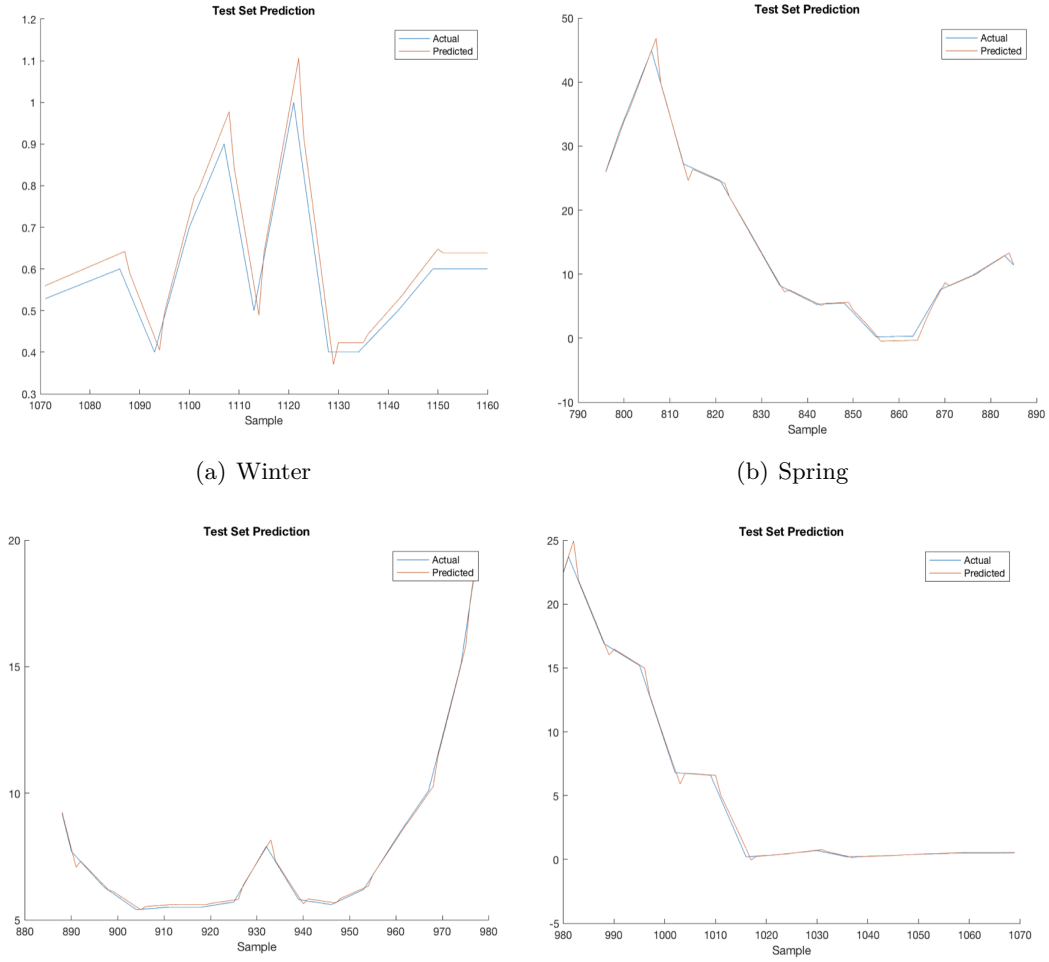
important for their water management. Biological studies are importante to assess uses of water in reservoirs due ti their close relation to the effects of algal blooms. Enhanced phytoplankton growth is a major concern for policy and management particularly when the reservoir is used for recreation, aquaculture or potable supplies.

As stated before, we used a set of satellite images from cefni reservoir, and then cut the images, 32x32 pixels, using imcrop, funcion from Matlab in the collecting the samples.

We used transformed images, by applying wavelet transform to train and test a NN with the three algorithms explained above (ANFIS, RBF and LSE). The output parameters of the inference system are Algae values, turbidity and suspended solids.

Our results showed the application of a method to estimate chemical and physical parameters, using Neural and Fuzzy Neural Network. Satellite imagens, band 2, band 3, band 4 and band 8 from Sentinel_2 demonstrated good acurrace in the predictions, being analyzed in the summer, autumn, winter and spring seasons, for the years 2015 to 2017 considered here in this study.

Neural and Fuzzy Neural Network demonstrated good results between observed and estimated after atmospheric corrections in satellites images. Better results were observed in the summer and spring season, when there are few clouds in the region, providing more efficient analysis of satellite imagery.



(a) Winter (b) Spring

(c) Summer (d) Autumn

Figure 9.5. RBF for Algae: Winter, Spring, Summer, Autumn from 2015 to 2017

The techniques proposed in the present study are noteworthy since they can aid in evaluating important parameters, allows for identification possible changes in ecosystems, being relevant in environmental management and policy decision-making processes.

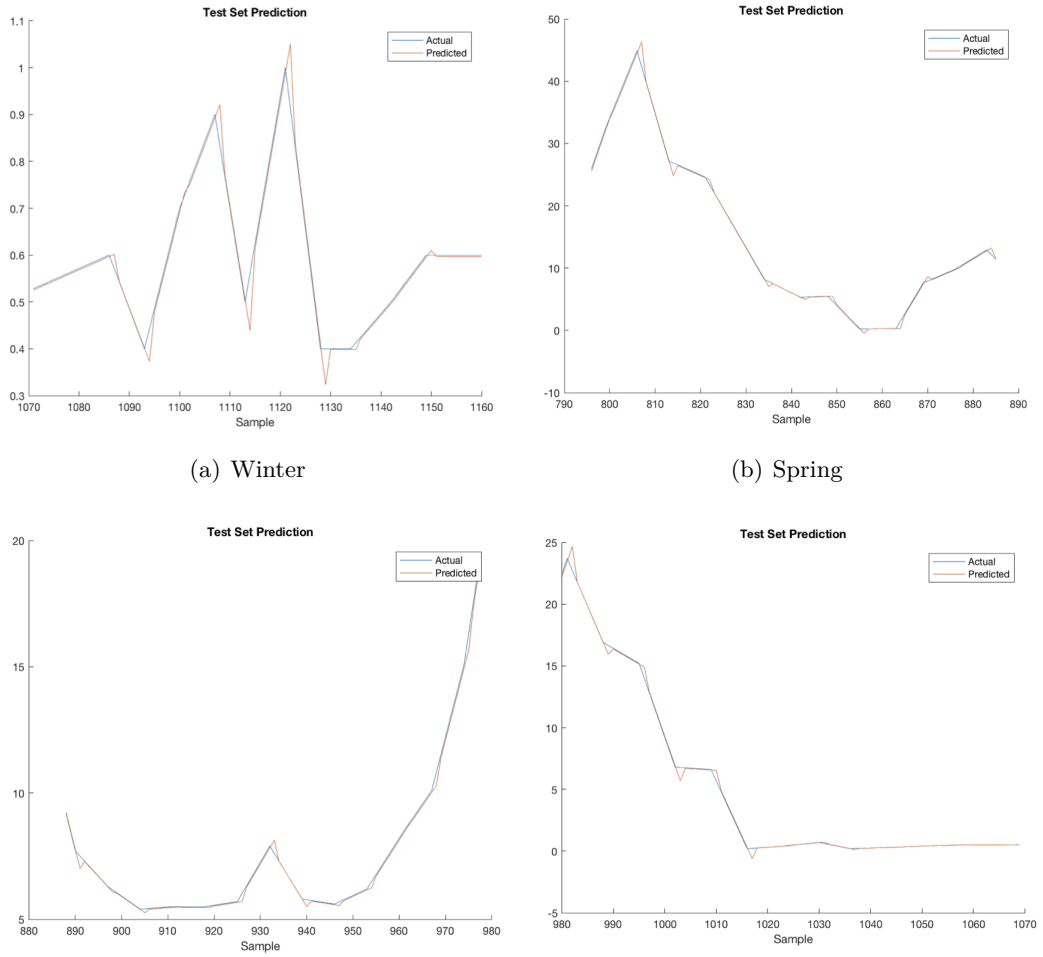
The proposed method showed good accuracy using images from Sentinel_2 satellite images, that provides color images of 10 m resolution. The results show that employment satellite images to estimate these parameters can be in a important tool for the systematic monitoring of water quality in reservoirs.

9.3 Results for Tucurui Reservoir

The integration of water quality parameters is essential in environmental monitoring and very important for decision-making. Advanced techniques to manage are required in complex evaluation process. We here propose a ANN hybrid model to assess reservoir water quality using Remote Sensing and Wavelet Transform.

Surface water quality is a major environmental concern as it is a main source of fresh water for human consumption governed by the complex anthropogenic activities and natural processes[158].

The changes in the water body ecosystem especially in rivers and lakes have a major impact



(a) Winter (b) Spring

(c) Summer (d) Autumn

Figure 9.6. LSE for Algae: Winter, Spring, Summer, Autumn from 2015 to 2017

on human welfare and the aquatic environment [159]. Continuous deposition of solid waste material and contaminants in the water of lakes and rivers has become a global health concern as these are a major source of water supply for human consumption and domestic purposes [160].

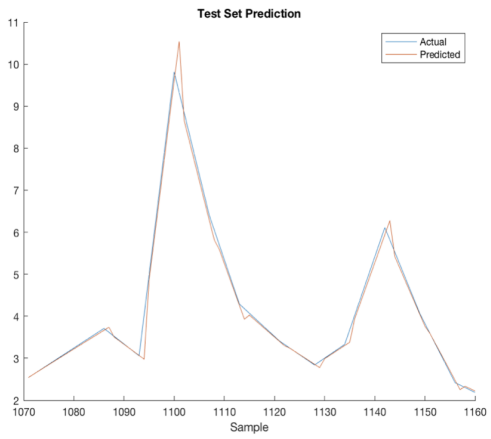
In Brazil due to a large demand for electricity, there is a growing number of hydroelectric power plants in the water reservoirs to generate energy. Monitoring the reservoirs is important for the decision-making process.

The parameters selected for analysis are strongly related to the water quality monitoring, such Chlorophyll-a, TSS and transparency, PTotal, FeTotal, Turbidity, Transparency, Fe2, PO4, Fe3. For example Clorophyll-a levels are responsible for the photosynthetic process and reflect the phytoplankton biomass in the ecosystem, while water transparency allows for estimations regarding the depth of the photic zone, i.e. the vertical depth of sunlight penetration in the water column, which indicates the level of photosynthetic activity in the reservoir [161] and, TSS can, among other effects, cause damage to aquatic life by settling at the bottom of the reservoir, destroying organisms and retaining bacteria and organic waste by promoting anaerobic decomposition [161], [29]. In this sense, these characteristics are routinely used to measure the trophic status of lakes.

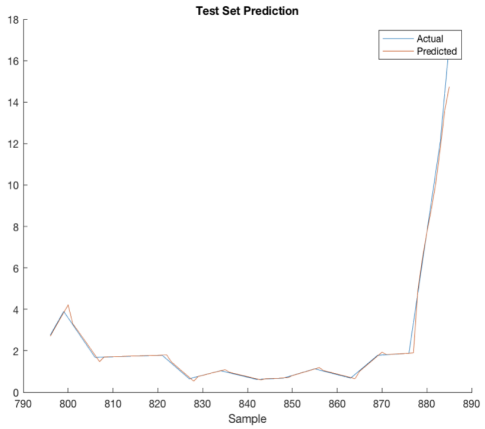
These parameters, as mentioned previously, require extensive fieldwork in collecting the

Table 9.2. Turbidity: Prediction results for two years

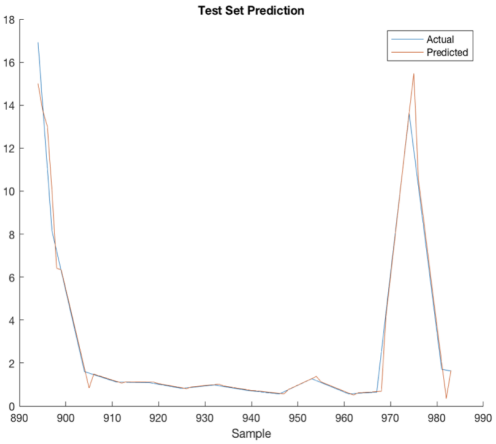
	NMSE - TURBIDITY			
Prediction Model	Summer	Autumn	Winter	Spring
LSE training	0.00688	0.00818	0.00874	0.00720
LSE test	**0.01688**	0.05864	0.01361	**0.00701**
ANFIS training	0.00277	0.00483	0.00521	0.00445
ANFIS test	0.02610	**0.04985**	**0.01318**	0.01143
RBF training	0.00228	0.00362	0.00392	0.00446
RBF test	0.13449	0.10193	0.01686	0.14174



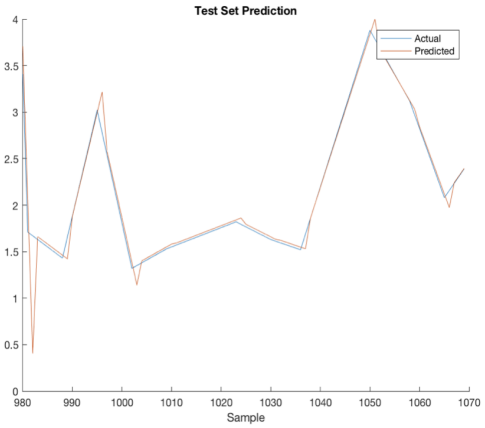
(a) Winter



(b) Spring



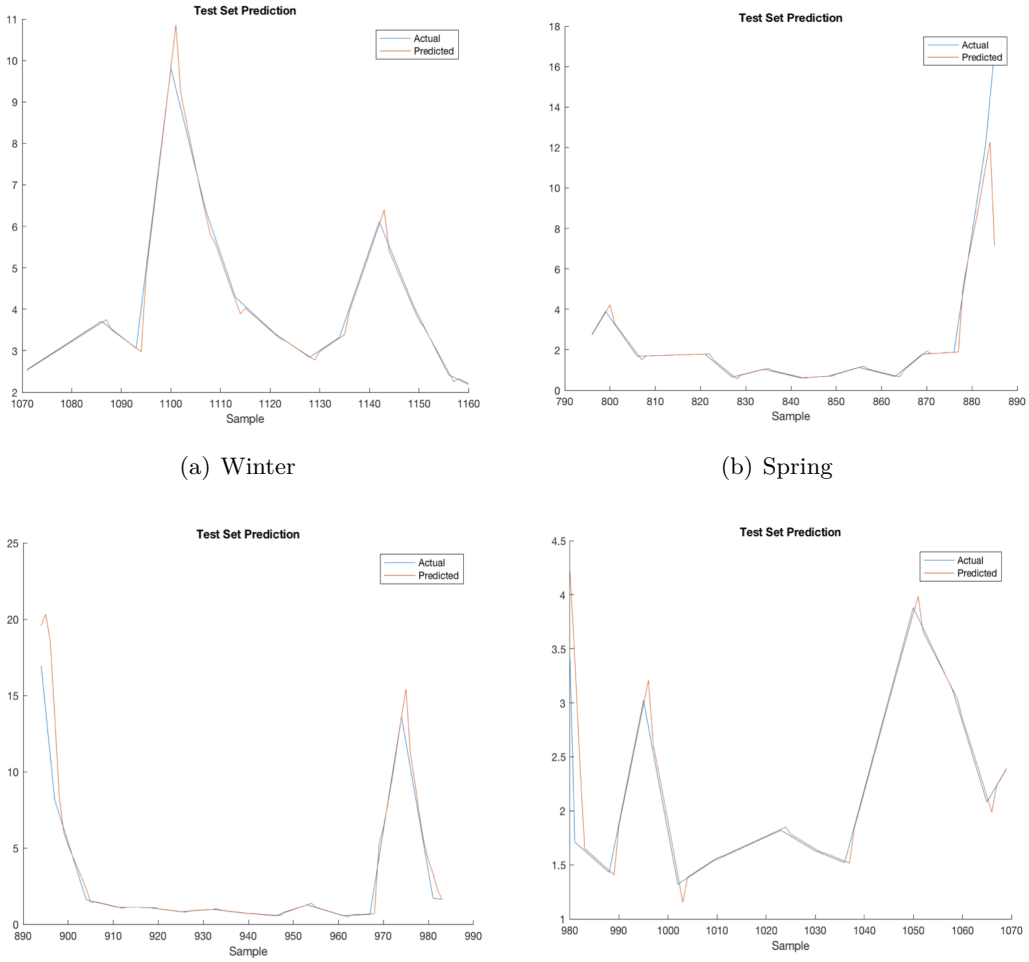
(c) Summer



(d) Autumn

Figure 9.7. ANFIS for Turbidity: Winter, Spring, Summer, Autumn from 2015 to 2017

samples, which are then analyzed by trained personnel in laboratory conditions. Usually, the applied measurement techniques for some of these parameters are sophisticated, such as determination of clorophyll-a by spectrophotometry after extraction with hot ethanol and



(a) Winter (b) Spring

(c) Summer (d) Autumn

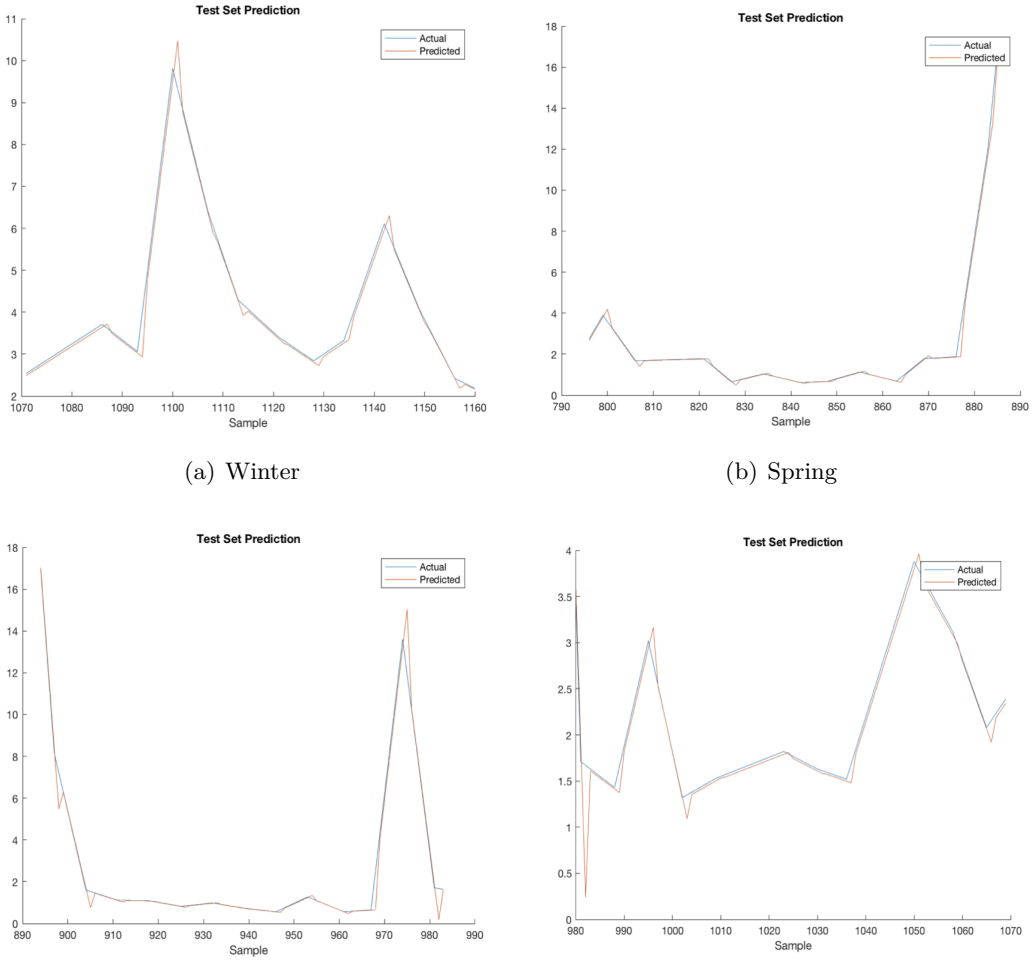
Figure 9.8. RBF for Turbidity: Winter, Spring, Summer, Autumn from 2015 to 2017

turbidity by nephelometric method, then scattered angle from the beam directed at the water sample [162].

In this research we propose the use of remote sensing techniques for the monitoring and prediction of water quality parameters. Nowadays, with advanced hardware and software, processing and analyzing satellite images have become easier and less costly, alongside the improvement of spatial and temporal resolution of the satellite imagery, in addition to algorithm optimization, has led to the acknowledgment of the incredible potential of this technique to monitor and, consequently, improve, water quality [163].

Global estimations of chlorophyll-a concentrations for example, have been investigated by several marine-specific satellite missions for the visible wavelength range, including SeaWiFS, MERIS and the planned Ocean and Land Colour Imager [164], [165].

Trophic classifications have been obtained based on these methods, by applying IRS-1C satellite imaging in monitoring chlorophyll-a content, for example, in specific water bodies, such as lakes, while anthropogenic impacts have also been demonstrated using these techniques, such as in the very recent study conducted in the Amazon that analyzed 40 years of Landsat-MSS/TM/OLI images to monitor the impacts of mining activities near the Tapajós River and observed that TSS concentrations were directly related to the mining activities [166].



(a) Winter (b) Spring

(c) Summer (d) Autumn

Figure 9.9. LSE for Turbidity: Winter, Spring, Summer, Autumn from 2015 to 2017

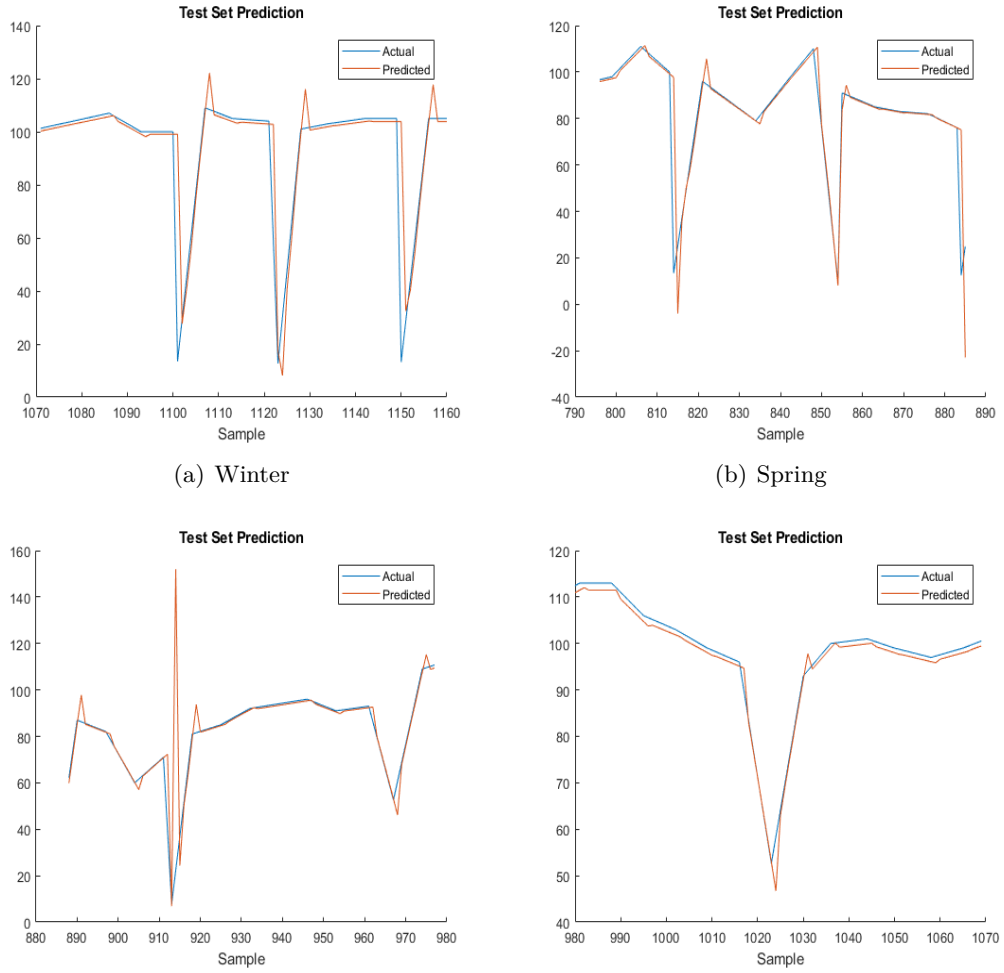
The coupling of remote sensing with other statistical and computation techniques has, increasingly, been applied and proven valid in monitoring water quality. The alternative method described in the present study with wavelet transformation of the remote sensing images and analysis by ANN, thus, would contribute positively to laboratory analysis in the determination of the mentioned parameters, which is obviously advantageous for numerous reasons. Thus, the application of predictive methods as the one proposed in the present study are of interest since they provide alternative methods to obtain good accuracy and are less expensive than the procedures presently used in environmental monitoring programs conducted in these ecosystems.

Accordingly, the proposal established herein is to work with past data from previous years, classified per hydrological cycle, to predict the values for future hydrological cycles, hence, the fact that we used data from previous hydrological cycles, from 2007 to 2014 in the Tucurui reservoir, and data from the 2014 hydrological cycle for validation. The ANN training results of the wavelet transformed remote sensing images for each sampling stations and the four stages of the well- defined hydrological cycle were considered adequate, with low mean square errors (MSE) displayed in Table 9.4 for the following sampling stations: C1, C2, M1, M3, MBB, MIP, ML, MP, MP in the Tucurui Reservoir - Amazon region.

Following ANN training, the methodology was validated by comparing laboratory- obtained

Table 9.3. Suspended Solids: Prediction results for two years

NMSE - SUSPENDED SOLIDS				
Prediction Model	Summer	Autumn	Winter	Spring
LSE training	0.31519	0.32596	0.33320	0.28567
LSE test	**0.15483**	0.04095	0.46958	0.48409
ANFIS training	0.08527	0.09662	0.09554	0.09411
ANFIS test	0.66929	**0.01965**	**0.36611**	0.33910
RBF training	0.13151	0.14208	0.13034	0.12952
RBF test	0.16748	0.04208	0.36657	**0.31466**



(a) Winter

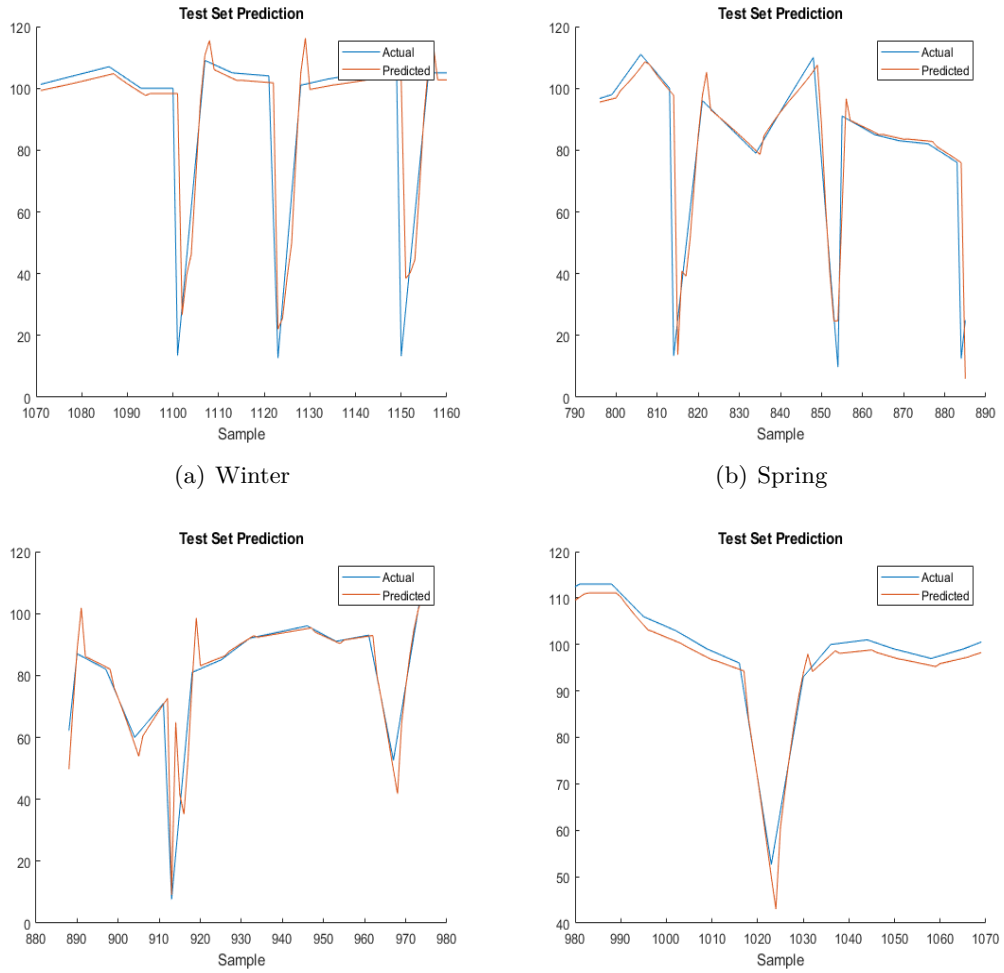
(b) Spring

(c) Summer

(d) Autumn

Figure 9.10. ANFIS for Suspended Solids: Winter, Spring, Summer, Autumn from 2015 to 2017

results for chlorophyll-a levels, TSS and transparency with the results obtained by the proposed methodology for the year 2014. Table 9.5 displays the approximation errors between the values obtained in the laboratory and those calculated by the proposed methodology per sampling



(a) Winter (b) Spring

(c) Summer (d) Autumn

Figure 9.11. RBF for Suspended Solids: Winter, Spring, Summer, Autumn from 2015 to 2017

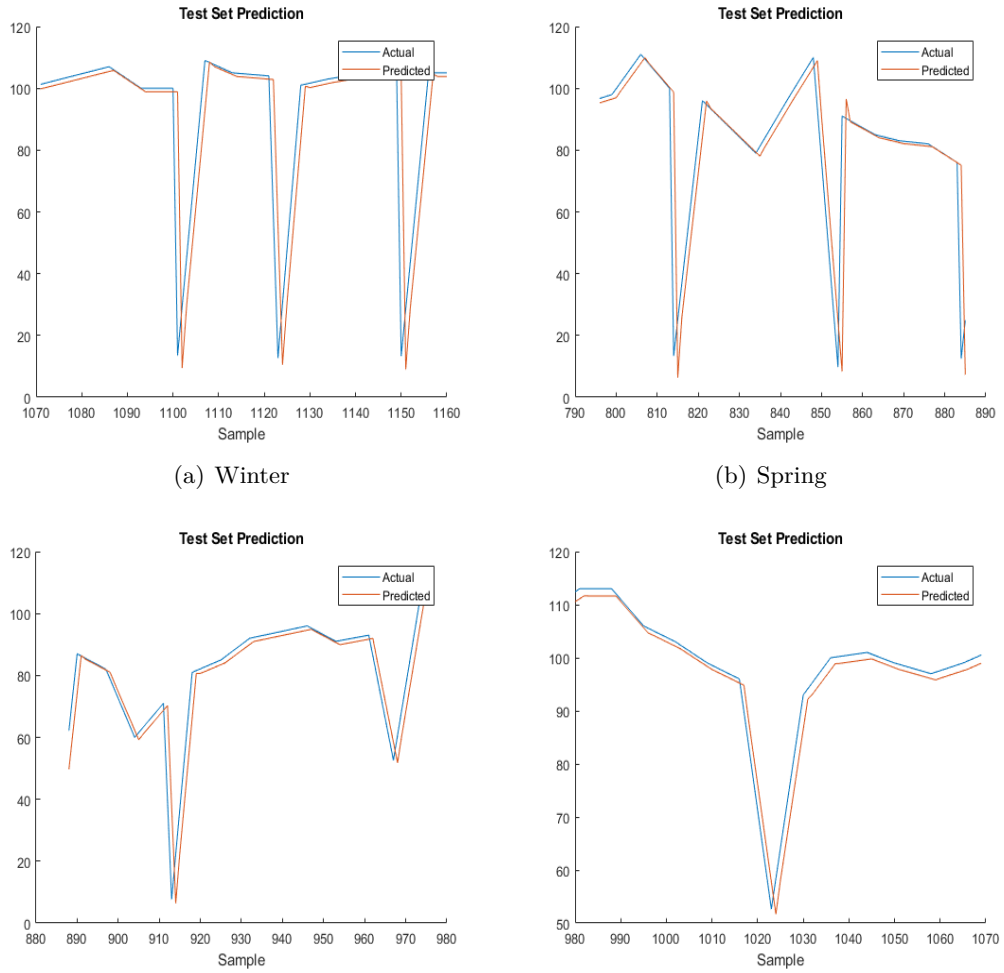
station and hydrological cycle for the year 2014. The approximation errors were calculated as follows 9.6 :

$$Err(i) = V\_lab(i) - Y(i) \tag{9.6}$$

Where:
V_lab (i): Is the value obtained in the laboratory for this parameter
Y(i): Is the output variable obtained by the ANN.

The figure 9.13 shows the validation results for 2014 for Chlorophylla in the C1, C2, MBB, MR sampling stations, with the laboratory results being the “observed values” and those obtained by wavelet transformation of the remote sensing images and subsequent analysis by ANN, proposed herein, being the “estimated values”, regarding chlorophyll-a, total suspended solids and transparency. The X-axes of figures represent the hydrological cycles (1, 2, 3 and 4, respectively, the full, emptying, dry and filling stages). The Y-axis represents the quantitative value of the analyzed parameter in a given hydrological cycle.

The figure 9.15 shows the validation results for 2014 for TSS in the M1, M3, MIP, MJV, ML, MP sampling stations.



(a) Winter (b) Spring

(c) Summer (d) Autumn

Figure 9.12. LSE for Suspended Solids: Winter, Spring, Summer, Autumn from 2015 to 2017

The figure 9.14 shows the validation results for 2014 for Transparency in the M1, M3, MBB, ML, MP, MR sampling stations.

As previously stated, we propose remote monitoring of the reservoir using LandSat7 to predict the physico-chemical parameters of the water in seven (7) points as can be seen in fig 7.2. The following sampling stations C1, C2, M1, M3, MBB, MJV, MIP were selected.

The collection of water in these points were done periodically 4 times a year corresponding to the hydrological cycle: full, empyting, dry and filling. The different cycles are a consequence of the differences in rainfall during the year and they influence a lot the water monitoring. Hydroelectric plants currently constitute an indispensable component for supplying renewable energy. However, this reservoir, as the others in the Amazon region, has had several impacts on the ecosystem: loss of biodiversity of terrestrial and aquatic fauna and flora, high concentration of organic matter in the water bottom due to vegetation inundation, chemical changes in the water downstream, large volume of anoxic water in the reservoir and downstream, loss of water quality (low dissolved oxygen, high conductivity, low pH, high content of dissolved and particulate, organic matter), high concentration of aquatic macrophytes and reduction of fisheries downstream.

The reservoir has impacted also on the human settlements in the area by weakening physical infrastructure, decreasing efficiency in land use, creating resettlement problems and influencing

Table 9.4. Mean Square Errors (MSE) in the ANN training conducted in the present study

		MSE			
Parameters	Station	Full	Emptying	Dry	Filling
C1	Transparency	2.73×10^{-24}	2.34×10^{-23}	4.60×10^{-22}	4.65×10^{-23}
	TSS	1.29×10^{-24}	3.57×10^{-23}	3.72×10^{-22}	9.18×10^{-25}
	Chlorophylla	$3.49 \times^{1} 0{-}22$	4.31×10^{-23}	3.93×10^{-24}	2.4×10^{-23}
C2	Transparency	1.43×10^{-22}	3.51×10^{-21}	4.08×10^{-22}	1.58×10^{-21}
	TSS	1.88×10^{-23}	8.44×10^{-23}	1.68×10^{-23}	6.52×10^{-23}
	Chlorophylla	8.69×10^{-24}	5.24×10^{-24}	1.84×10^{-23}	2.25×10^{-23}
M1	Transparency	1.14×10^{-22}	9.09×10^{-23}	1.47×10^{-22}	1.90×10^{-22}
	TSS	9.89×10^{-23}	9.95×10^{-22}	2.57×10^{-21}	1.18×10^{-22}
	Chlorophylla	1.63×10^{-24}	9.49×10^{-23}	2.25×10^{-24}	4.74×10^{-24}
M3	Transparency	1.45×10^{-21}	1.21×10^{-23}	1.32×10^{-22}	6.02×10^{-23}
	TSS	2.32×10^{-24}	9.94×10^{-23}	2.14×10^{-23}	8.16×10^{-25}
	Chlorophylla	7.00×10^{-24}	3.44×10^{-23}	3.05×10^{-24}	2.71×10^{-22}
MBB	Transparency	4.14×10^{-22}	5.59×10^{-22}	$1.20 \times 10{-}23$	5.00×10^{-23}
	TSS	3.22×10^{-24}	9.94×10^{-24}	6.05×10^{-24}	2.55×10^{-22}
	Chlorophylla	2.92×10^{-25}	1.22×10^{-22}	1.33×10^{-23}	3.65×10^{-24}
MIP	Transparency	2.41×10^{-21}	1.25×10^{-22}	1.77×10^{-22}	2.27×10^{-21}
	TSS	5.97×10^{-25}	6.10×10^{-22}	6.74×10^{-23}	6.18×10^{-22}
	Chlorophylla	9.75×10^{-24}	8.09×10^{-23}	3.83×10^{-24}	4.03×10^{-24}
ML	Transparency	3.65×10^{-23}	3.36×10^{-21}	1.19×10^{-22}	1.99×10^{-22}
	TSS	1.27×10^{-22}	2.24×10^{-22}	2.15×10^{-22}	2.06×10^{-22}
	Chlorophylla	2.47×10^{-24}	5.43×10^{-24}	1.99×10^{-24}	1.48×10^{-24}
MP	Transparency	3.86×10^{-22}	2.51×10^{-22}	2.64×10^{-22}	3.08×10^{-22}
	TSS	3.02×10^{-24}	2.04×10^{-22}	8.44×10^{-25}	1.85×10^{-23}
	Chlorophylla	1.33×10^{-22}	7.57×10^{-24}	7.61×10^{-24}	4.28×10^{-24}
MR	Transparency	2.76×10^{-23}	2.34×10^{-24}	4.20×10^{-22}	4.09×10^{-21}
	TSS	1.35×10^{-22}	4.92×10^{-23}	3.99×10^{-23}	8.55×10^{-24}
	Chlorophylla	3.16×10^{-24}	8.84×10^{-24}	2.47×10^{-23}	2.10×10^{-22}

mining operations on the reservoir itself [167]. For these reasons, the area is of highly interest and water monitoring is one of the important things that have to be in place to ensure the sustainability of the reservoir.

The ANN training results for the monitoring of this important reservoir by sampling stations and hydrological cycle water are shown in the table 9.9, the values are considered low, mean square errors (MSE) for neural network training as follows for Chlorophyll-a, Transparency and Total Suspended Solids.
The Relative Errors were calculated by the equation (9.7) and showed in 9.11 :

$$RelativeError(E_r) = \frac{|X_e - X_o|}{X_o} \tag{9.7}$$

where:
Xe: Estimated Value
Xo: Observed Value

Figures 9.13, 9.14, 9.15 shows the validation results for 2014 for C1, C2, M1, M3, MBB, MJV and MIP sample stations, with the laboratory results being the “observed values” and those

Table 9.5. Approximation errors of the proposed method for 2014 per sampling station, evaluated parameter and hydrological cycle.

FULL				
Station	Variable	Error	Lab	Ann
C1	Transparency	0.2814	1.4	1.6814
	TSS	2.6139	1.33	3.9439
	Chlorophylla	5.9924	2.38	8.7924
C2	Transparency	0.2192	2.8	3.0192
	TSS	0.0823	1.4	1.3177
	Chlorophylla	0.8967	4.76	5.6567
M1	Transparency	0.5216	1.8	2.3216
	TSS	1.1455	1.6	2.7455
	Chlorophylla	6.0581	2.62	8.6781
M3	Transparency	0.328	1.2	0.0872
	TSS	3.9784	5	8.9784
	Chlorophylla	0.8267	5.95	6.7767
MBB	Transparency	0.1356	2.2	2.3356
	TSS	0.2081	2	1.7919
	Chlorophylla	1.4084	3.81	2.4016
MIP	Transparency	0.0627	0.5	0.5627
	TSS	16.4403	24	40.4403
	Chlorophylla	1.92	5.78	3.86
ML	Transparency	0.4832	1.4	1.8832
	TSS	0.1003	4.1	4.2003
	Chlorophylla	4.678	5.155	0.477
MP	Transparency	0.097	2.7	2.603
	TSS	1.6706	1.6	3.2706
	Chlorophylla	0.3362	6.19	6.5262
MR	Transparency	0.2057	2.5666	2.3609
	TSS	1.7169	2.8	1.0831
	Chlorophylla	1.2658	8.57	7.3042

obtained by wavelet transformation of the remote sensing images and subsequent analysis by ANN, proposed herein, being the "estimated values", regarding chlorophyll-a, total suspended solids and transparency. The X-axes of figures represent the hydrological cycles (1, 2, 3 and 4, respectively, the full, emptying, dry and filling stages). The Y-axis represents the quantitative value of the analyzed parameter in a given hydrological cycle.

As the results given in table 9.10 and 9.11 show the errors between expected and observed values are quite low. In particular, the bests results were obtained during the dry season cycle 3 (September-October-November) for Total Suspended Solids. This period corresponds to the less cloudy period in the region. This facilitate the analysis based on satellite images and allows to obtain most accurate results.

In general, the errors are considered low and neural network showed good results and can aid in the evaluation of physico-chemical parameters, which in turn allows the identification of possible anthropogenic impacts, being relevant in environmental management and in political decision-making processes.

The results estimated by the method proposed in the present study, when compared with those observed in the laboratory proved extremely close to each other, demonstrating adequate

Table 9.6. Approximation errors of the proposed method for 2014 per sampling station, evaluated parameter and hydrological cycle.

EMPTYING				
Station	Variable	Error	Lab	Ann
C1	Transparency	0.2792	2.8	3.0792
	TSS	1.0023	1.4	2.4023
	Chlorophylla	0.6872	3.81	4.4972
C2	Transparency	0.5621	3.45	2.8879
	TSS	0.8448	1	1.8448
	Chlorophylla	2.5848	9.045	6.4602
M1	Transparency	0.0009	4.2	4.2009
	TSS	0.2141	0.4	0.1859
	Chlorophylla	2.0696	1.9	3.9696
M3	Transparency	0.4106	3	2.5894
	TSS	0.125	1.2	1.075
	Chlorophylla	0.7833	5.47	6.2533
MBB	Transparency	0.6815	4.3	3.6185
	TSS	0.9171	1	1.9171
	Chlorophylla	0.1938	3.57	3.7638
MIP	Transparency	0.3955	1.9	1.5045
	TSS	5.3545	1.8	5.3345
	Chlorophylla	0.29309	3.09	6.0209
ML	Transparency	0.5521	3.2	2.6479
	TSS	0.1969	1.4	1.5969
	Chlorophylla	1.8254	5.47	7.2954
MP	Transparency	0.9233	2.9	1.9767
	TSS	0.7272	1.5	2.2272
	Chlorophylla	0.1882	6.305	6.1168
MR	Transparency	0.5146	3.066	2.552
	TSS	0.3851	1.3	1.6851
	Chlorophylla	2.478	5.83	8.308

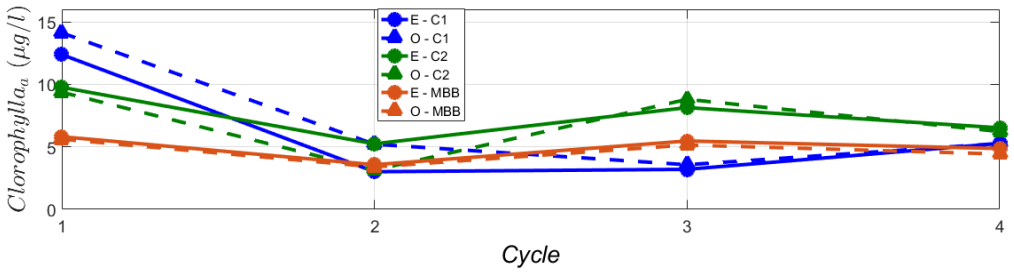


Figure 9.13. Predicting Chlorophylla-a Levels in hydroelectric power plant reservoir by wavelet transformation of spectral bands for sample station: C1 - Caraipé 1, C2 - Caraipé 2, MBB - Breu Branco; E = Estimated; O = Observed

efficiency of the proposed method. As the relationship between the estimated data and in-situ data is reliable, it would therefore be possible to estimate and chart the water quality of this reservoir back to the first satellite images obtained for the area, which would be interesting in order to track anthropogenic impacts rates throughout the years, from 1984 until today.

Table 9.7. Approximation errors of the proposed method for 2014 per sampling station, evaluated parameter and hydrological cycle.

DRY				
Station	Variable	Error	Lab	Ann
C1	Transparency	0.4642	3	2.5358
	TSS	0.0085	1.2	1.1915
	Chlorophylla	2.7352	4.76	2.0248
C2	Transparency	0.61	1.6	2.21
	TSS	1.823	4	3.177
	Chlorophylla	4.3263	12.14	7.8137
M1	Transparency	0.3704	4.2666	4.637
	TSS	0.637	0.8	0.163
	Chlorophylla	1.9406	1.9	0.0406
M3	Transparency	0.3566	2.7	2.3434
	TSS	0.7951	0.8	1.5951
	Chlorophylla	1.1384	4.76	5.8984
MBB	Transparency	1.0036	3.6	4.6036
	TSS	0.4702	1.8	1.3298
	Chlorophylla	0.3239	5.47	5.1461
MIP	Transparency	1.2495	1.1	2.3495
	TSS	1.788	3.4	1.612
	Chlorophylla	5.1001	3.81	1.2901
ML	Transparency	0.0592	1.8	1.8592
	TSS	0.7169	3.5	2.7831
	Chlorophylla	3.4678	9.64	113.1078
MP	Transparency	0.2012	1.9	2.1012
	TSS	3.5004	4.6	8.1004
	Chlorophylla	1.1722	12.61	11.4378
MR	Transparency	0.8117	2.3333	1.5216
	TSS	0.9255	2.4	1.4745
	Chlorophylla	1.0767	4.52	3.4433

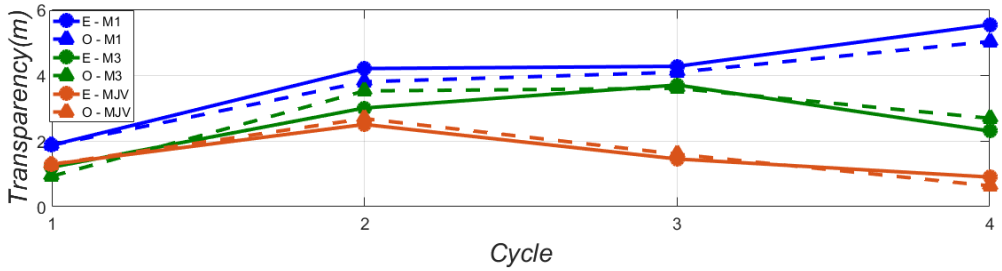


Figure 9.14. Predicting Transparency in hydroelectric power plant reservoir by wavelet transformation of spectral bands for sample station: M1 - Upstrem 1, M3 - Upstrem 3 , MJV - Jacunda Velho; E = Estimated; O = Observed

Potential sources of error when using remote sensing images can be due to varying atmospheric conditions. These include bad weather conditions, such as clouds, which affect the amount of incoming solar radiation reaching the water surface and the fraction of light leaving the water surface that reaches the satellite sensor[141]. In fact, clouds have been shown to interfere

Table 9.8. Approximation errors of the proposed method for 2014 per sampling station, evaluated parameter and hydrological cycle.

FILLING				
Station	Variable	Error	Lab	Ann
C1	Transparency	0.4097	2.8	2.3903
	TSS	2.5042	1.8	4.3042
	Chlorophylla	1.1741	6.66	5.4859
C2	Transparency	0.7294	1.95	1.2206
	TSS	3.5761	3.8	7.3761
	Chlorophylla	14.9382	13.39	28.3282
M1	Transparency	2.3255	1.5333	3.8558
	TSS	0.0816	2.3	2.3816
	Chlorophylla	0.1989	2.62	2.8189
M3	Transparency	0.0499	1.3	1.2501
	TSS	1.1787	5.2	4.0213
	Chlorophylla	1.6353	6.9	8.5353
MBB	Transparency	0.3015	1.6	1.9105
	TSS	0.0929	2.2	2.2929
	Chlorophylla	0.08846	2.86	3.7446
MIP	Transparency	0.2286	0.8	0.5714
	TSS	12.6127	24.75	37.3627
	Chlorophylla	1.8653	5.71	3.8447
ML	Transparency	0.0279	0.8	0.8279
	TSS	1.0154	9.2	8.1846
	Chlorophylla	3.2846	5.95	9.2346
MP	Transparency	0.1187	0.95	1.0687
	TSS	2.2222	8.45	6.2278
	Chlorophylla	4.0724	10.83	14.9024
MR	Transparency	0.2366	2.5	2.2634
	TSS	0.6972	2.5	1.8029
	Chlorophylla	2.5624	7.14	9.7024

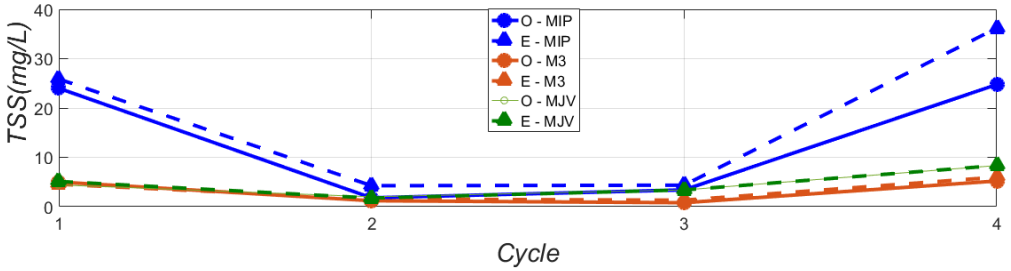


Figure 9.15. Predicting Total Suspended Solids in hydroelectric power plant reservoir by wavelet transformation of spectral bands for sample station: MIP - Ipixuna , M3 - Upstrem 3 , MJV - Jacunda Velho; E = Estimated; O = Observed

significantly in the monitoring of physico-chemical water characteristics and novel methods are being developed to surpass this issue, such as a model based on the ratio of green and blue band reflectance considering the bio-optical property of chlorophyll-a combined to ordinary kriging, which model was highly capable of predicting the chlorophyll-a concentration in regions covered

Table 9.9. Mean Square Errors (MSE) in the ANN training conducted in the present study

MSE					
Parameters	Station	Full	Emptying	Dry	Filling
CHLOROPHYLLa	C1	1.25×10^{-23}	3.70×10^{-07}	1.35×10^{-22}	6.68×10^{-24}
	C2	4.44×10^{-22}	3.97×10^{-24}	5.61×10^{-22}	2.73×10^{-07}
	MBB	1.27×10^{-22}	5.37×10^{-09}	1.33×10^{-23}	2.04×10^{-22}
TRANSPARENCY	MJV	3.78×10^{-06}	3.72×10^{-08}	2.43×10^{-22}	1.11×10^{-22}
	M1	5.39×10^{-08}	2.78×10^{-07}	5.83×10^{-07}	1.88×10^{-10}
	M3	1.14×10^{-08}	9.67×10^{-09}	3.15×10^{-22}	4.67×10^{-21}
TSS	M3	4.63×10^{-07}	2.43×10^{-08}	9.88×10^{-23}	3.83×10^{-23}
	MJV	3.10×10^{-05}	7.70×10^{-09}	1.31×10^{-21}	3.80×10^{-23}
	MIP	4.89×10^{-24}	5.06×10^{-09}	3.01×10^{-04}	1.66×10^{-22}

Table 9.10. Approximation errors for 2014 per sampling station, evaluated parameter and hydrological cycle.

MSE Validation by Cycle					
Parameters	Station	Full	Emptying	Dry	Filling
CHLOROPHYLLa	C1	1.1593	17.1529	0.2679	5.4940
	C2	0.1555	4.3905	0.4366	0.0889
	MBB	0.3346	0.0070	0.8778	0.1564
TRANSPARENCY	MJV	0.1789	0.0736	0.0156	0.0828
	M1	0.0010	0.4436	0.4957	0.0272
	M3	0.0966	0.2470	0.2272	0.2318
TSS	M3	0.0444	1.1343	0.0006	1.6069
	MJV	0.0106	1.1881	0.0471	0.2483
	MIP	1.1363	0.0135	11.3284	0.7024

by clouds and thus, effective in monitoring water quality in tropical shallow waters[168].

In addition, sun glint, the specular reflection of light from water surfaces towards the satellite sensor, is also a serious confounding factor for remote sensing of water column properties and benthos [165] .

Retrieval of information such as chlorophyll-a content, benthic features or bathymetry in these cases requires both high measurement sensitivity and a robust algorithm that can separate and remove the effect of glint [165]. Sun glint correction methods have been previously applied, such as use of wind speed and direction, application of neural networks [169], scaling depending on the brightest and darkest points of the images [170], use of the depth of the 760 nm oxygen absorption band [171] or methods using predictions of reflection based on water surface models (applied in operational ocean color data processing spatial resolutions of 100–1,000 m) and those that use in-scene information with the assumption of no near-infrared wavelength radiance leaving the water surface (applied to high resolution images of coral reefs and other shallow waters with pixel sizes of around 1–10 m) [34], [172].

Novel methods such as applying neural networks to separate the effects of the aerosol scattering, water-leaving radiance and glint are also being developed[165].

Table 9.11. Relative Error (Er) per sampling station, evaluated parameter and hydrological cycle.

Relative Error					
Parameters	Station	Full	Emptying	Dry	Filling
CHLOROPHYLLa	C1	0.0800	0.5791	0.1392	0.3128
	C2	0.0421	0.6663	0.0751	0.0479
	MBB	0.1106	0.0229	0.2067	0.0886
TRANSPARENCY	MJV	0.4824	0.0979	0.0793	0.2423
	M1	0.0180	0.1885	0.1416	0.0307
	M3	0.3495	0.1421	0.1479	0.1731
TSS	M3	0.0440	0.4702	0.0303	0.1960
	MJV	0.0229	1.1978	0.0682	0.0647
	MIP	0.0425	0.0174	0.4975	0.0328

In the present study, slightly different values between estimated and in-situ data were observed in some cases, attributed to either the presence of clouds or sun glint in the remote sensing images, corroborating previous studies [141].

This should, thus, be taken into account when applying this type of methodology to environmental monitoring of reservoirs, even though differences were very slight, and corrections to these issues were done using the Dark Object Subtraction (DOS) method, providing better results in this research.

Even though atmospheric conditions influenced estimated and in-situ data in the present study, the proposed method is still shown to be reliable in comparison to other studies that also suffered atmospheric interferences, such as a study that evaluated the performance of images obtained from the sensor Operational Land Imager (OLI) onboard the Landsat-8 satellite in determining Chl-a concentrations and aiming at classifying a Brazilian tropical reservoir in the state of São Paulo with regard to trophic status, that showed reasonable results but impaired performance due to atmospheric influences [39].

On the other hand, the use of Support Vector Machines in conjunction with a Radial Base function using TM/Landsat-5 time-series images showed more interesting results, albeit applied to observe differences in Land Use and Land Cover in a Hydroelectric system located between the states of Rio de Janeiro and São Paulo, also in Brazil [40].

Another study estimated the coloured dissolved organic matter absorption coefficient at 440 nm at another Brazilian reservoir in São Paulo using operational land imager (OLI)/Landsat-8 images and created distribution maps based on the adjusted algorithm. The authors of that study were able to adequately analyze this inland water body by the proposed method, but state that future research is needed to confirm if this model can be used in other reservoirs [38].

We intend to overcome these issues in this research and apply the methodology in other water reservoirs, in order to validate the proposal in a broader and concise way. This is also the case in the present study, where future studies are still required since, even though the data was very satisfactory regarding eutrophic categorization, there is no information of the application of this model in other ecosystems, which shall be the basis for future studies in order to demonstrate wider applicability.

9.4 Final Considerations for Tucurui Reservoir.

The present study in the Tucurui reservoir, Brazil, demonstrated the application of wavelet Neural Network for estimating Chlorophyll-a levels, Transparency and Total Suspended Solids using concentration of the water samples collected in the Amazon region. Satellite images, landsat7, ETM + sensor, band 1 (TM1), band 2 (TM2), band 3 (TM3) and band 4 (TM4) were used to train the ANN by hydrological cycle (full, emptying, dry, fillying) for 4 years.

A time series was analyzed and parameters were predicted with good acuracy considering the well-defined seasonal characteristics of the region. The method resulted in satisfactory approximations of laboratory results regarding the same water samples.

The neural network demonstrated good results between observed and estimated after Atmospheric corrections in satellites images. The ANNs showed in the results are useful to estimate these concentrations using remote sensing and wavelet transform. Therefore, the techniques proposed and applied in the present study are very importante since they can aid in evaluating important physico-chemical parameters, which, in turn, allows for identification of possible anthropogenic impacts, being relevant in environmental management and policy decision-making processes.

It is clear that the method proposed in this study is specific for reservoirs such as Tucuruí, which have well defined hydrological cycles, although nothing prevents this method from being used in the monitoring of waters of other reservoirs, which, in turn, must consider the seasonal differences from each region.

This research contributes the evaluation of different methods accuracy in estimating of physico-chemical parameters, from multispectral satellite images. Future studies will be conducted regarding atmospheric interferences and corrections and different ecosystem characteristics. The method proposed allows the identification of possible anthropogenic impacts, being relevant in environmental management, in order to mitigate these impacts and attempt the recovery of degraded water bodies.

Chapter 10

Conclusion and Future Research

Contents

10.1 Conclusions

We used ecological and optical approaches in spatial and temporal contexts to map physico-chemical parameters in two important water reservoirs: the Tucuruí hydroelectric reservoir (THR), Brazil and Cefni, UK.

Our main objective was to investigate whether the combination between water limnology, Neural Networks and satellite imagery is a suitable approach for monitoring water quality variables.

The overall conclusion is that the neural networks and wavelet transform are adequate to estimate physico-chemical concentrations. Thus, this algorithm may be applied to this study area, using Landsats and Sentinel satellites images.

The neural networks were trained using satellite images of spectral band 1, band 2, band 3 and band 4 that has sensitivity to the presence of particulate material in water bodies. The proposed method estimated good results, confirming that this approach is useful to analise water quality variables.

The WANN received as input a pixel vector of the satellite images. A sampling station was initially chosen for analysis and the images were cut at the water collection points, in size 32x32 pixels, corresponding to an array containing 1024 pixels. Then, the Wavelet Transform was applied, with only one level of decomposition, resulting in a matrix array of 16x16 pixels for each of the following three components: Horizontal (H), vertical (V) and diagonal (D).

The conversion of the arrays to the H, V and D components to their respective column-matrices was performed, and subsequently a concatenation of the three arrays (each containing 256 pixels) was executed, generating a vector with 768 column size (256 x 3). The images from water sampling area were decomposed via wavelet into its three wavelet components.

Tests were conducted considering the image representations isolated for each wavelet component, with satisfactory results. However, when the input data of the three wavelet components was considered, the approximations were even better, which motivated the choice of this arrangement in the proposed solution.

The results obtained showed that method for predicting of these parameters has sensitivity to the presence of suspended particulate matter with satisfactory between the estimated values

and the expected values.

The factorial analysis also confirmed the correlation between the parameters in the first depth of the Secchi disc, PTotal, FeTotal, Turbidity, Transparency, Chlorophill, Fe2, Total Suspended Solids, PO4, Fe2 and Fe3 were related to Factor 1, since these present great reflectance and good absorption of energy by the sensors of the satellites.

In general, the errors are considered low and Neural Network Wavelet showed good results and can aid in the evaluation of physico-chemical parameters, which in turn allows the identification of possible anthropogenic impacts in water reservoirs, being relevant in environmental management and in political decision-making processes.

The bests results were obtained during the dry season. The dry period has few clouds in the region, allowing a good analysis of satellite images. As the main drawback of this study was using a reduced number of satellite images, we recommend that other parameters studies in this area ensure that their ecological functioning is carefully considered when attempting to map occurrence using limited satellite imagery.

Moreover, the goal of this study was not to quantify the parameters directly. Instead, this approach meant estimating the values based on the water collections made on site using satellite images.

It is clear that the proposed method is specific to the Tucuruí and Cefni, although nothing prevents this method to be used in monitoring waters from other reservoirs, as we intend to do in the another lakes, which in turn should result in less expenditures on environmental monitoring processes, providing more efficiency in the monitoring of water quality, thus contributing to the processes of environmental monitoring.

The techniques proposed and applied in the present study can aid in evaluating important physico-chemical parameters, which, in turn, allows for identification of possible anthropogenic impacts, being relevant in environmental management and policy decision-making processes, in order to mitigate these impacts and attempt the recovery of degraded water bodies.

10.2 Future Directions of the Research

As a future work, further study on the bio-optical properties of Amazonian reservoir waters would be beneficial to local water management in order to understand the water quality issues in these areas.

To propose other computational techniques to evaluate water quality parameters, as well as to evaluate other physicochemical parameters, verifying and estimating the concentrations of these elements in surface reflectance.

Another possibility of future work in the Tucurui reservoir is the development of a software that can allocate the collection stations according to the representativity of the concentrations of the physico-chemical parameters by means of satellite images and clustering techniques, addressing techniques of the statistical and computational intelligence.

The proposed method can be applied to other ecosystems, other reservoirs of water, allowing to predict the water quality parameters, showing a worldwide relevance. But it is necessary, first and foremost, to investigate atmospheric interferences and analyze the different characteristics of the ecosystems, as it was developed in this research.

Appendix A

Water collection

Algal samples for identification are collected by a full column vertical net haul using a 57µm mesh net in the case of boat samples, and are therefore concentrated, and "dipped" by a Patalas sampler for fixed location samples and are therefore not concentrated. The determinand named "NET=0 DIP=1 ALGAE SAMPLE ?DIP" is used to differentiate between the two states. As the determinand name suggests, a net sample = 0, and a dip sample = 1. All samples are analysed at a centralised laboratory. All the results including the field readings are entered into the laboratory system, and are then archived. If the result value is null, a dip sample may be assumed.

Zooplankton samples are vertical net hauls (VNHs), using a 140 µm net. A collected sample consists of two VNHs added together, thus the biomass has to be calculated appropriately using the area of the net mouth (0.5 m diameter) to determine biomass per m2.

Zooplankton results are given as dry mass. This is measured after the sample has been separated into copepods and cyclopids, thus there are two values, ZOO BIOMASS CLADOCERA and ZOO BIOMASS COPEPODS. In some basins an additional techniques was used to analyse for zooplankton retained on a 1mm mesh net. In all cases the areal value should be multiplied by the full depth of the reservoir, in order to derive the biomass per m3. That value varies with level.

All West London reservoirs are filled by pumping from the river Thames via tunnels. Some reservoirs are artificially mixed during spring and summer by forcing the inlet water through jets arranged at angles in relation to the horizontal. These are typically 0 degree, 22.5 degree and 45 degree inlet, although not all reservoirs have them. The most used is the 0 degree. Mixing is used to prevent thermal stratification and to transfer algae out of the photic zone into the deeper parts of the reservoir.

In the event of low river flows or in cases of pollution, pumping may not be possible. In these circumstances, some reservoirs have submerged aeration devices can be used to prevent thermal stratification.

Data results table

Tag Number	Date	Time	Determinand code	Determinand name	Units	Qu.	Value
000006114930	18/08/2015	11:30:00	5745	ALGAE ID + ACFOR 1ST COMMON	Coded	#	62.4
000006114930	18/08/2015	11:30:00	5746	ALGAE ID + ACFOR 2ND COMMON	Coded	#	753.3
000006114930	18/08/2015	11:30:00	5747	ALGAE ID + ACFOR 3RD COMMON	Coded	#	792.3
000006114930	18/08/2015	11:30:00	5748	ALGAE ID + ACFOR 4TH COMMON	Coded	#	563.2
000006114930	18/08/2015	11:30:00	5749	ALGAE ID + ACFOR 5TH COMMON	Coded	#	112.2
000006114930	18/08/2015	11:30:00	5750	ALGAE ID + ACFOR 6TH COMMON	Coded	#	391.1
000006114930	18/08/2015	11:30:00	5751	ALGAE ID + ACFOR 7TH COMMON	Coded	#	1301.1
000006114930	18/08/2015	11:30:00	5752	ALGAE ID + ACFOR 8TH COMMON	Coded	#	91.1
000006114930	18/08/2015	11:30:00	5753	ALGAE ID + ACFOR 9TH COMMON	Coded	#	0
000006114930	18/08/2015	11:30:00	5754	ALGAE ID + ACFOR 10TH COMMON	Coded	#	0

Figure A.1. Data Results Table

Algal ID codes table

Result number	Genus	Type	ACFOR	ACFOR letter	1º Filter impact	Blocker?
62.4	Anabaena	Blue-Green	Common	C		
753.3	Melosira	Filamentous Diatom	Frequent	F	Moderate	!
792.3	Microcystis	Blue-Green	Frequent	F		
563.2	Fragilaria	Diatom	Occasional	O	Minor	**
112.2	Aphanizomenon	Blue-Green	Occasional	O		
391.1	Cosmarium	Green	Rare	R		
1301.1	Volvox	Colonial Green	Rare	R		
91.1	Ankyra	Unicellular green	Rare	R		
No alga						
No alga						

Figure A.2. Algal ID codes table

Bibliography

[1] A. Araghi, M. Mousavi-Baygi, J. Adamowski, C. Martinez, and M. van der Ploeg, "Forecasting soil temperature based on surface air temperature using a wavelet artificial neural network," *Meteorological Applications*, vol. 24, no. 4, pp. 603–611, 2017.

[2] S. Kumar, M. K. Tiwari, C. Chatterjee, and A. Mishra, "Reservoir inflow forecasting using ensemble models based on neural networks, wavelet analysis and bootstrap method," *Water resources management*, vol. 29, no. 13, pp. 4863–4883, 2015.

[3] E. Kaneko, H. Aoki, and M. Tsukada, "Image-based path radiance estimation guided by physical model," in *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*. IEEE, 2016, pp. 6942–6945.

[4] T. Blakey, A. Melesse, M. C. Sukop, G. Tachiev, D. Whitman, and F. Miralles-Wilhelm, "Developing benthic class specific, chlorophyll-a retrieving algorithms for optically-shallow water using seawifs," *Sensors*, vol. 16, no. 10, p. 1749, 2016.

[5] G. Lantzanakis, Z. Mitraka, and N. Chrysoulakis, "Comparison of physically and image based atmospheric correction methods for sentinel-2 satellite imagery," in *Perspectives on Atmospheric Sciences*. Springer, 2017, pp. 255–261.

[6] L. Shi, Z. Mao, P. Chen, S. Han, F. Gong, and Q. Zhu, "Comparison and evaluation of atmospheric correction algorithms of quac, dos and flaash for hico hyperspectral imagery," in *SPIE Remote Sensing*. International Society for Optics and Photonics, 2016, pp. 999 917–999 917.

[7] A. Kabata-Pendias, *Trace elements in soils and plants*. CRC press, 2010.

[8] O. E. Atobatele and O. A. Ugwumba, "Seasonal variation in the physicochemistry of a small tropical reservoir (aiba reservoir, iwo, osun, nigeria)," *African Journal of Biotechnology*, vol. 7, no. 12, 2008.

[9] M. W. Song, P. Huang, F. Li, H. Zhang, K. Z. Xie, X. H. Wang, and G. X. He, "Water quality of a tributary of the pearl river, the beijiang, southern china: implications from multivariate statistical analyses," *Environmental monitoring and assessment*, vol. 172, no. 1, pp. 589–603, 2011.

[10] F. Yenilmez, F. Keskin, and A. Aksoy, "Water quality trend analysis in eymir lake, ankara," *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 36, no. 5, pp. 135–140, 2011.

[11] R. Nazir, M. Khan, M. Masab, H. U. Rehman, N. U. Rauf, S. Shahab, N. Ameer, M. Sajed, M. Ullah, M. Rafeeq *et al.*, "Accumulation of heavy metals (ni, cu, cd, cr, pb, zn, fe) in the soil, water and plants and analysis of physico-chemical parameters of soil and water collected from tanda dam kohat," *Journal of Pharmaceutical Sciences and Research*, vol. 7, no. 3, pp. 89–97, 2015.

[12] J. Delgado, J. L. Cereijo-Arango, D. García-Morrondo, C. Cillero-Castro, A. Muñoz-Ibáñez, and R. Juncosa-Rivera, "Physical and chemical limnology of the abegondo-cecebre reservoir, a coruña, nw spain," in *EGU General Assembly Conference Abstracts*, vol. 18, 2016, p. 14201.

[13] B. Gülcü-Gür and S. Tekin-Özan, "The investigation of heavy metal levels in water and sediment from işikli lake (turkey) in relation to seasons and physico-chemical parameters," 2017.

[14] Y. Cavalcante, R. Hauser-Davis, A. Saraiva, I. Brandão, T. Oliveira, and A. Silveira, "Metal and physico-chemical variations at a hydroelectric reservoir analyzed by multivariate analyses and artificial neural networks: Environmental management and policy/decision-making tools," *Science of the Total Environment*, vol. 442, pp. 509–514, 2013.

[15] F. A. Prado, S. Athayde, J. Mossa, S. Bohlman, F. Leite, and A. Oliver-Smith, "How much is enough? an integrated examination of energy security, economic growth and climate change related to hydropower expansion in brazil," *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 1132–1136, 2016.

[16] "Aneel.agência nacional de energia elétrica banco de dados de geração 2014," http://<www.aneel.gov.br/, acessed in 15/08/2018.

[17] J. C. Derisio, *Introdução ao control de poluição ambiental.* Oficina de Textos, 2016.

[18] J. G. Tundisi and T. M. Tundisi, *Limnologia.* Oficina de textos, 2016.

[19] F. A. Khan and A. A. Ansari, "Eutrophication: an ecological vision," *The botanical review*, vol. 71, no. 4, pp. 449–482, 2005.

[20] A. O. Olaniran, K. Naicker, and B. Pillay, "Assessment of physico-chemical qualities and heavy metal concentrations of umgeni and umdloti rivers in durban, south africa," *Environmental monitoring and assessment*, vol. 186, no. 4, pp. 2629–2639, 2014.

[21] S. D. P. Yadav, K. Mishra, N. K. Chaudhary, and P. Mishra, "Assessing physico-chemical parameters of potable water in dhankuta municipality of nepal," *Science*, vol. 3, no. 2, pp. 17–21, 2015.

[22] F. Soltani, R. Kerachian, and E. Shirangi, "Developing operating rules for reservoirs considering the water quality issues: Application of anfis-based surrogate models," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6639–6645, 2010.

[23] A. A. Najah, A. El-Shafie, O. A. Karim, and O. Jaafar, "Water quality prediction model utilizing integrated wavelet-anfis model with cross-validation," *Neural Computing and Applications*, vol. 21, no. 5, pp. 833–841, 2012.

[24] A. Najah, A. El-Shafie, O. A. Karim, and A. H. El-Shafie, "Performance of anfis versus mlp-nn dissolved oxygen prediction models in water quality monitoring," *Environmental Science and Pollution Research*, vol. 21, no. 3, pp. 1658–1670, 2014.

[25] M. Bonansea, M. C. Rodriguez, L. Pinotti, and S. Ferrero, "Using multi-temporal landsat imagery and linear mixed models for assessing water quality parameters in río tercero reservoir (argentina)," *Remote Sensing of Environment*, vol. 158, pp. 28–41, 2015.

[26] M. Bonansea and R. Fernandez, "Remote sensing of suspended solids concentration in a reservoir with frequent wildland fires on its watershed," *Water Science and Technology*, vol. 67, no. 1, pp. 217–223, 2012.

[27] N. Karakaya, F. Evrendilek, G. Aslan, K. Gungor, and D. Karakas, "Monitoring of lake water quality along with trophic gradient using landsat data," 2012.

[28] B. R. Neto, R. Hauser-Davis, T. Lobato, A. Saraiva, I. Brandão, T. Oliveira, and A. Silveira, "Estimating physicochemical parameters and metal concentrations in hydroelectric reservoirs by virtual sensors: A case study in the amazon region," *Computer Science and Engineering*, vol. 4, no. 2, pp. 43–53, 2014.

[29] E. Tebbs, J. Remedios, and D. Harper, "Remote sensing of chlorophyll-a as a measure of cyanobacterial biomass in lake bogoria, a hypertrophic, saline–alkaline, flamingo lake, using landsat etm+," *Remote Sensing of Environment*, vol. 135, pp. 92–106, 2013.

[30] P. L. Brezonik, L. G. Olmanson, J. C. Finlay, and M. E. Bauer, "Factors affecting the measurement of cdom by remote sensing of optically complex inland waters," *Remote Sensing of Environment*, vol. 157, pp. 199–215, 2015.

[31] L. Cai, D. Tang, and C. Li, "An investigation of spatial variation of suspended sediment concentration induced by a bay bridge based on landsat tm and oli data," *Advances in Space Research*, vol. 56, no. 2, pp. 293–303, 2015.

[32] X. Giam and J. D. Olden, "Quantifying variable importance in a multimodel inference framework," *Methods in Ecology and Evolution*, 2015.

[33] L. G. Olmanson, P. L. Brezonik, J. C. Finlay, and M. E. Bauer, "Comparison of landsat 8 and landsat 7 for regional measurements of cdom and water clarity in lakes," *Remote Sensing of Environment*, 2016.

[34] B. A. Franz, S. W. Bailey, N. Kuring, and P. J. Werdell, "Ocean color measurements with the operational land imager on landsat-8: implementation and evaluation in seadas," *Journal of Applied Remote Sensing*, vol. 9, no. 1, pp. 096 070–096 070, 2015.

[35] W. Zhu and Q. Yu, "Inversion of chromophoric dissolved organic matter from eo-1 hyperion imagery for turbid estuarine and coastal waters," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 6, pp. 3286–3298, 2013.

[36] T. Dube, O. Mutanga, K. Seutloali, S. Adelabu, and C. Shoko, "Water quality monitoring in sub-saharan african lakes: a review of remote sensing applications," *African Journal of Aquatic Science*, vol. 40, no. 1, pp. 1–7, 2015.

[37] T. Rajaee and S. Khani, "Comment on "performance of anfis versus mlp-nn dissolved oxygen prediction models in water quality monitoring a. najah & a. el-shafie & oa karim & amr h. el-shafie. environ sci pollut res (2014) 21: 1658-1670"," *Environmental Science and Pollution Research*, vol. 23, no. 1, pp. 938–940, 2016.

[38] E. Alcântara, N. Bernardo, F. Watanabe, T. Rodrigues, L. Rotta, A. Carmo, M. Shimabukuro, S. Gonçalves, and N. Imai, "Estimating the cdom absorption coefficient in tropical inland waters using oli/landsat-8 images," *Remote Sensing Letters*, vol. 7, no. 7, pp. 661–670, 2016.

[39] F. S. Y. Watanabe, E. Alcântara, T. W. P. Rodrigues, N. N. Imai, C. C. F. Barbosa, and L. H. d. S. Rotta, "Estimation of chlorophyll-a concentration and the trophic state of the barra bonita hydroelectric reservoir using oli/landsat-8 images," *International journal of environmental research and public health*, vol. 12, no. 9, pp. 10 391–10 417, 2015.

[40] S. Martins, N. Bernardo, I. Ogashawara, and E. Alcantara, "Support vector machine algorithm optimal parameterization for change detection mapping in funil hydroelectric reservoir (rio de janeiro state, brazil)," *Modeling Earth Systems and Environment*, vol. 2, no. 3, p. 138, 2016.

[41] C. Dona, J. M. Sanchez, V. Caselles, J. A. Domínguez, and A. Camacho, "Empirical relationships for monitoring water quality of lakes and reservoirs through multispectral images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 5, pp. 1632–1641, 2014.

[42] I. Chorus and J. Bartram, *Toxic cyanobacteria in water: a guide to their public health consequences, monitoring and management.* CRC Press, 1999.

[43] W. W. Carmichael, "Health effects of toxin-producing cyanobacteria:"the cyanohabs"," *Human and ecological risk assessment: An International Journal*, vol. 7, no. 5, pp. 1393–1407, 2001.

[44] M. Y. Cheung, S. Liang, and J. Lee, "Toxin-producing cyanobacteria in freshwater: A review of the problems, impact on drinking water safety, and efforts for protecting public health," *Journal of Microbiology*, vol. 51, no. 1, pp. 1–10, 2013.

[45] T. W. Davis, D. L. Berry, G. L. Boyer, and C. J. Gobler, "The effects of temperature and nutrients on the growth and dynamics of toxic and non-toxic strains of microcystis during cyanobacteria blooms," *Harmful algae*, vol. 8, no. 5, pp. 715–725, 2009.

[46] C. Imrie, S. Durucan, and A. Korre, "River flow prediction using artificial neural networks: generalisation beyond the calibration range," *Journal of hydrology*, vol. 233, no. 1-4, pp. 138–153, 2000.

[47] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks:: The state of the art," *International journal of forecasting*, vol. 14, no. 1, pp. 35–62, 1998.

[48] A. C. Teodoro, F. Veloso-Gomes, and H. Goncalves, "Retrieving tsm concentration from multispectral satellite data by multiple regression and artificial neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 5, pp. 1342–1350, 2007.

[49] D. Tomassi, D. Milone, and J. D. Nelson, "Wavelet shrinkage using adaptive structured sparsity constraints," *Signal Processing*, vol. 106, pp. 73–87, 2015.

[50] J. M. B. Protázio, "Spatial pattern analysis applied to plant ecology," in *Center for Tropical Marine Ecology. Zentrum für Marine Tropenökologie (ZMT).* University of Bremen, 2007.

[51] S. Mallat, *A wavelet tour of signal processing: the sparse way.* Academic press, 2008.

[52] T. MathWorks, "Matlab documentation," *Disponıvel em*, 2016.

[53] N. P. Majozi, M. S. Salama, S. Bernard, D. M. Harper, and M. G. Habte, "Remote sensing of euphotic depth in shallow tropical inland waters of lake naivasha using meris data," *Remote sensing of environment*, vol. 148, pp. 178–189, 2014.

[54] S. Dlamini, I. Nhapi, W. Gumindoga, T. Nhiwatiwa, and T. Dube, "Assessing the feasibility of integrating remote sensing and in-situ measurements in monitoring water quality status of lake chivero, zimbabwe," *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 93, pp. 2–11, 2016.

[55] N. Pahlevan, S. Sarkar, B. Franz, S. Balasubramanian, and J. He, "Sentinel-2 multispectral instrument (msi) data processing for aquatic science applications: Demonstrations and validations," *Remote Sensing of Environment*, vol. 201, pp. 47–56, 2017.

[56] M. Chawira, T. Dube, and W. Gumindoga, "Remote sensing based water quality monitoring in chivero and manyame lakes of zimbabwe," *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 66, pp. 38–44, 2013.

[57] J. Ndungu, B. C. Monger, D. C. Augustijn, S. J. Hulscher, N. Kitaka, and J. M. Mathooko, "Evaluation of spatio-temporal variations in chlorophyll-a in lake naivasha, kenya: remote-sensing approach," *International journal of remote sensing*, vol. 34, no. 22, pp. 8142–8155, 2013.

[58] I. L. de Sousa Brandão, C. M. Mannaerts, W. Verhoef, A. C. F. Saraiva, R. S. Paiva, and E. V. da Silva, "Using synergy between water limnology and satellite imagery to identify algal blooms extent in a brazilian amazonian reservoir," *Sustainability*, vol. 9, no. 12, p. 2194, 2017.

[59] D. P. Roy, J. Li, H. K. Zhang, L. Yan, H. Huang, and Z. Li, "Examination of sentinel-2a multi-spectral instrument (msi) reflectance anisotropy and the suitability of a general method to normalize msi reflectance to nadir brdf adjusted reflectance," *Remote Sensing of Environment*, vol. 199, pp. 25–38, 2017.

[60] T.-M. Perivolioti, A. Mouratidis, D. Bobori, G. Doxani, and D. Terzopoulos, "Monitoring water quality parameters of lake koronia by means of long time-series multispectral satellite images," *AUC GEOGRAPHICA*, vol. 52, no. 2, pp. 176–188, 2017.

[61] N. Pettorelli, W. F. Laurance, T. G. O'Brien, M. Wegmann, H. Nagendra, and W. Turner, "Satellite remote sensing for applied ecologists: opportunities and challenges," *Journal of Applied Ecology*, vol. 51, no. 4, pp. 839–848, 2014.

[62] M. J. McCarthy, K. E. Colna, M. M. El-Mezayen, A. E. Laureano-Rosario, P. Méndez-Lázaro, D. B. Otis, G. Toro-Farmer, M. Vega-Rodriguez, and F. E. Muller-Karger, "Satellite remote sensing for coastal management: A review of successful applications," *Environmental management*, vol. 60, no. 2, pp. 323–339, 2017.

[63] N. Johnson, C. Revenga, and J. Echeverria, "Managing water for people and nature," *Science*, vol. 292, no. 5519, pp. 1071–1072, 2001.

[64] J. Meriluoto, L. Spoof, and G. A. Codd, *Handbook of cyanobacterial monitoring and cyanotoxin analysis*. John Wiley & Sons, 2017.

[65] T. Malthus and D. George, "Airborne remote sensing of macrophytes in cefni reservoir, anglesey, uk," *Aquatic Botany*, vol. 58, no. 3-4, pp. 317–332, 1997.

[66] C. Gibson, "The algicidal effect of copper on a green and a blue-green alga and some ecological implications," *Journal of applied ecology*, pp. 513–518, 1972.

[67] H. N. Silva, G. Laneve, A. Rosato, and M. Panella, "Retrieving chlorophyll-a levels, transparency and tss concentration from multispectral satellite data by using artificial neural networks," in *Progress in Electromagnetics Research Symposium-Fall (PIERS-FALL), 2017*. IEEE, 2017, pp. 2876–2883.

[68] G. Osório, J. Matias, and J. Catalão, "Short-term wind power forecasting using adaptive neuro-fuzzy inference system combined with evolutionary particle swarm optimization, wavelet transform and mutual information," *Renewable Energy*, vol. 75, pp. 301–307, 2015.

[69] A. Rosato, R. Altilio, R. Araneo, and M. Panella, "Prediction in photovoltaic power by neural networks," *Energies*, vol. 10, no. 7, p. 1003, 2017.

[70] M. Nazeer and J. E. Nichol, "Selection of atmospheric correction method and estimation of chlorophyll-a (chl-a) in coastal waters of hong kong," in *Earth Observation and Remote Sensing Applications (EORSA), 2014 3rd International Workshop on*. IEEE, 2014, pp. 374–378.

[71] N. Bernardo, F. Watanabe, T. Rodrigues, and E. Alcântara, "Atmospheric correction issues for retrieving total suspended matter concentrations in inland waters using oli/landsat-8 image," *Advances in Space Research*, 2017.

[72] L. Congedo, "Semi-automatic classification plugin for qgis," 2013.

[73] L. Congedo, L. Sallustio, M. Munafò, M. Ottaviano, D. Tonti, and M. Marchetti, "Copernicus high-resolution layers for land cover classification in italy," *Journal of Maps*, vol. 12, no. 5, pp. 1195–1205, 2016.

[74] G. Chander, B. L. Markham, and D. L. Helder, "Summary of current radiometric calibration coefficients for landsat mss, tm, etm+, and eo-1 ali sensors," *Remote sensing of environment*, vol. 113, no. 5, pp. 893–903, 2009.

[75] S. N. Goward, P. E. Davis, D. Fleming, L. Miller, and J. R. Townshend, "Empirical comparison of landsat 7 and ikonos multispectral measurements for selected earth observation system (eos) validation sites," *Remote Sensing of Environment*, vol. 88, no. 1, pp. 80–99, 2003.

[76] R. R. Irish, "Landsat 7 science data users handbook," *National Aeronautics and Space Administration, Report*, vol. 2000, pp. 415–430, 2000.

[77] T. G. Jones, "Climate change and dissolved organic carbon: Impacts on drinking water supplies," Ph.D. dissertation, University of Wales, Bangor, 2006.

[78] Z. Zhu, S. Wang, and C. E. Woodcock, "Improvement and expansion of the fmask algorithm: Cloud, cloud shadow, and snow detection for landsats 4–7, 8, and sentinel 2 images," *Remote Sensing of Environment*, vol. 159, pp. 269–277, 2015.

[79] T. Kutser, D. C. Pierson, K. Y. Kallio, A. Reinart, and S. Sobek, "Mapping lake cdom by satellite remote sensing," *Remote Sensing of Environment*, vol. 94, no. 4, pp. 535–540, 2005.

[80] T. Perkins, S. M. Adler-Golden, M. W. Matthew, A. Berk, L. S. Bernstein, J. Lee, and M. Fox, "Speed and accuracy improvements in flaash atmospheric correction of hyperspectral imagery," *Optical Engineering*, vol. 51, no. 11, p. 111707, 2012.

[81] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[82] Y. Meyer, "Ondelettes, fonctions splines et analyses graduées," *Lectures given at the University of Torino, Italy*, vol. 9, 1986.

[83] R. Ansari, C. Guillemot, and J. Kaiser, "Wavelet construction using lagrange halfband filters," *IEEE Transactions on Circuits and Systems*, vol. 38, no. 9, pp. 1116–1118, 1991.

[84] P. Dutilleux, "An implementation of the "algorithme à trous" to compute the wavelet transform," in *Wavelets*. Springer, 1990, pp. 298–304.

[85] E. Hernández and G. Weiss, *A first course on wavelets*. CRC press, 1996.

[86] T.-W. Kim and J. B. Valdés, "Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks," *Journal of Hydrologic Engineering*, vol. 8, no. 6, pp. 319–328, 2003.

[87] J. Adamowski and K. Sun, "Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds," *Journal of Hydrology*, vol. 390, no. 1-2, pp. 85–91, 2010.

[88] M. Shoaib, A. Y. Shamseldin, and B. W. Melville, "Comparative study of different wavelet based neural network models for rainfall–runoff modeling," *Journal of hydrology*, vol. 515, pp. 47–58, 2014.

[89] S. Djerbouai and D. Souag-Gamane, "Drought forecasting using neural networks, wavelet neural networks, and stochastic models: case of the algerois basin in north algeria," *Water resources management*, vol. 30, no. 7, pp. 2445–2464, 2016.

[90] R. Maheswaran and R. Khosa, "Comparative study of different wavelets for hydrologic forecasting," *Computers & geosciences*, vol. 46, pp. 284–295, 2012.

[91] R. S. Stanković and B. J. Falkowski, "The haar wavelet transform: its status and achievements," *Computers & Electrical Engineering*, vol. 29, no. 1, pp. 25–44, 2003.

[92] L. Congedo, "Semi-automatic classification plugin documentation," *Release*, vol. 4, no. 0.1, p. 29, 2016.

[93] A. Elzwayie, A. El-shafie, Z. M. Yaseen, H. A. Afan, and M. F. Allawi, "Rbfnn-based model for heavy metal prediction for different climatic and pollution conditions," *Neural Computing and Applications*, pp. 1–13, 2016.

[94] J.-W. Choi, J.-H. Han, C.-S. Park, D.-G. Ko, H.-I. Kang, J. Y. Kim, Y.-J. Yun, H.-H. Kwon, and K.-G. An, "Nutrients and sestonic chlorophyll dynamics in asian lotic ecosystems and ecological stream health in relation to land-use patterns and water chemistry," *Ecological Engineering*, vol. 79, pp. 15–31, 2015.

[95] V. Kumar, A. Sharma, A. Chawla, R. Bhardwaj, and A. K. Thukral, "Water quality assessment of river beas, india, using multivariate and remote sensing techniques," *Environmental monitoring and assessment*, vol. 188, no. 3, p. 137, 2016.

[96] V. Markogianni, E. Dimitriou, and I. Karaouzas, "Water quality monitoring and assessment of an urban mediterranean lake facilitated by remote sensing applications," *Environmental monitoring and assessment*, vol. 186, no. 8, pp. 5009–5026, 2014.

[97] N. T. B. Phuong, V. P. D. Tri, N. B. Duy, N. C. Nghiem *et al.*, "Remote sensing for monitoring surface water quality in the vietnamese mekong delta: The application for estimating chemical oxygen demand in river reaches in binh dai, ben tre," *VIETNAM JOURNAL OF EARTH SCIENCES*, vol. 39, no. 3, pp. 256–268, 2017.

[98] L. Chen, M. Jamal, C. Tan, and B. Alabbadi, "A study of applying genetic algorithm to predict reservoir water quality," *International Journal of Modeling and Optimization*, vol. 7, no. 2, p. 98, 2017.

[99] F. Halal, P. Pedrocca, T. Hirose, A.-M. Cretu, and M. B. Zaremba, "Remote-sensing based adaptive path planning for an aquatic platform to monitor water quality," in *Robotic and Sensors Environments (ROSE), 2014 IEEE International Symposium on.* IEEE, 2014, pp. 43–48.

[100] E. Sharaf El Din, Y. Zhang, and A. Suliman, "Mapping concentrations of surface water quality parameters using a novel remote sensing and artificial intelligence framework," *International Journal of Remote Sensing*, vol. 38, no. 4, pp. 1023–1042, 2017.

[101] C. Doña, N.-B. Chang, V. Caselles, J. M. Sánchez, A. Camacho, J. Delegido, and B. W. Vannah, "Integrated satellite data fusion and mining for monitoring lake water quality status of the albufera de valencia in spain," *Journal of environmental management*, vol. 151, pp. 416–426, 2015.

[102] N.-B. Chang, S. Imen, and B. Vannah, "Remote sensing for monitoring surface water quality status and ecosystem state in relation to the nutrient cycle: A 40-year perspective," *Critical Reviews in Environmental Science and Technology*, vol. 45, no. 2, pp. 101–166, 2015.

[103] Z. Yang, M. Reiter, and N. Munyei, "Estimation of chlorophyll-a concentrations in diverse water bodies using ratio-based nir/red indices," *Remote Sensing Applications: Society and Environment*, vol. 6, pp. 52–58, 2017.

[104] Y. Liu, M. A. Islam, and J. Gao, "Quantification of shallow water quality parameters by means of remote sensing," *Progress in physical geography*, vol. 27, no. 1, pp. 24–43, 2003.

[105] H. Jiao, Y. Zha, J. Gao, Y. Li, Y. Wei, and J. Huang, "Estimation of chlorophyll-a concentration in lake tai, china using in situ hyperspectral data," *International Journal of Remote Sensing*, vol. 27, no. 19, pp. 4267–4276, 2006.

[106] J. P. Cannizzaro and K. L. Carder, "Estimating chlorophyll a concentrations from remote-sensing reflectance in optically shallow waters," *Remote Sensing of Environment*, vol. 101, no. 1, pp. 13–24, 2006.

[107] N. Usali and M. H. Ismail, "Use of remote sensing and gis in monitoring water quality," *Journal of sustainable development*, vol. 3, no. 3, p. 228, 2010.

[108] X.-G. Xing, D.-Z. Zhao, Y.-G. Liu, J.-H. Yang, P. Xiu, and L. Wang, "An overview of remote sensing of chlorophyll fluorescence," *Ocean Science Journal*, vol. 42, no. 1, pp. 49–59, 2007.

[109] E. Robert, M. Grippa, L. Kergoat, S. Pinet, L. Gal, G. Cochonneau, and J.-M. Martinez, "Monitoring water turbidity and surface suspended sediment concentration of the bagre reservoir (burkina faso) using modis and field reflectance data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 52, pp. 243–251, 2016.

[110] R. W. Gentry, J. McCarthy, A. Layton, L. D. McKay, D. Williams, S. R. Koirala, and G. S. Sayler, "Loading at or near base flow in a mixed-use watershed," *Journal of environmental quality*, vol. 35, no. 6, pp. 2244–2249, 2006.

[111] S. J. Giovannoni, J. C. Thrash, and B. Temperton, "Implications of streamlining theory for microbial ecology," *The ISME journal*, vol. 8, no. 8, p. 1553, 2014.

[112] P. Galès and B. Baleux, "Influence of the drainage basin input on a pathogenic bacteria (salmonella) contamination of a mediterranean lagoon (the thau lagoon-france) and the survival of this bacteria in brackish water," *Water Science and Technology*, vol. 25, no. 12, pp. 105–114, 1992.

[113] J. van der Bliek, P. G. McCornick, and J. Clarke, *On target for people and planet: setting and achieving water-related sustainable development goals*. IWMI, 2014.

[114] P. S. Chavez Jr, "An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data," *Remote sensing of environment*, vol. 24, no. 3, pp. 459–479, 1988.

[115] P. S. Chavez *et al.*, "Image-based atmospheric corrections-revisited and improved," *Photogrammetric engineering and remote sensing*, vol. 62, no. 9, pp. 1025–1035, 1996.

[116] J. A. Sobrino, J. C. Jimenez-Munoz, and L. Paolini, "Land surface temperature retrieval from landsat tm 5," *Remote Sensing of environment*, vol. 90, no. 4, pp. 434–440, 2004.

[117] D. G. Hadjimitsis, C. Clayton, and V. Hope, "An assessment of the effectiveness of atmospheric correction algorithms through the remote sensing of some reservoirs," *International Journal of Remote Sensing*, vol. 25, no. 18, pp. 3651–3674, 2004.

[118] I. M. McCullough, C. S. Loftin, and S. A. Sader, "Combining lake and watershed characteristics with landsat tm data for remote estimation of regional lake clarity," *Remote Sensing of Environment*, vol. 123, pp. 109–115, 2012.

[119] N. M. Nasrabadi, "Pattern recognition and machine learning," *Journal of electronic imaging*, vol. 16, no. 4, p. 049901, 2007.

[120] F. Nie, J. Yuan, and H. Huang, "Optimal mean robust principal component analysis," in *International Conference on Machine Learning*, 2014, pp. 1062–1070.

[121] R. A. Johnson, D. W. Wichern *et al.*, *Applied multivariate statistical analysis*. Prentice-Hall New Jersey, 2014, vol. 4.

[122] K. Bhaskar and S. Singh, "Awnn-assisted wind power forecasting using feed-forward neural network," *IEEE transactions on sustainable energy*, vol. 3, no. 2, pp. 306–315, 2012.

[123] A. Yadav and K. Sahu, "Wind forecasting using artificial neural networks: A survey and taxonomy," *International Journal of Research In Science & Engineering*, vol. 3, 2017.

[124] M. Panella, L. Liparulo, and A. Proietti, "A higher-order fuzzy neural network for modeling financial time series," in *Neural Networks (IJCNN), 2014 International Joint Conference on*. IEEE, 2014, pp. 3066–3073.

[125] C.-N. Ko and C.-M. Lee, "Short-term load forecasting using svr (support vector regression)-based radial basis function neural network with dual extended kalman filter," *Energy*, vol. 49, pp. 413–422, 2013.

[126] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[127] J.-S. Jang, "Anfis: adaptive-network-based fuzzy inference system," *IEEE transactions on systems, man, and cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.

[128] D. Lowe, "Multi-variable functional interpolation and adaptive networks," *Complex Systems*, vol. 2, pp. 321–355.

[129] G. P. Zhang, "Time series forecasting using a hybrid arima and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.

[130] M. Panella, A. Rizzi, and G. Martinelli, "Refining accuracy of environmental data prediction by mog neural networks," *Neurocomputing*, vol. 55, no. 3-4, pp. 521–549, 2003.

[131] A. Mellit and A. M. Pavan, "A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected pv plant at trieste, italy," *Solar Energy*, vol. 84, no. 5, pp. 807–821, 2010.

[132] B. d. Mérona, A. Juras, G. Mendes Dos Santos, and I. Cintra, "Os peixes ea pesca no baixo rio tocantins: vinte anos depois da uhe tucuruí," 2010.

[133] T. Lobato, R. Hauser-Davis, T. Oliveira, A. Silveira, H. Silva, M. Tavares, and A. Saraiva, "Construction of a novel water quality index and quality indicator for reservoir water quality evaluation: A case study in the amazon region," *Journal of hydrology*, vol. 522, pp. 674–683, 2015.

[134] R. Deus, D. Brito, M. Mateus, I. Kenov, A. Fornaro, R. Neves, and C. Alves, "Impact evaluation of a pisciculture in the tucuruí reservoir (pará, brazil) using a two-dimensional water quality model," *Journal of hydrology*, vol. 487, pp. 1–12, 2013.

[135] E. Espíndola, T. Matsumura-Tundisi, A. Rietzler, and J. Tundisi, "Spatial heterogeneity of the tucuruí reservoir (state of pará, amazonia, brazil) and the distribution of zooplanktonic species," *Revista Brasileira de Biologia*, vol. 60, no. 2, pp. 179–194, 2000.

[136] E. J. Knight and G. Kvaran, "Landsat-8 operational land imager design, characterization and performance," *Remote Sensing*, vol. 6, no. 11, pp. 10 286–10 305, 2014.

[137] S. Gilmore, A. Saleem, and A. Dewan, "Effectiveness of dos (dark-object subtraction) method and water index techniques to map wetlands in a rapidly urbanising megacity with landsat 8 data," in *Research@ Locate'15*. http: http://SunSITE. Informatik. RWTH-Aachen. DE/Publications/CEUR-WS/., 2015, pp. 100–108.

[138] M. A. Spanner, L. L. Pierce, D. L. Peterson, and S. W. Running, "Remote sensing of temperate coniferous forest leaf area index the influence of canopy closure, understory vegetation and background reflectance," *TitleREMOTE SENSING*, vol. 11, no. 1, pp. 95–111, 1990.

[139] M. E. Jakubauskas, "Thematic mapper characterization of lodgepole pine seral stages in yellowstone national park, usa," *Remote sensing of environment*, vol. 56, no. 2, pp. 118–132, 1996.

[140] P. S. Chavez Jr, "Radiometric calibration of landsat thematic mapper multispectral images," *Photogrammetric Engineering and Remote Sensing*, vol. 55, no. 9, pp. 1285–1294, 1989.

[141] P. Brezonik, K. D. Menken, and M. Bauer, "Landsat-based remote sensing of lake water quality characteristics, including chlorophyll and colored dissolved organic matter (cdom)," *Lake and Reservoir Management*, vol. 21, no. 4, pp. 373–382, 2005.

[142] D. G. Hadjimitsis, G. Papadavid, A. Agapiou, K. Themistocleous, M. Hadjimitsis, A. Retalis, S. Michaelides, N. Chrysoulakis, L. Toulios, and C. Clayton, "Atmospheric correction for satellite remotely sensed data intended for agricultural applications: impact on vegetation indices," *Natural Hazards and Earth System Sciences*, vol. 10, no. 1, pp. 89–95, 2010.

[143] S. Karthick, S. Aanand, A. Srinivasan, and P. Jawahar, "Assessment of seasonal variations in water quality parameters of kadamba tank in thoothukudi district of tamil nadu, india," *IJAR*, vol. 2, no. 11, pp. 329–334, 2016.

[144] C. E. Boyd, "ph, carbon dioxide, and alkalinity," in *Water Quality*. Springer, 2015, pp. 153–178.

[145] S. Deekae, J. Abowei, and J. Alfred-Ockiya, "Seasonal variation of some physical and chemical parameters of luubara creek, ogoni land, niger delta, nigeria," *Research Journal of Environmental and Earth Sciences*, vol. 2, no. 4, pp. 208–215, 2010.

[146] P. N. Bande and S. Nandedkar, "Survey of water quality measurements sensors," *International Journal of research in Engineering and Technology*, vol. 5, no. 06, 2016.

[147] S. Pal, D. Das, and K. Chakraborty, "Colour optimization of the secchi disk and assessment of the water quality in consideration of light extinction coefficient of some selected water bodies at cooch behar, west bengal," *International Journal of Multidisciplinary Research and Development*, vol. 2, no. 3, pp. 513–518, 2015.

[148] E. T. Harvey, S. Kratzer, and P. Philipson, "Satellite-based water quality monitoring for improved spatial and temporal retrieval of chlorophyll-a in coastal waters," *Remote Sensing of Environment*, vol. 158, pp. 417–430, 2015.

[149] Y. Park, K. H. Cho, J. Park, S. M. Cha, and J. H. Kim, "Development of early-warning protocol for predicting chlorophyll-a concentration using machine learning models in freshwater and estuarine reservoirs, korea," *Science of the Total Environment*, vol. 502, pp. 31–41, 2015.

[150] P. De Wilde, "The gap between predicted and measured energy performance of buildings: A framework for investigation," *Automation in Construction*, vol. 41, pp. 40–49, 2014.

[151] R. Richter and D. Schläpfer, "Atmospheric/topographic correction for satellite imagery," *DLR report DLR-IB*, pp. 565–01, 2005.

[152] P. S. Addison, *The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance*. CRC press, 2017.

[153] M. Shoaib, A. Y. Shamseldin, B. W. Melville, and M. M. Khan, "Hybrid wavelet neural network approach," in *Artificial Neural Network Modelling*. Springer, 2016, pp. 127–143.

[154] L. Xu and S. Liu, "Study of short-term water quality prediction model based on wavelet neural network," *Mathematical and Computer Modelling*, vol. 58, no. 3, pp. 807–813, 2013.

[155] A. Sandryhaila and J. M. Moura, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 80–90, 2014.

[156] A. X. Patel, P. Kundu, M. Rubinov, P. S. Jones, P. E. Vértes, K. D. Ersche, J. Suckling, and E. T. Bullmore, "A wavelet method for modeling and despiking motion artifacts from resting-state fmri time series," *Neuroimage*, vol. 95, pp. 287–304, 2014.

[157] K. H. Kim and S. J. Kim, "Neural spike sorting under nearly 0 db signal-to-noise ratio using nonlinear energy operator and artificial neural-network classifier," 2000.

[158] H. Jarvie, B. Whitton, and C. Neal, "Nitrogen and phosphorus in east coast british rivers: speciation, sources and biological significance," *Science of the Total environment*, vol. 210, pp. 79–109, 1998.

[159] C. Christia, G. Giordani, and E. Papastergiadou, "Assessment of ecological quality of coastal lagoons with a combination of phytobenthic and water quality indices," *Marine pollution bulletin*, vol. 86, no. 1, pp. 411–423, 2014.

[160] R. Costanza, R. d'Arge, R. De Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R. V. O'neill, J. Paruelo *et al.*, "The value of the world's ecosystem services and natural capital," *nature*, vol. 387, no. 6630, pp. 253–260, 1997.

[161] I. Onyema and R. Popoola, "The physico-chemical characteristics, chlorophyll a levels and phytoplankton dynamics of the east mole area of the lagos harbour, lagos," *Journal of Asian Scientific Research*, vol. 3, no. 10, pp. 995–1010, 2013.

[162] S. Koponen, J. Pulliainen, K. Kallio, and M. Hallikainen, "Lake water quality classification with airborne hyperspectral spectrometer and simulated meris data," *Remote Sensing of Environment*, vol. 79, no. 1, pp. 51–59, 2002.

[163] X. Guan, J. Li, and W. G. Booty, "Monitoring lake simcoe water clarity using landsat-5 tm images," *Water resources management*, vol. 25, no. 8, pp. 2015–2033, 2011.

[164] C. Huang, J. Zou, Y. Li, H. Yang, K. Shi, J. Li, Y. Wang, X. Chena, and F. Zheng, "Assessment of nir-red algorithms for observation of chlorophyll-a in highly turbid inland waters in china," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 93, pp. 29–39, 2014.

[165] S. Kay, J. D. Hedley, and S. Lavender, "Sun glint correction of high and low spatial resolution images of aquatic scenes: a review of methods for visible and near-infrared wavelengths," *Remote Sensing*, vol. 1, no. 4, pp. 697–730, 2009.

[166] F. L. Lobo, M. P. Costa, and E. M. Novo, "Time-series analysis of landsat-mss/tm/oli images over amazonian waters impacted by gold mining activities," *Remote Sensing of Environment*, vol. 157, pp. 170–184, 2015.

[167] J. G. Tundisi, M. A. Santos, and C. F. S. Menezes, "Tucuruí reservoir and hydroelectric power plant," *Sharing Experiences and Lessons Learned in Lake Basin Mangement, Burlington, Vermont. Management Experiences and Lessons Learned Brief*, vol. 1, pp. 1–20, 2003.

[168] N. T. T. Ha, K. Koike, and M. T. Nhuan, "Improved accuracy of chlorophyll-a concentration estimates from modis imagery using a two-band ratio algorithm and geostatistics: As applied to the monitoring of eutrophication processes over tien yen bay (northern vietnam)," *Remote Sensing*, vol. 6, no. 1, pp. 421–442, 2013.

[169] R. Doerffer, H. Schiller, J. Fischer, R. Preusker, and M. Bouvet, "The impact of sun glint on the retrieval of water parameters and possibilities for the correction of meris scenes," in *Proceedings of the 2nd MERIS/(A) ATSR User Workshop, ESA SP-666,(ESA, Frascati, Italy, 2008)*, 2008.

[170] E. J. Hochberg, S. Andrefouet, and M. R. Tyler, "Sea surface correction of high spatial resolution ikonos images to improve bottom mapping in near-shore environments," *IEEE transactions on geoscience and remote sensing*, vol. 41, no. 7, pp. 1724–1729, 2003.

[171] T. Kutser, E. Vahtmäe, and J. Praks, "A sun glint correction method for hyperspectral imagery containing areas with non-negligible water leaving nir signal," *Remote Sensing of Environment*, vol. 113, no. 10, pp. 2267–2274, 2009.

[172] H. Fukushima, M. Toratani, H. Murakami, P.-Y. Deschamps, R. Frouin, and A. Tanaka, "Evaluation of adeos-ii gli ocean color atmospheric correction using simbada handheld radiometer data," *Journal of oceanography*, vol. 63, no. 3, pp. 533–543, 2007.