

DEVELOPMENT OF A DECISION-SUPPORT TOOL FOR MANAGING DRINKING
WATER RESERVOIR BY USING MACHINE LEARNING AND DEEP LEARNING
METHODS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS OF
THE MIDDLE EAST TECHNICAL UNIVERSITY
BY

SERKAN ÖZDEMİR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF INFORMATION SYSTEMS

DECEMBER 2023

Approval of the thesis:

**DEVELOPMENT OF A DECISION-SUPPORT TOOL FOR MANAGING DRINKING
WATER RESERVOIR BY USING MACHINE LEARNING AND DEEP LEARNING
METHODS**

Submitted by SERKAN ÖZDEMİR in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Information Systems Department, Middle East Technical University
by,

Prof. Dr. Banu Günel Kılıç
Dean, **Graduate School of Informatics**

Prof. Dr. Altan Koçyiğit
Head of Department, **Information Systems**

Prof. Dr. Sevgi Özkan Yıldırım
Supervisor, **Information Systems Dept., METU**

Dr. Muhammad Yaquub
Co-Supervisor, **Environmental Engineering Dept.,
Kumoh National Institute of Technology**

Examining Committee Members:

Prof. Dr. Nazife Baykal
Information Systems Dept., METU

Prof. Dr. Sevgi Özkan Yıldırım
Information Systems Dept., METU

Prof. Dr. P. Erhan Eren
Information Systems Dept., METU

Assist. Prof. Dr. Tuna Hacaloğlu
Information Systems Engineering Dept., Atılım
University

Assist. Prof. Dr. Banu Yüksel Özkaya
Industrial Engineering Dept., Hacettepe University

Date: 19.12.2023

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Serkan Özdemir

Signature : _____

ABSTRACT

DEVELOPMENT OF A DECISION-SUPPORT TOOL FOR MANAGING DRINKING WATER RESERVOIR BY USING MACHINE LEARNING AND DEEP LEARNING METHODS

Özdemir, Serkan

Ph.D, Department of Information Systems

Supervisor: Prof. Dr. Sevgi Özkan Yıldırım

Co-Supervisor: Dr. Muhammad Yaqub

December 2023, 136 pages

Global climate change has led to large fluctuations in lake levels in recent years, due to both changing meteorological parameters and intensive water use. A shift in input or output variables can easily alter the water balance equation and move water levels in the opposite direction. To understand the continuing trend and to create an action plan for dramatic water balance and water quality management, scientists use a variety of models to analyze several variables recorded. In this thesis, the predictive models used for the climatic and hydrologic variables are discussed and their relationships to Lake Water Level (LWL) and water quality are presented. Based on the technological progress, three different types of algorithms a) Naive Method, b) Artificial Neural Networks (ANN) and finally c) Recurrent Neural Network (RNN) models are used to predict water level in lakes. The prediction results from the thesis show that Long Short Term Memory (LSTM) has the highest accuracy with respect to the Root Mean Squared Error (RMSE) evaluation metric. The models were also compared with the performance of the Naïve Method, and the results show that ANN and RNN algorithms are superior in prediction accuracy as the prediction horizon increases. The prediction performances were assessed with Diebold Mariano Test to decide significant differences. It also reveals the water quality of the lake is highly correlated with temperature and evaporation. The models and evaluation metrics are constructed to build a prototype of decision support tool in order water managers to use in operational transactions.

Keywords: lake water level, deep learning, time series, water quality, decision support tool

ÖZ

MAKİNE ÖĞRENMESİ VE DERİN ÖĞRENME YÖNTEMLERİNİ KULLANARAK İÇME SUYU REZERVUARININ YÖNETİMİ İÇİN BİR KARAR DESTEK ARACININ GELİŞTİRİLMESİ

Özdemir, Serkan

Doktora, Bilişim Sistemleri Bölümü

Tez Yöneticisi: Prof. Dr. Sevgi Özkan Yıldırım

Ortak Tez Yöneticisi: Dr. Muhammad Yağub

Aralık 2023, 136 sayfa

Küresel iklim değişikliği, hem değişen meteorolojik parametreler hem de yoğun su kullanımı nedeniyle son yıllarda göl seviyelerinde büyük dalgalanmalara yol açmıştır. Girdi veya çıktı değişkenlerindeki bir değişim, su dengesi denklemini kolayca değiştirebilir ve su seviyelerini ters yönde hareket ettirebilir. Devam eden eğilimi anlamak ve dramatik su dengesi ve su kalitesi yönetimi için bir eylem planı oluşturmak amacıyla, bilim insanları kaydedilen çeşitli değişkenleri analiz etmek için çeşitli modeller kullanmaktadır. Bu tezde, iklimsel ve hidrolojik değişkenler için kullanılan tahmin modelleri tartışılmakta ve bunların Göl Su Seviyesi (LWL) ve su kalitesi ile ilişkileri sunulmaktadır. Teknolojik gelişmelere bağlı olarak, göllerdeki su seviyesini tahmin etmek için üç farklı algoritma türü a) Naïve Yöntem, b) Yapay Sinir Ağları (YSA) ve son olarak c) Tekrarlayan Sinir Ağı (RNN) modelleri kullanılmıştır. Tezden elde edilen tahmin sonuçları, Kök Ortalama Kareysel Hata (RMSE) değerlendirme metriğine göre Uzun Kısa Süreli Belleğin (LSTM) en yüksek doğruluğa sahip olduğunu göstermektedir. Modeller ayrıca Naïve Yönteminin performansı ile karşılaştırılmış ve sonuçlar, tahmin periyodu arttıkça YSA ve RNN algoritmalarının tahmin doğruluğunda üstün olduğunu göstermiştir. Tahmin performansları, anlamlı farklılıklara karar vermek için Diebold Mariano Testi ile değerlendirilmiştir. Ayrıca bu tez, gölün su kalitesinin sıcaklık ve buharlaşma ile yüksek oranda ilişkili olduğunu ortaya koymaktadır. Modeller ve değerlendirme ölçütleri, su yöneticilerinin operasyonel işlemlerde kullanmaları için bir karar destek aracı prototipi oluşturmak üzere geliştirilmiştir.

Anahtar Sözcükler: göl su seviyesi, derin öğrenme, zaman serisi, su kalitesi, karar destek aracı

To My Family

ACKNOWLEDGMENTS

First and foremost, I would like to sincerely thank my supervisor Prof. Dr. Sevgi Özkan Yıldırım. I am so appreciative of her creative thoughts. She consistently offers new perspectives and shares her experiences on the study. She has been very patient with me and has shown me warm support, wise counsel, and time. I also express my gratitude to my co-supervisor Dr. Muhammad Yaqub for his invaluable help, advice, insights, and knowledge sharing during this study.

I am very appreciative for my Thesis Monitoring Committee Members Prof. Dr. Nazife Baykal and Assist. Prof. Dr. Tuna Hacaloğlu for their valuable feedbacks. They opened new perspectives in this thesis and helped it to become useful and effective tool in the future.

I am grateful to Assoc. Prof. Dr. Edo Abraham and Dr. Ties Van der Heijden for their insightful criticism that enhances the thesis by taking a different approach. I also thank Prof. Dr. Tanju Karanfil and Dr. Eda Göz Semizer for their efforts in enhancing deep learning models and conducting assessments through weekly intercontinental meetings. I am appreciative that the Turkish Meteorological Office and State Hydraulic Works allowed me to use their dataset for this thesis study. I am thankful to the Samsung Innovation Campus for assisting me regarding interface design for AI initiatives. I also want to express my gratitude to all of my friends, instructors and relatives for their unwavering belief in me and support.

Lastly, I would like to sincerely thank my parents, Saim Özdemir and Türkan Özdemir, for their unwavering support, love, encouragement, and extraordinary patience. I could never have achieved this accomplishment without their support, affection, and faith in me.

TABLE OF CONTENTS

ABSTRACT	iv
ÖZ.....	v
DEDICATION	vi
ACKNOWLEDGMENTS.....	vii
TABLE OF CONTENTS	viii
LIST OF TABLES	xi
LIST OF FIGURES.....	xii
LIST OF ABBREVIATIONS	xiv
CHAPTERS	
1. INTRODUCTION.....	1
2. LITERATURE REVIEW.....	5
2.1. Lake Water Level	5
2.1.1. Regression Based Models	11
2.1.2. Mathematical Models	16
2.1.3. ANN Models	26
2.1.4. Decision Tree Models	30
2.1.5. Fuzzy Logic Models.....	32
2.1.6. Deep Learning Models.....	33
2.2. Overview of LWL Data-Driven Modeling.....	34
2.3. Effects of Climate Change on Lakes	35
2.4. Factors that Affect Lake Water Quality	39
2.5. Decision Support Tool for Lake Water Level	41
2.6. Literature Summary and Knowledge Gaps	42
3. RESEARCH METHODOLOGY	43
3.1. Research Questions and Objectives.....	43
3.2. Case Study Area	44

3.3.	Dataset Description	45
3.4.	Stationary Data Test	47
3.5.	Data Preprocessing	48
3.6.	Model Descriptions	51
3.6.1.	ANN	52
3.6.2.	LSTM	53
3.6.3.	GRU	54
3.6.4.	Stacked LSTM	55
3.6.5.	Bidirectional LSTM	55
3.7.	Sequence Creation	56
3.8.	Normalization	57
3.9.	Hyperparameters	58
3.10.	Evaluation Metrics	59
3.11.	Water Quality Indicator	61
3.12.	Decision Support Tool.....	62
4.	LAKE WATER LEVEL PREDICTION	63
5.	EVALUATION OF CLIMATIC CONDITIONS ON MICROCYSTIN VARIATIONS IN WATER COLUMN.....	83
6.	DECISION SUPPORT TOOL.....	89
6.1.	General Structure of the Tool.....	89
6.2.	Components and Abilities	91
6.2.1.	Dashboard	91
6.2.2.	Manual Data Entry	95
6.2.3.	Prediction	100
7.	DISCUSSION	103
8.	CONCLUSION.....	107
	REFERENCES.....	109
	APPENDICES	129
	APPENDIX A	129
	APPENDIX B	133
	CURRICULUM VITAE	135

LIST OF TABLES

Table 1: List of regression-based models and their input-output variables, evaluation metrics, and size of data.....	15
Table 2: List of mathematical models and their input-output variables, evaluation metrics, and size of data.....	23
Table 3: List of ANN models and their input-output variables, evaluation metrics, and data size.....	29
Table 4: List of decision tree models and their input-output variables, evaluation metrics, and data size	31
Table 5: List of fuzzy logic models and their input-output variables, evaluation metrics, and data size	33
Table 6: List of deep learning models and their input-output variables, evaluation metrics, and data size	34
Table 7: Descriptive Statistics for features in the dataset	45
Table 8: Augmented Dickey-Fuller test for the dataset	48
Table 9: Hyperparameter optimization setup.....	59
Table 10: Optimized hyperparameter values of algorithms.....	59
Table 11: The performance of Naïve Method, ANN, and RNN based algorithms for predicting LWL with increasing time intervals, RMSE results	63
Table 12: Benchmark performance comparison of algorithms, figures indicates improvement over Naïve Method	64
Table 13: 1 Day Diebold Mariano Test results	76
Table 14: 5 Days Diebold Mariano Test results.....	77
Table 15: 10 Days Diebold Mariano Test results.....	77
Table 16: 20 Days Diebold Mariano Test results.....	78
Table 17: 30 Days Diebold Mariano Test results.....	78
Table 18: 45 Days Diebold Mariano Test results.....	79
Table 19: 60 Days Diebold Mariano Test results.....	80
Table 20: 90 Days Diebold Mariano Test results.....	80
Table 21: 120 Days Diebold Mariano Test results.....	81
Table 22: Forecast difference results of Naïve Method, ANN and RNN algorithms based on Diebold Mariano test for increasing day intervals from day-1 to day-120.....	82
Table 23: Library Versions in Web Application.....	90

LIST OF FIGURES

Figure 1: Applied SLR methodology in the study	6
Figure 2: Number of publications on LWL and RWL prediction.....	7
Figure 3: LWL and RWL study areas	8
Figure 4: LWL and RWL articles' journal names.....	8
Figure 5: LWL and RWL prediction models	9
Figure 6: LWL and RWL evaluation metrics.....	10
Figure 7: Evaluation metrics in terms of algorithms.....	10
Figure 8: Lake Sapanca area and its catchment with river basins.....	44
Figure 9: Time series plots of daily meteorological data, water withdrawals and LWL for Lake Sapanca from 11 October 2012 through 31 December 2020.	47
Figure 10: Bar chart of features with missing values.....	50
Figure 11: Missingness Matrix of Features.....	51
Figure 12: ANN Model	52
Figure 13: ANN Model with multiple layers	53
Figure 14: Basic structure of LSTM algorithm.....	53
Figure 15: Basic structure of GRU algorithm	54
Figure 16: Stacked LSTM Model.....	55
Figure 17: Bidirectional LSTM Model	56
Figure 18: Flowchart of LWL Modelling	61
Figure 19: 1 day ahead prediction results.....	66
Figure 20: 5 days ahead prediction results	67
Figure 21: 10 days ahead prediction results	68
Figure 22: 20 days ahead prediction results	69
Figure 23: 30 days ahead prediction results	70
Figure 24: 45 days ahead prediction results	72
Figure 25: 60 days ahead prediction results	73
Figure 26: 90 days ahead prediction results	74
Figure 27: 120 days ahead prediction results	75
Figure 28: Variable Importance	82
Figure 29: Vertical distribution of microcystin concentrations measured monthly intervals between 21 March 2019 to 12 April 2023 in Lake Sapanca.	85
Figure 30: Box and whisker plot of microcystin concentrations at different water depths of Lake Sapanca during 21 March 2019 to 12 April 2023.....	86
Figure 31: Linear trend of microcystin concentration in the water column at different depths from surface to 20 m during 21 March 2019 to 12 April 2023	87
Figure 32: SRC between microcystin and meteorological parameters	88
Figure 33: LWL Management Decision Support Tool MVC Design Pattern.....	89

Figure 34: Navigation of Application	91
Figure 35: Latest average values in dashboard	92
Figure 36: Radar chart in dashboard	92
Figure 37: Line chart in dashboard	93
Figure 38: Feature Distributions in Dashboard.....	94
Figure 39: Detailed Data View in Dashboard.....	94
Figure 40: Manual data entry	95
Figure 41: Show Current Data in Manual Data Entry.....	96
Figure 42: Add Input Values in Manual Data Entry	97
Figure 43: After Add button effects	97
Figure 44: Edit Values in Manual Data Entry.....	98
Figure 45: After Edit Button Effects.....	99
Figure 46: Delete a row in Manual Data Entry	100
Figure 47: LWL Prediction with LSTM Algorithm in Prediction Component	101
Figure 48: Actual versus predicted values graph	102

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ANFIS	Adaptive Neuro Fuzzy Inference System
ANN	Artificial Neural Network
ANNBP	Adaptive Neural Network Backpropagation
AOFF	Amount of Forecast Factor
ARMAX	Auto Regressive Moving Average with Exogenous Input
BART	Bayesian Additive Regression Trees
BLR	Bayesian Linear Regression
BDTR	Boosted Decision Tree Regression
BHRM	Bayesian Harmonic Regression Models
CART	Classification and Regression Trees
CC	Chance Constrained
CE	Copula Entropy
CNN	Convolutional Neural Network
DFA	Dynamic Factor Analysis
DFR	Decision Forest Regression
DIC	Deviance Information Criterion
DL	Deep Learning
DOC	Degrees of Confidence
DWL	Dam Water Level
DWT- iNARX	Discrete Wavelet Transformation-Improved Nonlinear Autoregressive with Exogenous Inputs Network
DTC	Decision Tree Classifier
DTR	Decision Tree Regression
EF	Efficiency Index
ELM	Extreme Learning Machine
ETC	Extra Tree Classifier
ETR	Extra Tree Regression
FCNN	Fully Connected Neural Network
FeFLOW	Finite Element Subsurface FLOW
FNN	Fuzzy Neural Networks
FVCOM	Finite-Volume 3D Ocean Model

GAM	Generalized Additive Model
GBR	Gradient Boosting Regression
GEV	Generalized Extreme Value
GLM	Generalized Linear Model
GNB	Gaussian Naïve Bayes
GP	Gaussian Process
GPR	Gaussian Process Regression
GPU	Graphics Processing Unit
GRACE	Gravity Recovery and Climate Experiment
GRNN	Generalized Regression Neural Network
GRU	Gated Recurrent Unit
HTC	Hoeffding Tree Classifier
HAT-C	HAT Classifier
HAT-R	HAT Regression
HSS	Heidke Skill Score
HTR	Hoeffding Tree Regression
KGE	Kling–Gupta efficiency
KNC	K-Neighbors Classifier
KNN	K-Nearest Neighbor
KNR	K-Neighbors Regression
LBRM	Large Basin Runoff Model
LCC	Linear Correlation
LLC	Lake Level Change
LMI	Legates and McCabes Index
LoR	Logistic Regression
LR	Linear Regression
LSTM	Long Short Term Memory
LWB	Lake Water Balance
LWL	Lake Water Level
M5P	M5 Pruned
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MARE	Mean Absolute Relative Error
MARS	Multivariate Adaptive Regression Spline
MBE	Mean Biased Error
MCC	Multiple Correlation Coefficient

ML	Machine Learning
MLP	Multilayer Perceptron
MLP-R	Multilayer Perceptron Regression
MLR	Multiple Linear Regression
MPMR	Minimax Probability Machine Regression
MRE	Mean Relative Error
MSE	Mean Squared Error
MODM	Multi-Objective Decision Model
MODSIM	Water Balance Network Model
NBS	Net Basin Supply
NIC	Normalized Information Contribution
NN	Neural Network
NNR	Neural Network Regression
NOAH- MP LSM	Noah-Multiparameterization Land Surface Model
NRMSE	Normalized Root Mean Square Error
NSC	Nash-Sutcliffe Coefficient
NSE	Nash–Sutcliffe Efficiency
PBIAS	Percent Bias
PLSR	Partial Least Squares Regression
PREVAH	Precipitation-Runoff-Evapotranspiration Hydrotope
QAC	Quality Assessment Checklist
R	Coefficient of Correlation
R²	Coefficient of Determination
RAE	Residual Absolute Error
RAM	Random Access Memory
RAPID	Routing Application for Parallel Computation of Discharge
RBNN	Radial Basis Neural Networks
RCP	Representative Concentration Pathway
REPT	Reduced Error Pruning Tree
RF	Random Forest
RFC	Random Forest Classifier
RFR	Random Forest Regression
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
RSD	Residual Standard Deviation
RSE	Residual Standard Error

RSOS	Residual Sum of Squares
RSS	Regression Square Sum
RVM	Relevance Vector Machine
RWI	Refined Willmott's Index
RWL	Reservoir Water Level
RT	Random Tree
SD	Standard Deviation
SLR	Systematic Literature Review
SRC	Spearman's Rank Correlation
SRPF	Statistical Regression Predictive Function
SS	Sum of Squares
SVM	Support Vector Machine
SVR	Support Vector Regression
SVR-GWO	Support Vector Regression -Grey Wolf Optimization
SWAT	Soil and Water Assessment Tool
SWGM	Surface-Water-Groundwater Model
U.S.	United States
VIC	Variable Infiltration Capacity
WBCM	Water Budget Conceptual Model
WBM	Water Balance Model
WNN	Wavelet Neural Network
WSHM	Watershed-Scale Hydrologic Model

CHAPTER 1

1. INTRODUCTION

Water quality and quantity difficulties vary from place to place and are influenced by a variety of elements, including climatic, geographic, geological, social, and economic. In addition, ongoing global warming and meteorological patterns are likely to disrupt the temporal and spatial balance of water, leading to freshwater scarcity and impeding the achievement of the United Nations Sustainable Development Goals around the world. Modeling studies suggest that by 2050 there will be a paradigm shift in the distribution of freshwater on the planet (Paul and Elango, 2018; Castillo-Botón et al., 2020). Therefore, a sound water management plan, developed using reliable forecasting models, is essential for implementing sustainable water use and conserving water resources in a given basin or region.

Turkey experiences frequent droughts that significantly reduce surface and groundwater resources, including wetlands and lakes. (Soylu Pekpostalci et al., 2023; Yeşilköy and Şaylan, 2022). Drought conditions affect standing water bodies when there is a reduction in surface runoff and in stream inputs. Droughts typically coincide with hot weather which causes evaporation to increase significantly during dry periods. The effects of drought include a decrease in water levels in what is usually a very fertile littoral zone. This can leave aquatic fauna (e.g. mussels, snails, and flora) stranded in the area. The increased water temperature associated with drought can lead to stratification, increased salinity, and reduced oxygen levels. In some cases, the combination of high temperatures with low oxygen may lead to the extinction of fish species (Bond et al., 2008).

Uncontrolled drinking water supplies and inadequately managed reservoirs pose a significant threat to developing and densely populated cities. Lake Sapanca, for example, is an important source of fresh water supply for the cities of Sakarya and Kocaeli and is also used by several bottled water companies for commercial purposes. The prospects of the reservoir appear to be affected by climate change and recent droughts, which could negatively impact several parts of the region and its ecosystems (Duru, 2017). Because of the multitude of factors that affect the surface of a lake, one of the most critical hydrologic problems is to estimate the water level of a lake before it reaches its threshold. Hydrologic models have certain limitations on accurate predictions due to the complex nature of

hydrologic and meteorological variables, as well as the temporal and spatial characteristics of each watershed (Zounemat-Kermani, 2021; Chang and Chang, 2006). Therefore, it is essential to develop more reliable predictive models that can accurately and reliably estimate the water level of a lake.

There are two different approaches for Lake Water Level (LWL) prediction in the literature. The most prominent approach follows the physical process, and the emerging approach is called a data-driven methodology that focuses on historical data sets to predict future values. The physical process, often achieved by solving hydrodynamic equations, forms the basis for the analysis of the LWL in physics-based methods. Lai et al. (2011) used a combined hydrodynamic analysis model on the middle Yangtze River to calculate changes in the water balance caused by water storage. Wu et al. (2014) investigated the effects of TGD on the water level of Lake Poyang, located in the lower reaches of the Yangtze River, using physical model experiments. Data-driven methods simulate the LWL in addition to the factors affecting it using scientific computer models. Different types of models have been developed to promote specific cases. For instance, Liu et al. (2015) evaluated the model with Support Vector Regression (SVR) and Adaptive Neuro-Fuzzy Inference System (ANFIS). They presented a multivariate conditional model based on copulas to predict water level and improve spatial precipitation estimation. Wang et al. (2017) applied SVR to simulate causality between LWL and the quantity of water discharged from the reservoir. Statistical methods and Artificial Intelligence (AI) techniques are two common data-driven approaches to solve LWL prediction problems (Zhang et al., 2018). These methods include multiple regression, pattern recognition, Neural Network (NN) techniques, time series methods, and probability features (Bourdeau et al., 2019).

In the last decade, a variety of contemporaneous techniques have been applied to compare the predictive power of algorithms. For example, Ghorbani et al. (2010) studied the ability of the Genetic Programming and Artificial Neural Network (ANN) to predict LWL in Australia and reported accurate predictions with good agreement. Talebizadeh and Moridnejad (2011) used ANN and ANFIS to predict the LWL at Lake Urmia in Iran. In another study, NN, neural fuzzy, and GP models were applied to estimate the LWL on a daily basis (Kisi et al., 2012). The results showed that each of the three models accurately predicted the LWL. Buyukyildiz et al. (2014) developed a series of AI models, Multilayer Perceptron (MLP), hybridized SVR with Particle Swarm Optimization (PSO), Radial Basis Neural Network (RBNN), and ANFIS to predict LWL. Their results show that the hybrid SVR-PSO model is a reliable prediction model. Similarly, for three upstream rivers on the east coast of Malaysia, water levels for the next five hours were successfully estimated using ANN (Lukman et al., 2017). To predict the LWL, Yadav and Eliza (2017) used a Support Vector Machine (SVM) and Wavelet. The results of the study showed that the model implemented to predict future values of the reservoir was more accurate compared to regression models. Despite the successful attempts to use Machine Learning (ML) methods in these studies, there are certain inherent limitations in the algorithms used in the literature (Ozdemir et al., 2023). For instance, in ANNs, the rules that could explain underlying methods are not given. In terms of fuzzy logic, setting precise, fuzzy

enrollment limitations and parameters can be difficult and the fuzzy justification isn't always correct. Regression models show that as the number of variables increases, their accuracy decreases. The regression models work better when there are fewer variables. Lastly, training a Deep Learning (DL) model requires a lot of computing power, which leading to the need for powerful Graphics Processing Unit (GPU)s and a large amount of Random Access Memory (RAM). Another potential drawback is an overfitting issue that arises when a model performs poorly on newly untrained data after being overtrained on training data.

The most common time series prediction model with statistical analysis used by scientists to predict lake level is the Autoregressive Integrated Moving Average (ARIMA) model (Yu et al., 2017; Viccione et al., 2020). It can be expressed in several ways, including Moving Average (MA), Autoregressive Average (AR), hybrid AR or MA, known as Autoregressive Moving Average (ARMA) or Seasonal Autoregressive Integrated Moving Average (SARIMA) (Azad et al., 2022). The SARIMA model, on the other hand, has the advantage of requiring fewer model features to explain the structure of time series that exhibit non-stationarity in seasons and between seasons (Fang and Lahdelma, 2016). Unlike MLs, which often require multiple features as input, this is an important simplification (Viccione et al., 2020). The ANN algorithm is a widely used ML method for water flow modeling, water quality assessment, and water level prediction in hydrology and water resources (Lukman et al., 2016; Altunkaynak et al., 2007; Nouri et al., 2019; Adhikary et al., 2018;). In addition, some research has presented a hybrid ANN-ARIMA model (Khandelwal et al., 2015; Phan and Nguyen, 2020).

A review of the above research papers shows that various approaches to predicting LWL produce distinct outcomes and estimation uncertainties. Recently, some scholars have used time series techniques to predict various areas such as energy prices, stock prices, and corporate sales forecasts that are critical to the global economy (Sethia and Raut, 2019; Anupa et al., 2021), including weather, environment, hydrology, and geological phenomena (Ebtehaj et al., 2019, Xiang and Demir, 2020). Almost all of them concluded that the time series forecasting methods provide more accurate results compared to the benchmark models.

The Recurrent Neural Network (RNN)-based DL approach is proposed in this thesis as a state-of-the-art technique to solve the above problems in the study of LWL, which would improve the prediction performance. DL networks, which differ from conventional approaches in that they allow computer models consisting of numerous layers to learn representations of data consisting of multiple levels of abstraction, replicate the functioning of the human brain (Chen et al., 2018). The DL approach has been used for object recognition, speech recognition, visual object recognition, genomics, and drug discovery (LeCun et al., 2015). The extraordinary success of supervised RNN-based DL algorithms for conducting recognition studies directed to use the RNN-based algorithms in multivariate time series studies. The LWL studies also have time series data due to its nature and attracts hydrologists to exploit the power of these DL algorithms in their future time series prediction studies. However, the application of DL models for LWL is limited

and is the focus of this thesis to overcome several drawbacks of available approaches for LWL prediction, such as the large number of input variables and their uncertainty. The work motivation in this thesis is to provide effective prediction technique for water managers to handle drinking water supply availability in lakes before reaching an alarming level. The limited water supply in lakes not only cause frequent drought experiences and water shortage, but also cause a decrease in water quality.

In this thesis, novel gated RNN-based algorithms are used to build a model that can predict future LWL to support drought mitigation and reservoir management. In addition, this study aims to help fill the gap in the literature regarding the selection of DL models and the evaluation of the performance of LWL prediction algorithms by using the Naïve Method Benchmark and the Diebold-Mariano test. There is no study in the literature that focuses on the comparison between algorithms for multivariate prediction studies with different future time periods.

CHAPTER 2

2. LITERATURE REVIEW

This chapter consists of six sections. The first section focuses on background information on LWL with a Systematic Literature Review (SLR) conducted in October 2021. The second section gives information for comparison of studies in literature in terms of similarities and differences. The third section focuses on effect of climate change on lakes and the studies in literature that investigates this area. The fourth section gives literature information for the factors that affect lake water quality by considering microcystin an algal toxins produced by *Plankthotrix Rubescens*. The fifth section introduces the Decision Support Tool studies for LWL that are investigated by previous scholars. The last section summarizes the studies in literature and remarks the knowledge gaps.

2.1. Lake Water Level

A comprehensive review of the literature was conducted to assess the current state of knowledge regarding LWL prediction and forecasting methods. By carefully formulating search queries and inclusion/exclusion specifications, the SLR accurately evaluates the breadth of the current literature. It is also reproducible and transparent. Identification of data sources and formulation of search parameters are the starting point of SLR. Next, publications that meet the objectives of the study are either included in the final sample or excluded. Finally, relevant data are extracted from the entire text of the sample, and the results are analyzed and reported. Despite the fact that a large sample was extracted from the SLR, not all important references may have been included. In some cases, certain references were used for discussion but did not appear in the SLR results.

After defining the scope of the research, the research questions are developed and the keywords of the systematic search are retrieved. These keywords, which are: (lake AND water AND level AND predict*), (reservoir AND water AND level AND predict*), (lake AND water AND level AND forecast*), and (reservoir AND water AND level AND forecast*), were extracted at least once from the title, abstract, or keywords. To narrow the search, they were first refined using specific inclusion and exclusion criteria. In the first phase of the study, only published research from open access book chapters, journals,

and conference proceedings were considered. However, studies written entirely in a language other than English were not included in this analysis.

A Quality Assessment Checklist (QAC) has been established to assess the eligibility of each study for second and third stages. The checklist includes items such as: a) Does the study focus on predicting water levels? b) Does the study relate to a reservoir where water levels were predicted? (c) Does the publication of the study disclose the methodology?

Once the review methodology has been determined, there are other criteria for selecting studies that are relevant to the objective of the study. The following three criteria are applied in the study:

- I. The research must focus on a natural or artificial freshwater lake (i.e., excluding groundwater, rivers, oceans, or similar environments);
- II. Research models must focus on LWLs, dynamics, or fluctuations (e.g., research that predicts flood hazards, drought-related conditions, economic impacts, human populations, or other animal populations are excluded from consideration); and
- III. Models must be able to anticipate LWLs, validate observed data, and/or provide qualitative justification for them.

A thorough search of ScienceDirect, Web of Science, and Scopus databases was conducted on October 3, 2021. There were 1470 articles in phase 1 (Figure 1).

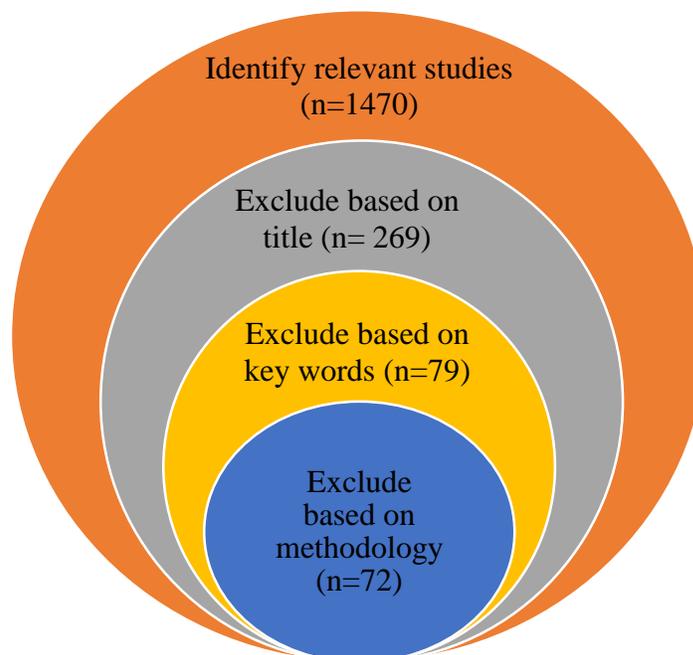


Figure 1: *Applied SLR methodology in the study.* Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684

269 and 72 articles were still available after QAC was applied to the second and third stage data, respectively (Figure 1).

The literature background shows an increasing interest in the research of LWL and Reservoir Water Level (RWL) prediction, especially in light of the recent impacts of climate change. The number of publications over the past 15 years also shows an increase (Figure 2).

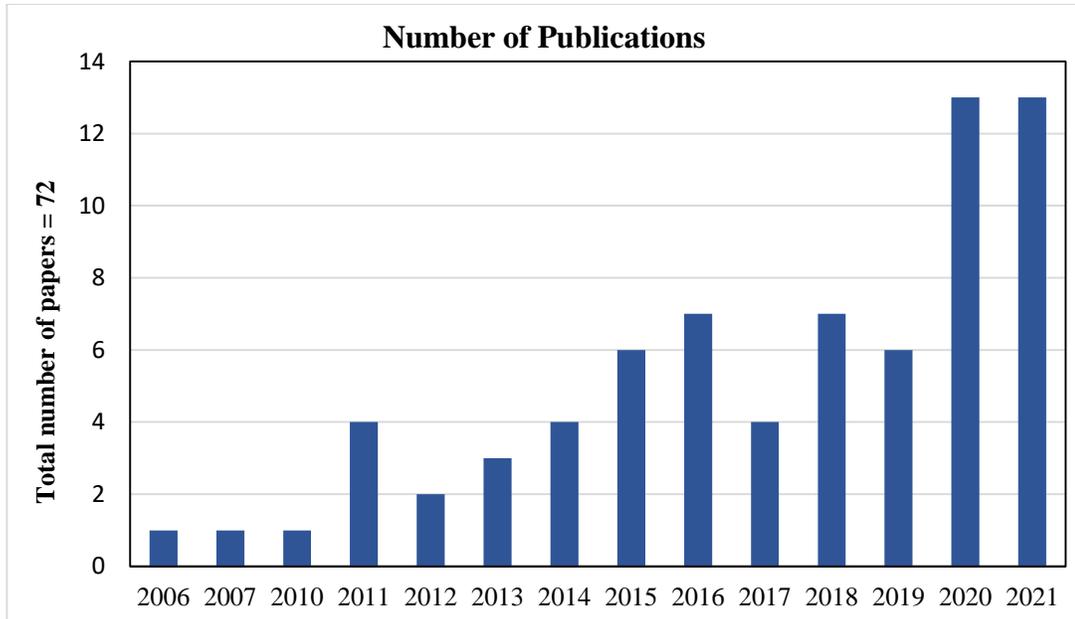


Figure 2: *Number of publications on LWL and RWL prediction.* Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684.

The number of publications in Figure 2 shows a large increase in research on LWL and RWL forecasting, especially for 2020 and 2021. The impacts of climate change and recent water shortages may be one reason for research on LWL and RWL forecasting. Figure 3 below shows how the research region differs in terms of the geographic areas where climate change impacts have the greatest impact on demographics.

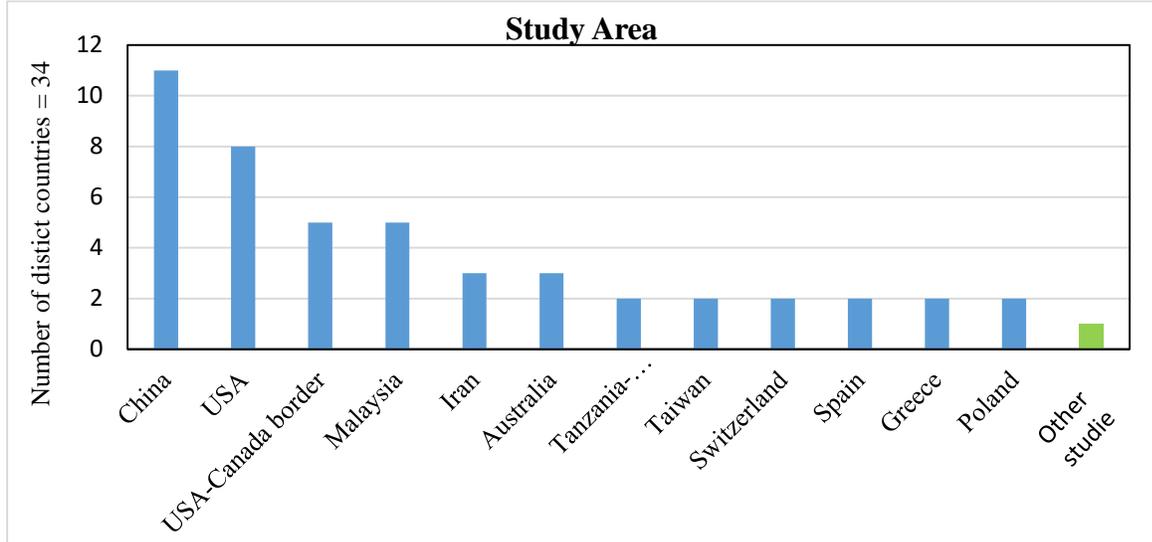


Figure 3: *LWL and RWL study areas*. Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684.

Most studies have focused on lakes on the United States (U.S.)-Canada border and in China or the United States. The literature also examined lakes from various parts of the world. The 17 other studies examined reservoirs in Europe, while 13 dealt with them in Asia. Nine study regions in Africa, three in Australia, and two lake levels in South America were modeled. Because a study may examine more than one lake or reservoir to compare, Figure 3 shows a higher number of study regions than studies. In addition, although the studies in the literature cover a 15-year period (2006-2021), the length of the datasets used in the studies varies depending on the methods used to collect the data and is generally longer than 15 years.

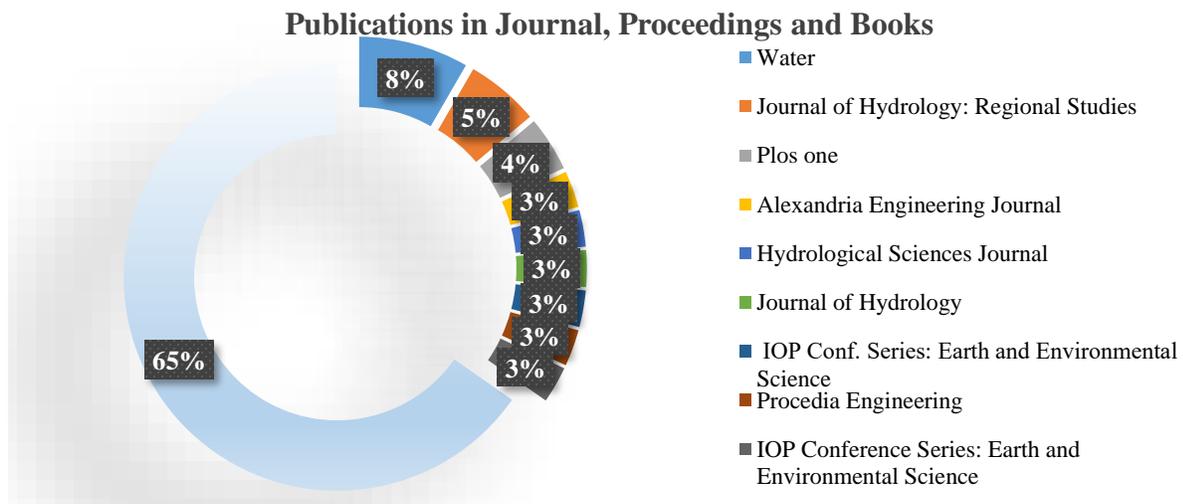


Figure 4: *LWL and RWL articles' journal names*. Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684.

Scientists publish their articles and proceedings in prestigious journals and conferences. The Figure 4 show the proportion of publications in this SLR that compare entire studies. "Water" is the journal with the most publications for this type of research. In addition, "Plos One" and "Journal of Hydrology: Regional Studies" publish a large number of studies. In the field of water research, the other studies are almost equally represented in the different journals. The results of the figure show that the listed journals have a strong emphasis on studies related to LWL and RWL forecasting and are most likely to include related topics in their publications.

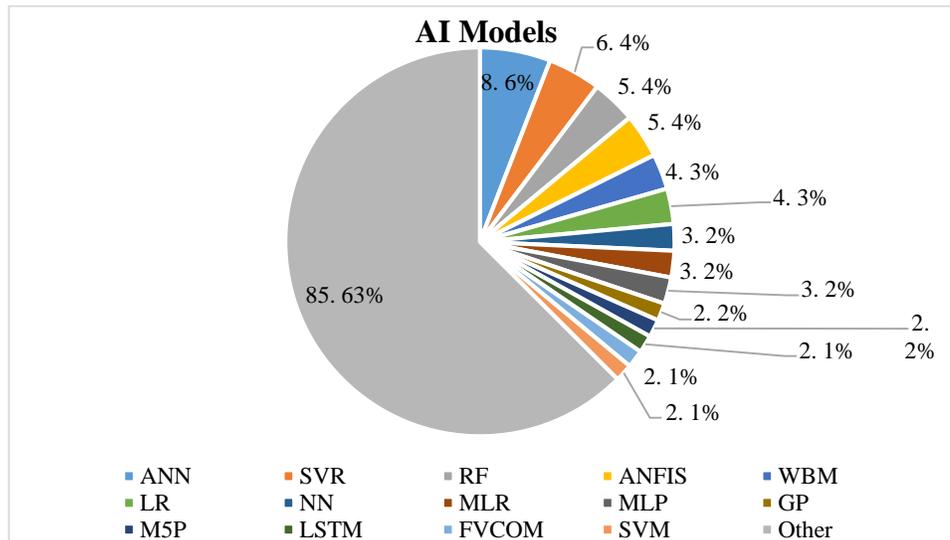


Figure 5: *LWL and RWL prediction models*. Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684.

Most scientists have chosen ANNs to predict future LWL. SVR, Random Forest (RF) and Artificial Neuro-Fuzzy Inference System (ANFIS) are some of the other most commonly used techniques. Figure 5 shows that the algorithm used in the literature is not dominant or widely accepted. Moreover, the positive results of NN-based prediction models encourage researchers to develop prediction models utilizing these algorithms and DL techniques (Yuan et al., 2022).

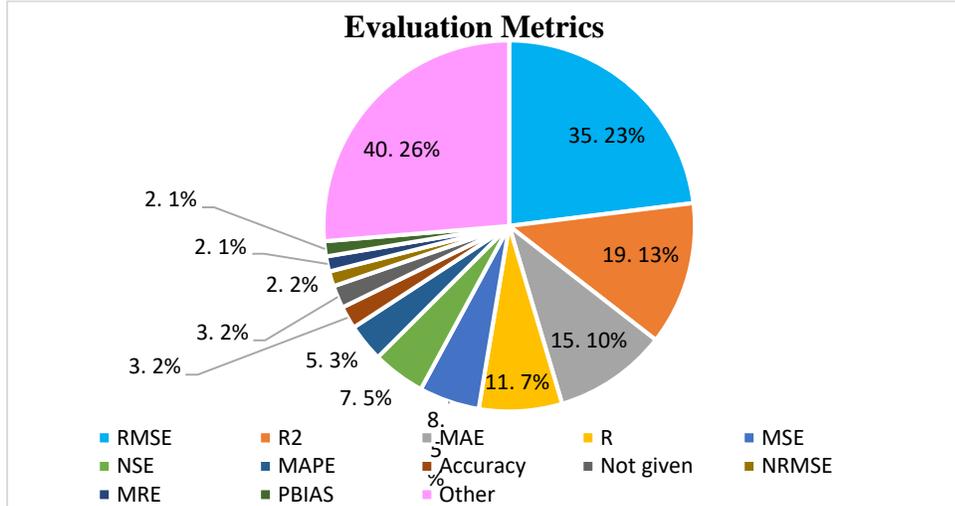


Figure 6: *LWL and RWL evaluation metrics*. Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684.

As shown in Figure 6, almost in every three research, researchers used Root Mean Squared Error (RMSE) as the evaluation metric. Coefficient of Determination (R^2) and Mean Absolute Error (MAE) are the second and third most common evaluation metrics in the literature, respectively. Despite the fact that these algorithms are most commonly used by researchers, most of them analyze results using multiple evaluation metrics rather than just one. Increasing the credibility of the results of evaluation measures could be one of the reasons for not limiting oneself to one evaluation metric.

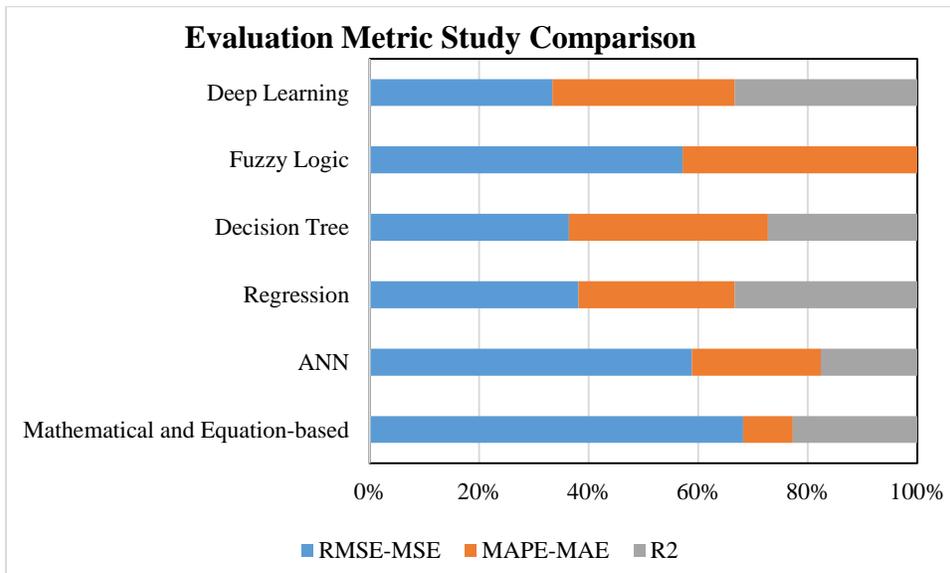


Figure 7: *Evaluation metrics in terms of algorithms*. Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684.

Most of the evaluation metrics for mathematical and equation-based models, ANN and fuzzy logic, and in published results for AI models are RMSE and Mean Squared Error (MSE) (Figure 7). R^2 is almost never used in fuzzy logic models and is almost exclusively chosen in regression and DL models. Another problem is that less than 10% of studies use the Mean Absolute Percentage Error (MAPE)-MAE to assess the effectiveness of equation-based and mathematics-based models.

2.1.1. Regression Based Models

Zhu et al. (2017) developed Linear Regression (LR) model to predict the water level of Lake Chad, which is located in Nigeria, Cameroon, Niger, and the border with Chad. The area of the lake is estimated to be 15000 km². The dataset contains 288 data lines estimated from 24 years of monthly data. The model forecasts Lake Water Balance (LWB) using direct precipitation, evaporation, lake inflow from the Chari/Logone River system, and lake outflow, which includes both surface water runoff to the northern basin and lake seepage in the form of groundwater discharge, as input variables. The evaluation metric used to test the model, on the other hand, is Correlation Analysis (R). The results of the study reveals that the water level of the lake is increasing steadily for 0.5 cm per year. Furthermore, there is a variation of water levels throughout the year which is 1.38 m in terms of seasons.

Lin et al. (2015) studied a combination of the Routing Application for Parallel Computation of Discharge (RAPID) model, the Noah- Multiparameterization Land Surface Model (NOAH-MP LSM), and Multiple Linear Regression (MLR) for lake level prediction. The study area is Lake Buchanan, which is located in Texas, USA. The area of Lake Buchanan is 90 km². The authors used a data set with 108 data rows estimated from 9 years of monthly accumulated data. The model predicts Lake Level Change (LLC) with input variables such as inflow rate from the main stream and tributaries as simulated by RAPID, the outflow rate (in m³/s); P(t), precipitation, and evaporation. The model is tested using R^2 , RMSE, and NRMSE as performance measures. According to the results, runoff can be simulated at any scale required for lake level modeling.

García Molinos et al. (2015) investigated Bayesian Harmonic Regression Models (BHRM) to predict LWL in natural Irish lakes located in Ireland. The dataset used in this study includes an estimated 456 rows of data accumulated from 38 years of monthly data. The model predicts LWL by again using water level. The performance of the model was tested using the Deviance Information Criterion (DIC). From the results, it can be concluded that it is possible to group lakes in terms of their annual seasonality and their inter-annual and inter-decadal cycles.

Castillo-Botón et al. (2020) proposed SVR and Gaussian Processes (GPs) for predicting the RWL of the Belesar reservoir. The reservoir has an area of 19.1 km² and is located in Spain. In the study, the Dam Water Level (DWL) in hm³ was set as the initial parameter and an attempt was made to predict this water level by using flow (m³/s) and elevation

(m) upstream and on tributaries, and precipitation (mm) as input features. The total size of the dataset is 2704 for all parameters. RMSE and MAE were chosen as the evaluation metrics for this study. The results show that the model allows short- and long-term prediction and analysis of reservoir levels, which contributes to hydropower management.

Kenda et al. (2020) analyzed LWLs in their study using different techniques such as LR, Decision Tree Regression (DTR), Random Forest Regression (RFR), Gradient Boosting Regression (GBR), Partial Least Squares Regression (PLSR), Extra Tree Regression (ETR), SVR, Multilayer Perceptron Regression (MLP-R), K-Neighbors Regression (KNN), Hoeffding Tree Regression (HTR), HAT Regression (HAT-R), Logistic Regression (LoR), Decision Tree Classifier (DTC), Extra Tree Classifier (ETC), Random Forest Classifier (RFC), SVC, K-Neighbors Classifier (KNC), Perceptron, Gaussian Naïve Bayes (GNB), Hoeffding Tree Classifier (HTC), and HAT Classifier (HAT-C). The study is conducted in Slovenia. The dataset was generated from government sources and has a size of 2555 data rows estimated from 7 years of daily data. The study predicts LWL using humidity, precipitation intensity, temperature, dew point, precipitation type, pressure, raw hourly weather forecast data, cloud cover and daytime as input variables. The performance of the model is tested using R^2 , RMSE, and MAPE. In their results, the authors claim that the proposed models outperform streaming methods such as standard batch and incremental ML techniques.

M Dawam and Ku-Mahamud (2019) used a normalization and multiple regression model for their LWL prediction study. The study area is Timah Tasoh Reservoir, which is located in Malaysia. The authors developed the model using a dataset consisting of 501 line data with a daily accumulated dataset. The model predicts RWL based on the amount of precipitation and changes in RWL using the sliding window technique. R and R^2 are the performance evaluation methods used in this study. According to the results extracted from the evaluation metrics, the combination of the RF and WC data gives the best input simulation for the multiple regression model.

To predict the water levels of Hongjiadu Reservoir in China, Liu et al. (2017) employed a Multi-Objective Decision Model (MODM) using a Statistical Regression Predictive Function (SRPF). The dataset, which includes a total of 63 entries, was constructed annually. The model uses inflow, water level at the beginning of the year, power generation and power yield as inputs, and outputs the water level of the reservoir. This study uses a number of evaluation metrics, including Mean Relative Errors (MRE), Multiple Correlation Coefficient (MCC), Amount of Forecast Factor (AOFF), Sum of Squares (SS), Residual Standard Deviation (RSD), Residual Sum of Squares (RSOS), Regression Square Sum (RSS), and Degrees of Confidence (DOC). The findings suggest two strategies for managing water resources in the face of their erratic changes. Statistical regression is an approach that can be used in situations where there is a consistent influx. MODM is another technique that can be used in cases of significant contention.

To predict LWLs in Malaysia, Sapitang et al. (2020) used Boosted Decision Tree Regression (BDTR), Decision Forest Regression (DFR), Bayesian Linear Regression

(BLR), and Neural Network Regression (NNR). The 12531 data rows in the dataset were gathered each day. In addition to rainfall and water level, the models can anticipate lake level based on inputs such as water level and rainfall. The performance of the models was evaluated using the metrics MAE, RMSE, R^2 , Residual Absolute Error (RAE), and Residual Standard Error (RSE). The results show that the LWLs can be predicted using any model. However, the evaluation metrics showed that the BLR model performed better than the other models.

Hu et al. (2018)(a) studied the prediction of LWL with the SVR model. The study location is selected in China, where the lake surface area is 2600 km². The estimated size of the data set is 1095, which was accumulated from 3 years of daily data. In order to predict LWL, the study utilizes precipitation data. The evaluation metrics in this study are RMSE and MAE. According to the results, the model is preferable in forecasting without including rainfall data.

Hu et al. (2018)(b) analyzed LWL prediction using the SVR model. The surface area of the lake in this study is 2623 km² and the location is based in China. The dataset consists of 1095 data rows which was estimated from daily data series of 3 years. The study develops a model to predict LWL using flow rate and outflow discharge as input variables. The performance of the model is evaluated using RMSE and R^2 . The results of the study highlights that the historical water levels of the lake are an important variable for predicting the current water level.

Mohammadi et al. (2020) proposed SVR and Support Vector Regression -Grey Wolf Optimization (SVR-GWO) models in order to forecast LWLs. The study area is located between the Peru-Bolivia border and has a surface area of 58 km². The dataset contains 522 data series accumulated monthly. The study forecasts LWL by using the historical LWL values. The performance of the model is tested using RMSE, MAE and R^2 . The results show that there are six scenarios to develop the combination of models. Compared with other methods, the RF preprocessing method provides the best performance in finding the best input combination. In addition, all the models are well-suited forecasting tools that provide 1-month water level forecasts.

Bonakdari et al. (2019) offered the Minimax Probability Machine Regression (MPMR), Relevance Vector Machine (RVM), Gaussian Process Regression (GPR), and Extreme Learning Machine (ELM) models for their LWL prediction study. The case study is applied to a 2402 km² lake on the U. S.-Canada border. The dataset contains 1152 rows of data estimated from 96 years of monthly data collected. Using the LWL with lags as an input variable, the models predict the lake levels. Metrics such as R^2 , MAE, RMSE, Legates and McCabes Index (LMI), Refined Willmott's Index (RWI), and Nash-Sutcliffe Coefficient (NSC) are used to assess the performance of the models. The results show that all models are very good at estimating LWL. The models in the study show that MPMR has the best predictive performance.

The study focuses on a mechanistic model for LLC and combines it with a land surface model to increase the accuracy of the results in the first regression model for 2015 LWL prediction (Lin et al., 2015). However, García Molinos et al. (2015) underline the significance of the seasonality effect in predictions of water level dynamics. A change in water volume in Chad Lake was predicted by Zhu et al. (2017) utilizing a series of remote sensing observations. The study's astounding conclusion suggests that most of the water losses in the entire lake are related to evaporation, while the losses in the southern pool are related to the outflow.

Regression models from the past never made an effort to compare their model to other prediction models. For LWL prediction, Liu et al. (2017) employed the statistical regression function and MODM models. They discovered that the MODM could be most useful in situations where there is a significant discrepancy in reservoir function, but the statistical regression function model is beneficial when the dataset size is large and the lake's inflow is consistent.

Hu et al. (2018)(a) were the first to predict LWL with only one external parameter-precipitation. They discovered that the only information needed to predict water levels is precipitation data, which also helps to avoid both over- and underestimation of precipitation amount and magnitude. The same authors conducted a second experiment with different input settings for the same lake in the same year. They found that historical LWL data is the most important input parameter for LWL prediction. As in the case of Hu et al. (2018)(a), M Dawam and Ku-Mahamud (2019) chose precipitation as the only external input variable; however, the combination of precipitation and historical LWL data provided the best results. They also discovered that normalization of the dataset had a significant impact on MLR results. In addition to these studies, Sapitang et al. (2020) presented two scenario-based forecasting models, one of which included discharge, LWL, and precipitation as inputs. The models were also tested over a range of time horizons, from one day to seven days. The authors discovered that the sole external input that is a reliable predictor of LWL is precipitation, in agreement with the findings of earlier researchers.

Both short- and long-term prediction models were put forth by Castillo-Botón et al. (2020). They discovered that precipitation and dam outputs are less important for prediction in the short term, while upstream and tributary flow are very effective characteristics. Therefore, the results of this study contradict those of Hu et al. (2018) (a) and M Dawam and Ku-Mahamud (2019), who found that using only precipitation data and precipitation & LWL produced the best forecast results. The most thorough regression analysis, conducted by Kenda et al. (2020), used 21 models to forecast the water level and surface area of a single lake. Consequently, the conclusion that batch regression algorithms outperform incremental regression algorithms could only come from this study. Similar to the study by García Molinos et al. (2015), Mohammadi et al. (2020) conducted experiments in which they simply used the LWL as an input parameter. However, they focused on the effect of temporal lag on LWL prediction rather than the

effect of seasonality. They discovered that utilizing 1, 2, 3, or 4 time lags is the only way to get the best results. Regression-based models are summarized in Table 1.

Table 1: *List of regression-based models and their input-output variables, evaluation metrics, and size of data.* Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684.

AI Model	Input	Output	Scale	Evaluation Metrics	No. of Observation	Ref
LR	Direct Precipitation, Evaporation, Inflow, Outflow	LWB	Monthly	R	288	(Zhu et al., 2017)
Combination of RAPID, NOAH-MP LSM and MLR	Inflow Rate, Outflow Rate, Precipitation, and Evaporation	LLC	Monthly	R ² , RMSE, NRMSE	108	(Lin et al., 2015)
BHRM	LWL Height, Flow, and Precipitation	LWL	Monthly	DIC	456	(García Molinos et al., 2015)
SVR and GPs		DWL	Weekly	MAE	2704	(Castillo-Botón et al., 2020)
LR, DTR, RFR, GBR, PLSR, ETR, SVR, MLP-R, KNR, HTR, HAT-R, LoR, DTC, ETC, RFC, SVC, KNC, Perceptron, GNB, HTC, HAT-C	Precipitation Probability, Precipitation Intensity, Precipitation Type, Temperature, Cloud Cover, Dew Point, Humidity, Pressure, and Daytime	LWL	Daily	R ² , RMSE, MAPE	2555	(Kenda et al., 2020)
MLR	Rainfall and Changes in RWL	RWL	Daily	R, R ² , DOC, AOFF, SS, RSOS, RSS, MCC, RSD,	501	(M Dawam and Ku-Mahamud, 2019)
MODM and SRPF	Year-Start Water Level, Inflow, Power Output and Power Generation	RWL	Annual	MRE, MAE, RMSE,	63	(Liu et al., 2017)
BDTR, DFR, BLR and NNR	Rainfall, Water level, and Sent Out	LWL	Daily	R ² , RAE and RSE	12531	(Sapitang et al., 2020)
SVR	Precipitation	LWL	Daily	MAE	1095	(Hu et al., 2018)(a)
SVR	Flow rate, Outflow Discharge, LWL	LWL	Daily	RMSE, R ²	1095	(Hu et al., 2018)(b)

Table 1 (cont.)

SVR and GWO	SVR-	LWL	LWL	Monthly	RMSE, MAE, R ²	522	(Mohammadi et al., 2020)
MPMR, GPR, ELM	RVM,	LWL	LWL	Monthly	R ² , MAE, RMSE, LMI, RWI and NSC	1152	(Bonakdari et al., 2019)

2.1.2. Mathematical Models

Dinka (2020) used the lake capacity curve equation to predict the water level of Lake Basaka in Ethiopia. The lake has a surface area of 500 km². The size of the dataset is 264, accumulated from 22 years of monthly data rows. The study predicts LWL by using lake stage, area and volume as input parameters. The performance of the model was evaluated by several metrics, namely RMSE, MAE, R², and Nash–Sutcliffe Efficiency (NSE). The results of the study indicate that the observed and simulated lake levels are compatible with an NSE value of 0.98 for the monthly basis.

Paul et al. (2019) developed a two-dimensional, depth-averaged model for Lake Victoria water level prediction study. The lake is located between the border of Tanzania, Uganda and Kenya and has a total surface area of 68800 km². The dataset consists of 612 data rows estimated from 51 years of monthly data. The output variable of the dataset is mean water level, while the input variables are determined as evaporation, precipitation, river inflow and outflow. Accuracy was chosen as the metric to evaluate the performance of the model. From the results, it can be inferred that the measurements generated by the numerical model in CM agree well with the calculated water levels.

Li et al. (2013) investigated the prediction of the LWL using mathematical models for the reflection of sunlight and thermal radiation from the lake. Hulun Lake, located in China, was selected as the study area. The lake has a surface area of 2054 km². The authors used a dataset with 41 data rows randomly arranged in time. The model predicts the adjusted water depth by considering the water depth with bands. The evaluation metrics for the performance of the model are R and RMSE. It is concluded from the results that observed water levels are compatible with the predicted values. In addition, the model can be used reliably in cold and arid areas.

In their study, Abbaspour et al. (2012) investigated an unstructured grid Finite-Volume 3D Ocean Model (FVCOM) for the prediction of LWLs. The study is based on Lake Urmia, which is located in Iran. The dataset contains 10950 data rows estimated from 30 years of daily data collection. In this study, the proposed model predicts the LWL using precipitation, evaporation, river flow and discharge as input variables. The performance of the model was tested using RMSE as the evaluation metric. The results reveal that the water level conditions in the area are diminishing. Furthermore, Lake Urmia is increasingly expected to dry up in about 10 years if the dry period in the area continues like this.

In their study, Talsma et al. (2016) proposed a nonlinear predictive control model to predict the water level of Lake IJssel. The lake is located in the Netherlands. The surface areas of the region in Lake IJssel is 1193, 737, and 62 km² for the three lakes used in this study. The number of observations and assessment metrics are not reported in this study. On the other hand, the dataset contains daily LWL data as output and the inflow of the IJssel River, the inflow of the Vecht River, the exchange with the regional water systems, rainfall and evaporation forecasts as input variables. The results indicate that the implementation of predictive models has not only a technical aspect, but also implementation and communication aspects that require additional attention.

In their study, Lofgren and Rouhana (2016) used the Large Basin Runoff Model (LBRM) to forecast the water levels in the Laurentian Great Lakes, which are situated between the US and Canada borders. 684 data rows, which make up the dataset for this study, have been estimated from 57 years' worth of monthly data. With temperature, energy, Priestley-Taylor, and Clausius-Clapeyron adjustments as input variables, the model predicts the level of water in lakes. The model's performance was assessed using statistical significance. Based on evaluation measures, the results show that a 99.98% significance level for the water level prediction may be achieved. The exceptionally high significance value increases ET's sensitivity to air temperature.

Gillies et al. (2015) analyzed Great Salt Lake LWL prediction by developing Observed Model, Tree Ring Model and Tree-ring model 2. The region of lakes resides in the United States. The authors used a dataset that consists of 17 years of data. However, the size of the data is not given. The output variable of the model is change in lake level, while the input variables are change in lake level and tree ring reconstructed change in lake level. The performance evaluation metric of the model is RMSE. The results reveal that using tree-ring reconstructed data in addition to observed data helps reduce the RMSE score and improves predicted results.

Hirsch et al. (2014) explored RWL dynamics by developing a model as a depth-volume relationship. The study area is located in Switzerland. The dataset size is 365, which was estimated from 1 year of daily data. The authors predict water level fluctuation with depth, volume, and the slope of the reservoir basin as input variables. They evaluated the performance of the model using NSC. The results show that the hydro-economic model that the authors developed helps to understand the requirements of environmental water levels and their predicted values.

In their study, Li et al. (2014) used a spatiotemporal pattern model to estimate the water level in lakes. The investigation is conducted in China's Tibetan Lakes. 14600 data rows, estimated using 40 years of daily cumulative data, make up the dataset. With average temperature, solar radiation, precipitation, wind speed, and vapor pressure as input variables, the analysis predicts changes in LWLs. R served as the performance metric for assessing the model. According to the findings, glacier melt has very little impact on the region's predicted water availability; however, permafrost degradation significantly does.

Paynter and Nachabe (2011) investigated the Generalized Extreme Value (GEV) model to predict LWLs. The study covers various lakes, all of which are based in the USA. The dataset size is not given. However, there are four different lakes used in the study. The average data size is 60 out of the annual data accumulation. The model forecasts LWL with maximum and minimum lake levels, as well as flood and drought stages. The performance of the model is evaluated by drawing a quantile-quantile plot. It can be determined from the results that the lakes in the research don't have a significant prediction trend for future flood or drought return levels unless a starting lake stage is obtained.

Bertone et al. (2017) studied the Monte Carlo approach combined with nonlinear threshold autoregressive models built on Matlab software. The study was applied in Australia to a lake that has a surface area of 35.2 km². The authors used a dataset with a size of 624 data rows, estimated from 12 years of weekly accumulated data. The model predicts the storage volume of the reservoir using rainfall, main river inflow, and gross volume variation as input variables. The evaluation metric used in this study is accuracy. The authors claim from the results that the model is beneficial in order to quantify depletion rates, and treatment operators can take different actions based on climate conditions.

Hussain et al. (2021) developed a prediction model of Gravity Recovery and Climate Experiment (GRACE) which is applied in the Indus Basin located in Pakistan. The dataset used in this study has 78 data rows accumulated from 13 years of bimonthly data. The model sets terrestrial water storage as the output variable while evapotranspiration and precipitation as the input variables. The performance of the developed model was tested using Spearman's Rank Correlation (SRC) as an evaluation metric. It is concluded from the results that terrestrial water storage has a decreasing trend, but it is not statistically significant. In addition, the water storage was best predicted by using soil moisture and snow water equivalents as input parameters.

Guinaldo et al. (2021) used Mass-Lake model in order to contribute to the literature in the LWL prediction area. The researchers' study area is located between the Tanzania, Uganda, and Kenya borders. The dataset in this study contains 360 data rows estimated from 30 years of monthly data. The study forecasts prognostic net water storage by using over-lake precipitation, over-lake evaporation, runoff, drainage over the runoff mask, inflow entering the lake from the tributaries, lake outflow, and the contribution of the lake-groundwater fluxes as an input. The performance evaluation metrics to test the model were selected as Kling-Gupta efficiency (KGE), Normalized Information Contribution (NIC), and R. In the results, the model simulated the LWL variations with discharge values, and there is an improvement comparing with the baseline model.

Magyar et al. (2021) proposed Dynamic Factor Analysis (DFA) in order to predict Neusiedlersee LWLs. The lake is located on the Austria-Hungary border. The dataset consists of 180 data rows estimated from 15 years of monthly data. In the model, fluctuation in LWL was determined as an output variable, while evapotranspiration and

precipitation were determined as input variables. The model was tested by using p and MSE as evaluation metrics. According to the results extracted from evaluation metrics, the data analysis tool set used in this study can predict water table variations significantly, which helps to determine the big picture in the case of climate change effects.

Fry et al. (2020) investigated the Net Basin Supply (NBS) model for the Laurentian Great Lakes, which are situated between the USA and Canada border, in order to research the LWL forecast. For the dataset used in this investigation, the size of the data is not provided. In addition, the input is not stated clearly, and the output is chosen as the seasonal water budget. The dataset's temporal span covers three periods, each lasting one, three, and six months. The Heidke Skill Score (HSS) was used to assess the model's performance. The results, according to the authors, indicate that it is possible to anticipate monthly average water levels using a six-month time lag. To improve the model forecast output, however, some expertise and relevant operator selection are needed for the model interpretation.

Jahani et al. (2016) analyzed Chance Constrained (CC) optimization model for the LWL prediction study. The number of observations is not explicitly given for this study, but there are different time horizons. The study predicts reservoir capacity by using the index of season, random inflow to the reservoir, release from the reservoir, and reservoir storage at the beginning of the season as input variables. The model was tested using reliability as an evaluation metric. The results reveal that in the event of inflow limitations on reservoirs in the future, the reservoir water budget will be negatively affected. Furthermore, above a certain threshold level, the inflow information is not helpful to forecast future water levels.

Myakisheva et al. (2021) applied the ARIMA model in order to forecast Lake Ladoga water level, which is located in Russia. The dataset contains 1476 data rows, estimated from 123 years of monthly data. It forecasts LWL by taking historical LWLs into account. The performance of the model was tested using RMSE. Results indicate that the ARIMA model can be applied to large lakes.

Jiang et al. (2021) employed the Discrete Wavelet Transformation-Improved Nonlinear Autoregressive with Exogenous Inputs Network (DWT-INARX) in order to conduct their study. The study area is Taihu Basin, which is located in China. The reservoir has a surface area of 2336.8 km². The dataset in this study consists of 9855 data rows, estimated from 27 years of daily data. The authors used LWL as output while discharge and precipitation were input variables. The performance evaluation metrics in this study are NSE, RMSE, and MRE. The results show that the simulation models produce better accuracy results over a shorter time horizon comparing with longer periods, such as more than 3 days. Among the three models used in this study, DWT-iNARX gives the best performance, while the BP model's performance is not that satisfactory.

Haque et al. (2021) used a 2D hydrodynamic model to predict the LWL in the Inner Niger Delta, Mali. In the model, an estimated 13140 data rows were used, accumulated from 36 years of occurrence in a daily manner. The model also used discharge as an input in order

to predict LWL. The evaluation metrics consist of Relative Error (RE), the NSC, and the RMSE. According to the results extracted from evaluation metrics, hysteresis inclusion improved the simulation results.

Li et al. (2016) (b) developed a reservoir hydrological model for the Miyun Reservoir based in Beijing, China. The reservoir has a surface area of 188 km². The model is developed using a dataset that has an estimated 4745 data rows from 13 years of daily data. It predicts reservoir volume by using gauged surface inflow, ungauged surface inflow, total outflow, surface area, direct precipitation, evaporation, net groundwater source, and unit conversion factor as input. The authors of this study selected RMSE as an evaluation metric. The results reveal that the water level of the reservoir can be predicted as high as 0.93 NSE.

Montroull et al. (2013) studied the Variable Infiltration Capacity (VIC) hydrology model in order to forecast Iberá wetlands' water levels in Argentina. The basin has a total of 12000 km² surface area. The authors developed the model with a dataset that consists of 3650 rows of data with 10 years of daily accumulated data. The only variable in the dataset is LWL. The evaluation metric in this study is Normalized Root Mean Square Error (NRMSE). Results indicate that water levels in Ibera wetlands have a potential to increase in the 21st century. In addition, the increase could be higher during the summer season compared to the winter season.

Croley (2006) investigated forecasting LWL with several models, which are Deterministic Hydrology Forecasts, Probabilistic Hydrology Forecasts, Probabilistic Meteorology Outlooks. The study focused on the Great Lakes Basin area between the U.S. and Canada borders. The dataset was generated from the Great Lakes Environmental Research Laboratory, which includes 444 rows of data among 37 years of monthly dataset. The model predicts LWL by using precipitation, runoff, and evaporation as input. The author tested the model using bias (m), skill, correlation, and RMSE evaluation metrics. It is concluded from the results that the model enables a daily probabilistic outlook that takes advantage of real-time available data.

The Delft3D-Flow model was proposed by Ouni et al. (2020) to forecast and model the water levels of Ichkeul Lake in Tunisia. Data on LWLs and bed roughness make up the model's dataset. Nevertheless, the dataset's size is not specified. By accounting for the roughness of the lake bed, it forecasts the water level in lakes. RMSE is the assessment metric used to assess the Delft3D-Flow model's performance. Based on the RMSE value of 0.027 m, the Delft3D-Flow model exhibits good predictive power for water level.

Haddout et al. (2018) applied the FVCOM model for water level prediction in Aguelmam Sidi Ali Lake, which is located in the Middle Atlas, Morocco. The dataset has a total of 35 years of annual data. The model forecasts LWL by using precipitation, evaporation, and runoff discharges as input variables. The model is tested by using RMSE, MAE, NSE, goodness-of-fit (R^2), and Percent Bias (PBIAS) as evaluation metrics. The results reveal that the model results are compatible with the observed data. In addition, it can be

predicted from the simulation that there is a high probability that the lake will dry in 20 years if the conditions in the region stay the same.

Taminskas et al. (2013) analyzed the lake level prediction topic by developing the Finite Element Subsurface FLOW (FeFLOW) 5.0 model. The study area focuses on Lake Cedasas, based in northeast Lithuania. The authors work on a dataset that tries to predict LWL with changes in water balance due to climate change variable. The dataset size and evaluation metrics are not given for this study. The dataset temporal is determined at random. The results show that a mathematical model can be used to extract a simulation of surface water resources.

Rodríguez-Rodríguez et al. (2012) explored the prediction model of the Water Budget Conceptual Model (WBCM), which is applied in Medina Playa Lake located in Southwest Spain. The surface area of the lake is 14.78 km². The dataset contains 72 data rows out of 6 years of monthly data collection. In the model, LWL is predicted by direct precipitation onto the lake's surface, groundwater flow, base flow in streams, and subsurface runoff as input variables. The evaluation metric for the model is R². It can be inferred from the results that the water level can be predicted with the proposed model with some discrepancies.

Zappa et al. (2014) employed the Precipitation-Runoff-Evapotranspiration Hydrotope (PREVAH) model in order to assess the surface water level in Switzerland. The number of observation and evaluation metrics is not given in this study. The model took precipitation, evaporation, soil moisture, litter moisture, water temperature, ground water level, runoff, and snow water equivalent into account in order to predict surface water level. The results reveal that prototyping and experience on tests give extra benefit for early drought recognition in Switzerland.

Person et al. (2007) applied the three-dimensional Surface-Water-Groundwater Model (SWGM) for Crow Wing Watershed located in Minnesota, USA. The total surface area of the basin is 14,000 km². The dataset consists of 600 data rows accumulated from 50 years of monthly data. The authors predict LWL fluctuations by using input variables such as runoff, evapotranspiration, infiltration, streamflow, and groundwater hydrodynamics. The evaluation metric is determined as the accuracy between simulation and observed data. It is concluded from the results that in the short-term climate timeframe, upland parts of the watershed fluctuate higher compared with the lowest parts.

Voulanas et al. (2021) used the FeFLOW model in order to forecast the water level for the Kastoria basin located in Western Macedonia, Greece. The basin has a total surface area of 33 km². The dataset has a total of 216 data rows, estimated from monthly 18-year data. The input variables of the dataset include rainfall, surface runoff, evaporation, discharge, or the flow volume at the basin's outlet, while the output variable is LWL. The performance of the model is tested by using evaluation metrics such as RMSE, R, and Standard Deviation (SD). The authors claim from the results that the diminishing effect

of sub-surface flow and direct precipitation recharge and rise in evaporation is going to affect the total outflow of the basin in a negative way in future projections for all models.

Chen et al. (2020) investigated the Water Balance Model (WBM), based on GR4J, in order to forecast water levels in Thirlmere Lakes National Park, located in New South Wales, Australia. The surface areas of lakes are 11.5, 4.7, 10.2, 12.8, and 10.1 for the five lakes in the region. The size of the dataset is not given for this study. However, the variables in the dataset include precipitation, runoff, and evapotranspiration as inputs and change in lake storage as output. The authors used R^2 as an evaluation metric to test the model. The model is developed using a MATLAB software program. The results in this study highlight that the ground-filtering approach improves the performance of lake level prediction by as much as 70% thanks to its ability to reduce the vertical error in the model.

Mtilatila et al. (2020) studied water balance for water level prediction in Lake Malawi and the Shire River, which are located in Tanzania. The water surface of the area is 29,600 km². The dataset size is 516, which is estimated from 43 years of monthly accumulated data. The model estimates LWL change by using precipitation and evaporation variables. The evaluation metric to test the model's performance is R. The results indicate that increases in temperature and decreases in precipitation significantly increase drought conditions in the region.

Cai et al. (2016) developed WBM in order to use it on Hulun Lake, located in China. The surface area of the lake is around 2000 km². The dataset contains 372 data rows accumulated out of 31 years of monthly data. In the model, LWL is forecasted by taking precipitation, mean temperature, wind speed, relative humidity, sunshine duration from 1960 to 2014, river discharge, and evaporation into account. The performance of the model is tested by calibration. In the results, it can be inferred that the model is capable of forecasting fluctuations in LWL.

Ricko et al. (2011) analyzed WBM in various tropical lakes located in Nigeria. The model predicts LWL by only using the rainfall variable as an input. The dataset consists of 5840 data rows with an accumulated 16 years of daily data. The evaluation metrics for the model are selected as R and RMSE. According to the results, the output of this research can help predict weather in the medium term.

Ahn et al. (2016) proposed the Water Balance Network Model (MODSIM) and a Watershed-Scale Hydrologic Model (WSHM)-based Soil and Water Assessment Tool (SWAT) for LWL prediction. The study is done in the Geum River Basin, located in South Korea. The basin has a surface area of 9645.5 km². The dataset size is 2190, which is estimated from 6 years of daily data. In the model, DWL is predicted by taking air temperature, precipitation, relative humidity, wind speed, and sunshine hours as input variables. The study used R^2 , the NSE, and the RMSE as evaluation metrics. The results show that water shortage is expected to be 38.2% in the 2040s, 38.2% in the 2080s for the Representative Concentration Pathway (RCP) 4.5 scenario, 21.3% in the 2040s, and 22.1% in the 2080s for the RCP 8.5 scenario.

Morgan et al. (2019) investigated WBM in order to predict LWLs in Australia. The dataset size and evaluation metrics are not given in this study. However, the temporal of the data is “daily”. In addition, the WBM predicts LWL by using annual average runoff volume, equilibrium pit lake surface area, predicted equilibrium water level, and WBM equilibrium water level as input variables. The results of the study indicate that the water level of lakes is highly sensitive in terms of evaporation and runoff.

Some academics preferred the WBM model over all other mathematical models, which improves the model's accuracy performance. Ricko et al. (2011) made the first attempt at WBM by using various rainfall data to modify input parameters while keeping other model input parameters constant. On the other hand, Rodríguez-Rodríguez et al. (2012) ran the model by including groundwater flow, baseflow in streams, and subsurface runoff as input factors in addition to rainfall data. Ahn et al. (2016) built their model using a daily dataset that included inputs for air temperature, precipitation, relative humidity, wind speed, and daylight hours. With only 372 data rows in their dataset, Cai et al. (2016) were still able to forecast LWL using the input parameters of precipitation, mean temperature, wind speed, relative humidity, sunshine duration, river discharge, and evaporation. The only study that never used rainfall as an input parameter was Morgan et al.'s (2019). Other researchers (Chen et al., 2020; Mtilatila et al., 2020) predicted LLC using precipitation, runoff, and evapotranspiration as input factors.

As specific equational models, researchers created various models. Despite the fact that various models share input parameters, some generate unique parameters solely for their models. Li et al. (2014) employed vapor pressure as an input variable, while Gillies et al. (2015) used tree ring reconstructed change in lake level. On the other hand, Ouni et al. (2020) choose to include lake bed roughness as one of their input variables. Mathematical and Equation-based models are presented in Table 2.

Table 2: *List of mathematical models and their input-output variables, evaluation metrics, and size of data.* Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684.

AI Model	Input	Output	Scale	Evaluation Metrics	No. of Observation	Ref
Lake Capacity Curve Equation Two-Dimensional Depth-Averaged Model	Lake Stage, Area, Volume	LWL	Monthly	RMSE, MAE, NSE, R ²	264	(Dinka, 2020)
	Evaporation, Precipitation, River Inflow and Outflow	Mean Water Level	Monthly	Accuracy	612	(Paul et al., 2019)

Table 2 (cont.)

Sunlight Reflection and Lake Thermal Radiation	Water Depth with Bands	Adjusted Water Depth	Random	R, RMSE	41	(Li et al., 2013)
FVCOM Non-Linear Model	Precipitation, Evaporation, River and Runoff Discharge	LWL	Daily	RMSE	10950	(Abbaspour et al., 2012)
Predictive Control	Inflow, Exchange with Water Systems, Rainfall and Evaporation	LWL	Daily	Not Given	Not Given	(Talsma et al., 2016)
LBRM Observed Model, Tree Ring Model, Tree-Ring Model 2	Temperature Adjustment, Energy Adjustment, Priestley–Taylor, Clausius–Clapeyron	LWL	Monthly	Statistical Significance	684	(Lofgren and Rouhana, 2016)
Depth-Volume Relationship	LLC, Tree Ring Reconstructed LLC	LLC	Annual	RMSE	17	(Gillies et al., 2015)
Spatio-Temporal Pattern	Depth, the Volume and the Slope of the Reservoir Basin Mean Temperature, Precipitation, Solar Radiation, Wind Speed, and Vapor Pressure	Water Level Fluctuation	Monthly	NSC	365	(Hirsch et al., 2014)
GEV	Maximum and Minimum LWL, the Flood and Drought Stages.	LWL	Annual	Quantile-Quantile Plot	60	(Li et al., 2014)
Monte Carlo	Rainfall, Main River Inflow, and Gross Volume Variation	Storage Volume	Weekly	Accuracy	624	(Paynter and Nachabe, 2011)
GRACE	Precipitation, Soil Moisture and Snow Water Equivalent	Terrestrial Water Storage	Two monthly	SRC	78	(Bertone et al., 2017)
Mass-Lake Model	Precipitation, Evaporation, Runoff, Drainage, Inflow, Outflow, and the Contribution of the Lake–Groundwater Fluxes	Prognostic Net Water Storage	Monthly	KGE, NIC, R	360	(Hussain et al., 2021)
DFA	Evapotranspiration and Precipitation	Fluctuation in LWL	Monthly	p, MSE	180	(Guinaldo et al., 2021)
NBS CC Optimization Model	Not Explicit	Seasonal Water Budget	1,3 and 6 Months	HSS	Not Given	(Magyar et al., 2021)
ARIMA	Index of season, Inflow, Release from Reservoir, and Reservoir Storage	Reservoir Capacity	Monthly	Reliability	Not Given	(Fry et al., 2020)
	LWL	LWL	Monthly	RMSE	1476	(Jahani et al., 2016)
						(Myakishev et al., 2021)

Table 2 (cont.)

DWT-iNARX 2D Hydrodynamic Model	Discharge and Precipitation	LWL	Daily	NSE, RMSE, and MRE	9855	(Jiang et al., 2021)
Reservoir Hydrological Model	Discharge Gauged Surface Inflow, Ungauged Surface Inflow, Total Outflow, Surface Area, Precipitation, Evaporation, Net Groundwater Source, Unit Conversion Factor	LWL	Daily	RE, NS and RMSE	13140	(Haque et al., 2021)
VIC Deterministic and Probabilistic Hydrology Forecasts Delft3D-Flow	LWL	LWL	Daily	NSE	4745	(Li et al., 2016)(b) (Montroull et al., 2013)
	Precipitation, Runoff, Evaporation	LWL	Monthly	Bias, Skill, Correlation, RMSE	444	(Croley, 2006)
	Lake Bed Roughness	LWL	Not Given	RMSE, MAE, NSE, R ² and PBIAS	Not Given	(Ouni et al., 2020)
FVCOM	Precipitation, Evaporation, Runoff Discharges	LWL	Annual		35	(Haddout et al., 2018)
FeFLOW	Change in Water Balance due to climate change	LWL	Random	Not Given	Not Given	(Taminkas et al., 2013)
WBCM	Precipitation, Groundwater Flow, Baseflow in Streams and Subsurface Runoff	LWL	Monthly	R ²	72	(Rodríguez-Rodríguez et al., 2012)
PREVAH	Precipitation, Evaporation, Soil Moisture, Litter Moisture, Water temperature, Ground Water Level, Runoff, Snow Water Equivalent	Water Level Surface Water	Changes according to Variable	Not Given	Not Given	(Zappa et al., 2014)
SWGM	Runoff, Evapotranspiration, Infiltration, Streamflow and Groundwater Hydrodynamics	Lake-Level Fluctuations	Monthly	Accuracy	600	(Person et al., 2007)
FeFLOW Water Balance GR4J Model	Rainfall, the Surface Runoff, Evaporation and Discharge or the Flow Volume at the Basin's Outlet	LWL Change in Lake Storage	Monthly	RMSE, R, SD	216	(Voulanas et al., 2021)
	Precipitation, Runoff, and Evapotranspiration	Lake Storage	Monthly	R ²	Not Given	(Chen et al., 2020)
WBM	Precipitation and Evaporation	LLC	Monthly	R	516	(Mtilatila et al., 2020)

Table 2 (Cont.)

WBM	Precipitation, Mean Temperature, Wind Speed, Relative Humidity and Sunshine Duration, River Discharge, Evaporation	LWL	Monthly	Calibration	372	(Cai et al., 2016)
WBM	Precipitation, Evaporation, Effective Catchment Area, Lake area, and time lag	LWL	Daily	R, RMSE	5840	(Ricko et al., 2011)
MODSIM, WSHM	Air Temperature, Precipitation, Relative Humidity, Wind Speed and Sunshine Hours	DWL	Daily	R ² , NSE, RMSE	2190	(Ahn et al., 2016)
WBM	Runoff Volume, Equilibrium Surface Area, Equilibrium Water Level, WBM Equilibrium Water Level	LWL	Daily	Not Applicable	Not Given	(Morgan et al., 2019)

2.1.3. ANN Models

Mislan et al. (2018) developed Adaptive Neural Network Backpropagation (ANNBP) in their study. The study area is Lake Cascade Mahakam, which is located in Indonesia. The dataset consists of 1008 data rows. It forecasts max LWL, min LWL, and average LWL by using historical max LWL, min LWL, and average LWL as input variables. The evaluation metrics for the performance of the model are MSE and MAPE. According to the results extracted from evaluation metrics, the 240-5-1 NNBP architecture reached a 9.7% MSE score, which indicates the model can be used as a prediction method for future LWLs.

Wang et al. (2018) studied the combination of Copula Entropy (CE) with Wavelet Neural Network (WNN) in order to work on the LWL study. The study area was determined to be the Taihu Lake Basin, which is located in China. The dataset contains 3650 data rows, which are estimated from 10 years of daily data. In this study, the output is determined as LWL, while the inputs are rainfall and LWL. The authors used RMSE as an evaluation metric to test the model. The results reveal that the hybrid model consisting of WNN and CC surpassed CE ANN, Linear Correlation (LCC) WNN and LCC ANN models, which contribute to future decision-making during water level fluctuations.

For a study on LWL prediction, Young et al. (2015) looked into three-dimensional hydrodynamic models, ANN, Auto Regressive Moving Average with Exogenous Input (ARMAX), and combined hydrodynamic and ANN. The study is conducted in a Taiwanese Alpine Lake. There are 7296 observations in the dataset, with hourly temporal resolution. Using precipitation, outflow discharge, and input discharge, the model forecasts the level of water in lakes. The R, RMSE, and MAE are used to assess the model's performance. The hydrodynamic model can forecast water levels during the calibration step but not during the validation stage, according to the results. When

compared to the hydrodynamic model, the ANN and ARMAX models performed better than the others.

Mpallas et al. (2011) used ANFIS and ANN in their study on the prediction of water levels in Kerkini Lake. The territory is limited to the Greek portion of the lake, which has a surface area of 6472 km². The 252 data rows in the study's dataset were estimated using monthly data collected over a 12-year period. The models are created with Visual Basic computer tools and use rainfall, lake evaporation, evapotranspiration in the Strymonas basin, water flow from Bulgaria, and water consumption from human activity as input variables to predict the level of water in lakes and runoff. Reduced-MSE and R were used to test the model's performance. The study's models are deemed to be adequate predictive instruments for predicting lake levels based on the findings. The models' performance outcomes are fairly comparable to one another.

Jaafar et al. (2010) analyzed nonlinear ANN in order to study LWL prediction. The study area was determined to be Sungai Gumum, Tasik Chini Pahang, which is located in Malaysia. The lake has a surface area of 49.8 km². The size of the dataset is not given, but the data was collected monthly. The model forecasts LWL by using LWL with lags as an input variable. The evaluation metric to test the model's performance is R². According to the results that are extracted from evaluation metrics, 89% of water fluctuation can be predicted by using an ANN model with streamflow as an independent variable.

NN and ANFIS were investigated by Páliz Larrea et al. (2021) in their study on LWL prediction. The study is conducted in the northern Ecuadorian reservoir of Salve Faccha. 2920 data rows, calculated from 8 years of daily data, make up the dataset used in this study. In this study, rainfall is the input and LWL is the output, which is determined 1-6 days in advance. R, the Nash Index, and the RMSE were used to assess the model's performance. The findings demonstrate that rainfall is not a reliable indicator of RWL. Evaluation metric results indicated that the highest performing models were the NN with t+4 and the ANFIS with t+6.

In order to apply algorithms using the Radial Basis Function (RBF) in LWL prediction studies, Latif et al. (2021) proposed ANN. The research area used by the authors was the Malaysian Klang Gate Reservoir. The estimated 132 data rows in the dataset come from 11 years of monthly data accumulation. Using inflow, dam release, and reservoir beginning and end storage, the model forecasts water losses from the reservoir. As a performance statistic, RMSE was used to assess the model's performance. The findings show that the ANN model has an RMSE score of 20.07%, which is sufficient to estimate water levels. The model offers additional information on water losses, final storage, and variations in water level, which is beneficial for reservoir operations.

Üneş et al. (2015) used ANN for the LWL prediction study. The study uses a lake with a surface area of 53.45 km², which is located in the U.S. The size of the dataset is 2272, accumulated from daily data. The model determines LWL by using LWL with different time lags as an input variable. The performance of the model is evaluated using MSE,

Mean Absolute Relative Error (MARE), and R. It can be inferred from the results that ANN models predict water levels better than conventional prediction models.

Piasecki et al. (2018) also studied ANN in order to effectively forecast LWLs. The study uses a small glacial lake in Poland with a surface area of 1.16 km². The dataset consists of monthly accumulated data; however, the dataset size is not given. The model forecasts LWL by using water level, binary variables, evaporation, and precipitation as input variables. The model's performance was evaluated using MAE, RMSE, MSE, R², and MAPE. The authors claim from the results that the ANN prediction of the LWL prediction is satisfactory. The model's performance can be increased even more using the WT preprocessing method. However, meteorological variables couldn't forecast water fluctuations according to insignificant results.

Piasecki et al. (2017) employed ANN, MLP, and MLR to predict LWLs in their study. The authors use Lake Serwy in northeastern Poland with a surface area of less than 100 km² in this study. The dataset contains 11680 data rows, which are estimated from 32 years of daily accumulated data. The model uses LWL as an output variable, while maximal and minimal temperature (Tmax, Tmin), wind speed, vertical circulation, and water level from previous periods as input variables. The performance evaluation metrics of the model are MAPE, RMSE, R², and Mean Biased Error (MBE). In the results, the model performed well in small lakes based in temperate climate locations. The model outperformed MLR models according to evaluation metrics.

Ashaary et al. (2015) investigated NN in their LWL prediction study. The size of the dataset is 5779, accumulated from daily data. The model predicts RWL by using historical RWL as an input variable. The model's performance was evaluated using MSE. It can be determined from the results that the best performance of prediction was achieved when 2 days of delay were used and the architecture of the model was set as 4-17-1.

Ishak et al. (2011) created a NN-based decision support model to facilitate the investigation of LWLs. 3041 data rows total, gathered from daily data, make up the dataset. The model uses rainfall as an input variable to anticipate RWL. This study used training, testing, and validation errors as evaluation metrics. The findings show that the NN model in this study did well in decision models as well as predicting.

Most of the time (Mislán et al., 2018; Wang et al., 2018; Jaafar et al., 2010; Üneş et al., 2015; Piasecki et al., 2018; Ashaary et al., 2015), the authors that predicted LWL with ANN models chose LWL as their input variable. Other metrics used in the studies by Wang et al. (2018), Young et al. (2015), Mpallas et al. (2011), Páliz Larrea et al. 2021, Latif et al. 2021, Piasecki et al. (2018), and Piasecki et al. (2017) included rainfall, discharge, evaporation, temperature, and wind speed alone or in combination with other parameters.

Table 3: List of ANN models and their input-output variables, evaluation metrics, and data size. Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684.

AI Model	Input	Output	Temporal Scale	Evaluation Metrics	No. of Observation	Ref
ANNBP	Max LWL, Min LWL, Average LWL	Max LWL, Min LWL, Average LWL	Monthly	MSE, MAPE	1008	(Mislan et al., 2018) (Wang et al., 2018)
CE, WNN Hydrodynamic Model, ANN, ARMAX and Combined Hydrodynamic and ANN	Rainfall and LWL Inflow Discharge, Outflow Discharge, Precipitation, Rainfall, Lake Evaporation, Strymonas Basin Evapotranspiration, Water Flow and Water Consumption	LWL	Daily	RMSE	3650	(Young et al., 2015)
ANN, ANFIS		LWL and Runoff	Hourly	MAE, RMSE and R	7296	(Mpallas et al., 2011) (Jaafar et al., 2010) (Páliz Larrea et al., 2021)
Nonlinear ANN	LWL with Lags	LWL	Monthly	MSE and R	252 Not Given	(Jaafar et al., 2010) (Páliz Larrea et al., 2021)
NN and ANFIS	Rainfall	LWL	Daily	R ² , R, Nash Index and RMSE	2920	(Páliz Larrea et al., 2021)
ANN	Inflow, the Release of Dam, Initial and Final Storage of the Reservoir	Water Losses from the Reservoir	Monthly	RMSE, MSE, MARE and R	132	(Latif et al., 2021)
ANN	LWL with Different Time Lags	LWL	Daily	MAE, RMSE, MSE, R ² , MAPE	2272	(Üneş et al., 2015)
ANN	LWL, Binary Variables, Evaporation, Precipitation	LWL	Monthly	MSE, R ² , MAPE	Not Given	(Piasecki et al., 2018)
ANN, MLP and MLR	Maximal and Minimal Temperature, Wind Speed, Vertical Circulation and LWL	LWL	Daily	MAPE, RMSE, R ² , MBE	11680	(Piasecki et al., 2017) (Ashaary et al., 2015)
NN	RWL	RWL	Daily	MSE	5779	(Ishak et al., 2011)
NN	Rainfall	RWL	Daily	Testing and Validation Error	3041	(Ishak et al., 2011)

Although NN or ANN was the only model used in the majority of studies to predict the outcome, some researchers added additional models to compare it to the ANN model, such as CE, ARMAX, ANFIS, MLP, MLR, and the hydrodynamic model (Wang et al., 2018; Young et al., 2015; Mpallas et al., 2011; Páliz Larrea et al., 2021; Ashaary et al., 2015). Young et al. (2015) discovered that, when compared to the hydrodynamic model, the MAE and RMSE results from the ANN and ARMAX models were lower. However, Mpallas et al. (2011) and Páliz Larrea et al. (2021) showed that when they compared the ANN and ANFIS models, the results were remarkably similar, indicating that the models may be utilized interchangeably. ANN models are presented in Table 3.

2.1.4. *Decision Tree Models*

An RF model was suggested by Guyennon et al. (2021) for their study on LWL prediction. In central Italy, near Lake Bracciano, the study is conducted. The dataset consists of 812 rows of monthly data accumulation. Precipitation, temperature, surface evaporation, wind speed, relative humidity, atmospheric pressure, sun radiation, and abstraction are used as input variables in the model to estimate LWL. RMSE and MAE are the performance assessment methods employed in this study. The assessment measures' results show that the RF model can accurately predict LWL despite the incompleteness of the data under study. The study's air temperature and short-term precipitation are yet unfinished. On the other side, the water level varied the most during prolonged rains.

In order to estimate future LWLs, Wang and Wang (2020) investigated the GP, MLR, MLP, M5 Pruned (M5P) model tree, RF, and K-Nearest Neighbor (KNN). They conducted their experiment at the U.S.-Canada border using the Lake Erie dataset. 4380 data rows are expected to be included in the dataset based on 12 years of daily data. Precipitation, air temperature, shortwave and longwave radiation, wind speed, and relative humidity are used as input variables in the study to forecast LWL. RMSE and MAE were the evaluation metrics employed in this investigation. According to the findings, MLR and M5P have the best performance ratings when compared to evaluation criteria for process-based models.

In order to test several models on the LWL prediction area, Nhu et al. (2020) evaluated the M5P, RF, Random Tree (RT), and Reduced Error Pruning Tree (REPT) models. The study is taking place in Iran's 8.9 km² Zrebar Lake, which is where the study is taking place. Using daily data over six years, the dataset size was calculated to be 2190. Previous-day water level (t-1), (t-1)+(t-2), (t-1)+(t-2)+(t-3), (t-1)+(t-2)+(t-3)+(t-4), and (t-1)+(t-2)+(t-3)+(t-4)+(t-5) are used as input variables in the study to forecast LWL. RMSE, MAE, R², PBIAS, the ratio of the RMSE to the SD of measured data (RSR), and visual frameworks (Taylor diagram and box plot) were used to assess how well the study performed. Results show that the one-day lag input produced the best performance. Additionally, as the lag time increased, the performance of the outcome declined. The M5P model performed the best among the models, while the REPT model performed the poorest.

Obringer et al. (2018) examined a variety of models to predict urban RWL in the USA, including the Generalized Linear Model (GLM), Generalized Additive Model (GAM), Multivariate Adaptive Regression Spline (MARS), Classification and Regression Trees (CART), Bagged CART, RF, SVM, Bayesian Additive Regression Trees (BART), and Null (Mean-Only) model. A 51-year period of daily data was used to estimate the dataset's 18615 data rows. In this study, streamflow, dew point, population, soil moisture, ENSO, discharge, humidity, water use, and precipitation are used as the inputs, and RWL is determined as the output. R^2 was used as an evaluation statistic to test the model. The findings indicate that RF is the most accurate predictive model for predicting LWL. The main factors influencing the outcome are streamflow, city population, and the El Nio/Southern Oscillation (ENSO) index.

Table 4: *List of decision tree models and their input-output variables, evaluation metrics, and data size.* Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684.

AI Model	Input	Output	Temporal Scale	Evaluation Metrics	No. of Observation	Ref
RF	Precipitation, Temperature, Surface Evaporation, Wind Speed, Relative Humidity, Atmospheric Pressure, Solar Radiation, Abstraction	LWL	Monthly	RMSE, MAE	812	(Guyennon et al., 2021)
GP, MLR, MLP, M5P, RF, KNN	Precipitation, Air Temperature, Shortwave Radiation, Longwave Radiation, Wind Speed, and Relative Humidity	LWL	Daily	RMSE, MAE	4380	(Wang and Wang, 2020)
M5P, RF, RT, REPT	LWL	LWL	Daily	RMSE, MAE, R^2 , PBIAS, RSR, Taylor Diagram and Box Plot	2190	(Nhu et al., 2020)
GLM, GAM, MARS, CART, Bagged CART, RF, SVM, BART	Streamflow, Dew Point, Population, Soil Moisture, ENSO, Discharge, Humidity, Water Use, Precipitation	RWL	Daily	R^2	18615	(Obringer et al., 2018)
RF, ANN, SVR, LR	LWL, Discharge and Time Lags	LWL	Daily	R^2 , RMSE	20805	(Li et al., 2016)(a)
CatBoost	Soil Texture, Geology, Topography Based Characteristics, Water Body Proximity, Land Cover, and Outputs from a Hydrological Simulation with the DK-Model	Depth of the Upper most Water Table	Daily	MAE	10950	(Koch et al., 2021)

In their study, Li et al. (2016) (a) used RF, ANN, SVR, and LR to apply LWL prediction. In China's Poyang Lake, where the study is conducted. In the dry season, the lake's surface area is 1000 km², whereas during the rainy season, it is 4000 km². Based on 57 years'

worth of daily data, the dataset has an estimated 20805 data rows. Using water level, outflow, and time lags as input variables, the study forecasts LWL. The R^2 and RMSE measures were used to assess the model's performance. The findings of the evaluation measures showed that RF performed the best for LWL prediction. The Yangtze River's discharge and the prior water level were found to be the most useful variables for the model when the variable's importance was also examined.

In their LWL prediction investigation, Koch et al. (2021) examined CatBoost (Gradient Boosting Decision Tree). The case study is utilized in a lake with a 43000 km² surface area in Denmark. Using daily data acquired over 30 years, the dataset size of 10950 is approximated. In this study, the inputs are soil texture, geology, topography-based characteristics, proximity to water bodies, land cover, and results from a hydrological simulation using the DK model, with the output being the depth of the topmost water table. MAE is the model's performance evaluation statistic. The findings demonstrate the proposed model's high accuracy in predicting water table variability. Decision tree based models are presented in Table 4.

2.1.5. Fuzzy Logic Models

ANFIS-SO, ANFIS-FA, ANFIS-PSO, MLP-SO, MLP-FA, and MLP-PSO models were created by Ehteram et al. (2021) to examine the LWL prediction area. An Iranian lake was chosen as the study region. 168 rows of data are calculated in the dataset for this study using 14 years of monthly data. Temperature and precipitation, which are lagged Seven input combinations, according to PCa, are used as input variables in the study to predict LWL. RMSE is the performance assessment metric used in this investigation. The results indicate that the ANFIS-SO model, when combined with rainfall and temperature inputs, produced the best results. On the other hand, the input variable representing the rainfall lag periods across six months had poor model performance.

In their investigation on LWL prediction, Üneş et al. (2019) employed ANFIS, SVM, RBNN, and Generalized Regression Neural Networks (GRNN). The investigation is conducted in a 70-km² lake in the United States of America. 2272 data rows total, gathered from daily data, make up the dataset. The study uses RWL with lag durations as an input variable to estimate future RWL. The MSE, MAE, and R metrics were used to assess the model's performance. The results suggest that when it comes to forecasting RWL, ANFIS models outperform auto-regressive models (AR), auto-regressive moving averages (ARMA), MLR models, and AI models.

In order to investigate upcoming LWLs, Tsao et al. (2021) used Fuzzy Neural Networks (FNN) with a multi-stage architecture. The investigation is being conducted at Taiwan's Techihydropower plant, which has a surface area of 1235.73 km². The estimated data size from two years of hourly cumulative data is 17520. The study's input comprises meteorological rainfall data, rainfall observation data, water level, and power generation, while the output is decided by reservoir inflow and water level. The model's performance is evaluated using the following metrics: MAE, RMSE, and MSE. Based on the study's

findings, the authors assert that the model might efficiently use water resources. Given that the model performs well in heavy rains, it may also have a good effect on the management of water plants.

Valizadeh et al. (2014) studied the LWL prediction area using ANFIS with membership functions. A lake in Malaysia with a surface area of 1290 km² is the subject of the case study. Estimated from 11 years of daily data, the dataset has 4015 data rows. Using rainfall and LWL as input variables (with varying time delays for each), the study forecasts LWL. The MAPE, RMSE, and MAE were used to assess the model's performance. The results show high model performance because there is a strong fit between the actual and predicted data. Fuzzy logic based models are presented in Table 5.

Table 5: *List of fuzzy logic models and their input-output variables, evaluation metrics, and data size.* Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684.

AI Model	Input	Output	Temporal Scale	Evaluation Metrics	No. of Observation	Ref
ANFIS, MLP	Temperature and Rainfall	LWL	Monthly	RMSE	168	(Ehteram et al., 2021)
ANFIS, SVM, RBNN, GRNN	RWL	RWL	Daily	MSE, MAE and R	2272	(Üneş et al., 2019)
FNN	Meteorological Rainfall Data, Rainfall Observation Data, Water Level and Power Generation	Reservoir Inflow and RWL	Hourly	MAE, RMSE, MSE	17520	(Tsao et al., 2021)
ANFIS	Rainfall and LWL	LWL	Daily	RMSE, MAPE, MAE	4015	(Valizadeh et al., 2014)

2.1.6. Deep Learning Models

In order to carry out their experiment on the LWL prediction study, Liang et al. (2018) examined Long Short Term Memory (LSTM). The Three Gorges Dam in China was chosen as the study area. 4015 data rows, calculated from ten years of daily data, make up the dataset. The daily average precipitation, daily average discharge at Xiangtan Station, daily average discharge at Taojiang Station, daily average discharge at Taoyuan Station, daily average discharge at Jinshi Station, and daily average TGD discharge are used as input variables in the study to forecast LWL. R² and RMSE were the evaluation measures used to assess the performance of this investigation. The findings indicate that the LSTM DL model performs better than the SVM model.

The Convolutional Neural Network (CNN) was studied by Damova et al. (2020) to estimate future LWLs. Three lakes in Bulgaria were the subject of the study: Kyrdjali, Studen Kladenec, and Ivaylovgrad. The dataset's estimated 3650 rows were gathered over

a ten-year period from daily data accumulation. The study uses soil moisture, air temperature, skin temperature, vegetation index, precipitation (liquid and solid), and in-situ measurements of water balance parameters as input variables to forecast RWL. Using MinMAE, the study's performance was assessed. The suggested model works better than independent GIS models and non-semantic solutions, according to the results.

In order to predict LWLs located in the USA, Costa Nogueira Jr. et al. (2021) employed LSTM and Fully Connected Neural Networks (FCNN). The temporal extent of the accumulated data is daily; however, the magnitude of the data is unknown. In this study, temperature, density, and the eastward and northward water velocity are the inputs, and the output is calculated as LWL. An average error was employed as the evaluation metric in this research. The results show that both models produce good enough results to allow for accurate LWL prediction. The performance of the LSTM and FCNN models was comparable to one another. LSTM requires additional learning parameters in order to achieve comparable outcomes. DL based models are presented in Table 6.

Table 6: List of deep learning models and their input-output variables, evaluation metrics, and data size. Adapted from A systematic literature review on Lake water level prediction models, by Ozdemir et al., 2023, *Environmental Modelling & Software*, 105684.

AI Model	Input	Output	Temporal Scale	Evaluation Metrics	No. of Observation	Ref
LSTM	TGD Discharge, Daily Average Precipitation	LWL	Daily	R ² , RMSE	4015	(Liang et al., 2018)
CNN	Precipitation, Soil Moisture, Air Temperature, Skin Temperature, Vegetation Index, and In-situ Measurements of Water Balance Characteristics	RWL	Daily	MinMAE	3650	(Damova et al., 2020)
FCNN LSTM	Temperature, Density, and Northward and Eastward Water Velocity	LWL	Daily	Average Error	Not Given	(Costa Nogueira Jr et al., 2021)

2.2. Overview of LWL Data-Driven Modeling

All of the reviewed research made use of mathematical models. In order to predict lakes or RWLs, the researchers either developed their own formula or produced a combination of other mathematical models. Generally speaking, historical LWLs are used in feature selection. Multivariate features were employed in several studies to forecast future water levels. Furthermore, RMSE, R², and MAE are often used as evaluation metrics and account for roughly 46% of all SLR research.

Studies provide useful data for measuring the water supplies in drinking water reservoirs, but they do not integrate the procedures used to make decisions about the water management system with the prediction of water supplies. Most of the study's attention is on making predictions about floods or droughts.

Depending on when they were published, the modeling methodologies change. While earlier studies used mathematical (Person et al., 2007; Paynter and Nachabe, 2011) and regression-based models (Lin et al., 2015; García Molinos et al., 2015), more recent studies typically concentrate on ML (Guyennon et al., 2021; Koch et al., 2021) and DL techniques (Damova et al., 2020; Costa Nogueira Jr. et al., 2021). The significant improvement is the result of recent advances in DL and ML approaches. In order to predict LWL and RWL for future periods, research on ML and DL techniques is still open.

Though some researchers select particular algorithms for their research, these algorithms are chosen based on a variety of requirements, taking into account both their benefits and drawbacks.

2.3. Effects of Climate Change on Lakes

Singh et al. (2019) investigated urban lakes in India because urban lakes are rapidly becoming vulnerable to water budget and quality due to urbanization, climate change, and anthropogenic pollutant inputs. To safeguard lakes in the long run, it is important to accurately identify and treat the impact of these driving variables on their hydrology. They looked at Sukhna Lake, a northern Indian urban lake that has recently seen recurrent drying up. Many theories, including those involving anthropogenic activities, climate change, changes in land use, and other natural processes, were put forth in an attempt to identify the variables influencing the lake's condition and its drinking water budget. A hydrologic model, lake-catchment data, and meteorological data were used to thoroughly investigate these theories. For the experimental analysis, historical data on temperature, groundwater, wind, lake inflows, rainfall, lake physical features, catchment land uses, and soil texture were examined. A temporal trend analysis of pertinent elements was conducted to identify significant drivers of hydrological changes and test the hypotheses. Prioritizing the main variables influencing the yearly lake water budget between 1971 and 2013 involved a sensitivity analysis that focused on the inflows and outflows from the lake. The SCS-curve number was used to calculate the lake's annual inflow, or catchment run-off, using a rainfall-run-off model. The evaporation loss was estimated in order to quantify the outflow. The findings of their investigation suggest that, in contrast to urbanization and climate change, the construction of check dams and the siltation process in the catchment were the primary factors influencing the hydrology of the lake and its recent alterations.

Su et al. (2019) reported lakes that are sensitive indicators of climate change on the Tibetan Plateau that respond quickly to climate change. The performance of the lake model, extended by straightforward parameterizations about the salinity impact, for brackish lakes was evaluated in that study using the surface meteorological parameter dataset, lake temperature at the surface data, and directly collected data. It also revealed the reaction of thermal circumstances, radiation, and heat balance. The findings showed how well the FLake captures seasonal fluctuations in both the inner thermal structure of the lake and the lake surface temperature. Over the past few decades, the simulated lake temperature at the surface has increased. It has a negative correlation with the speed of the

wind and downward shortwave radiation and a positive correlation with air temperature and downward longwave radiation. The dimictic lake with two overturn episodes in late spring and fall was revealed by the simulated inner thermodynamic structure. Although the bottom temperatures exhibited no discernible trend, even slightly declining from 1989 to 2012, the surface and average water temperatures of the lake increased dramatically between 1979 and 2012. Wintertime saw the biggest warming of the air and lake surface temperatures. The later ice-on and earlier ice-off trends that were reproduced in the lake due to climate change had a major impact on the inter annual and seasonal fluctuations in radiation and heat flow. While overall longwave radiation and sensible heat flux somewhat decline, the annual average of net shortwave rays and latent heat transfer clearly increases. As a result of the ice's decreasing duration, the temperature differential between the lake and the air rose during both times.

Odongo et al. (2019) reported the effect of land use cover changes on the hydrological regime in Lake Naivasha Basin in Kenya, which has experienced significant water level changes in recent decades. They studied the land use conversions of grassland and croplands that have influenced evaporation and crop transpiration within the Lake Naivasha Basin from 2003 to 2012. They have used the evapotranspiration data sets for modeling using the Surface Energy Balance System. According to their findings, evapotranspiration increased by up to 12% when grasslands were converted to crops, but it decreased by about 4% when crops were converted back to grasslands. They came to the conclusion that while recently farmed agricultural land can raise local water demands, farm abandonment could reduce water loss and ultimately raise water availability.

The effects of climate change on plains' water resources and the interplay between surface water (SW) and groundwater (GW) components were examined by Guevara-Ochoa et al. (2020). In order to quantify the spatiotemporal dynamics of the water equilibrium and GW-SW relationships for the upper creek basin of Del Azul, which resides in the province of Buenos Aires, this study aims to apply a paired hydrological-hydrogeological model according to changing climate scenarios. Two scenarios using the local climate model that were simulated for 2020–2050 were compared with the baseline scenario that was run in the simulation and calibrated and verified for the years 2003–2015. Study topics include temperature anomalies, groundwater discharge, precipitation, recharge, soil moisture, flow, evapotranspiration, head level, and reserves, both annually and monthly. The GW-SW interaction's spatiotemporal abnormalities were examined. Wet and dry periods were examined using the yearly water balance and the standardized precipitation index. The spatiotemporal structure of the GW-SW relationship and the balance of water will be considerably altered by climate change, according to simulation studies. These displayed annual, seasonal, and monthly fluctuations. With the exception of soil moisture, they indicate an increase in the majority of the water balance elements by the middle of the twenty-first century. In relation to GW-SW interactions, it is anticipated that the aquifer's average annual flow to the stream will rise nearly 24% alongside RCP 8.5 and by 5% with RCP 4.5. It is anticipated that at RCP 4.5, the recharge through the river into the aquifer will rise by 12%, whereas at RCP 8.5, it will fall by 5%. Regarding the SPI associated with the water level balance over the period 2020–2050, variations are anticipated in the

duration and frequency of wet and dry periods for the two scenarios. RCP 4.5 exhibits a small number of wet seasons, but they are more severe and persistent over time, while RCP 8.5 displays a low frequency of dry periods but a high degree of permanence and severity. Groundwater levels could be altered by climate change primarily through changes in recharge, which would reverse the GW-SW flow in some areas within the stream by increasing or lowering the release of groundwater to the stream.

Woolway et al. (2021) used satellite observations and a computational model that utilizes long-term in situ temperature measurements of the surface to investigate variations in water heatwaves for numerous lakes globally between 1901 and 2099. They demonstrate that toward the end of the 21st century, heatwaves in lakes will get hotter and last longer. Under the scenario of high greenhouse gas emissions, the average heat wave intensity in lakes is predicted to rise from 3.7 ± 0.1 to 5.4 ± 0.8 degrees Celsius, and their average duration will witness a significant increase from 7.7 ± 0.4 to 95.5 ± 35.3 days, in comparison to the historical period spanning from 1970 to 1999. The heatwave intensity and length will rise to 4.0 ± 0.2 °C and 27.0 ± 7.6 days, respectively, under the low-greenhouse gas emission scenario. At both historical and future timelines, surface heat waves were less powerful but more persistent in deep lakes (as much as 60 meters deep) compared to shallow lakes. In the 21st century, heat waves from lakes will last for multiple seasons as they warm, with some lakes experiencing year-round heat waves. Lakes that experience heat waves are likely to experience severe physical and chemical changes that will compound the harmful impacts of long-term lake warming. They came to the conclusion that heatwaves could change the species composition of lakes by straining the resilience of aquatic ecosystems and species. The biodiversity of lakes as well as the significant environmental and economic advantages that lakes offer to humanity may be in danger as a result.

Schulz et al. (2020) looked at how Lake Urmia, one of the world's largest hypersaline lakes with distinctive biodiversity, might be affected by climate change. The lake's water level has dropped significantly over the past few decades, endangering the ecosystem's ability to function. There is a contentious discussion on the causes of this decline, with the two most likely theories being mishandling of the water supplies or changes in the climate. In one particular study, they measured the components of Lake Urmia's water budget and examined their temporal history and interactions over the previous 50 years. Through this thorough analysis, it has been demonstrated that changes in Lake Urmia's water level over the studied period were mostly caused by changes in the climate. In contrast to the remainder of the water from the surface input quantities, agricultural extraction of water volumes was noteworthy given the prevailing meteorological circumstances. The lake volume would be significantly affected by shifts in water from agriculture withdrawal, and these changes might either balance the lake or force it to completely collapse.

Mengistu et al. (2021) used the SWAT, which was calibrated and verified using observed flow of streams data collected at four gauging stations in the basin, to analyze the effects of changing climates regarding the hydrological regime within the Upper Blue Nile (Abay) River Basin. Climate change projections showed mean annual temperature

increases and precipitation decreases in most parts of the basin. The hydrological regime of the basin is anticipated to be impacted by these changes, which were evaluated by executing the SWAT model under previous (1981–2010) and future (2040–2069, 2070–2099) climate scenarios. According to the findings, potential evapotranspiration could rise from the baseline era by as much as 27% by the end of the twenty-first century. There could be a 14% rise in surface runoff. Nevertheless, the basin's overall water supply could not rise due to the rise in surface runoff. Rather, the basin's overall water production is projected to drop by –1.7 through –6.5%, as well as –10.7 through –22.7% for models. It is also anticipated that base flow's share of the basin's overall water supply will decrease to 11.4% towards the last decade of the twenty-first century from 41.3% throughout the period of the baseline. The loss in the basin's overall water production is partially explained by the reduction in baseflow. The management of water in the basin will be greatly impacted by such modifications to the hydrologic equilibrium.

In their study, Ashagre et al. (2023) assessed and quantified the consequences of climate change as well as future intended abstraction of water for irrigation scenarios in Lake Ziway, Ethiopia, since the lake's water balance has been threatened by both climate change and ongoing human activity. They did this by using a large number of Global and Regional Climate Models for input in the Hydrologiska Byråns Vattenbalansavdelning hydrological model. They evaluated the effects of climate factors for the years 2021–2050 and 2051–2080. Power transform and variation scaling techniques were used to adjust for biases in the temperature and rainfall data. The irrigation water requirements of the main crops produced in the research region were calculated using the FAO CROPWAT model. The findings suggest that rising temperatures and rainy-season runoff levels are anticipated in the future. In the worst-case scenario for climate change, the lake's level might decrease by 25 cm per year due to changes in the climate and the withdrawal of water for agriculture, leaving behind a surface area of 10 km² and a volume reduction of 101 mm³. In order to ensure long-term lake water consumption and level preservation, they recommended stringent monitoring procedures for water withdrawal as well as appropriate planning and lake ecology management approaches.

The ecosystems of freshwater lakes are essential to aquatic life and human requirements, and any change in the rates of water renewal in these lakes has profound ecological and social effects. Changes in the ratio of intake (precipitation) to outflow (evaporation, withdrawal) affect lakes' depth, area, and hydraulic residence time, as well as their water budget. Lakes all throughout the world are seeing a range of effects from climate change. Temperature variations, especially warming, are frequently the first things that spring to mind when discussing climate change. Variations in the volume and timing of precipitation are a significant aspect of a changing climate. The ecosystems of lakes can be dramatically impacted by changes in precipitation. According to the aforementioned publication, two significant and connected markers of weather and climate change in freshwater lakes are rising lake temperatures and water levels. Numerous factors affect water level, including precipitation amount, precipitation intensity, surface runoff, water-bearing rivers, drought, evaporation rates, and withdrawal of water from the lake and lake watershed for various purposes. Warmer surface water temperature is increasing rates of

evaporation, influencing the production of lake surface ice, and lengthening the evaporation season. Although human activities like water consumption from the lake and its catchment also affect the lake's water budget, climate change has an impact on lake levels. Those studies indicate the urgent need for a successful management plan prepared based on computer prediction models. Precipitation changes are far more difficult to model or predict in hydrologic or computer modeling than temperature variations. On any given day, the temperature fluctuations are often relatively small, only fluctuating by a few degrees. On the other hand, not only the total amount of precipitation but also the intensity is important for building models.

2.4. Factors that Affect Lake Water Quality

Blue-green algae/cyanoprokaryote blooms have become more prevalent worldwide in recent decades (Harke et al., 2016; Paerl and Otten, 2013). High cyanotoxin concentrations have been related to animal fatalities and a health risk to humans from recreational and drinking waters. (Carmichael et al., 2001; Azevedo et al., 2002; Backer et al., 2015). This is because some species have the potential to produce toxins that impact live-stocks. Hepatotoxins, neurotoxins, cytotoxins, dermatotoxins, and irritating toxins are only a few of the hazardous compounds that cyanobacteria are capable of producing (Westrick et al., 2010; Bláha et al., 2009). Several cyanobacterial species, including *Microcystis*, *Nodularia*, *Anabaena*, *Planktothrix*, *Aphanizomenon*, and *Cylindrospermopsis*, have been found to produce cyanotoxins (Bernard et al., 2017; Merel et al., 2013).

The limnological properties and lake contamination level were discussed in earlier works, but nothing was known about phytoplankton. Aykulu et al. (2006) looked at the compositions of phytoplankton and zooplankton and how they related to the water quality of the lake. The study has primarily focused on the pelagic zone and the deepest parts of the lake. The first sampling locations were 100 meters apart and located at the point closest to the shore.

A new and increasing hazard to water ecosystems is the regional effects of climate change (Liboriussen et al., 2005). Climate change has been considered a possible instrument for the future growth of dangerous cyanobacterial blooms because of factors such as higher temperatures, enhanced stratification, longer residence durations, and high nutrient loading, all of which favor cyanobacterial dominance in eutrophic environments. (Nöges et al., 2010; Paerl and Huisman, 2009; Joehnk et al., 2008). Studies (Teubner et al., 2006; Jacquet et al., 2005; Anneville et al., 2005) suggest that a combination of deeper P-depleted zones, reduced phosphorus loads, and enhanced water column stability is most likely responsible for the increase in metalimnetic *Planktothrix rubescens* in several lakes. However, since *P. rubescens* is a metalimnetic species and can obtain nutrients from the hypolimnion during the stratified season, giving it an advantage over other phytoplankton species, there is always competition among organisms for resources. This allows the species to reach high densities in the summer. *P. rubescens* typically experiences poor conditions during the mixing phase, which causes its population to decline and gives room

for other phytoplankton species to expand. On the other hand, a surface bloom may happen in the winter if *P. rubescens* conditions improve (Walsby et al., 2006).

Due to the lake's red tint, attributed to the surface scum of dead *P. rubescens* filaments, it had been believed to be polluted in 1997. Following this incident, *P. rubescens* populations that were metalimnetic were noted in the lake, and microcystin was found in the specimens (Albay et al., 2003). Since the lake is an important source of water for drinking, it is crucial to keep an eye on its cyanobacterial and cyanotoxin levels.

Between 2000 and 2010, the temperature of the air in the vicinity of Lake Sapanca grew only marginally (0.06 °C), but the lake's surface temperatures exhibited a more substantial increase (0.68 °C) than air temperatures. Based on Livingstone (1998), the air temperature is a trustworthy indicator of biotic activities like cyanobacteria behavior in the epilimnion. Cyanobacteria are impacted by air temperature both directly and indirectly. An increase in growth rate and a stabilization of the water column favor buoyant cyanobacteria.

Variations in temperature, glacier cover, winds, and precipitation are only a few of the ways that climate change affects these natural environments (Mooij et al., 2005). Lake Sapanca contains no ice because it is a lowland lake (31 m above sea level). *P. rubescens* was found at depths of 10 to 20 m, where the metalimnion temperature ranged from 9 to 18 °C; however, in 2007, when the water temperature was 11 °C, it obtained its greatest biomass levels. In a Sicilian reservoir, a similar top bloom was discovered by Naselli-Flores et al. (2007) in December at cooler temperatures of the water (9–10 °C). During November and December at Lake Zurich, apparent filaments above the surface were observed in conditions of mild weather and with temperatures of 6 to 8 °C (Walsby et al., 2006). According to Davis & Walsby (2002), *P. rubescens* would outcompete *P. agardhii* in the temperature range between 10–21 °C since it favors a cooler environment to develop.

Temperature, transmission of light, and nutrients had the biggest impacts on algal expansion among the influencing physical and chemical attributes (Özer et al., 2012). The prolonged retention period of the lake may be to blame for the spike in nutrient contents. This rise may be contributing to the *P. rubescens* surface bloom in Lake Sapanca. The presence of such cyanobacteria in Lake Mondsee's surface waters was also noted throughout the lake's eutrophication stage (Dokulil and Teubner, 2012).

P. rubescens, on the contrary, predominates the phytoplankton biomass in layered reoligotrophic waters, usually at a total phosphorus level of about 10 g/L. (Dokulil and Teubner, 2000). Accordingly, *P. rubescens* mass occurrences emerged when phosphate levels were substantially reduced in Lake Bourget, according to Jacquet et al. (2005). The lack of SRP inside the epilimnion will result in a negative impact on phytoplankton biomass and improve the formation of metalimnetic blooms by *P. rubescens* in the summer. Especially in the previous two decades, this was the scenario in the lake (Akcaalan et al., 2006). But only in February throughout the experiment period in Lake Sapanca was SRP under detection boundaries, suggesting that *P. rubescens* bloom might

have consumed all SRP at the time it began to form a bloom. It's interesting to note that SRP concentrations rose as *P. rubescens* biomass decreased, indicating a negative correlation between the two variables.

2.5. Decision Support Tool for Lake Water Level

Multiple interrelated environmental, social, hydrological, and administrative components make up water resource systems. Making comprehensive and integrated decisions is necessary for their management. The examination of the requirements of various sectors as well as the competing benefits and costs in the area are both necessary for system wide management. The development of a DSS that reflects the many components of the multipurpose reservoir system can lead to improved comprehension and management of this complex and intricate system.

DSSs are electronic tools that support rather than take over the process of deciding between options and putting those choices into action. Typically, they serve to make judgments better informed by offering information that might not otherwise be available. DSSs frequently use or include mathematical (computer) models of the system under study.

DSSs frequently use mathematical systems models as a framework and as an integral part of their design. A complete and accurate documentation or representation of the actual system is provided by the system model. The systems model is incorporated into the DSS to make it possible to evaluate alternate operations and enhance operational processes. The DSS offers suggested long-term schedules, short-term operational actions, and assessments of their potential long-term effects on the integrated system. To account for the system's complexity, assumptions, simplifications, and parameterization are applied to various system components (Loucks 1995). These models also depict the relationships and interactions between the system's components in a systematic and in-depth manner. The mathematical model is utilized to research, discover, and assess potential operations and decision-making enhancements (Loucks 1992).

There are numerous methods for creating a mathematical model that supports decision-making. The difficulties in designing a mathematical operational system are frequently dealt with using dynamic programming. Numerous studies (Karamouz and Mousavi 2003; Tejada-Guilbert et al. 1995; Zambelli et al. 2006; Kim and Palmer 1997; Contesse et al. 2003; Stedinger et al. 1984; Kim et al. 2007; Faber and Stedinger 2001) have demonstrated the versatility of stochastic dynamic programming (SDP) in order to deal with nonlinear, stochastic, and multireservoir systems. Using transition probabilities, which establish the predicted value of the goal function, SDP models can account for reservoir inflow-related uncertainty (Kim and Palmer 1997).

Hydropower system optimization has also been developed using linear programming (LP). Some benefits of employing LPs were outlined by Rani and Moreira (2010), including their adaptability for use with large-scale issues, convergence towards

worldwide best practices, and the abundance of both for-profit and free software options. Furthermore, stakeholders may easily understand the mathematical concept behind LPs and their objective functions. The power manufacturing units may be approached linearly because they are located over 1,000 feet under the surface of the reservoir, and variations in preservation have no effect on the system head. Given that this DSS demands 40–50 distinct solutions for each decision and that the linear program comprises 1,200 variables and 1,000 constraints in each iteration, it is crucial that large linear programs be solved efficiently.

Since the mid-1980s, interactive computer technologies and decision support systems (DSSs) have become widely used for researching complex water resource issues (Ahmad and Simonovic, 2006; Metcalfe et al., 2005; Turon et al., 2007; Watkins and McKinney, 1995). In a user-machine interface, a decision support system enables decision-makers to mix personal decisions with computerized output to provide useful information to aid in a decision-making process, according to Ahmad and Simonovic (2006). Such systems are capable of utilizing any information that is made available upon request to assist in the solution of all problems, whether they are structured, semi-structured, or unstructured. For problem-solving, they make use of quantitative models and database components. They play a crucial role in the approach taken by decision-makers to problem-solving and problem identification.

The general structure of an environmental decision support system (EDSS) consists of the following elements: (1) a clear layout to address a range of problems, both structured and unstructured; (2) a robust and intuitive user interface; and (3) data-driven models and algorithms that can easily incorporate data that changes rapidly (Segrera et al. 2003).

2.6. Literature Summary and Knowledge Gaps

The literature presented in this section gives detailed information regarding LWL, the input, the temporal scale for the prediction, the relation between LWL and water quality indicators, and a decision support tool. The LWL is found to be related to some indicators of water quality, such as cyanobacteria, heavy metals, and pH level. Therefore, the imminent water decrease in lakes would not only cause a drinking water supply shortage but also affect water quality. Another point is that the developed algorithms are only effective if they can be used by water managers during their operations. By evaluating the literature, it was found that:

- There is no study that compares the algorithm results with basic benchmark methods for LWL studies.
- There is no study to compare different future time periods for LWL prediction.
- There is no study to investigate the link between LWL, temperature, precipitation, light intensity, and evaporation with water quality, especially the cyanobacteria species of *Plankthotrix Rubescense* that is highly populated in Lake Sapanca.
- There is no study to reveal algorithmic performance in practical use for water managers.

CHAPTER 3

3. RESEARCH METHODOLOGY

3.1. Research Questions and Objectives

Understanding the effects of climatic variables and water usage on patterns of LWL fluctuations is crucial to estimating the ecological and financial consequences of climate change since it has a substantial impact on the natural hydrological cycle and exacerbates water scarcity.

There have been several initiatives to stop any potential drought or flood danger that could be harmful to society during a downturn in the economy. In general, the effects of human activities and variations in the climate on water supplies can be evaluated through the use of prior evaluations of historical data on watershed fluctuations across time. These kinds of hazards can be avoided in a number of ways, but implementing the proper strategies at the right moment is essential for their effectiveness. As a result, forecasting water levels is crucial to developing effective and efficient defenses against all forms of flood and drought danger. This thesis examines experimental DL-based LWL prediction models and suggests improvements to be made to them utilizing decision support tool algorithms in relation to particular hydrological and meteorological data. The research questions of this study can be listed as follows:

- Could LWL estimation algorithms with NN be improved with an ANN-based times series algorithm?
- Could LWL estimation algorithms with NN be improved with the LSTM-based times series RNN algorithm?
- Could LWL estimation algorithms with NN be improved with the Gated Recurrent Unit (GRU)-based times series RNN algorithm?
- Could LWL estimation algorithms with NN be improved with Stacked LSTM-based times series RNN algorithm?
- Could LWL estimation algorithms with NN be improved with Bidirectional LSTM-based times series RNN algorithm?
- Which ANN- or RNN-based DL algorithm should be selected in terms of performance improvement?

- Which prediction period should be selected in order to optimize prediction performance?
- Could ANN or RNN-based algorithms be used in order to create a warning and decision-support tool for water managers?

3.2. Case Study Area

The Lake Sapanca extends between latitudes of 40°41'–40°44' E and longitudes of 30°09'–30°20' N in the northwestern part of Turkey (Figure 8). It is located between two cities; the western part of the lake is in Kocaeli, and the eastern end is within the provincial border of Sakarya. It is a 16 km (east-west) and 5 km (north-south) long tectonic fresh water source that provides the drinking water needs of both cities. It has a surface area of 46 km² and a reasonable depth of 30 m. There are about 1.3 billion m³ in the lake. The greatest depth of the lake basin is 54 m, and its catchment area is 250 km² (Akkoyunlu and Akiner, 2012). The lake is surrounded by southern mountains and northern hills.

The transitional climate found in the Sapanca Basin is influenced by the Black Sea and Mediterranean climates. While the basin exhibits characteristics of both the Black Sea and Mediterranean climates, it may also display elements of a continental climate due to its interaction with an intermediary air system. Despite the warm and rainy winters experienced in the basin, summers are comparatively less hot and dry than what is typically observed in the Mediterranean region.

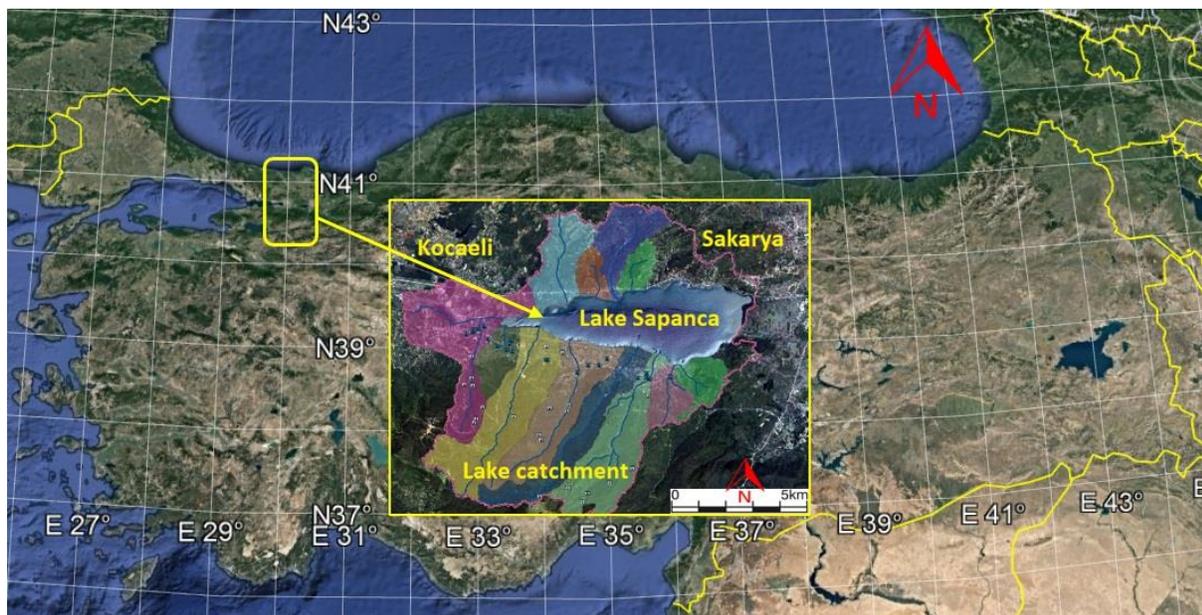


Figure 8: Lake Sapanca area and its catchment with river basins

Figure 8 depicts the catchment area of Lake Sapanca and its sub-basins, which consist of 12 streams that inflow Lake Sapanca. The lake has a controlled outflow with Cark Creek, which regulates the maximum lake level. The seasonal precipitation, water withdrawal,

and surface outflow result in inter-annual LWL variations of 2.28 m between 29.90 and 32.18 m above sea level. The lake is noteworthy because it supplies drinking water to the provinces of Kocaeli and Sakarya. It is also believed that the lake basin will eventually meet the bottled water needs of Istanbul. Although the basin area doesn't include any industrial regions, 23% of the basin area is used as cultivated land mainly covered by ornamentals and fruit orchard, and 9.5% is used as settlement land. Remaining basin area is covered by 65% forest land and 2.5% as natural land. However, the water quantity and quality of the Lake Sapanca basin have deteriorated because of the water demands of urban, agricultural, and industrial facilities. Turkey's average population growth rate is 0.8%; however, over the past 20 years, the population growth rate in the basin has risen from 1.5% to 3.5% (World Bank, 2021). The rapid population growth of the basin is adversely affecting the quantity and quality of water. Although the lake is in a transitional stage from oligotrophic to mesotrophic, its ecological status is deteriorating as the water level drops below the lake's surface discharge during droughts and point and nonpoint runoff flows in from numerous sources (Akkoyunlu and Akiner, 2012). The lake's ecological state deteriorates primarily as a result of unchecked agricultural operations and household wastewater leakages in the vicinity. In addition, the droughts that periodically occurred caused the lake's water quality and quantity to deteriorate (Duru, 2017).

3.3. Dataset Description

The descriptive statistics of the features in the dataset are presented in Table 7. It reveals the SD, minimum value, average value, maximum value, and quarter values for each feature. It is important to notice that the LWL has a minimum value of 29.27 m, which is below the threshold for a lake to sustain itself in order to supply healthy drinking water. Water managers should target not to reduce water levels below the threshold level, which is 30 m in the short and medium term.

Table 7: Descriptive Statistics for features in the dataset

	Mean	Std. Dev.	Min. Value	25%	50%	75%	Max. Value
Maximum Temperature (°C)	21.6	8.2	0.4	15.4	22.2	28.6	40.7
Minimum Temperature (°C)	11.7	6.6	-9.4	6.7	11.7	17.5	25.6
Average Temperature (°C)	16.0	7.1	-3.1	10.3	16.1	22.5	30.2
Precipitation (mm)	2.5	6.9	0.0	0.0	0	1.4	93.5
Withdrawal (m ³)	183,914.2	23,165.5	125,998.0	165,219.5	188,420.5	203,795.5	222,869.0
Water Level (m)	31.5	0.7	29.3	31.2	31.8	32.1	32.3

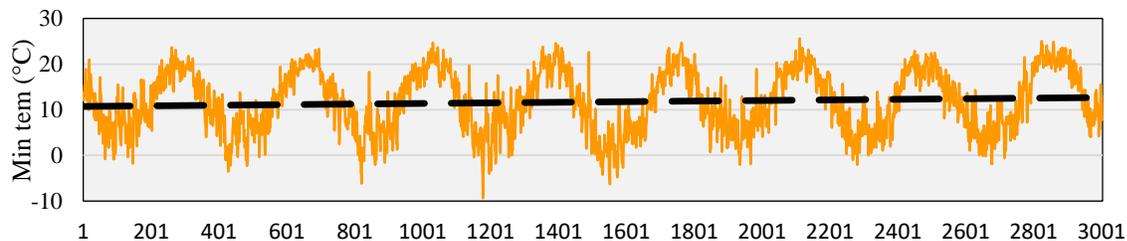
Several characteristics are used in the literature to evaluate future LWLs. The most commonly used features in the literature are precipitation (17%), LWL, and evaporation (Ozdemir et al., 2023). Other major features used by researchers include discharge (Jiang et al., 2021), temperature (Nourani et al., 2021), inflow (Tsao et al., 2021), streamflow

and humidity (Obringer and Nateghi, 2018), wind speed and solar radiation (Guyennon et al., 2021), volume, and area (Dinka, 2020).

The State Hydraulic Works and Turkish State Meteorological Service, through their river monitoring program for Lake Sapanca, provided the data examined for this study. LWL, maximum temperature, minimum temperature, average temperature, precipitation, and flow rates were the features that were supplied. Among those, withdrawal feature includes water withdrawal for industrial, agricultural and domestic use. Measurements were taken daily between 2012 and 2020, with occasional missing data. However, Figure 9 displays the data between 2020 and 2023 for LWL, although the data was not used for the experiments because of insufficient data from other features. The interpolation approach was used to fill in the missing data.

A sequence of values measured over a time step in continuous or discrete units of time is called a time series. Numerous studies have demonstrated the usefulness of time series prediction as a tool for early warning and control. Predicting potential modifications at points of observation throughout time is the goal of time series analysis. The data set employed in this investigation is a standard multivariate time series, as shown in Figure 9, which typically comprises water withdrawal from the reservoir along with real-valued LWL and climatic information. Looking more closely at the graph, one can see that there are irregularities in both the LWL and meteorological and hydrological data. However, the temperature data show only annual and seasonal patterns of variation. Distribution data from LWL are compatible to meteorological data, especially in annual precipitation. Additionally, water withdrawals show an increasing trend over time (Figure 9).

As shown in Figure 9, during prolonged drought, the LWL drops below the discharge elevation at the surface, which is 29.90 m above sea level in Lake Sapanca. Data from LWL indicate that in years of low precipitation, LWL decreases. Higher precipitation in the last decade (2015–2018) coincides with LWL above the lake's discharge elevation. In addition, higher maximum temperatures and low precipitation in recent years have reduced LWL to the surface runoff elevation during dry periods (Figure 9). Low precipitation also increases water demand, while low temperatures decrease water use. In addition, increasing population and industrialization are related to water withdrawal from the lake. Therefore, multivariate time series data that include freshwater demand and meteorological characteristics are critical for predicting LWLs.



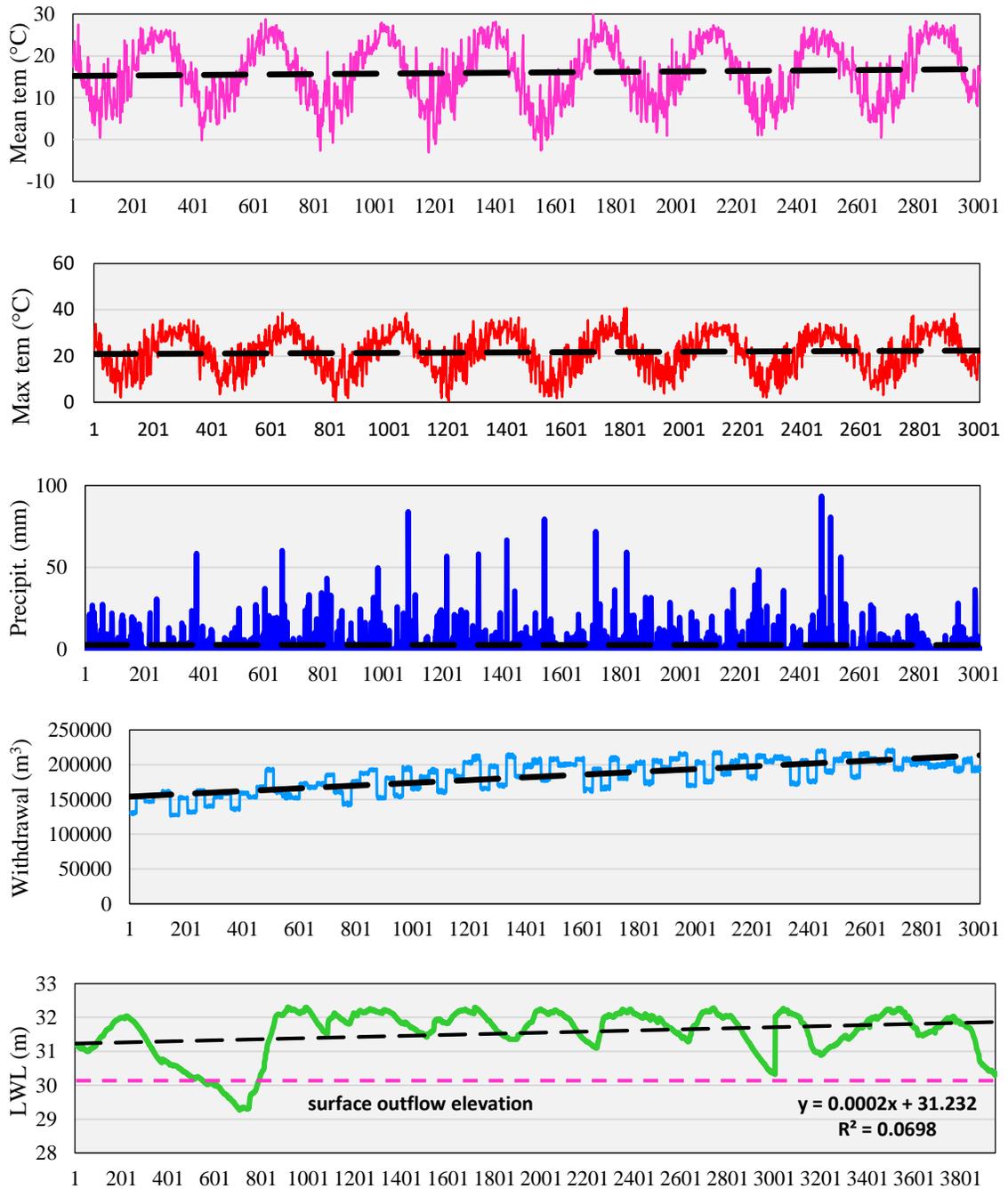


Figure 9: Time series plots of daily meteorological data, water withdrawals and LWL for Lake Sapanca from 11 October 2012 through 31 December 2020.

3.4. Stationary Data Test

When the variance, mean, SD, and covariance of a series remain constant or unaffected by time, it is said to be stationary. Stated differently, time series that lack a pattern or

seasonality are referred to as stationary. It is additionally referred to as the unit root test and is used to determine how much a trend dominates a time series. Among the various unit root tests available, the Augmented Dickey-Fuller test could be the most widely used. An autoregressive model is used to improve the information criteria over a range of different lag values.

The time series is not stationary (has significant temporal-dependent character) and could be expressed as a unit root, according to the null hypothesis of the test. The time series being steady is the alternate hypothesis (rejecting the null hypothesis).

The results are interpreted based on the p -value of the test. A p -value under a certain level signifies that the null hypothesis is rejected (stationary), and a p -value above the threshold of significance suggests that the null hypothesis has failed to be rejected (non-stationary).

- The data are non-stationary and have a unit root if the p -value is greater than 0.05, which means that the null hypothesis (H0) failed to be rejected.
- The null hypothesis (H0) is rejected if the p -value is below 0.05 because the data are stationary and lack a unit root.

Table 8: Augmented Dickey-Fuller test for the dataset

Test Statistic:	-22.776061
p-value:	0.041776
Critical Values:	1%: -3.423
	5%: -2.853
	10%: -2.467

To ascertain whether the time series is stationary, the Augmented Dickey Fuller test was employed, and the test results are shown in Table 8. There is strong evidence for the alternate hypothesis in this situation because the test statistic was found to be highly negative (-22.78) and the p-value was found to be below threshold (0.04). As a result, the null hypothesis was rejected, indicating that the series is probably stationary.

3.5. Data Preprocessing

The dataset was created on a daily basis with monthly stacks and converted to a time series format to be used as a predictive model. The dataset contains several missing points that prevent the model from running. Although the dataset has small gaps, some columns contain large blanks. The large gaps are located either at the beginning or at the end of the dataset. Therefore, these parts were removed from the dataset. Other missing data was interpolated using the linear method.

The easiest technique for estimating a function's value between any two known values is to utilize the linear interpolation method. Additionally, curve fitting with linear polynomials can be accomplished with the help of the linear interpolation formula (Equation 1). Essentially, utilizing the set of values, the interpolation method finds new

values for any function. To determine the unknown values in the table, this thesis applied the linear interpolation formula. There is frequently some association, and other values can be predicted with the aid of experiments conducted at a range of values. Interpolation is useful for determining the equation from the untabulated points. Interpolation can be used to estimate any desired value at a provided, existing coordinate position. It is very useful for numerous additional quantitative and scientific applications, such as data forecasting and market research.

$$y = y_1 + (x - x_1) \frac{(y_2 - y_1)}{(x_2 - x_1)} \quad (1)$$

Where x_1 and y_1 are the initial coordinates, x_2 and y_2 are the subsequence coordinates, x is the place where the interpolation is to be done, and y represents the interpolated value.

An outlier is an observation on a sample dataset that deviates from the general pattern. The outliers could point to anomalies, measurement variability, or experimental errors. Outliers have a major effect on the mean and SD of the dataset. These can yield results that are statistically erroneous. When outliers are present, most ML algorithms perform poorly. Therefore, it is better to find and remove outliers. When detecting anomalies like impossible negative values, where the expected value should always be positive, outliers are quite helpful.

This thesis used both the Interquartile Range (IQR) and expert opinion in order to detect erroneous data and outliers in the dataset. IQR divides a data set into quartiles in order to quantify variability. The information is divided into four equal sections and sorted in increasing order. The first, second, and third quartiles, or Q1, Q2, and Q3, are the values that divide the four equal sections.

- Q1 stands for the data's 25th percentile.
- Q2 stands for the data's 50th percentile.
- Q3 stands for the data's 75th percentile.

IQR stands for information quality ratio. It is calculated as in Equation 2. Outliers are the data points that are either above $Q3 + 1.5 \text{ IQR}$ or below $Q1 - 1.5 \text{ IQR}$.

$$IQR = Q3 - Q1 \quad (2)$$

Where IQR is the Interquartile Range, Q3 is the 3rd percentile, and Q1 is the 1st percentile.

No outlier was detected using the IQR method except precipitation and LWL features. (Appendix A) The data points in the range of 3.5 mm to 93.5 mm for precipitation and between 29.27 m and 29.93 m for LWL are detected as outliers for the IQR method. The values in these features were discussed with experts in the field and later determined not to omit related parts in the dataset since these values are likely to happen in these features. The data were used after the necessary preprocessing steps had been performed.

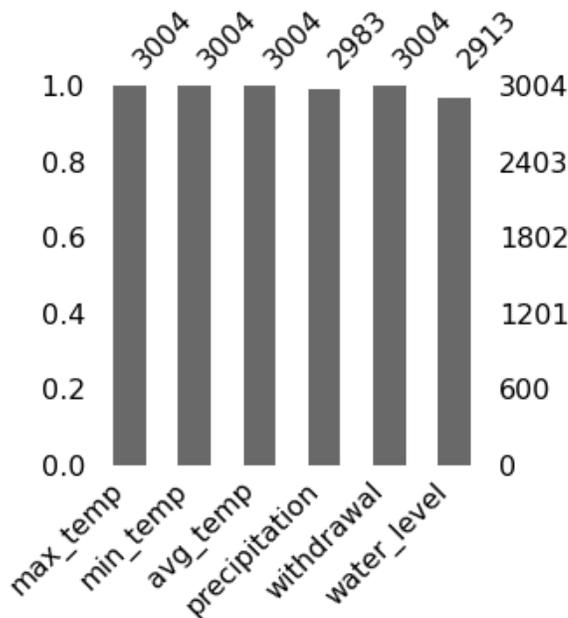


Figure 10: Bar chart of features with missing values

The bar chart reveals the number of missing parts of features. As it can be seen in Figure 10, maximum temperature, minimum temperature, average temperature, and withdrawal do not have any missing values. Only missing values are included in the features of precipitation (0.7%) and water level (3.03%).

Figure 11 demonstrates the Missingness Matrix of features in terms of their missingness pattern. As it can be seen from the figure, the missing parts for precipitation and water level don't overlap with each other. Therefore, it can be said that the pattern's missingness completely at random. The result in the figure justifies the interpolation method for the missing data does not cause any bias in the results.

Since time series are made of sequential data, the data should be used to specify training and test sets. From October 11, 2012, to December 31, 2020, daily meteorological and hydrological data were collected at the lake basin and utilized to train and test the algorithms. The best subsets to use for training and testing might be a problem for data-driven models. To discover the optimum pattern for the data and to increase the validity of the model outputs in unknown data, the dataset was separated into training, validation, and test sets with 60%, 20%, and 20% proportions, respectively. This allowed for the coverage of high and low values in the training and test subsets. To avoid overfitting and to include all seasons in the data set, dry and wet seasons were included in all sets created.

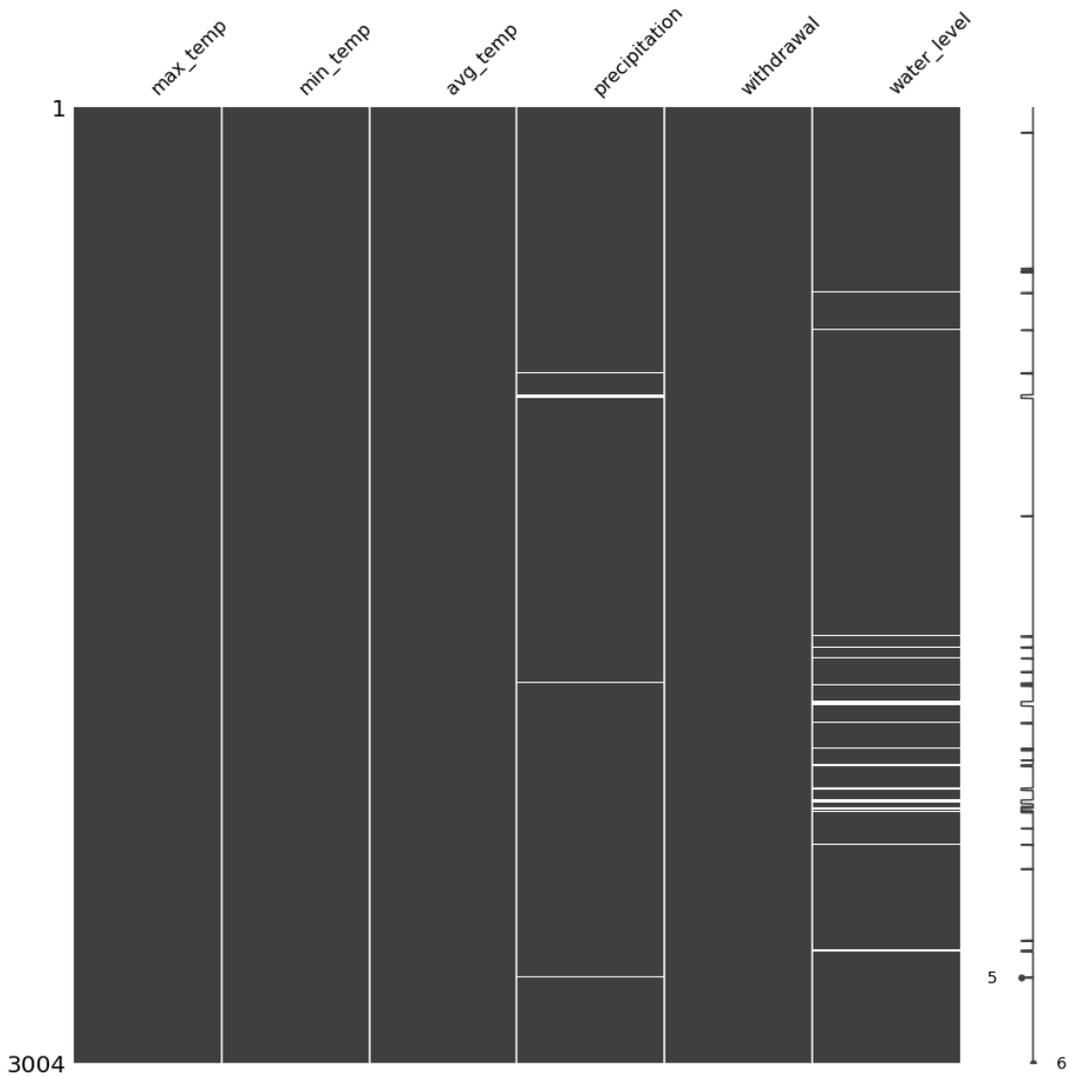


Figure 11: Missingness Matrix of Features

3.6. Model Descriptions

The RNN is the ancestor of gated recurrent-based networks. Due to time delays during the backpropagation error in the learning process of the RNN model network, it was established as a remedy for the gradient explosion problem. At each time step, gated RNN networks predict the label of an activity. To predict an activity label, any number of previous time steps can be merged. The gated RNN model networks have been shown to be significant models in the past and are capable of learning from sequential inputs. It can effectively learn from sequences of different lengths and capture long-term dependencies. There are ANN and four different gated RNN networks used in this thesis: LSTM, GRU, Stacked LSTM, and Bidirectional LSTM.

3.6.1. ANN

A massively parallel-distributed information processing system called an ANN mimics the function of the neuron network in the human brain. Human learning is a result of neurons, and ANNs employ this important feature for ML.

A NN is made up of several nodes, or basic processing units. Figure 12 depicts a basic and generic representation of a processing element. The mathematical functions and network architecture make up the ANNs. The architecture is made up of the arrangement of nodes in a specific way. Typically, the nodes are organized in layers that facilitate the flow of information from the input layer to the output layer. Between the input and output layers, there may be multiple hidden levels. The network's capacity to represent more complicated events is enhanced by the hidden layers.

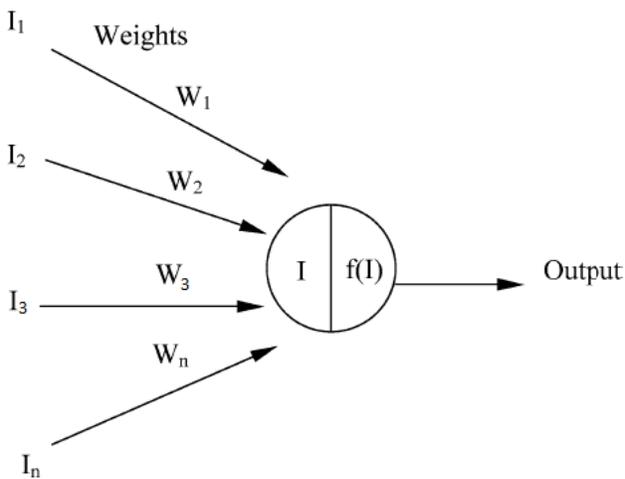


Figure 12: ANN Model

Figure 13 displays a four-layer feed-forward ANN together with a typical processing element. Although they are not connected to one another, the nodes in one layer are connected to those in the next. Every node computes a function based on its input and transmits the output to other units in the network that are connected. As a result, a node's output is tied to its inputs and matching weights. The weight of the connection determines how strong the signal is as it travels from one neuron to the next. Trial and error is typically used to determine the number of hidden levels and the number of nodes in each hidden layer.

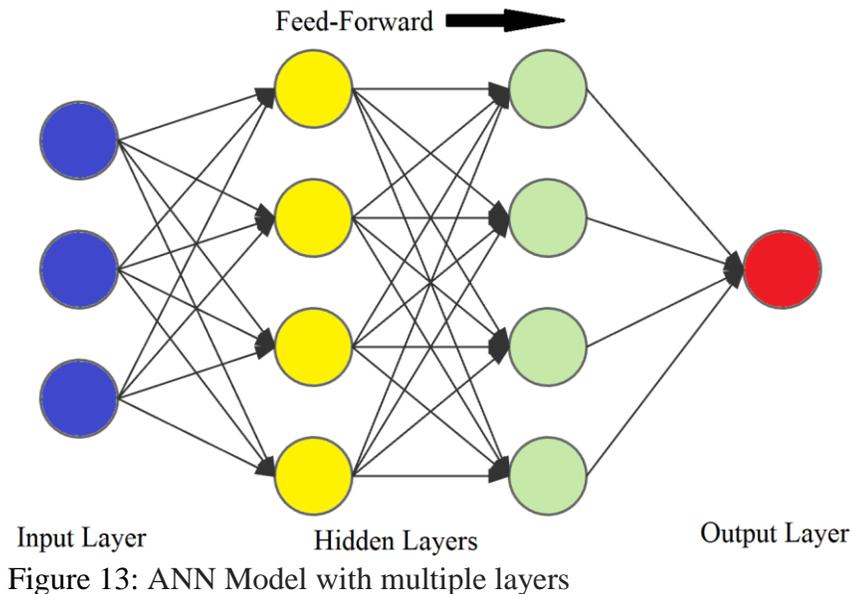


Figure 13: ANN Model with multiple layers

When an ANN first starts off, it knows nothing. The process of learning begins with data entry into the network's input layer. The weights are frozen once the learning process is finished. It uses a data set from performance testing for validation.

3.6.2. LSTM

A sequence of feedback loops and gates that are self-trained using the input data is used by LSTM. According to Zhang et al. (2018), the LSTM network is a particular architecture that is intended to simulate complex temporal and spatial sequences. By adding a gate mechanism, it can more correctly handle dependencies that are long-range. A collection of recurrently connected memory cells constitutes each of the multiple memory blocks that make up the LSTM network, which are connected via layers (Figure 14).

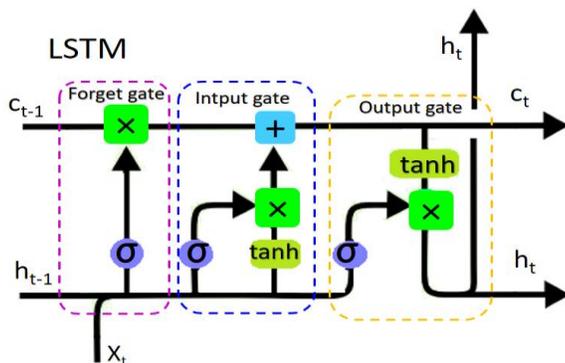


Figure 14: Basic structure of LSTM algorithm

LSTM has three multiplicative units: input, output, and forget gates. The input gate changes the information using the sigmoid function, regulatory filter, and hyperbolic

tangent function. The forget gate changes information of smaller importance. The output gate selects the useful information from the current cell. The LSTM layer uses the following mathematical operation to determine the output variable (Ozdemir et al., 2023):

$$\sigma(t) = \frac{1}{1+e^{-t}} \tanh(t) = \left(\frac{e^t - e^{-t}}{e^t + e^{-t}} \right) \quad (3)$$

$$f_t = \sigma(Wf(h_{t-1}, X_t) + B_f) \quad (4)$$

$$i_t = \sigma(Wi(h_{t-1}, X_t) + B_i) \quad (5)$$

$$o_t = \sigma(Wo(h_{t-1}, X_t) + B_o) \quad (6)$$

Where f_t forgotten variable, i_t input variable, o_t output variable. X_t indicates the values that the feature receives at t time, and h_{t-1} is the output cell of the previous cell. Inside the LSTM cell, memory is indicated by c_{t-1} . W is the weight matrix, and B is the term bias. The sigmoid function (σ), the hyperbolic tangent function (\tanh), processes the X_t variable and the h variable from the previous learning.

3.6.3. GRU

While GRU has two gates, it runs faster and requires less memory than LSTM. GRU is computationally more efficient than LSTM because the structure is simpler and more straightforward. The input gate and the forgetting gate are combined into one update gate and simplified (Figure 15).

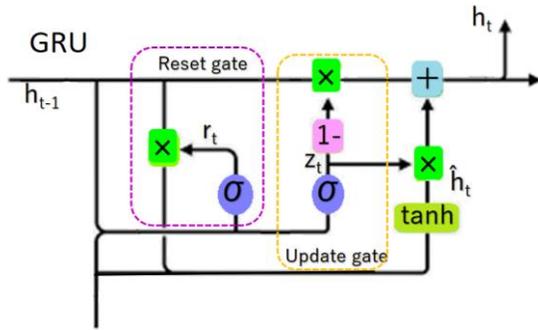


Figure 15: Basic structure of GRU algorithm

GRU contains one \tanh function and two activation functions. As a result, GRU has the same capacity for long-term memory formation as LSTM, but it also benefits from containing fewer variables and a quicker training rate. The output variables of GRU are calculated using the following equations (Ozdemir et al., 2023):

$$r = \sigma(W_r(h_{t-1}, X_t) + U_r X_t) \quad (7)$$

$$z = \sigma(W_z(h_{t-1}, X_t) + U_z X_t) \quad (8)$$

$$c = \tanh(W_c(h_{t-1} \times r) + U_c X_t) \quad (9)$$

$$h_t = (z * c) + ((1 - z) \times h_{t-1}) \quad (10)$$

Where, \tanh and σ are the hyperbolic tangent and logistic sigmoid functions, r and z are vectors for the activation values of the update and reset gates. W_r , U_r , W_z , U_z , W_c , U_c represent the weight matrix.

3.6.4. Stacked LSTM

Stacked LSTM is a variant of LSTM with multiple LSTM layers containing multiple memory cells that give the model the ability to capture the structure of time series and combine the learned representation of previous layers while providing a higher level of abstraction for the final results. This structure contributes to the model's ability to learn higher-level temporal representations but can lead to degradation problems due to the low convergence rate of the LSTM layers, although this error is different from the vanishing gradient problem.



Figure 16: Stacked LSTM Model

The model is deeper and more properly qualifies as a DL technique thanks to the stacked LSTM hidden layers. NNs' depth is credited with the method's success on a variety of difficult prediction issues. Thus, the stacked LSTM provides a dependable approach for challenging sequence prediction problems. An LSTM model with many LSTM layers is called a stacked LSTM structure (Figure 16). An output sequence, as opposed to a single value, is sent to the LSTM layer below from an LSTM layer above. To be more precise, one output for each input time step equals one output time step overall.

3.6.5. Bidirectional LSTM

Another variation of the LSTM is the Bidirectional LSTM, in which the input currents of the LSTM flow in both directions so that information from both the input and output sides can be used. This model's accuracy is increased by using both forward and reverse information, which also facilitates better learning over dependence on long-term data.

$$h_n = \text{LSTMforward} \left(in, \overset{\rightarrow}{h_{n-1}} \right) \otimes \text{LSTMbackward} \left(in, \overset{\leftarrow}{h_{n+1}} \right) \quad (11)$$

Where, h_n =new state, in =input, h_{n-1} =output of past state, h_{n+1} =output of future state, \otimes symbol represents the concatenation operation.

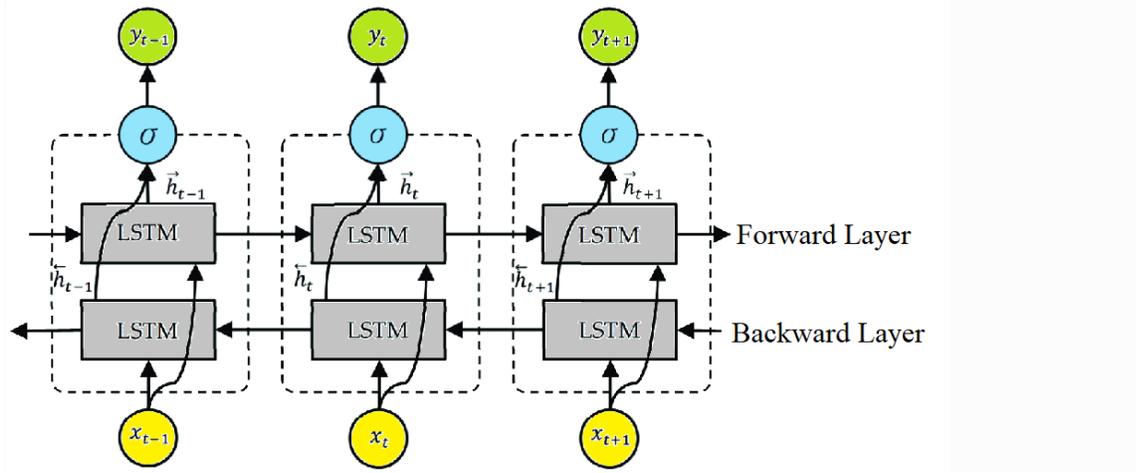


Figure 17: Bidirectional LSTM Model

Typically, the LSTM layers constrain and process sequence input in a unidirectional manner in order to capture the periodicity or unpredictability of the system. On the other hand, adding a backward LSTM layer could solve the issue and make the network bidirectional. As a result, the construction of an unfolded Bidirectional LSTM layer that links two distinct hidden layers to a single output layer, processing sequence data with a forward LSTM layer and a backward LSTM layer (Figure 17). The inputs generate the forward layer output series sequentially through time period h_{n-1} to h_{n+1} in a positive direction. On the other hand, the backward layer sequence is calculated using the reverse data throughout the same period. The final output, h_n , is obtained from the following equation after the layer outputs are calculated using the conventional LSTM Eqs. (3–6).

3.7. Sequence Creation

This work made use of supervised deep NN models that train the models using input data that is typically organized as a matrix. The time series comprises a collection of values arranged sequentially. DL models are unable to process values in their organic state. In order for the models to learn from past values of input and forecast potential outcome values, the data must be turned into input and output sequences. The following three tenets form the foundation of the sequence generation procedure:

- **Input Data Sequences Length:** the number of time-varying cycles in the data being provided into the model to forecast the output based on the relevant sequence in the context of time series analysis. Depending on the model structure being utilized and the hardware capacities needed to analyze larger amounts of data, the sequence size may vary.
- **Slide Window:** This method is often used to create sequences in a time series. Through sliding the sequence into x (sliding window size) steps, it enables the development of several sequences of identical length. For example, given that a time series consists of one thousand points of data and sequences of length ten are generated, slide a window ($w = 1$) with a step of 1 across the time series. As a

result, 991 sequences of length 10 would be produced, each with a unique time series beginning point.

- **Output Data Sequences Length:** Also referred to as the forecasting horizon, this parameter establishes how many future outcomes the model will forecast from the given input sequence. When it comes to LWL forecasting, a forecasting horizon of one time step (days) indicates that, depending on the input sequence length, the model will forecast the LWL values for the upcoming day.

Applying a window that slides in varying lengths, this thesis generates the sequences of inputs for each model design. The output duration of the sequence (forecast horizon), which is 1, indicates that the next day will be predicted using the previous time step values, and a forecast horizon of 30 means the next 30th day will be forecasted using the same input sequence. The sequence of input lengths is regarded as a training parameter that needs to be adjusted. The scope of this thesis adjusts the forecast horizon rather than the input sequence. Thus, the input sequence remains the same for all algorithms, and the forecast horizon is 5.

3.8. Normalization

In ML, data normalization is a technique that converts numerical data into scaled intervals of [0, 1], [-1, 1], or comparable values (Shanker et al., 1996). Equation 11 illustrates how this operates by taking the least value of a feature and dividing it by the feature's range. Normalization guarantees that each feature contributes equally to the model and keeps a small number of characteristics with big values from dominating it when the data contains many features of different scales (Atlas et al., 1989).

$$x_{normalized} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (12)$$

Time series prediction techniques require data normalization methods in order to function with consistent features. Data normalization is a crucial step in the data preprocessing procedure when dealing with LWL data because the ranges vary based on the units of measurement applied.

The min-max data scaling technique from the Sklearn library's MinMaxScaler function is used in this thesis. All of the quantitative values that are continuous in this thesis are changed to the range [-1, 1]. As a result, equation 4 needs to be modified. Equation 12 is used to scale the values of the features across a range of values [a,b], where a = -1 and b = 1, resulting in equation 13.

$$x_{normalized} = a + \frac{(x - x_{min})(b - a)}{x_{max} - x_{min}} \quad (13)$$

$$x_{normalized} = 2 \frac{x - x_{min}}{x_{max} - x_{min}} - 1 \quad (14)$$

3.9. Hyperparameters

The thesis uses the Tensorflow Keras libraries to implement the proposed different ANN and RNN-based networks, with Tensorflow as the backend. The implementation of ANN, LSTM, GRU, Stacked LSTM, and Bidirectional LSTM layers in the algorithm uses the sequential approach. There are several hyperparameters in the algorithms that need to be optimized to get the best score out of experiments, which are neuron number, epoch, batch size, number of layers, prediction period, loss function, and optimizer.

The neuron number is the total number of neurons the model will have. The output learning result may be impacted by the more hierarchical learning capacity achieved by adding more layers and various amounts of neurons in each layer.

An epoch, which is the entire number of training data iterations in a single cycle for the purpose of training the ML model, is when every piece of training data is used simultaneously. The performance of the model may be impacted by different iteration sizes.

The number of sequences needed to calculate the gradient during every update stage throughout training depends on the batch size. A larger batch size may result in poorer generalization but faster convergence. This entails striking a balance between utilizing the appropriate quantity of computational resources and producing quality outcomes. More GPU RAM is required for bigger batch sizes, as smaller ones could cause overfitting and slow down training.

The number of layers is generally referred to as the number of hidden layers. Model performance may be impacted by adding more layers and different numbers of neurons within each layer to increase hierarchical learning capacity.

The prediction period is the number of future periods that the time series prediction model will perform. The model performance would typically change drastically in terms of prediction period.

In a DL model, the difference between expected and actual values is quantified by a mathematical function called a loss function. It evaluates the model's effectiveness and directs the process of optimization by giving comments on how well the model fits the data.

Optimizers have an impact on the model's rate of convergence and are essential in assessing a DL model's training efficacy. Every optimizer has its own set of advantages and disadvantages, and no optimizer is ideal in every situation. The optimization setup trials are listed in Table 9.

Table 9: Hyperparameter optimization setup

Hyperparameters	Trials
Neuron Number	32, 64, 128
Epoch	50, 100, 250, 1000
Batch Size	32, 64, 128
Number of Layers	1, 2, 3
Prediction Period	1, 5, 10, 20, 30, 45, 60, 90, 120

The loss function is set as "MAE" and the optimizer as "Adam", since these hyperparameters do not have a significant impact on the performance of the algorithm. However, the hyperparameters such as neuron number, epoch, batch size, number of previous time steps, and number of layers are optimized. The hyperparameters in the algorithms that have the best performance are briefly listed in Table 10.

Table 10: Optimized hyperparameter values of algorithms

	ANN	LSTM	GRU	S. LSTM	B. LSTM
Neuron number	128	128	64	128	32
Epoch	250	100	100	100	50
Batch size	64	128	128	128	128
Number of layers	1	2	2	2	2
Prediction period	45	60	60	60	60

*S. LSTM: Stacked LSTM, B. LSTM: Bidirectional LSTM

The optimized hyperparameter values in Table 10 are different for the different algorithms. However, all RNN-based algorithms performed the best when the number of layers was 2 and the prediction period was 60 days. The loss functions of the model results can be seen in Appendix B.

3.10. Evaluation Metrics

A significant number of researchers in the literature prefer RMSE, MSE, MAPE, MAE, R^2 , or R as evaluation metrics to compare their algorithms with the base model or with other algorithms (Ozdemir et al., 2023). These metrics account for more than 50% of the evaluation metrics in the literature. On the other hand, there are less favorable evaluation metrics used by some researchers, including NSE (Dinka, 2020), accuracy (Paul et al., 2019), MRE (Jiang et al., 2021), and PBIAS (Nhu et al., 2020).

The goal of model performance evaluation is to verify the accuracy of the proposed model and determine the difference rate so that it can be used with confidence (Zheng, 2018). The evaluation metric chosen for this thesis is RMSE. The RMSE value shows the root of the squares of the average differences between predicted and observed values. Lower RMSE values indicate higher model performance and a better correlation between observed and predicted values. Equation (10), discussed in more detail below, was used as a performance measure in the evaluation of the model.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - Q_i)^2}{n}} \quad (15)$$

Where “ Σ ” stands for sum, P_i is the expected value for the dataset’s i^{th} observation, Q_i is the actual value for that observation, and “ n ” denotes sample size.

Although the performance of the model is evaluated using the RMSE value, the comparison of forecast accuracy is made using the Naïve Method. The Naïve Benchmark is one of the most commonly used method for comparing time series forecasting models because it is easy to compute and understand (Van der Heijden et al., 2021). In this approach, each forecast is equated to the last observed value for the intended time step. The performance of the algorithm was considered successful if the RMSE value was lower than the RMSE results of the Naïve Method. The reason for such a comparison is that the RMSE values for earlier time steps are always lower than for further time steps due to their proximity to the actual values. For this reason, the performance of further time periods (i.e., 60 days and 120 days) cannot be compared using only the RMSE value itself. To compare all time values, performance is evaluated using the percentage increase in RMSE compared to the RMSE score of the Naïve Method.

$$Y_t = Y_{t-n} \quad (16)$$

Where Y_t is forecast value at time t and Y_{t-n} is the value at the previous n^{th} day

To determine whether the proposed algorithms are successful enough to be used as a prediction method, the prediction results of the algorithms are also compared. Diebold-Mariano significance test was used to control the algorithm differences and their significance at p -value < 0.05 , as described by Van der Heijden et al. (2021). If the p -value of the test is less than 0.05, the prediction accuracies are significantly different from each other. This method suggests that a complex operation should only be suggested if it outperforms the benchmark by a substantial margin, rather than just having superior accuracy rates. This also implies that there is a significant difference between the prediction results and that the proposed algorithms cannot be used interchangeably.

A graphical representation of the entire modeling process with the flowchart applied to predict LWL in this thesis is shown in Figure 18.

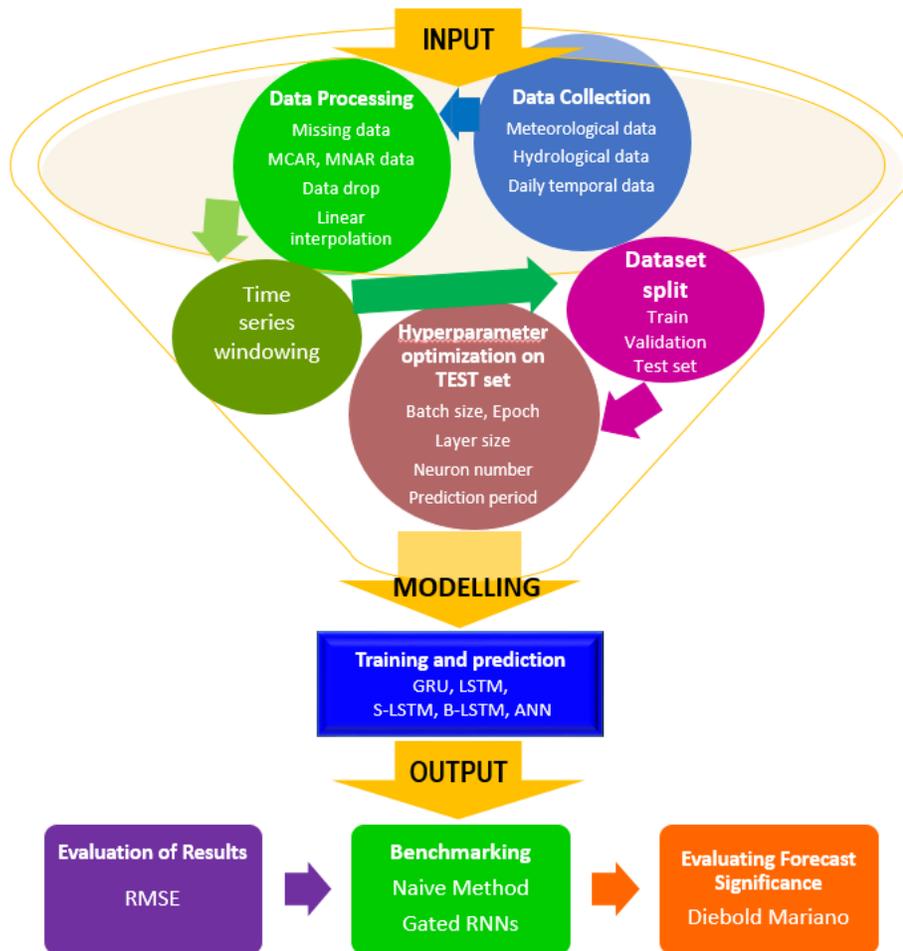


Figure 18: Flowchart of LWL Modelling

3.11. Water Quality Indicator

In freshwater habitats, temperature and light intensity are the most important meteorological factors determining algal photosynthesis and algal blooms. The toxin production behavior of freshwater algal species is strongly influenced by environmental conditions. As an indicator of biological water quality, monthly microcystin measurement data at various depths from the surface to 20 m during the period from March 2019 to April 2023 were subjected to statistical analysis.

Computer-based models require long-term data to make more reliable and accurate predictions for the future. Therefore, the Mann-Kendall trend analysis test has been used with environmental time series. The degree of disparity between data measured later and data measured earlier is examined by the Mann-Kendall test. After comparing every value measured later with every value measured earlier, $n(n-1)/2$ potential pairings of data are obtained, where n is the entire number of occurrences. In this test, the null hypothesis assumes that there is no trend, and the alternative hypothesis assumes that there is a trend.

Furthermore, SRC analysis was applied to determine the relationship between key meteorological parameters and the concentration of the cyanobacterial growth byproduct, microcystin, since microcystin concentrations did not follow a normal distribution. SRC measures the strength and direction of the association between the two variables that are ranked. In essence, it indicates how well a monotonic function could capture the association of two variables, or the level of monotonicity of that relationship. The range of the Spearman correlation coefficient is +1 to -1. A rank correlation of +1 denotes a perfect positive link; a rank correlation of zero denotes no association between ranks; and a rank correlation of -1 denotes a strong negative relationship between ranks.

3.12. Decision Support Tool

The eventual decision support tool that is built upon LWL prediction results for this thesis used the Streamlit development platform. Streamlit is an open-source application for developers that is fully integrated with the Python language and its libraries. The platform is highly embraced by data scientists because of its integration with the most popular data science platform, Python, and its easy use and free access. Another appealing side of the platform is that it provides free access to the cloud for deployment after initializing a local computer. The libraries that are used for the decision support tool are Pandas, Numpy, Scikit-Learn, Keras, TensorFlow, Matplotlib, Streamlit, Seaborn, Plotly, and OpenPyXL.

CHAPTER 4

4. LAKE WATER LEVEL PREDICTION

The thesis uses ANN and four different RNN-based DL algorithms to compare their forecasting accuracy from day 1 to day 120 ahead, based on RMSE values, Naïve Method and Diebold Mariano test results. The ANN, LSTM, GRU, Stacked LSTM, and Bidirectional LSTM algorithms were successfully trained and validated, and compared with test data consisting of 3004 lines to evaluate the reliability of the model for the unknown data set. Table 11 presents the performance of LWL prediction of investigated ANN and RNN algorithms from day 1 to day 120 ahead forecasting. These results show that all the investigated ANN and RNN algorithms have excellent prediction accuracy in the 1-day to 10-day-ahead prediction scenario with a RMSE values of < 0.1 m. On the other hand, the LSTM algorithm had the best value for training and testing in the 60-day scenarios with an RMSE value of 0.1762 m, while GRU showed the best performance in the 120-day scenarios with an RMSE value of 0.3838 m (Table 11). In contrast, the Stacked LSTM and Bidirectional LSTM models showed no additional performance in terms of prediction accuracy compared to LSTM. Since there are seasonal cycles of water levels, especially in summer and winter, this thesis did not investigate further than 120 days of prediction. After a certain time, the cycle would repeat itself and the Naive Method would perform unjustifiably well. Therefore, the investigation to compare 365 days could not be an appropriate comparison with other algorithms.

Table 11: The performance of Naïve Method, ANN, and RNN based algorithms for predicting LWL with increasing time intervals, RMSE results. (Metric is based on m)

Algorithm/Prediction Period	Naïve Method	ANN	LSTM	GRU	Stacked LSTM	Bidirectional LSTM
1 day	0.0134	0.0131	0.0162	0.0134	0.0171	0.0156
5 days	0.0484	0.0445	0.0514	0.0429	0.0494	0.0563
10 days	0.0875	0.0815	0.0799	0.0732	0.0890	0.0875
20 days	0.1551	0.1271	0.1227	0.1070	0.1289	0.1257
30 days	0.2168	0.1540	0.1356	0.1316	0.1221	0.1226
45 days	0.3139	0.1918	0.1775	0.1728	0.1769	0.1947
60 days	0.4041	0.2627	0.1762	0.2203	0.1976	0.1985
90 days	0.5652	0.3796	0.2879	0.3126	0.3003	0.3228
120 days	0.6973	0.4810	0.4586	0.3838	0.4275	0.3873

To summarize, considering its high accuracy compared to other advanced models, the LSTM is very efficient, especially for long-term predictions such as the 60-day forecast, due to the architectural advantages resulting from the process of parameter tuning and transfer to other tasks. GRUs are easier to train and faster to execute than LSTMs, but they may not be as effective at storing and accessing long-term dependencies. A Bidirectional

LSTM is better suited for applications that work offline, as it requires the subsequent timestamp in advance. On the other hand, the performance difference between stacked LSTMs and LSTMs comes from the additional dimensions for predicting the next value besides the time dimension.

Additionally, the RMSE, which describes the prediction error rate of time series algorithms, was compared with the Naïve Method and the algorithms that performed better than the Naïve Method were identified as successful algorithms for predicting future LWL values. The results of the Naïve Benchmark comparison of the algorithms are presented in Table 12 from day-1 to day-120 forecasting. The higher value for each analyzed algorithm for each prediction period indicates higher performance and good predictive power.

Based on the Naïve Benchmark, the performance of the algorithms increased up to the 60-days ahead predictions and then decreased at the 90- and 120-days ahead predictions. On average, the performance of GRU was higher for all time periods studied, whereas Stacked LSTM had a lower average performance value, followed by the Bidirectional LSTM algorithms.

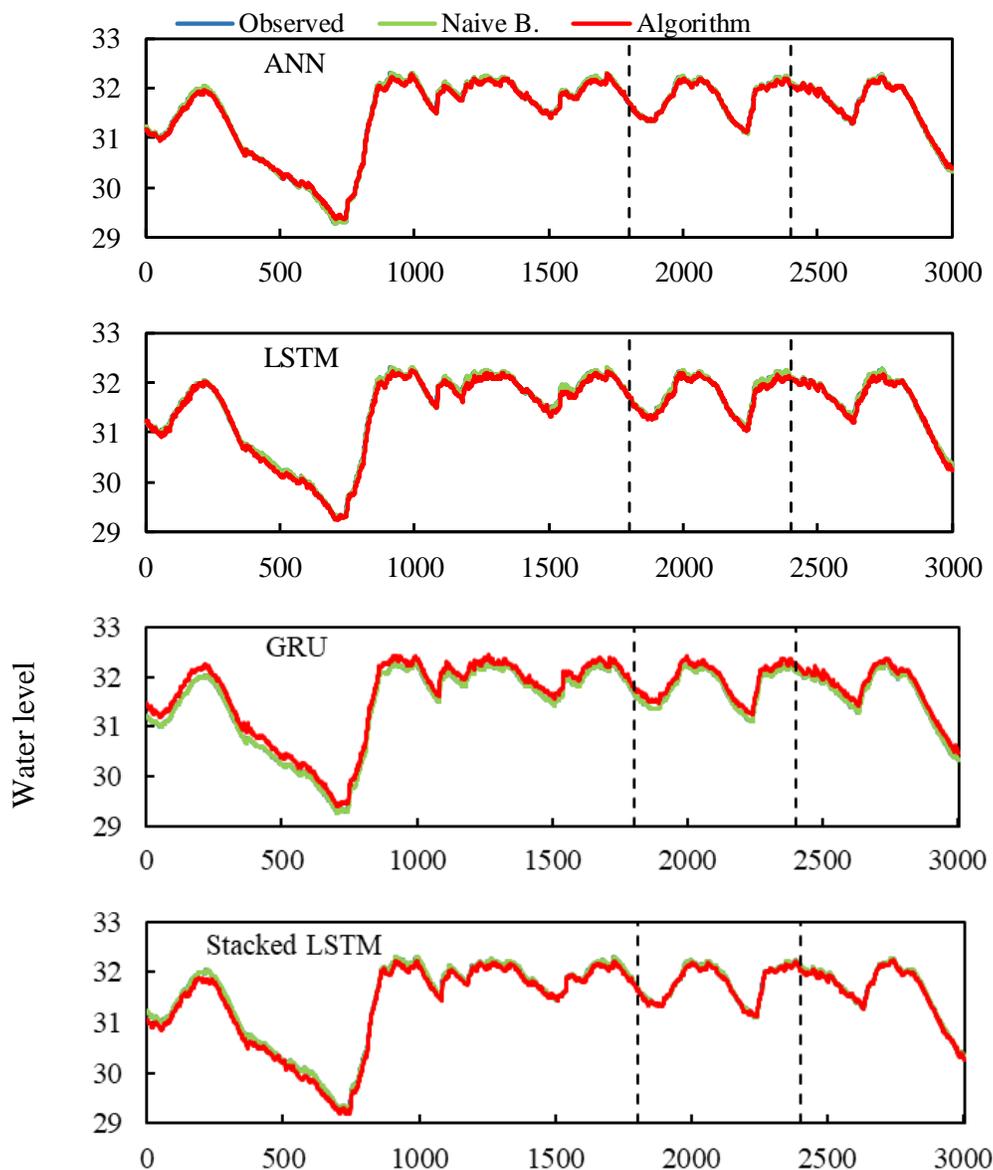
Table 12: Benchmark performance comparison of algorithms, figures indicates improvement over Naïve Method

Algorithm/Prediction Period	ANN	LSTM	GRU	Stacked LSTM	Bidirectional LSTM
1 day	2.26%	-18.92%	0.00%	-24.26%	-15.17%
5 days	8.40%	-6.01%	12.05%	-2.04%	-15.09%
10 days	7.10%	9.08%	17.80%	-1.70%	0.00%
20 days	19.84%	23.33%	36.70%	18.45%	20.94%
30 days	33.87%	46.08%	48.91%	55.89%	55.51%
45 days	48.29%	55.51%	57.98%	55.83%	46.87%
60 days	42.41%	78.55%	58.87%	68.64%	68.24%
90 days	39.29%	65.01%	57.55%	61.21%	54.59%
120 days	36.71%	41.30%	58.00%	47.97%	57.16%

The variability between the Naïve Benchmark comparison values is much more pronounced than the RMSE values (Table 12). The decreasing performance goes down to -24.26%, suggesting that it would be disadvantageous to use the RNN-based algorithm for predicting the specific time period. The results also show that the RMSE results of some algorithms are close to those of the Naïve Method, especially for the prediction of 5 and 10 days. Therefore, the algorithms will be tested even more if it is necessary to use these algorithms for future LWL values. The results show an increase in performance of at least 18.45% (Stacked LSTM) when the prediction horizon is set to 20 days or more. When compared to the Naive Method, LSTM showed the highest performance with an improvement of 78.55% over the Naive method when forecasting for 60 days. It is also

worth to note that ANN is the only algorithm that performed better than Naïve Method in 1-day prediction period.

The performance of ANN, LSTM, GRU, Stacked LSTM and Bidirectional LSTM for LWL, observed and estimated values compared to the Naïve Method for scenarios from day 1 to day 120 in advance are shown in Figures 19-27. It can be seen from Figure 19 that the observed and simulated lines are generally distributed closely for each investigated model, which shows that all RNN algorithms and ANN have high simulation performance on day 1. However, as forecasting time extend from day 1 to day 120, the observed, estimated and Naïve Method lines differ between the algorithms.



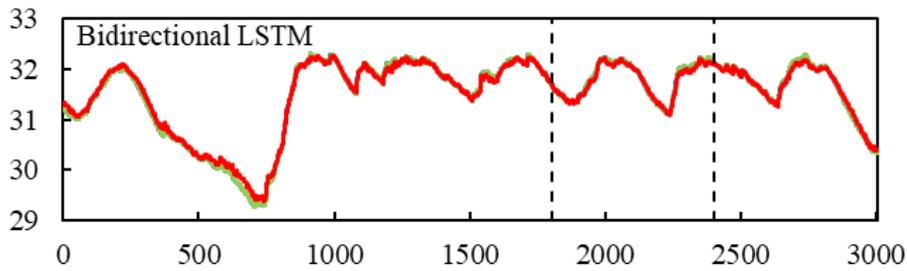
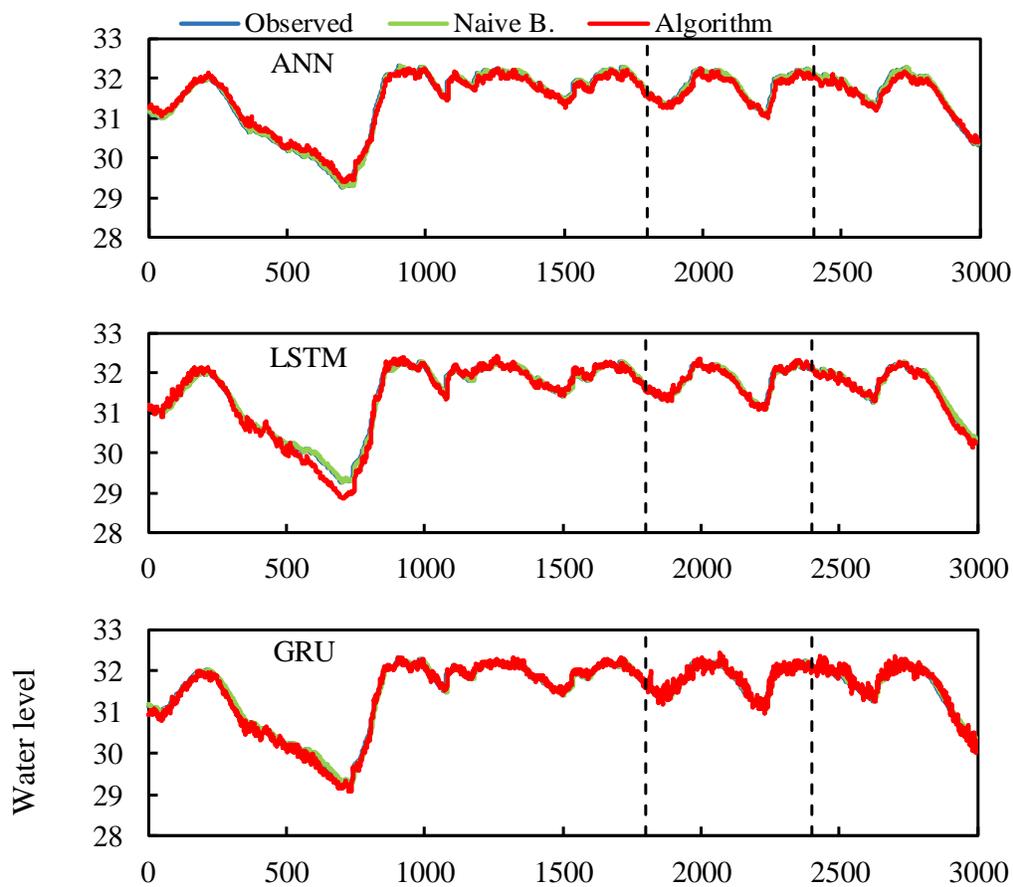


Figure 19: 1 day ahead prediction results

Figures 19 and 20 show the 1-day and 5-day prediction results of the gated RNN algorithms and ANN in comparison to the observed and Naïve Method values. The prediction results of all the studied algorithms are quite similar to each other and to the Naïve Method for 1-day and 5-day prediction (Figures 19–20), indicating good training, validation, and prediction.



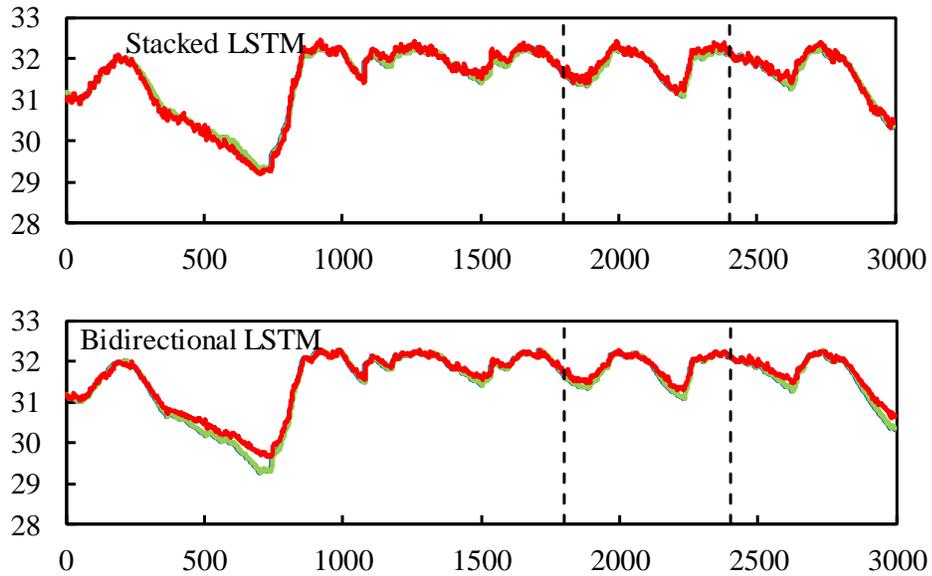
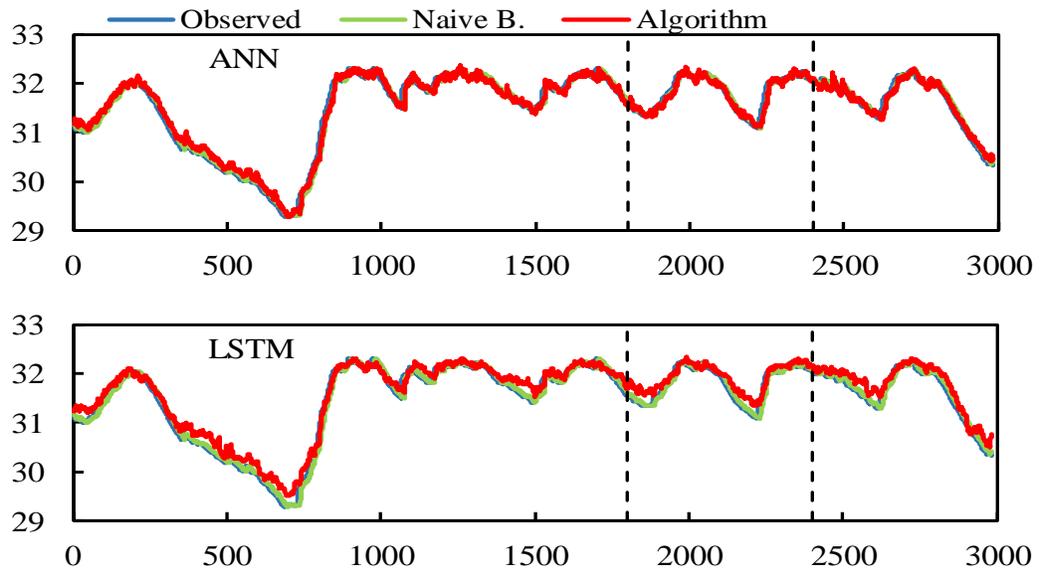


Figure 20: 5 days ahead prediction results

Figures 21 and 22 show the prediction results of the gated RNN algorithms for 10 and 20 days ahead and the comparison with the observed values and the Naïve Method. Compared to the Naïve method, all tested algorithms had a similar prediction trend for 10 days ahead, but all algorithms outperformed the Naïve method in prediction for 20 days ahead.



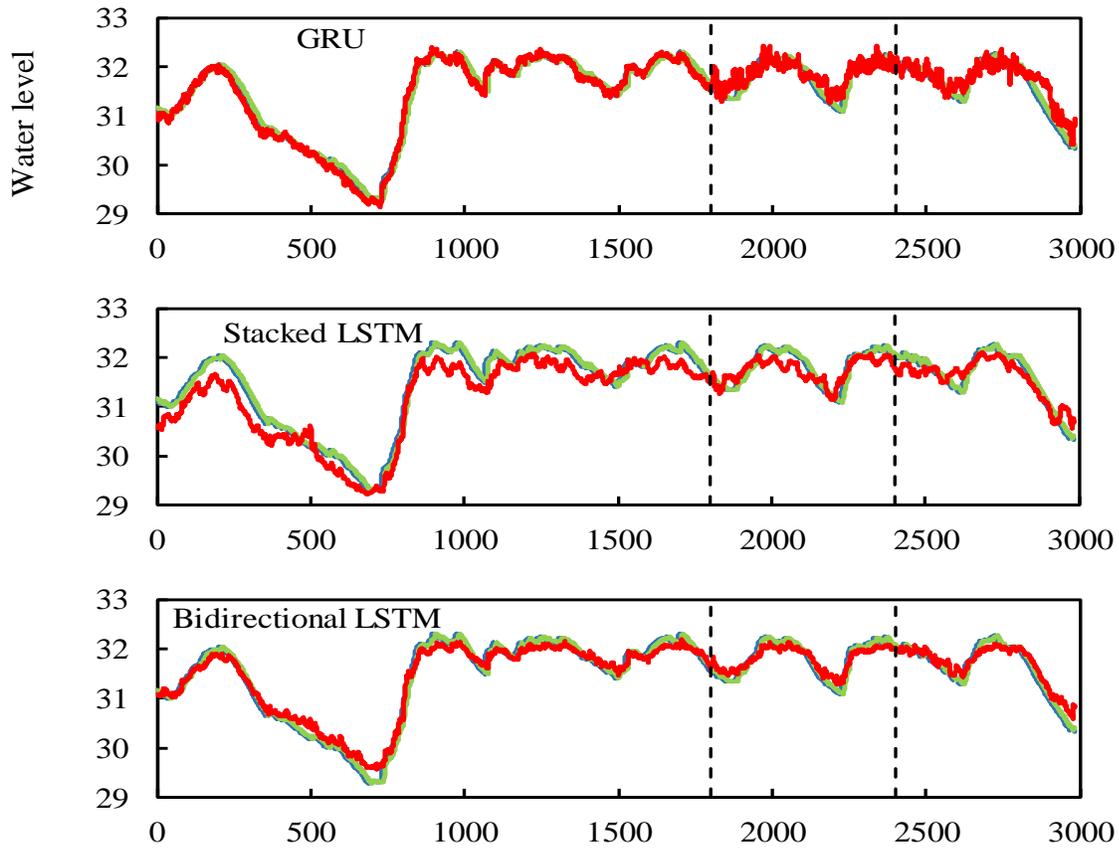
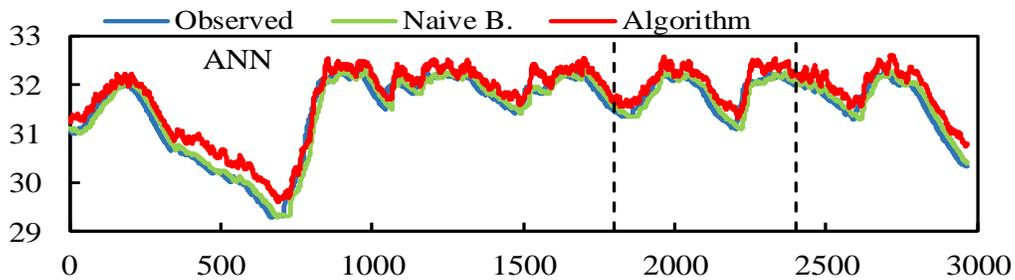


Figure 21: 10 days ahead prediction results

When forecasting 10 and 20 days ahead, the GRU achieved the best results (Figures 21–22), showing a lower RMSE (Table 11) and a higher performance improvement of the Naïve Method (Table 12).



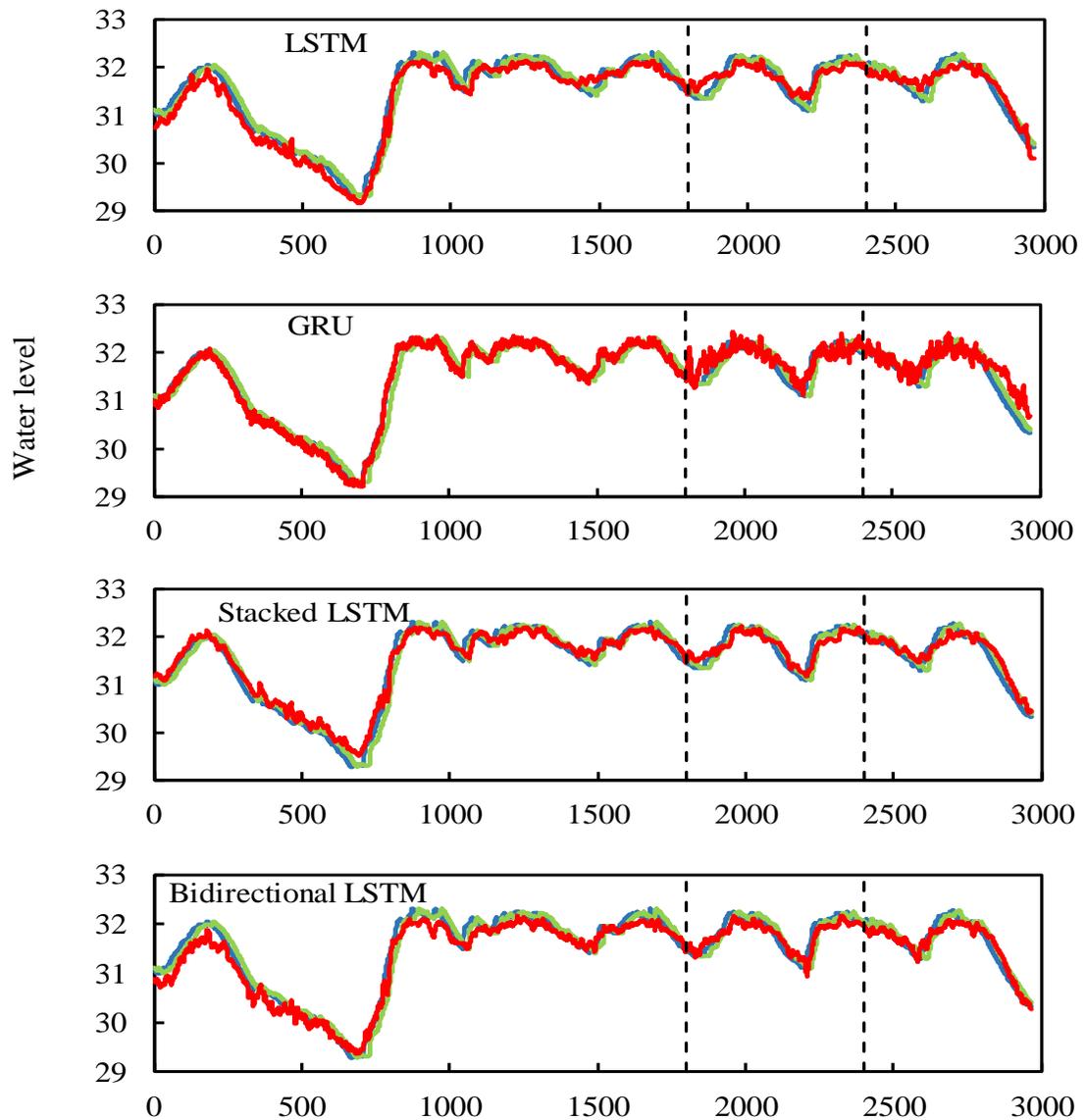


Figure 22: 20 days ahead prediction results

When comparing the performance results of the algorithms for day 30, Stacked LSTM and Bidirectional LSTM produced similar prediction performance to LSTM and GRU, whereas for day 45 prediction, GRU, Stacked LSTM, and LSTM algorithms produced similar performance to Bidirectional LSTM (Figures 23–24, Table 11).

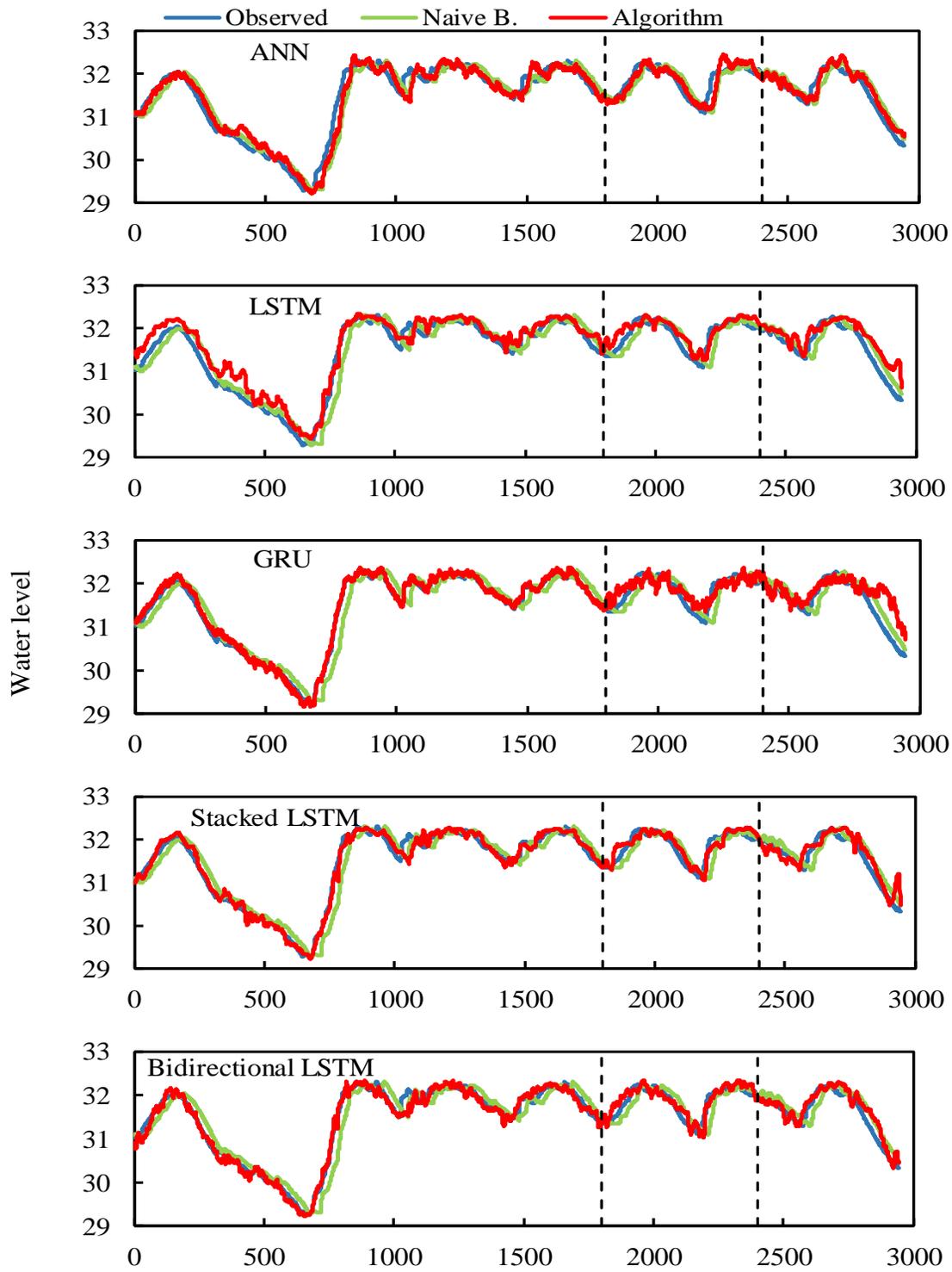
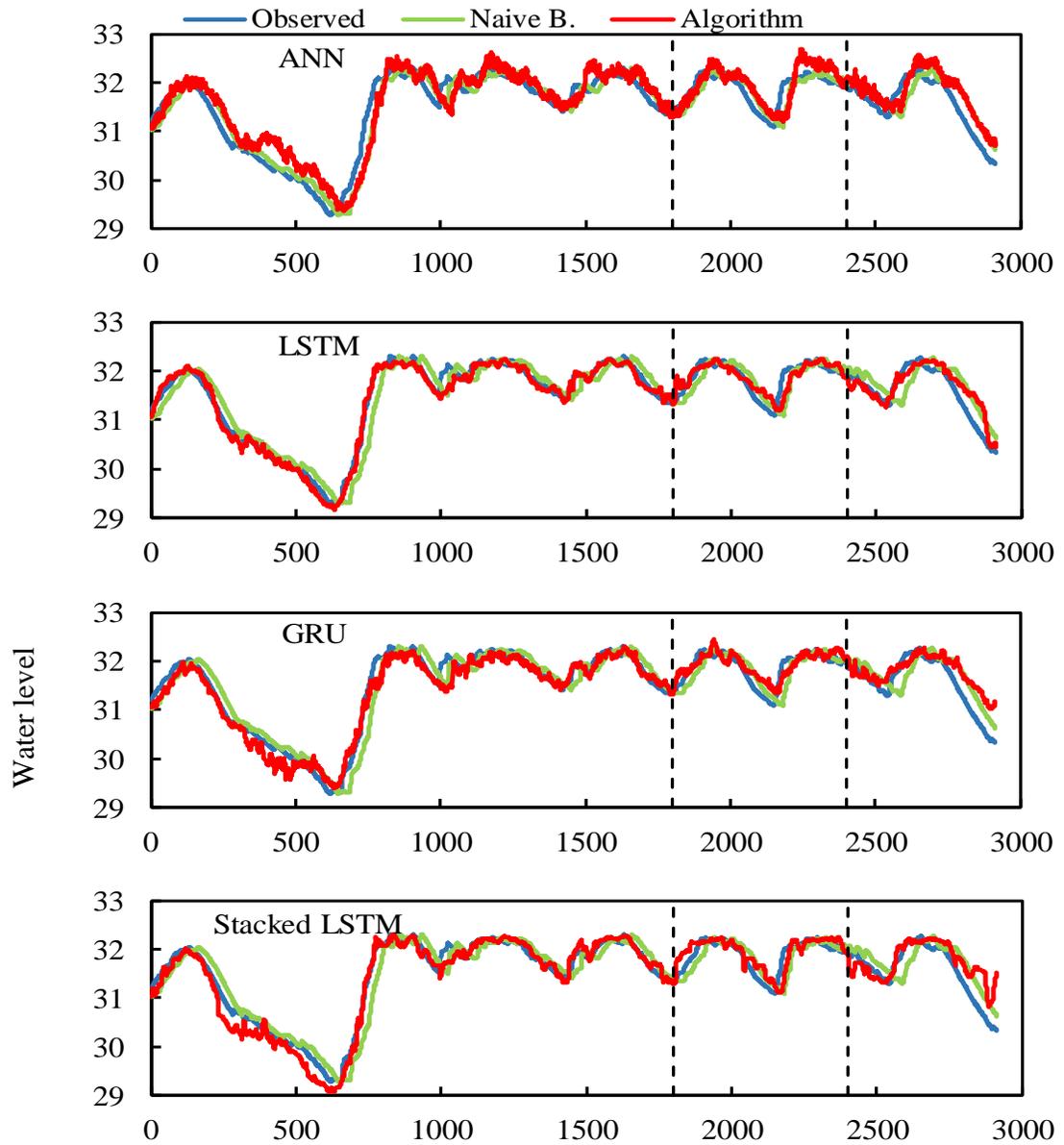


Figure 23: 30 days ahead prediction results

The 45-day prediction results in Figure 24 indicate the ANN algorithm starts to perform less than RNN-based algorithms and deviate from observed values visibly. The RNN-based algorithms don't have this many drastic changes in 45 days of prediction.



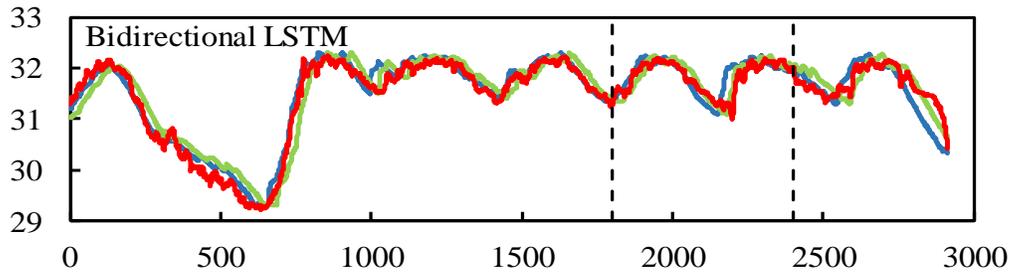
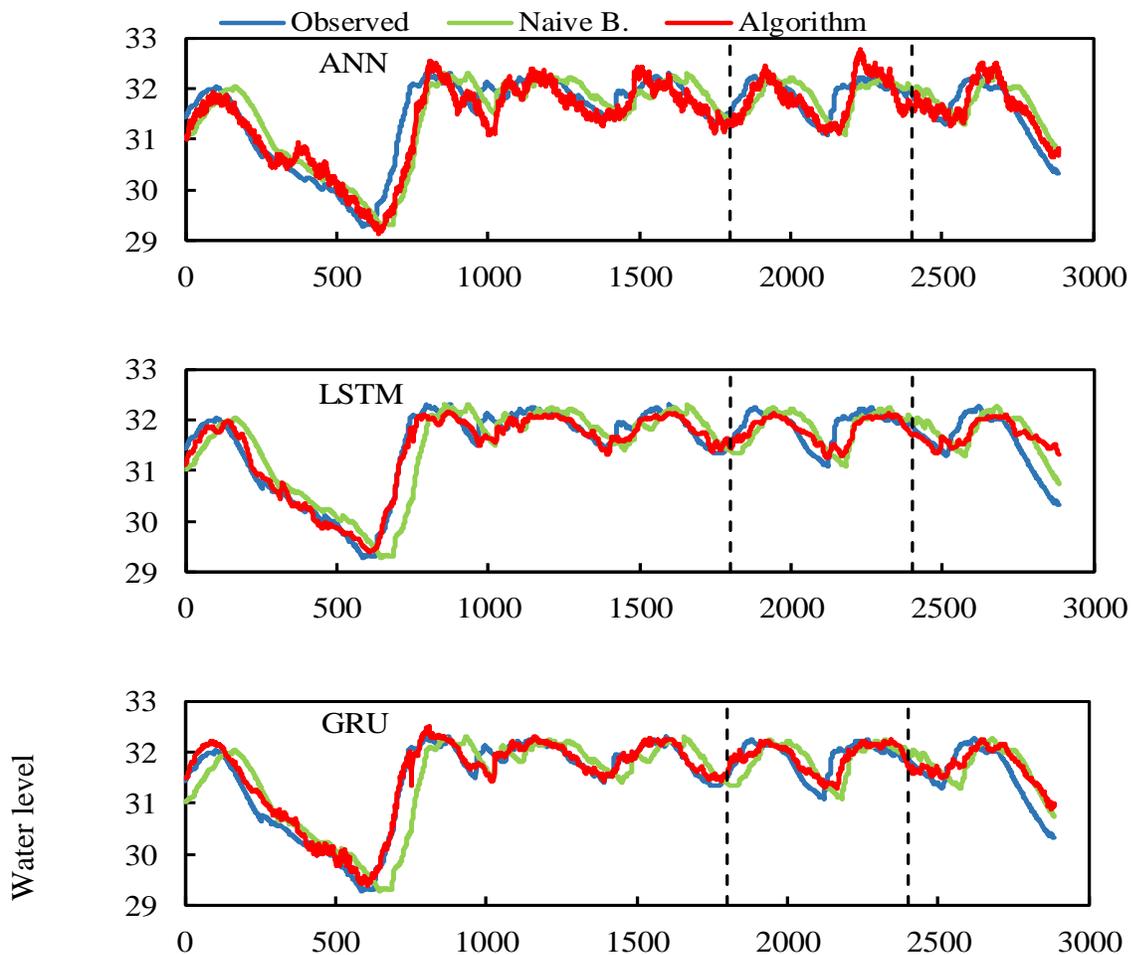


Figure 24: 45 days ahead prediction results

Figures 25 and 26 show the 60- and 90-day forecast results of the gated RNN algorithms and the comparison with the observed values and the Naive Method. The 60th day was the culminating point for the prediction performance of the tested algorithms, and LSTM performed better than LWL at the 60-day prediction based on the RMSE and Naïve Method values.



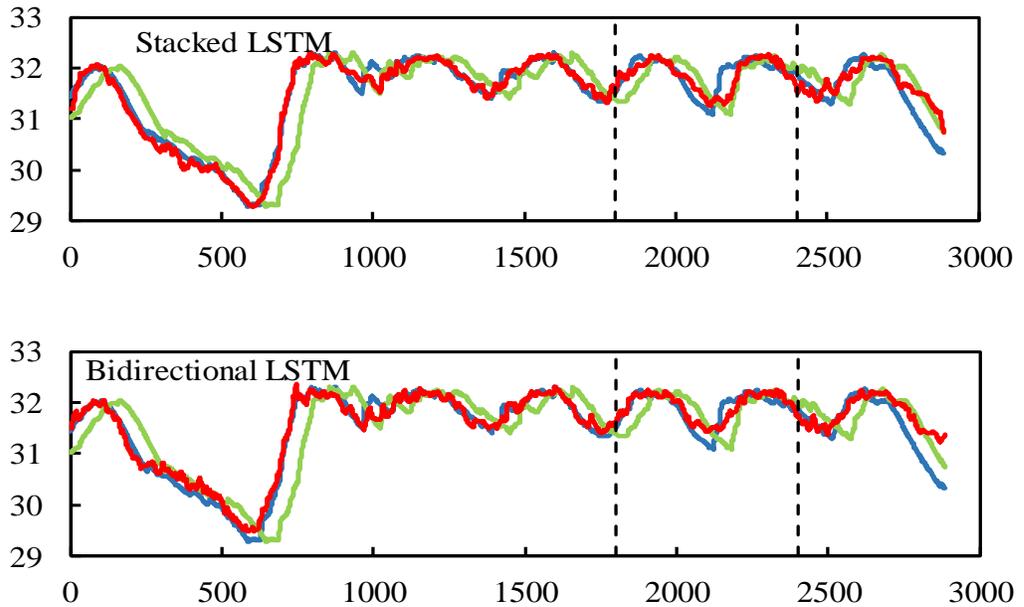
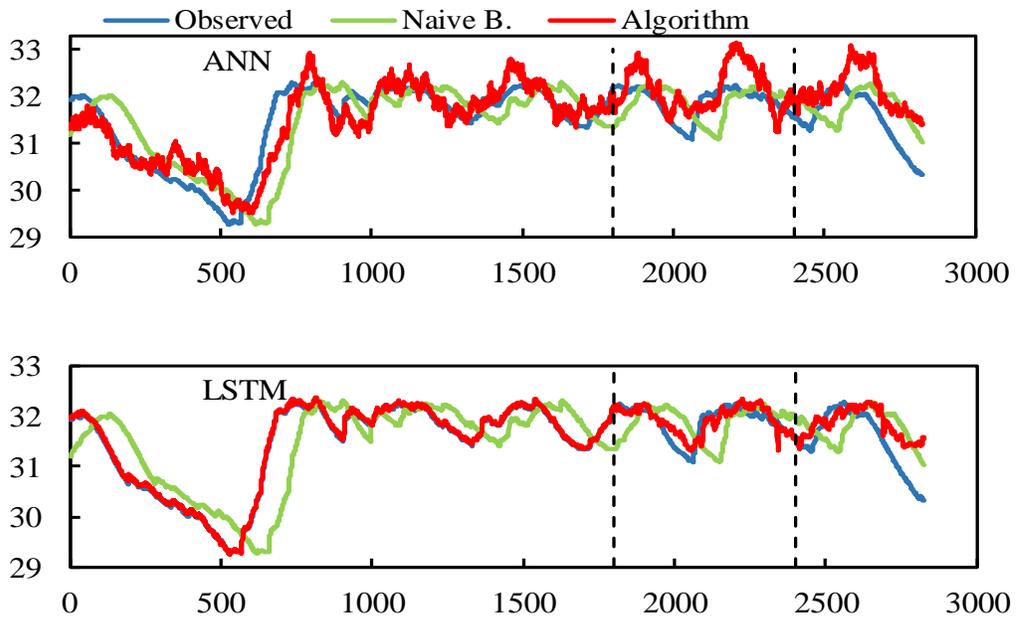


Figure 25: 60 days ahead prediction results

Although all tested algorithms performed well in 60-day prediction (Table 11), LSTM provided the closest prediction values to the observed values of LWL 60 days in advance than the other methods, as shown in Figure 25. The algorithms started to deviate more for 90-day prediction from observed values, as presented in Figure 26. The deviation is drastic for the ANN algorithm, which proves the algorithm performs not as well as RNN-based algorithms.



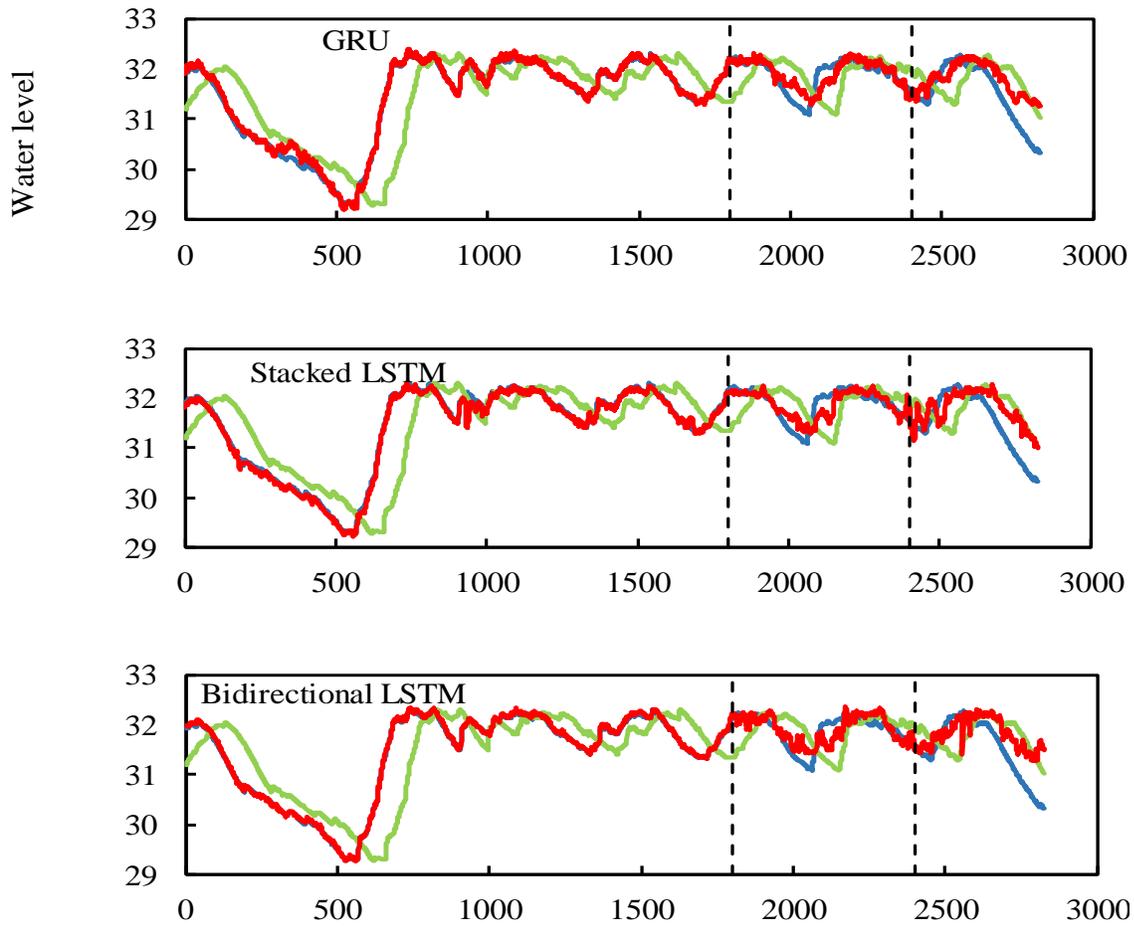


Figure 26: 90 days ahead prediction results

For the 120-day-ahead predictions, there was a significant decrease in values for the studied algorithms compared to the Naïve method, with the exception of GRU. Although the prediction performance was low, the GRU algorithm provided statistically similar prediction performance for days 60, 90, and 120. These results show that the GRU algorithm may still be superior to the other algorithms in terms of prediction accuracy with higher Naïve Benchmark values. The level of concordance between the actual value and the predicted value, however, is greater than the actual value and is not exceptionally high.

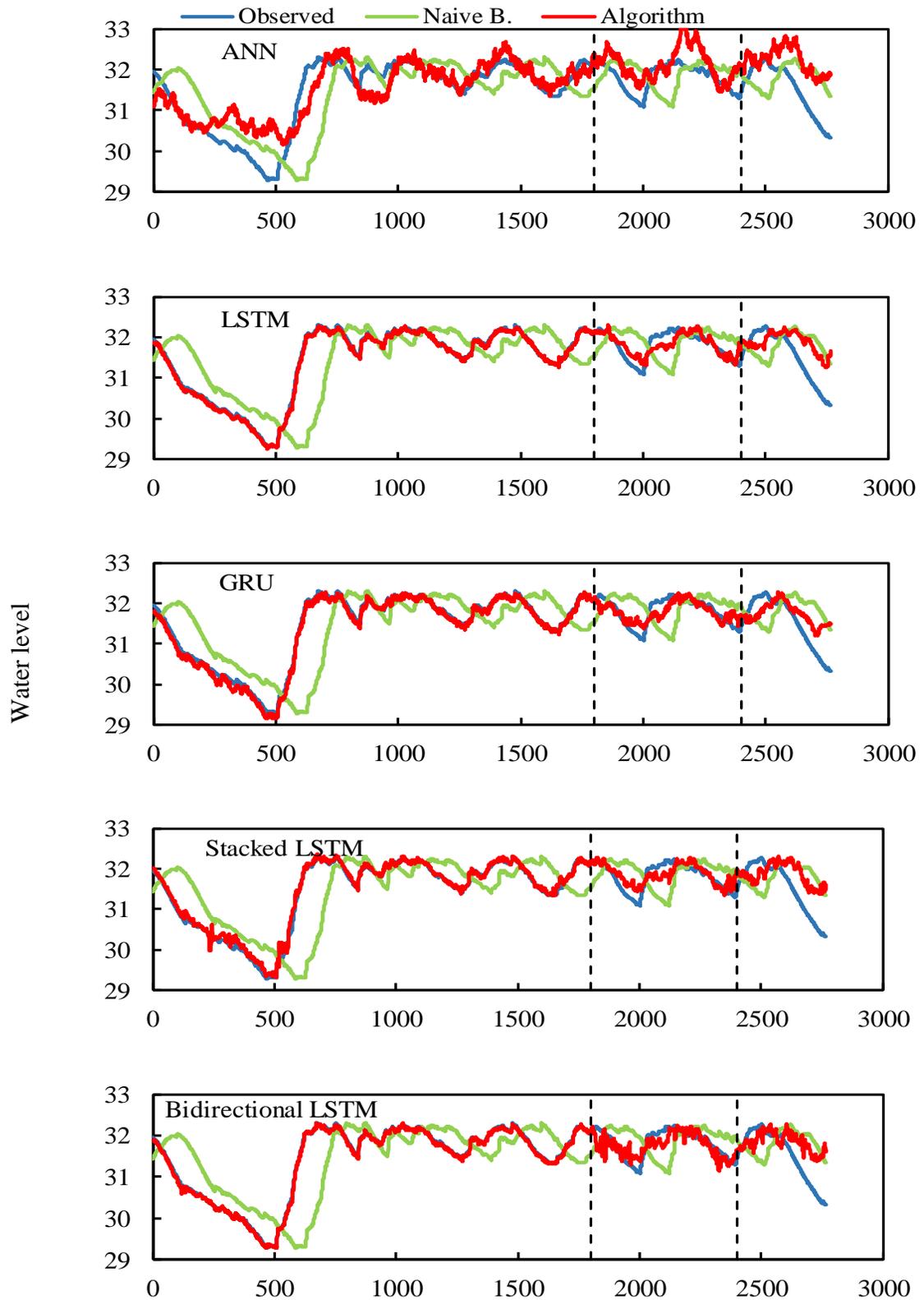


Figure 27: 120 days ahead prediction results

As a summary of Figures 19-27, the LSTM algorithm was the best-performing model because it had higher prediction accuracy and a smoother fitting curve than other models. Although there were some discrepancies due to significant discrepancies in the data set, the predicted results almost matched the actual results when predicting 60 days ahead. Among the proposed algorithms, the LSTM algorithm was clearly superior in tracking the nonlinear behavior of Lake Sapanca over a 60-day period, with the smallest RMSE (0.1762 m) and a higher performance ratio compared to the Naïve Method score (78.55%). Thus, when a model is needed for long-term forecasting LWL, the LSTM-based DL algorithm can help automate and manage LWL to implement more effective water management strategies for 60-day forecasting.

Table 13: 1 Day Diebold Mariano Test results (Numbers represent p -values)

Naïve Method-ANN	0.578	ANN-B. LSTM	0.523
Naïve Method-LSTM	0.122	LSTM-GRU	0.099
Naïve Method-GRU	0.005	LSTM- Stacked LSTM	0.326
Naïve Method-Stacked LSTM	0.485	LSTM- B. LSTM	0.608
Naïve Method-B. LSTM	0.261	GRU- Stacked LSTM	0.014
ANN-LSTM	0.264	GRU- B. LSTM	0.037
ANN-GRU	0.011	Stacked LSTM- B. LSTM	0.623
ANN-Stacked LSTM	0.878		

*B. LSTM: Bidirectional LSTM

As it can be seen in Table 13, Naïve Method-ANN, Naïve Method-LSTM, Naïve Method - Stacked LSTM, Naïve Method - Bidirectional LSTM, ANN-LSTM, ANN-Stacked LSTM, ANN-Bidirectional LSTM, LSTM-GRU, LSTM- Stacked LSTM, LSTM-Bidirectional LSTM and Stacked LSTM- Bidirectional LSTM forecasting comparisons are not significant in terms of Diebold Mariano test (p -value >0.05). Therefore, these algorithms can be used interchangeably. On the other hand, Naïve Method-GRU, ANN-GRU, GRU-Stacked LSTM, and GRU-Bidirectional LSTM algorithms have significantly (p -value <0.05) different forecasting capabilities, which indicates these algorithms shouldn't be used interchangeably. In this case, it is better to use the algorithm that has a lower RMSE score in Table 11.

Table 14 presents Diebold Mariano Test results for a 5-day prediction period. As it can be seen in the table, the forecasting capabilities are not significantly different for Naïve Method-ANN, Naïve Method-LSTM, Naïve Method-Stacked LSTM, ANN-LSTM, ANN-GRU, ANN-Stacked LSTM, ANN-Bidirectional LSTM, LSTM-GRU, LSTM-Stacked LSTM, LSTM-Bidirectional LSTM, GRU-Stacked LSTM, GRU-Bidirectional LSTM, Stacked LSTM-Bidirectional LSTM algorithm comparisons. However, Naïve Method-GRU and Naïve Method-Bidirectional LSTM are significantly different with a p -value <0.05 . Therefore, since the Naïve Method and the GRU algorithm for 5 days produced a p -value of 0.031, the GRU algorithm should be preferred instead of the Naïve Method to predict the next 5 days.

Table 14: 5 Days Diebold Mariano Test results (Numbers represent p -values)

Naïve Method-ANN	0.094	ANN-B. LSTM	0.233
Naïve Method-LSTM	0.310	LSTM-GRU	0.210
Naïve Method-GRU	0.031	LSTM-Stacked LSTM	0.728
Naïve Method-Stacked LSTM	0.181	LSTM-B. LSTM	0.062
Naïve Method-B. LSTM	0.007	GRU-Stacked LSTM	0.358
ANN-LSTM	0.474	GRU-B. LSTM	0.516
ANN-GRU	0.581	Stacked LSTM-B. LSTM	0.122
ANN-Stacked LSTM	0.710		

*B. LSTM: Bidirectional LSTM

Diebold Mariano Test results for the 10-day prediction are shown in Table 15. The table indicates the forecasting capabilities are not significant for the algorithm pairs of Naïve Method-ANN, LSTM-Stacked LSTM, LSTM-Bidirectional LSTM, GRU-Stacked LSTM, and Stacked LSTM-Bidirectional LSTM. On the contrary, Naïve Method-LSTM, Naïve Method-GRU, Naïve Method-Stacked LSTM, Naïve Method-Bidirectional LSTM, ANN-LSTM, ANN-GRU, ANN-Stacked LSTM, ANN-Bidirectional LSTM, LSTM-GRU, and GRU-Bidirectional LSTM have significantly different forecasting capabilities.

Table 15: 10 Days Diebold Mariano Test results (Numbers represent p -values)

Naïve Method-ANN	0.984	ANN-B. LSTM	0.012
Naïve Method-LSTM	0.007	LSTM-GRU	0.032
Naïve Method-GRU	0.000	LSTM-Stacked LSTM	0.244
Naïve Method-Stacked LSTM	0.000	LSTM-B. LSTM	0.878
Naïve Method-B. LSTM	0.011	GRU-Stacked LSTM	0.319
ANN-LSTM	0.008	GRU-B. LSTM	0.022
ANN-GRU	0.000	Stacked LSTM-B. LSTM	0.188
ANN-Stacked LSTM	0.000		

*B. LSTM: Bidirectional LSTM

Diebold Mariano Test results for the 10-day prediction are shown in Table 15. The table indicates the forecasting capabilities are not significant for the algorithm pairs of Naïve Method-ANN, LSTM-Stacked LSTM, LSTM-Bidirectional LSTM, GRU-Stacked LSTM, and Stacked LSTM-Bidirectional LSTM. On the contrary, Naïve Method-LSTM, Naïve Method-GRU, Naïve Method-Stacked LSTM, Naïve Method-Bidirectional LSTM, ANN-LSTM, ANN-GRU, ANN-Stacked LSTM, ANN-Bidirectional LSTM, LSTM-GRU, and GRU-Bidirectional LSTM have significantly different forecasting capabilities.

The best-performing algorithm, GRU (Table 11), is significantly different from other algorithms except Stacked LSTM. Thus, the GRU algorithm should be preferred to predict the next 10 days, and the Stacked LSTM algorithm could be used interchangeably.

Table 16: 20 Days Diebold Mariano Test results (Numbers represent p -values)

Naïve Method-ANN	0.055	ANN-B. LSTM	0.593
Naïve Method-LSTM	0.612	LSTM-GRU	0.000
Naïve Method-GRU	0.000	LSTM-Stacked LSTM	0.874
Naïve Method-Stacked LSTM	0.506	LSTM-B. LSTM	0.364
Naïve Method-B. LSTM	0.686	GRU-Stacked LSTM	0.001
ANN-LSTM	0.000	GRU-B. LSTM	0.000
ANN-GRU	0.000	Stacked LSTM-B. LSTM	0.287
ANN-Stacked LSTM	0.072		

*B. LSTM: Bidirectional LSTM

It is presented in Table 16 that Naïve Method-ANN, Naïve Method-LSTM, Naïve Method-Stacked LSTM, Naïve Method-Bidirectional LSTM, ANN-Stacked LSTM, ANN-Bidirectional LSTM, LSTM-Stacked LSTM, LSTM-Bidirectional LSTM, and Stacked LSTM-Bidirectional LSTM don't have significantly different forecasting capabilities. Therefore, these pairs can be used interchangeably in order to forecast the next 20 days. However, Naïve Method-GRU, ANN-LSTM, ANN-GRU, LSTM-GRU, GRU-Stacked LSTM, and GRU-Bidirectional LSTM are significantly different in terms of forecasting the next 20 days.

From Table 11, it can be seen that, on day-20 predictions, the best performance improvement comes from the GRU algorithm. Accordingly, the p -values are significant ($p < 0.05$) based on the Diebold-Mariano test, which confirms the superiority of GRU. Regarding the Naïve Method comparison (Table 12) and Diebold Mariano (Table 16) test results, only the GRU algorithm should be preferred to predict the LWL for the next 20 days.

Table 17: 30 Days Diebold Mariano Test results (Numbers represent p -values)

Naïve Method-ANN	0.055	ANN-B. LSTM	0.593
Naïve Method-LSTM	0.009	LSTM-GRU	0.004
Naïve Method-GRU	0.000	LSTM-Stacked LSTM	0.006
Naïve Method-Stacked LSTM	0.009	LSTM-B. LSTM	0.000
Naïve Method-B. LSTM	0.161	GRU-Stacked LSTM	0.000
ANN-LSTM	0.000	GRU-B. LSTM	0.000
ANN-GRU	0.000	Stacked LSTM-B. LSTM	0.202
ANN-Stacked LSTM	0.072		

*B. LSTM: Bidirectional LSTM

The results for Diebold Mariano tests indicate the Naïve Method-ANN, Naïve Method-Bidirectional LSTM, ANN-Stacked LSTM, ANN-Bidirectional LSTM, and Stacked LSTM-Bidirectional LSTM are not significantly different in terms of forecasting capability. It is significantly different for Naïve Method-GRU, Naïve Method-Stacked LSTM, ANN-LSTM, ANN-GRU, LSTM-GRU, LSTM-Stacked LSTM, LSTM-Bidirectional LSTM, GRU-Stacked LSTM, and GRU-Bidirectional LSTM.

When evaluating the Diebold-Mariano tests, the predictions of the best-performing Stacked LSTM (Table 11) algorithm compared to the Naïve Method, LSTM, and GRU have a p -value of less than 0.05, indicating that the Stacked LSTM algorithm should be chosen for predicting the next 30 days. However, since the p -value is greater than 0.05 compared to the Bidirectional LSTM, the Stacked LSTM algorithm can be used interchangeably with the Bidirectional LSTM algorithm.

Table 18: 45 Days Diebold Mariano Test results (Numbers represent p -values)

Naïve Method-ANN	0.815	ANN-B. LSTM	0.741
Naïve Method-LSTM	0.014	LSTM-GRU	0.003
Naïve Method-GRU	0.006	LSTM-Stacked LSTM	0.000
Naïve Method-Stacked LSTM	0.253	LSTM-B. LSTM	0.011
Naïve Method-B. LSTM	0.923	GRU-Stacked LSTM	0.541
ANN-LSTM	0.025	GRU-Bidirectional LSTM	0.662
ANN-GRU	0.443	Stacked LSTM-B. LSTM	0.295
ANN-Stacked LSTM	0.169		

*B. LSTM: Bidirectional LSTM

The Table 18 shows that Naïve Method-ANN, Naïve Method-Stacked LSTM, Naïve Method-Bidirectional LSTM, ANN-GRU, ANN-Stacked LSTM, ANN-Bidirectional LSTM, GRU-Stacked LSTM, GRU-Bidirectional LSTM and Stacked LSTM-Bidirectional LSTM pairs do not have significant difference when it comes to forecasting capability according to Diebold Mariano test. It differs for Naïve Method-LSTM, Naïve Method-GRU, ANN-LSTM, LSTM-GRU, LSTM-Stacked LSTM, and LSTM-Bidirectional LSTM because these pairs are significantly different in terms of forecasting ability.

Table 11 demonstrates that GRU performs better than the other algorithms in the 45-day forecast based on RMSE scores. Table 17 confirms once again that the prediction capabilities of the GRU algorithm have p -values less than 0.05 for the Naïve Method and LSTM. Furthermore, the p -values are more remarkable than 0.05 compared to ANN, stacked LSTM, and bidirectional LSTM. Therefore, the GRU algorithm can be used interchangeably with the ANN, Stacked LSTM, and Bidirectional LSTM algorithms in order to predict the next 45 days.

As Table 19 indicates, Naïve Method-Bidirectional LSTM, ANN-Stacked LSTM, GRU-Stacked LSTM, GRU-Bidirectional LSTM, and Stacked LSTM-Bidirectional LSTM pairs do not have statistical significance in order to forecast the next 60 days. It is the opposite for Naïve Method-ANN, Naïve Method-LSTM, Naïve Method-GRU, Naïve Method-Stacked LSTM, ANN-LSTM, ANN-GRU, ANN-Bidirectional LSTM, LSTM-GRU, LSTM-Stacked LSTM, and LSTM-Bidirectional LSTM pairs.

Table 19: 60 Days Diebold Mariano Test results (Numbers represent p -values)

Naïve Method-ANN	0.000	ANN-B. LSTM	0.002
Naïve Method-LSTM	0.003	LSTM-GRU	0.000
Naïve Method-GRU	0.012	LSTM-Stacked LSTM	0.000
Naïve Method-Stacked LSTM	0.009	LSTM-B.LSTM	0.017
Naïve Method-B. LSTM	0.187	GRU-Stacked LSTM	0.917
ANN-LSTM	0.000	GRU-B. LSTM	0.229
ANN-GRU	0.046	Stacked LSTM-B. LSTM	0.192
ANN-Stacked LSTM	0.058		

*B. LSTM: Bidirectional LSTM

According to the Naïve Benchmark performance comparison score, the LSTM algorithm gave the highest performance with 78.55% for the 60-day ahead predictions (Table 12). As a result, with a p -value of less than 0.05, the Diebold-Mariano test results likewise demonstrate that the LSTM algorithm performs significantly better in terms of both prediction accuracy and stability (Table 19). Considering the results of RMSE, the Naïve Benchmark, and the Diebold-Mariano test with a p -value of less than 0.05 for day-60 ahead prediction, it suggests that only the LSTM algorithm should be preferred to predict the next 60 days for more accuracy.

Table 20: 90 Days Diebold Mariano Test results (Numbers represent p -values)

Naïve Method-ANN	0.023	ANN-B. LSTM	0.000
Naïve Method-LSTM	0.000	LSTM-GRU	0.119
Naïve Method-GRU	0.000	LSTM- Stacked LSTM	0.437
Naïve Method- Stacked LSTM	0.000	LSTM-B. LSTM	0.940
Naïve Method-B. LSTM	0.000	GRU-Stacked LSTM	0.434
ANN-LSTM	0.000	GRU-B. LSTM	0.138
ANN-GRU	0.000	Stacked LSTM-B. LSTM	0.482
ANN-Stacked LSTM	0.000		

*B. LSTM: Bidirectional LSTM

It is revealed in Table 20 that LSTM-GRU, LSTM-Stacked LSTM, LSTM-Bidirectional LSTM, GRU-Stacked LSTM, GRU-Bidirectional LSTM, and Stacked LSTM-Bidirectional LSTM are not statistically different when it comes to forecasting the next 90 days. On the other hand, Naïve Method-ANN, Naïve Method-LSTM, Naïve Method-GRU, Naïve Method-Stacked LSTM, Naïve Method-Bidirectional LSTM, ANN-LSTM, ANN-GRU, ANN-Stacked LSTM, and ANN-Bidirectional LSTM are significantly different from each other.

In Table 11, the best-performing algorithm is LSTM, according to RMSE results. The Diebold Mariano test results suggest that the algorithm's forecasting capability is significantly different from the Naïve Method, ANN, and GRU. Therefore, the LSTM algorithm should be preferred to predict the next 90 days over these algorithms. However, since the results for LSTM- Stacked LSTM and LSTM-Bidirectional LSTM are not

significantly different ($p > 0.05$), the LSTM algorithm can be used interchangeably with Stacked LSTM and Bidirectional LSTM.

Table 21: 120 Days Diebold Mariano Test results (Numbers represent p -values)

Naïve Method-ANN	0.222	ANN-B. LSTM	0.000
Naïve Method-LSTM	0.000	LSTM-GRU	0.015
Naïve Method-GRU	0.000	LSTM-Stacked LSTM	0.995
Naïve Method-Stacked LSTM	0.000	LSTM-B. LSTM	0.752
Naïve Method-B. LSTM	0.000	GRU-Stacked LSTM	0.014
ANN-LSTM	0.000	GRU-B. LSTM	0.033
ANN-GRU	0.000	Stacked LSTM-B. LSTM	0.747
ANN-Stacked LSTM	0.000		

*B. LSTM: Bidirectional LSTM

Diebold Mariano Test results for 120 days of prediction are shown in Table 21. Accordingly, the pairs of Naïve Method-ANN, LSTM-Stacked LSTM, LSTM-Bidirectional LSTM, and Stacked LSTM-Bidirectional LSTM are not statistically different from each other to forecast the next 120 days. The statistically different pairs are Naïve Method-LSTM, Naïve Method-GRU, Naïve Method-Stacked LSTM, Naïve Method-Bidirectional LSTM, ANN-LSTM, ANN-GRU, ANN-Stacked LSTM, ANN-Bidirectional LSTM, LSTM-GRU, GRU-Stacked LSTM, and GRU-Bidirectional LSTM.

According to Table 21, it is clear that the implemented RNN algorithms provide a relatively accurate prediction pattern when the forecasting values are compared with the observed data for 120 days ahead prediction, even though the magnitude of the Naïve Benchmark scores is reduced compared to the 60 days ahead prediction. In Table 11, it can be seen that the GRU algorithm has the best performance against the Naïve Method for day-120 ahead prediction. Additionally, the algorithm has Diebold-Mariano test results with a p -value of less than 0.05 against other algorithms (Table 21), which suggests that the GRU algorithm should be preferred to predict the next 120 days of LWL.

The summary of all Diebold Mariano test results is presented in Table 22. From the obtained results for LWL prediction from day 1 to day 120, we can see that: (1) day-60 predictions is the most optimized LWL detection based on high Naïve Benchmark performance comparison values. (2) The best performance of the investigated algorithms can change in terms of selected prediction periods. (3) The LSTM algorithm can better predict LWL for a 60-day advance with higher accuracy, which allows water managers to take action. In addition, it is worth noting that Bidirectional LSTM and Stacked LSTM algorithms contribute little or no performance increase for short prediction periods of less than 20 days. (4) Lastly, FFNN-based algorithms perform better on short time periods and even surpass RNN-based methods and the Naïve Method, but the forecasting capability is not statistically significant, which suggests using the most primitive method that consumes less computing power in shorter time periods.

Table 22: Forecast difference results of Naïve Method, ANN and RNN algorithms based on Diebold Mariano test for increasing day intervals from day-1 to day-120 (p value ≤ 0.05 indicates the significance of the test results, green boxes indicate significantly different prediction results with distinct tones, red boxes indicate insignificant results).

	Day-1	Day-5	Day-10	Day-20	Day-30	Day-45	Day-60	Day-90	Day-120
Naïve Method- ANN	0.578	0.094	0.984	0.055	0.055	0.815	0.000	0.023	0.222
Naïve Method- LSTM	0.122	0.310	0.007	0.612	0.009	0.014	0.003	0.000	0.000
Naïve Method- GRU	0.005	0.031	0.000	0.000	0.000	0.006	0.012	0.000	0.000
Naïve Method- Stacked LSTM	0.485	0.181	0.000	0.506	0.009	0.253	0.009	0.000	0.000
Naïve Method- B. LSTM	0.261	0.007	0.011	0.686	0.161	0.923	0.187	0.000	0.000
ANN-LSTM	0.264	0.474	0.008	0.000	0.000	0.025	0.000	0.000	0.000
ANN-GRU	0.011	0.581	0.000	0.000	0.000	0.443	0.046	0.000	0.000
ANN-Stacked LSTM	0.878	0.710	0.000	0.072	0.072	0.169	0.058	0.000	0.000
ANN-B. LSTM	0.523	0.233	0.012	0.593	0.593	0.741	0.002	0.000	0.000
LSTM-GRU	0.099	0.210	0.032	0.000	0.004	0.003	0.000	0.119	0.015
LSTM- Stacked LSTM	0.326	0.728	0.244	0.874	0.006	0.000	0.000	0.437	0.995
LSTM- B. LSTM	0.608	0.062	0.878	0.364	0.000	0.011	0.017	0.940	0.752
GRU- Stacked LSTM	0.014	0.358	0.319	0.001	0.000	0.541	0.917	0.434	0.014
GRU- Bidirectional LSTM	0.037	0.516	0.022	0.000	0.000	0.662	0.229	0.138	0.033
Stacked LSTM- B. LSTM	0.623	0.122	0.188	0.287	0.202	0.295	0.192	0.482	0.747

*B. LSTM: Bidirectional LSTM

Among the features, the most important one that affects the output was determined as withdrawal by the Mutual Information technique conducted (Lv et al., 2020). The importance levels can be ordered as withdrawal, average temperature, minimum temperature, maximum temperature, and precipitation respectively (Figure 28).

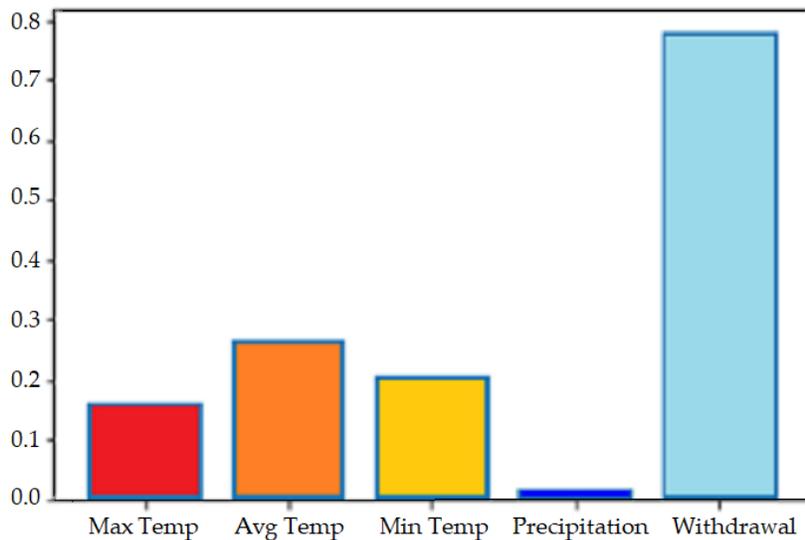


Figure 28: Variable Importance

CHAPTER 5

5. EVALUATION OF CLIMATIC CONDITIONS ON MICROCYSTIN VARIATIONS IN WATER COLUMN

Accurate LWL prediction is a necessity not only to prevent possible drought conditions but also possible water quality effect. Therefore, this thesis conducted extra work to observe the relationship between microcystin that previously observed during low LWL periods. In addition to LWL effect, this study experimented the relationship for maximum temperature, mean temperature, minimum temperature, precipitation, light intensity and evaporation. This experiment was conducted in order to reveal the importance to predict LWL in advance to take measurable actions in advance.

In recent years, cyanobacterial blooms in water reservoirs, including freshwater lakes, have become a serious problem worldwide. In natural freshwater, temperature and light intensity are meteorological parameters directly related to photosynthesis that promote algal growth. In addition to climatic conditions, various nutritional parameters such as phosphorus, nitrogen forms, and carbon sources favor the growth of various cyanobacterial species that are precursors for algal toxin microcystin production. Microcystin is an algal toxin produced by many genera of cyanobacteria. *Planktothrix* is a common producer in Lake Sapanca. Microcystin has been shown to have harmful effects on humans, animals, and even plants. Because of its relatively low volatility and short half-life (a few days to weeks), microcystin can linger in water bodies, with serious consequences for the environment, society, and economy. The presence of toxic microcystin is becoming one of the most important problems for water quality and health. Therefore, the World Health Organization has recommended a provisional guideline level of $1 \mu\text{g L}^{-1}$ for drinking water. Microcystin is routinely monitored in freshwater systems where cyanobacterial blooms occur. For instance, microcystin is monitored for Lake Sapanca monthly during the last decade during the summer months at five different depths because microcystin production is highly influenced by environmental conditions.

Various meteorological, nutritional, and anthropogenic factors such as storms, surface runoff, and point and diffuse pollution accelerate the increase of nutrients in aquatic ecosystems. The increase in phosphorus and nitrogen in water, as well as the increase in turbidity with full water turnover in the water column, leads to frequent cyanobacterial blooms, which in turn produce toxins. Although conventional drinking water treatment methods (e.g., coagulation, flocculation, and filtration) remove both cyanobacterial cells and intracellular microcystin, additional treatment is needed for algal byproducts that

increase treatment costs and are a public nuisance. Therefore, to understand and evaluate algal blooms and microcystin in water, it is critical to evaluate both meteorological parameters, nutrient enrichment, and aquatic ecosystem parameters that affect the phytoplanktonic biomass.

A lot of research has been done to predict the concentration of this cyanobacterial toxin in freshwater using statistical models, Bayesian methods, ML, and DL techniques. This is because microcystin is a crucial indicator of phytoplankton blooms and is resistant to decomposition and long-term persistence in aquatic environments. These studies highlight the importance of different environmental parameters, including particle organic matter, water temperature, total phosphorus, and dissolved oxygen. However, predicting the occurrence and elucidating the causes of microcystin using statistical or computer models does not always provide reliable results. Because of slow water exchange, ongoing global warming, and frequent anthropogenic interventions, the range of variation in water quality in a lake is extremely high, making it impractical to specify a general threshold for each environmental factor. In addition, microcystin levels measured over long periods of time in the past are not always sufficient to build computer models that use advanced estimation algorithms. To avoid such problems, many water researchers study water reservoirs by using an unsupervised classification algorithm. Estimating microcystin occurrence patterns from limited data as climate change progresses makes computerized prediction and analysis difficult. Therefore, to evaluate meteorological data on microcystin concentrations at five different depths in Lake Sapanca, we used regression models based on monthly monitoring data for the past five years to identify risk. Because there are few guidelines for total microcystins, we chose the safe concentration of microcystin suggested by the World Health Organization as the evaluation criterion (1.0 µg/L).

Empirical models' predictions are complex and possibly site-specific due to variability in regional environmental conditions, making them challenging to generalize across wide areas. Consequently, statistical methods that can effectively manage vast and diverse meteorological and biological data sets, establish causal links, and capture both linear and nonlinear correlations between variables are required.

Cyanobacterial blooms and their toxin product, microcystin, occur mainly in the metalimnion region. As a result, microcystin is abundant in the surface water layer, which is typically 0–10 m. However, upon cell lysis, it may mix with the water column. Particularly prior to and following bloom, abundance on the bottom of the water column occasionally exceeded what was present in the surface water. Due to this fact, microcystin was observed in the water column at various depths. Studies conducted on Lake Sapanca revealed that while microcystin was not detectable in the water column on top of 10 m, many variations were found as low as 20 m. However, the region between 15 and 25 m has the highest amount (Albay et al., 2003). Taking account of those findings, the microcystin concentrations at the surface (1 m, 5 m, 10 m, 15 m, and 20 m) are measured over the period of 2019–2023 to understand the changing meteorological situation on water quality affected by algal growth (Figure 29).

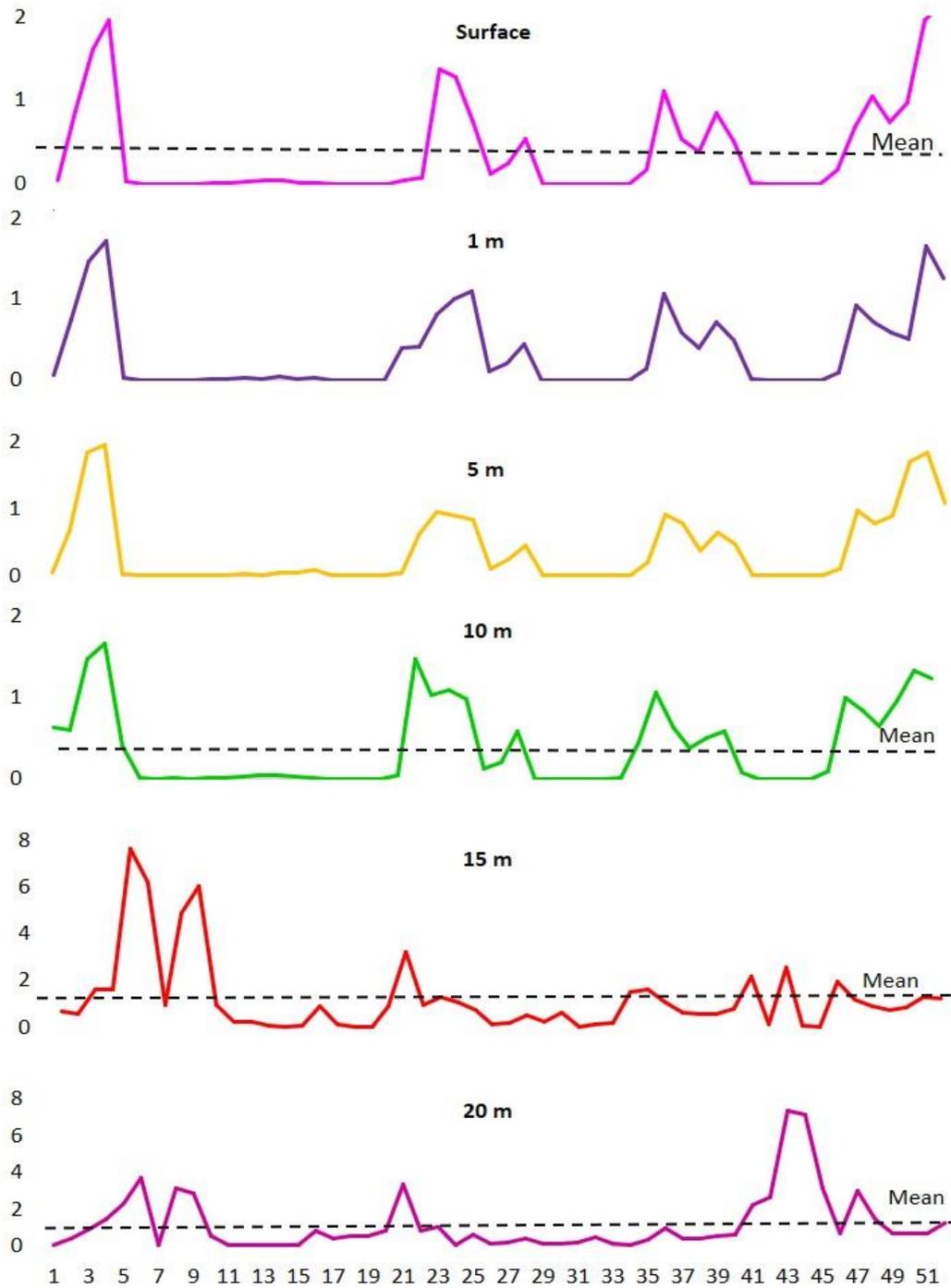


Figure 29: Vertical distribution of microcystin concentrations measured monthly intervals between 21 March 2019 to 12 April 2023 in Lake Sapanca.

Concentrations of microcystin ranged from 0.0 to 2.15 $\mu\text{g/L}$ on the surface; 0.0 to 1.71 $\mu\text{g/L}$ in 1m; 0.0 to 1.95 $\mu\text{g/L}$ in 5m; 0.0 to 1.66 $\mu\text{g/L}$ in 10m; 0.0 to 7.66 $\mu\text{g/L}$ in 15m; and 0.0 to 7.30 $\mu\text{g/L}$ in 20m during the monitoring period. The highest concentrations and extreme values were obtained from depths of 15 and 20 m (Figure 30). The highest concentrations were observed during the summer months (May–August) for depths of 15 and 20, respectively, but interestingly, the higher concentrations for surface water above 5 m were observed during the spring turnover (March–April) period. These results indicate the obvious effects of meteorological factors on cyanobacterial bloom formation and, consequently, microcystin production. However, seasonal effects, stratification, and lake turnover also have a significant impact on microcystin variation in the water column at different depths.

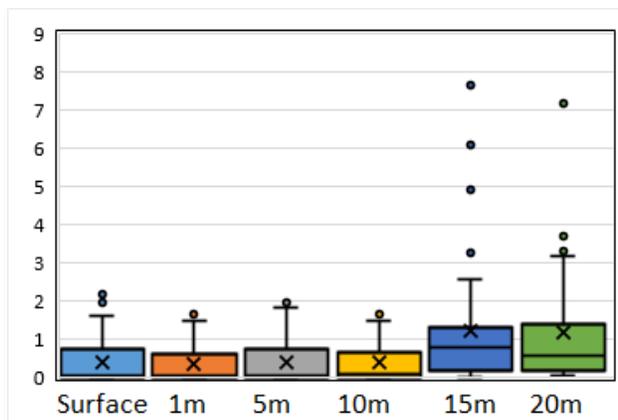


Figure 30: Box and whisker plot of microcystin concentrations at different water depths of Lake Sapanca during 21 March 2019 to 12 April 2023.

The microcystin concentration in all sampled depths showed approximately the same increase pattern over time, except for the samples collected from a depth of 15m (Figure 31). The variations in Figure 31 indicate there is an increasing trend of microcystin for surface water, 1m, 5m, 10m and 20 m depths. However, the trend is decreasing for 15 m depth. The microcystin level was almost similar at each depth of the first 10 m; however, significant differences were recorded in spring and autumn, specifically in vertical mixing periods. During the summer, the microcystin concentration stayed relatively low ($<0.5 \mu\text{g/L}$) or undetectable from May to October. The highest concentrations were observed during the winter period from November to April, with a significant fluctuation that coincided with the mixing period. By contrast, the microcystin concentrations were higher at sampling depths of 15 and 20 m. The microcystin is recorded at all sampling times during the experimental period. In general, the concentrations were below $2 \mu\text{g/L}$ for both sampling depths; however, the highest concentrations of around $8 \mu\text{g/L}$ were recorded during the summer stratification phase (June to August). For the two years 2020 and 2021, the microcystin concentration was the lowest (with $<3.31 \mu\text{g/L}$), especially for 2021 ($<1.61 \mu\text{g/L}$).

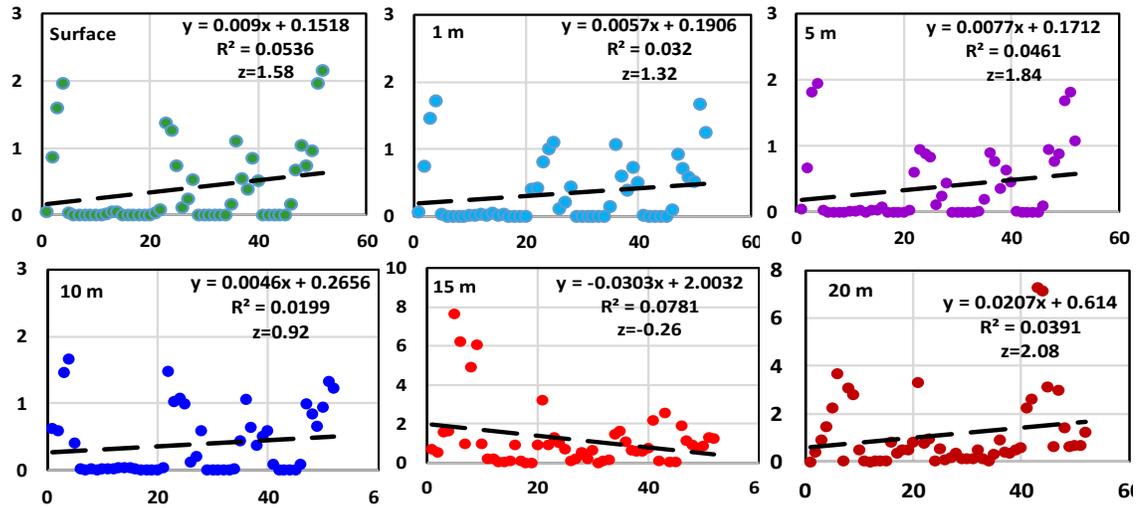


Figure 31: Linear trend of microcystin concentration in the water column at different depths from surface to 20 m during 21 March 2019 to 12 April 2023. (x-axis: Data rows in sequence, y-axis: Microcystin concentration).

The nonparametric Mann-Kendall test shows that the microcystin concentration decreases monotonically at a depth of 15 m and increases at other depths. However, only the microcystin trend at 20 m depth was significant at the 95% confidence limit with a z-value of 2.08 (Figure 31), indicating an increasing positive trend in the microcystin data time series that dominates at this depth.

Due to temporal and spatial variability, it is difficult to obtain sufficient input data needed for data-driven predictive models to analyze and learn the relationships between microcystin and meteorological parameters, i.e., temperature, precipitation associated with algal proliferation, and microcystin concentration. To better understand the changing meteorological parameters affecting microcystin concentration, Spearman correlations were evaluated using monthly microcystin data collected from raw water before water treatment. From Figure 32, the significant positive contribution of temperature to microcystin concentration is evident. Light intensity also has a positive effect on microcystin concentration. On the other hand, the water level of the lake had no significant effect on the microcystin concentration.

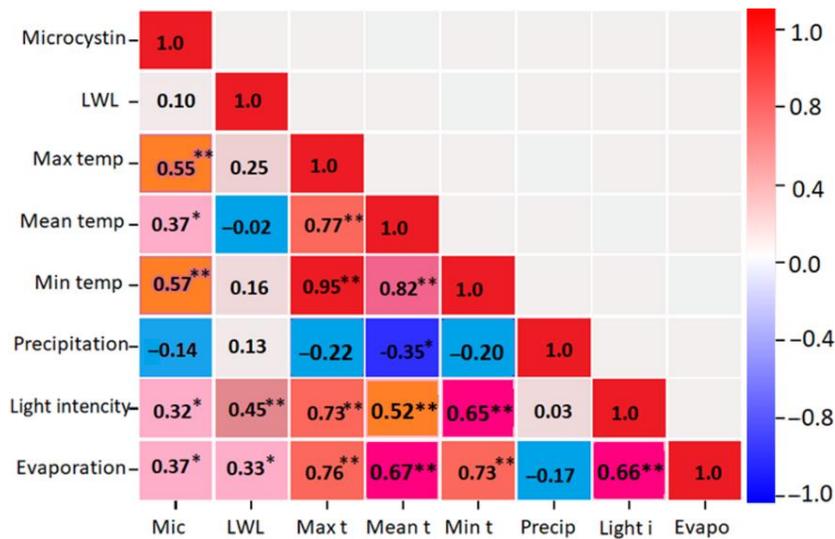


Figure 32: SRC between microcystin and meteorological parameters (** $p < 0.01$, * $p < 0.05$).

The degree of association differs in terms of features in Figure 32. The minimum temperature, maximum temperature, mean temperature, and evaporation have a moderate correlation with microcystin (Fowler et al., 2013). In addition, light intensity corresponds to a weak correlation. On the other hand, the LWL and precipitation have a very weak correlation with microcystin. From the results, it is better understood that the quality of water is rarely affected by the level of water. However, temperature, which is one of the indicators for predicting LWL, affects the water quality. Thus, it can be concluded that microcystin does not directly affect the water quality, but the effect is indirect when considering temperature values.

CHAPTER 6

6. DECISION SUPPORT TOOL

6.1. General Structure of the Tool

The decision support tool for LWL management using DL techniques is built using the Streamlit web-based prototype tool with the MVC (Model-View-Controller) software design pattern. Streamlit is an open-source and free web application platform to share data scripts created with the Python language on a decent application platform. It is the most popular web application for data scientists because of its well-suited integration with Python libraries and Plotly charts. It also allows for the deployment of the application on the community cloud platform. On the other hand, MVC (Model-View-Controller) is one of the most commonly used software development architecture that separates the application into three parts that are interconnected with each other. By doing this, it is possible to distinguish between internal representations of information and the methods by which information is offered to and received by users. The Model serves as the framework for the system design. It oversees the application's rules, logic, and data. The final output that the data is presented as is known as the View, which is represented by graphs, charts, and diagrams. The element that receives input and transforms it into commands for the 'Model' or 'View' is the Controller. The MVC pattern for the decision support tool is presented in Figure 33.

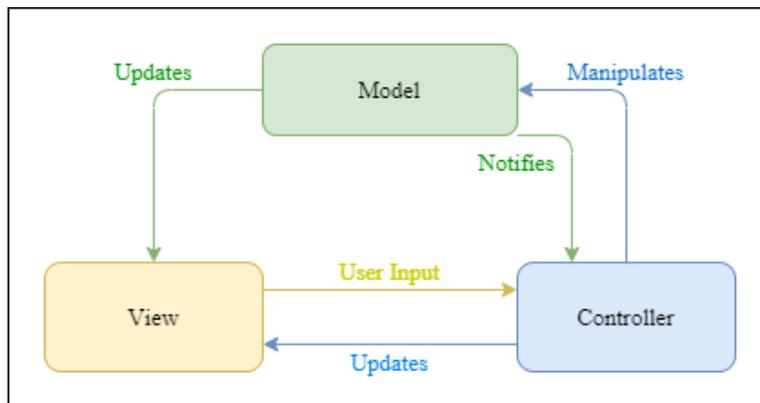


Figure 33: LWL Management Decision Support Tool MVC Design Pattern

The libraries and their versions are represented in Table 23 that is used for generating the LWL management decision support tool by using the Python programming methodology.

Table 23: Library Versions in Web Application

Libraries	Versions
Streamlit	1.21.0
Pandas	1.5.2
Numpy	1.24.1
Scikit-Learn	1.2.0
Keras	2.11.0
TensorFlow	2.11.0
Matplotlib	3.6.2
Seaborn	0.12.2
Plotly	5.16.1
OpenPyXL	3.1.2

The LWL management decision support tool is built using the Python programming language, with several libraries involved. The libraries used in the tool are Streamlit, Pandas, Numpy, Scikit-Learn, Keras, TensorFlow, Matplotlib, Seaborn, Plotly, and OpenPyXL.

The LWL management decision support tool aims to:

- To give a general snapshot of the lake with its hydrological and meteorological aspects.
- To assist water managers with their decisions regarding lake water use.
- To prevent the possible aftermath of climate change effects on drinking water supply
- To present late changes and possible actions in the future.
- To prevent excessive water use that might cause water shortages during drought seasons in the future.

The LWL management decision support tool has four main characteristics:

- Based on a web application
- Available to related units
- Availability of CRUD (Create, Read, Update, and Delete)
- Prediction for a 1, 2, and 4-month period

The LWL management decision support tool has three components:

- Dashboard
- Manual Data Entry
- Prediction

6.2. Components and Abilities

There are three components in the LWL management decision support tool: dashboard, manual entry, and prediction. The dashboard serves as a display of past and current trends in LWL and other determinative features. Furthermore, manual entry has the ability to enable user interaction with the application to feed input for the tool. Lastly, the predict component provides the anticipated future values in terms of 1, 2, and 4 months ahead. The component also functions as a warning to the water managers in case there is a water alert in the lake.

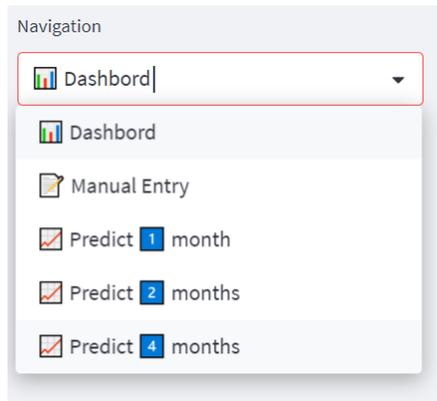


Figure 34: Navigation of Application

The navigation of the LWL management decision support tool is shown in Figure 34. It is designed as a sliding window, and when the user clicks one of the options, it immediately opens the selected page. In addition, when the user clicks one of the 'Predict' pages, the application automatically starts to predict the intended future period.

6.2.1. Dashboard

The dashboard component of the tool provides general structure, historical data analysis, and the latest view of the situation in the lake. The abilities in this component include the latest average values and their comparison to total averages, a radar chart, a line chart, the distribution of features, and the current dataset view.

a) Latest Average Values

The managers in water management for the lake require to see the latest changes before taking any action to modify the water level. In the dashboard, the first ability is to show the latest average values for the last week for LWL, precipitation, maximum temperature, average temperature, minimum temperature, and withdrawal. In addition, under the average values, it reveals the difference between the average value of last week and the total average value for each feature.

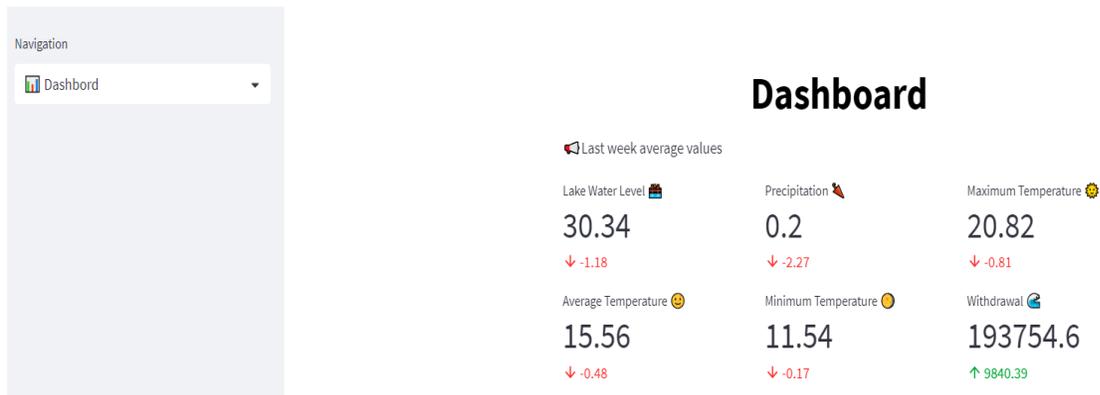


Figure 35: Latest average values in dashboard

The latest mean values and the difference between total means are shown in Figure 35.

b) Radar Chart

Another important point to consider for water managers is to compare the values between each variable. In order to compare the features that have similar mean values, a radar chart was created. Multivariate data are piled at the same central point on an axis in a radar chart. The chart has three or more radii, which are quantitative variables for comparison.

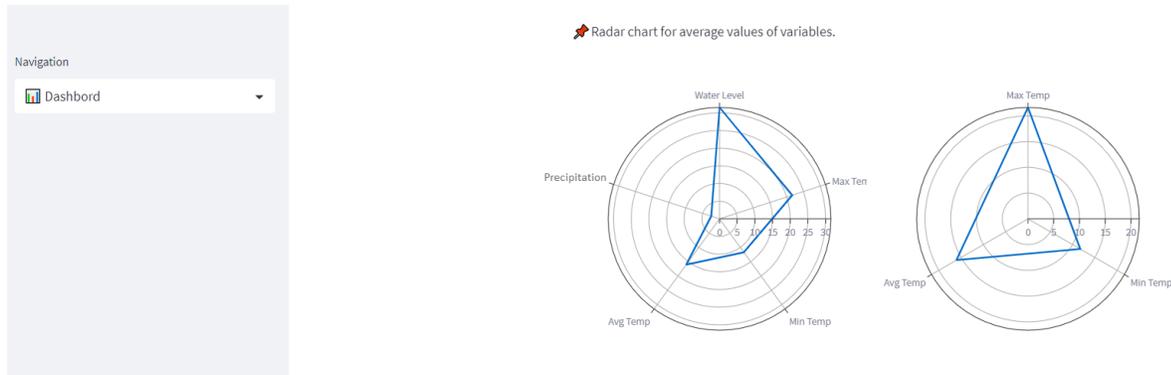


Figure 36: Radar chart in dashboard

As it can be seen in Figure 36, there are two radar charts demonstrated for water managers. The first radar chart compares average values for LWL, precipitation, maximum temperature, average temperature, and minimum temperature. On the other hand, the second chart compares only temperature features that have three sections: maximum, average, and minimum. The reason to create such two charts is because some features' average values are too low or too high, which prevents seeing the whole picture. This is also the reason not to add withdrawal to the radar chart since it has way too high values compared to other features.

c) Line Chart

It is important to see the trends of other features that affect LWL to see if the patterns overlap with each other. The potential benefits of trend comparison could be to see seasonal changes and to get detailed information in terms of feature importance for the output feature. In order to see the pattern between the output feature, LWL, and other features, a line chart was created. However, precipitation and withdrawal have such high values comparing with LWL, so the comparison between them becomes impossible with a line chart. Therefore, maximum temperature, average temperature, and minimum temperature were added to the line chart along with LWL.

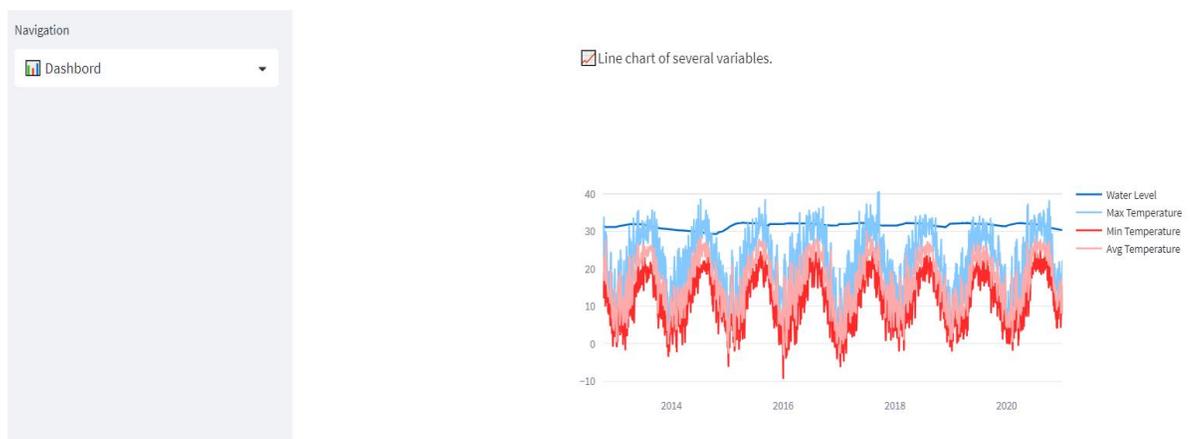


Figure 37: Line chart in dashboard

Figure 37 reveals the line chart that compares LWL with maximum temperature, average temperature, and minimum temperature. As it can be seen in the figure, the pattern between LWL and other features is visible. When all types of temperatures rise, the LWL decreases, and vice versa. The pattern becomes more drastic when the user zooms in on the graph in the application.

d) Distribution of Features

The line chart doesn't reveal all the distributions for the features in the dataset. Thus, the application has another ability to show feature distributions. The distribution graphs are all interactive, so the users can modify the parts to see from buttons such as '1m' for 1 month, '6m' for 6 months, 'YTD' for year-to-date, '1y' for 1 year, and 'all' to see the whole distribution. The application also allows users to select the intended part manually.

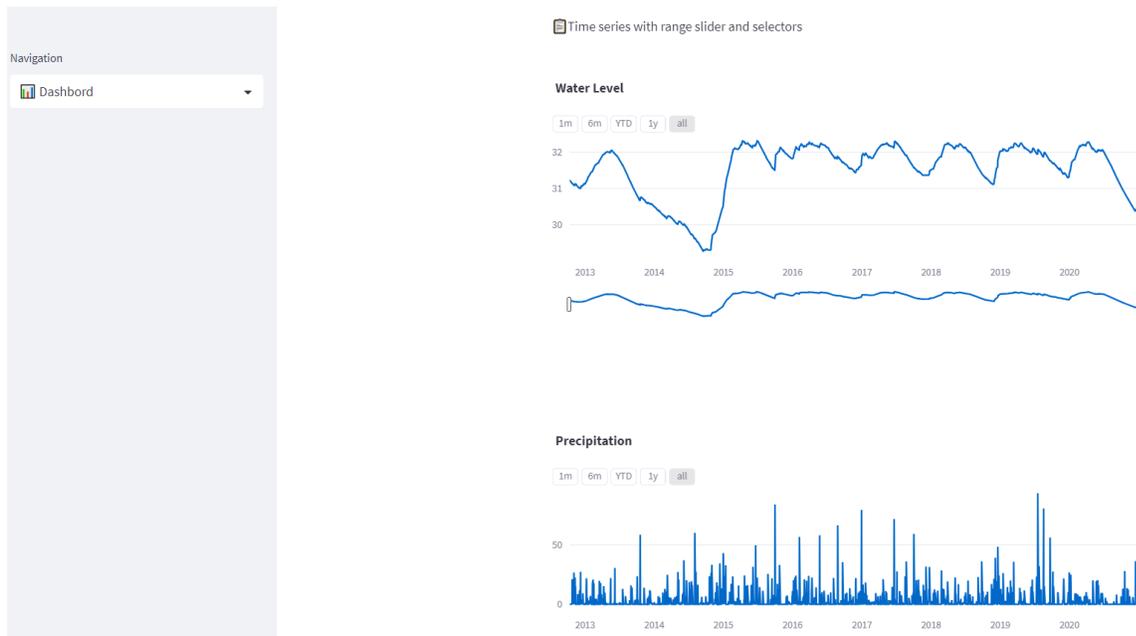


Figure 38: Feature Distributions in Dashboard

Figure 38 shows the distribution graph in the dashboard. All the features are displayed in a sequence, and the distributions can be compared in the same sequence.

e) Current Dataset View

Although the distributions and averages can be seen in the dashboard component, specific values for selected dates cannot be seen from the graphs. Consequently, the ‘Current Dataset View’ ability provides all the actual values in the dataset, considering related date value.

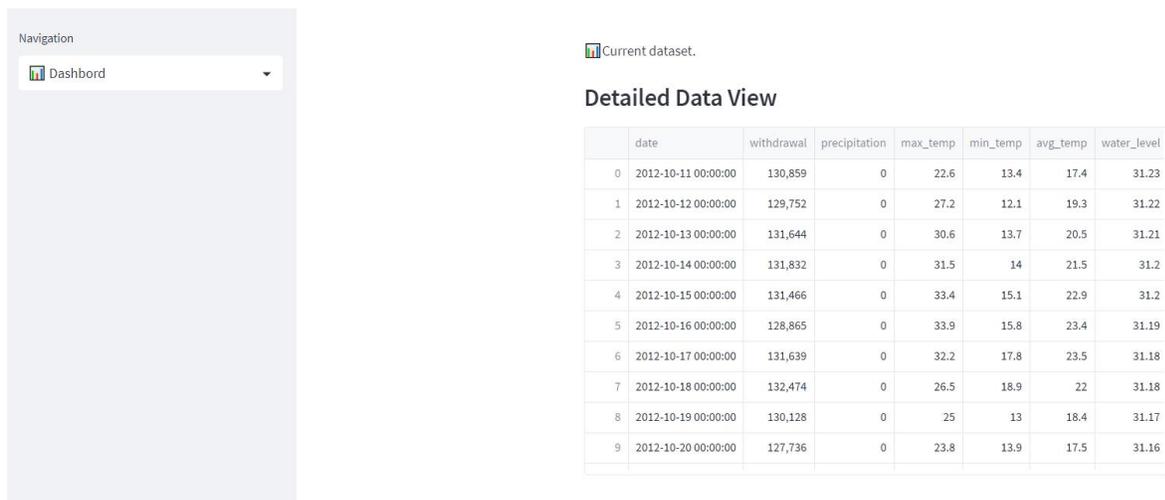


Figure 39: Detailed Data View in Dashboard

The Detailed Dataset View ability can be seen in Figure 39. The graph is shown in table format, considering all the features and their index numbers.

6.2.2. Manual Data Entry

The second component in the web application is ‘Manual Data Entry’. The component provides user interaction as the users can add, modify, or delete the values in the dataset. The abilities in this component include current data display, adding input values, modifying values in features, and deleting from the dataset. The main screen of the manual data entry component can be seen in Figure 40.

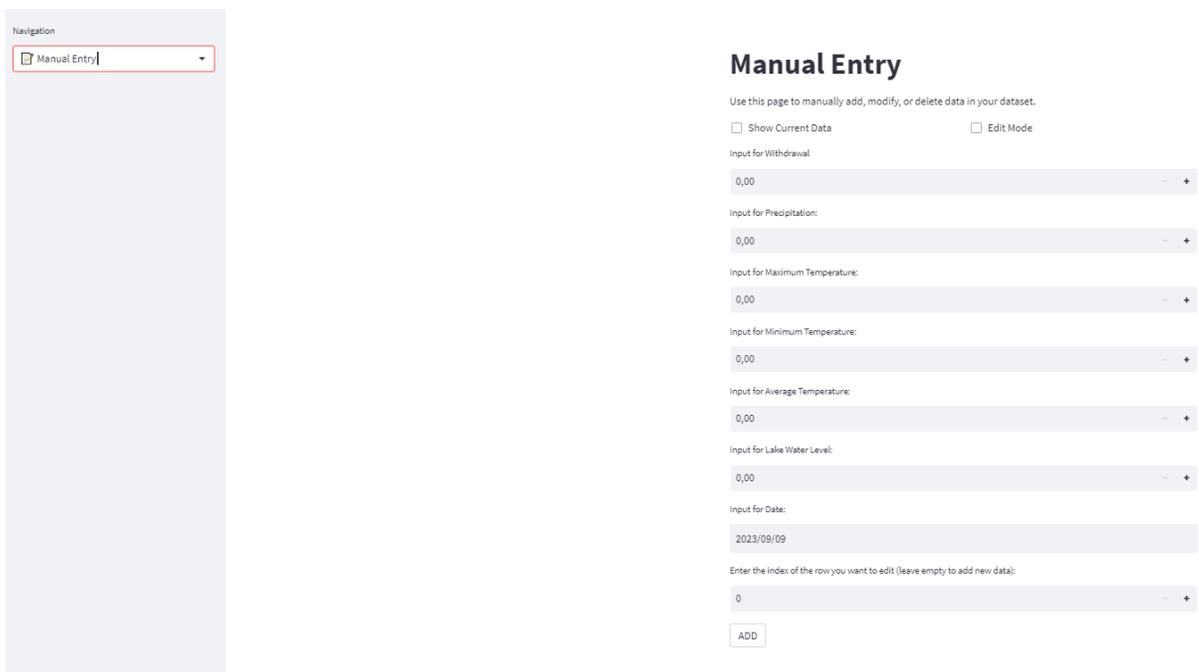


Figure 40: Manual data entry

a) Show Current Data

Although this component is designed for user interaction, there is one ability that doesn't allow user interaction, which is ‘Show Current Data’. This ability allows users to see the results of their actions, whether they add, edit, or delete anything from the dataset. The users can add specific values to the dataset without using this ability, but they must check the index number for the dates they want to edit or delete from the dataset.

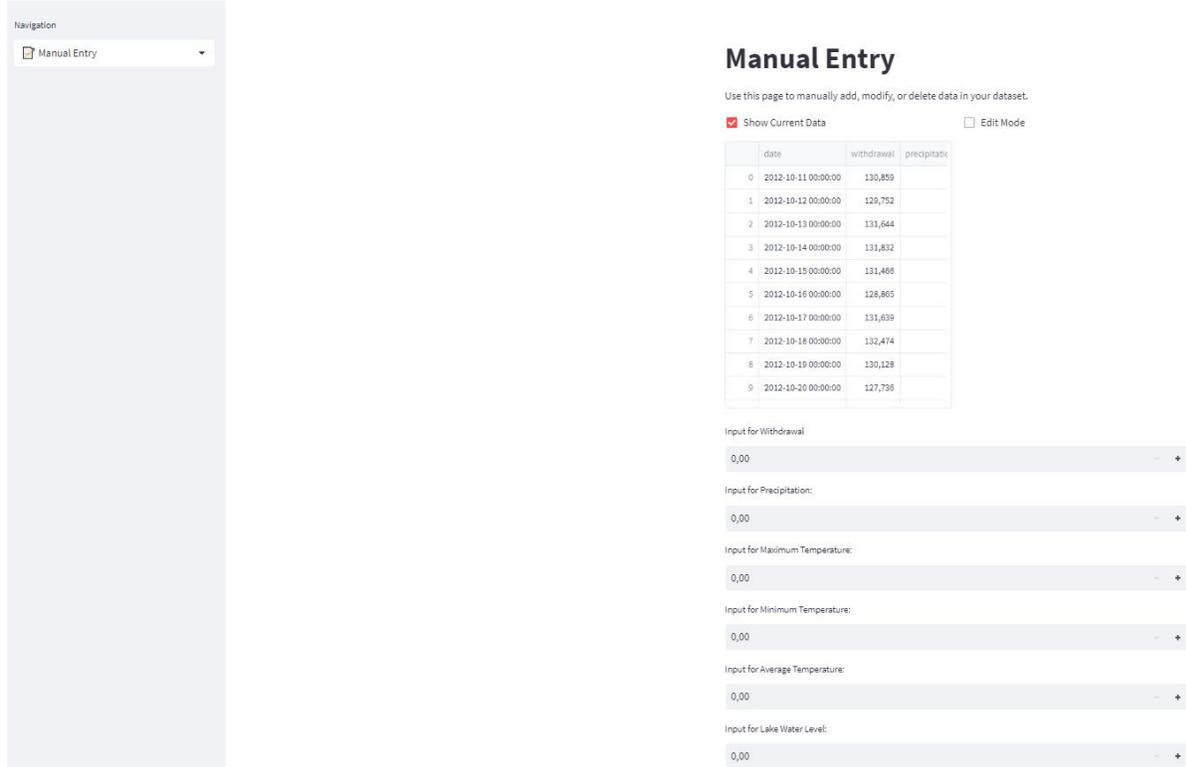


Figure 41: Show Current Data in Manual Data Entry

The Show Current Data ability can be seen in Figure 41. The users have the option to see the dataset by clicking the checkbox button.

b) Add Input Values for Features

The 'numberbox' in the application allows to take input values from the user. All the features are set to take only float-type numbers. However, input for date takes the input in date format. The features are determined as zero for the default value and today's date for the date input. The input value called 'Enter the index of the row you want to edit (leave empty to add new data)' doesn't have any effect on the add input ability in the application. The users either enter the values using the keyboard or using the increase or decrease buttons next to the 'numberbox'.

Navigation
Manual Entry

Manual Entry

Use this page to manually add, modify, or delete data in your dataset.

Show Current Data Edit Mode

Input for Withdrawal: 190000,00

Input for Precipitation: 0,00

Input for Maximum Temperature: 22,00

Input for Minimum Temperature: 8,00

Input for Average Temperature: 14,00

Input for Lake Water Level: 30,00

Input for Date: 2021/01/01

Enter the index of the row you want to edit (leave empty to add new data): 0

ADD

Figure 42: Add Input Values in Manual Data Entry

The ability to add input values can be seen in Figure 42. As it can be seen, there is an ‘ADD’ button that allows users to add values to the dataset. The user can add today’s values, or he or she can add the previous day’s value that was deleted beforehand.

Navigation
Manual Entry

Manual Entry

Use this page to manually add, modify, or delete data in your dataset.

Show Current Data Edit Mode

Input for Withdrawal: 190000,00

Input for Precipitation: 0,00

Input for Maximum Temperature: 22,00

Input for Minimum Temperature: 8,00

Input for Average Temperature: 14,00

Input for Lake Water Level: 30,00

Input for Date: 2021/01/01

Enter the index of the row you want to edit (leave empty to add new data): 0

ADD

Data added successfully!

Figure 43: After Add button effects

As soon as the user types all the necessary values for the input features, the input values are transferred to an Excel sheet that collects raw data on a cloud-based platform. The user gets a message as ‘Data added successfully!’ in order to be notified about the successful transaction (Figure 43).

c) *Edit Values in Features*

In any case, if a user types a mistaken value for the input, forgets to write the input value, or modifies the erroneous data in the Excel sheet that is stored on a cloud-based platform, the user can edit the values in a specific feature or the whole row altogether. The modified values are changed in the Excel sheet and saved immediately after the user clicks the ‘Edit button’. However, in order to prevent misclicks for the edit values, the user must check ‘Edit Mode’, and then the ‘EDIT’ button appears through the screen.

The screenshot shows a web interface for manual data entry. On the left is a navigation sidebar with a 'Manual Entry' menu item. The main content area is titled 'Manual Entry' and contains the following elements:

- A sub-header: 'Manual Entry'
- A description: 'Use this page to manually add, modify, or delete data in your dataset.'
- Two checkboxes: 'Show Current Data' (unchecked) and 'Edit Mode' (checked).
- Input fields with numerical values and range indicators (dash and plus signs):
 - Input for Withdrawal: 190000,00
 - Input for Precipitation: 0,00
 - Input for Maximum Temperature: 22,00
 - Input for Minimum Temperature: 8,00
 - Input for Average Temperature: 14,00
 - Input for Lake Water Level: 30,00
- Input for Date: 2021/01/01
- A text prompt: 'Enter the index of the row you want to edit (leave empty to add new data):' followed by an input field containing '0'.
- An 'EDIT' button.
- A checkbox for 'Delete Data' (unchecked).

Figure 44: Edit Values in Manual Data Entry

Edit values in the Manual Data Entry screen can be seen in Figure 44. As it can be seen in the figure, the user must type the index number for the intended row. The index numbers can always be seen by checking the ‘Show Current Data’ box.

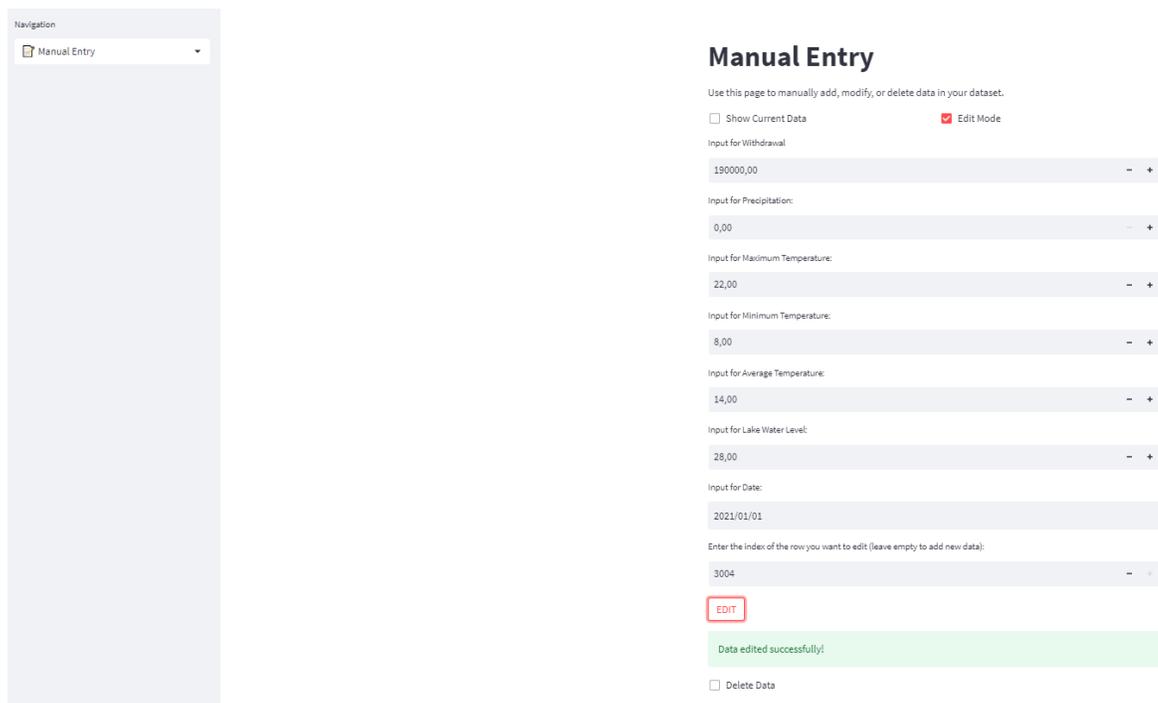


Figure 45: After Edit Button Effects

The decision support tool application always edits the whole row for the intended feature. Therefore, the features that don't require any modification must be written again with their actual values. After the necessary modifications have been done, the user gets a 'Data edited successfully message!' to make sure the transaction performed successfully (Figure 45).

d) Delete a Row from Dataset

There is deletion ability for the 'Manual Data Entry' component. The data in a row that needs to be deleted permanently must be done with careful consideration, so there is another checkbox button after the 'Edit Mode' checkbox. The user doesn't need to type anything into the features textbox; he or she only needs to type the index number that needs to be deleted.

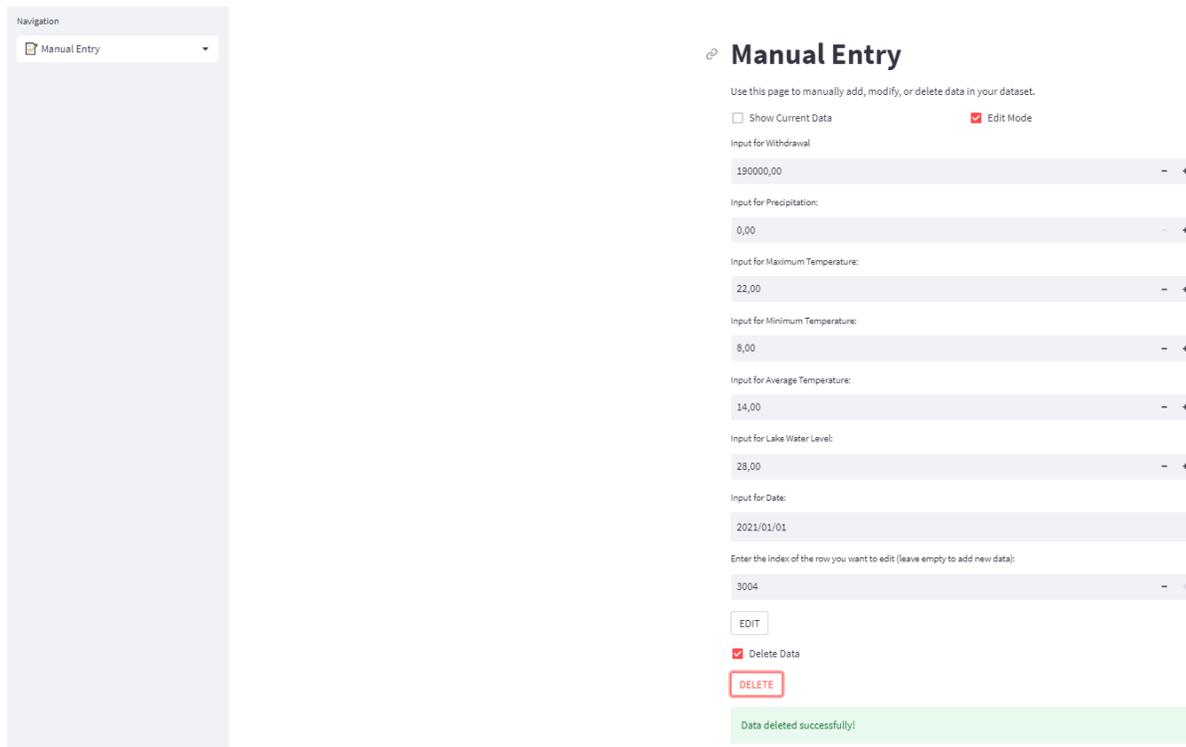


Figure 46: Delete a row in Manual Data Entry

The delete row in the Manual Data Entry screen can be seen in Figure 46. The application displays the ‘Data deleted successfully!’ message after the transaction is completed.

6.2.3. Prediction

The last component, which predicts the next period, is based on three parts: a 1-month prediction, a 2-month prediction, and a 4-month prediction.

a) LWL Prediction with LSTM Algorithm

The first ability in the ‘Prediction’ component is to use the LSTM algorithm to predict the next 1, 2, or 4 months. The LSTM algorithm is built on selected hyperparameters from Chapter 4 LWL that performed best on trials and experiments. Therefore, there is no additional hyperparameter tuning in the application since it increases the transaction duration. The user needs to wait for a substantial amount of time in order for the modeling to complete. The duration differs for the intended prediction period. It usually takes longer to predict further time periods. In order to activate the ability, the user only needs to select the predict page from the sliding windows in ‘Navigation’.

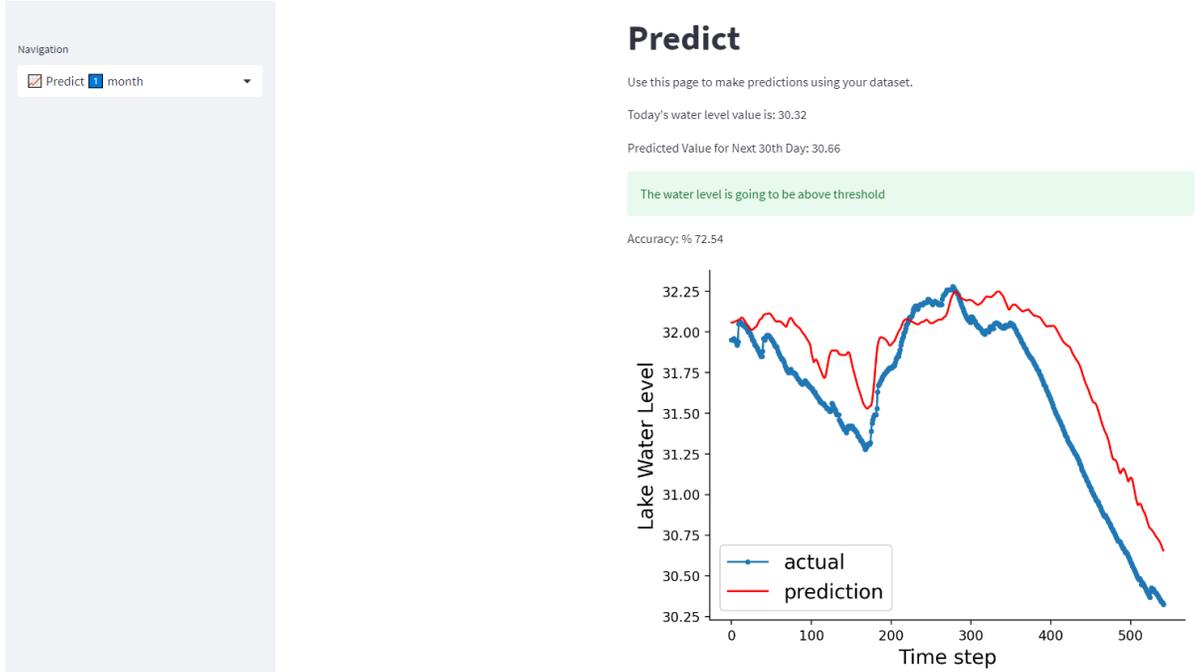


Figure 47: LWL Prediction with LSTM Algorithm in Prediction Component

The screen in Figure 47 reveals the results of 'Predict' ability. As it can be seen in the figure, the water level for today and the intended period of time are highlighted at the top.

b) Graph of Actual versus Prediction Results

The second ability that the 'Prediction' component provides is to give actual versus predicted values in terms of a test set that consists of the last 20% of the dataset. It is provided in order to help water managers comprehend the error rate that could divert water levels from the predicted value. The ability also provides an accuracy rate, which is calculated by reducing MAPE from 100%. The graph of the comparison between actual and predicted values is given in Figure 48.

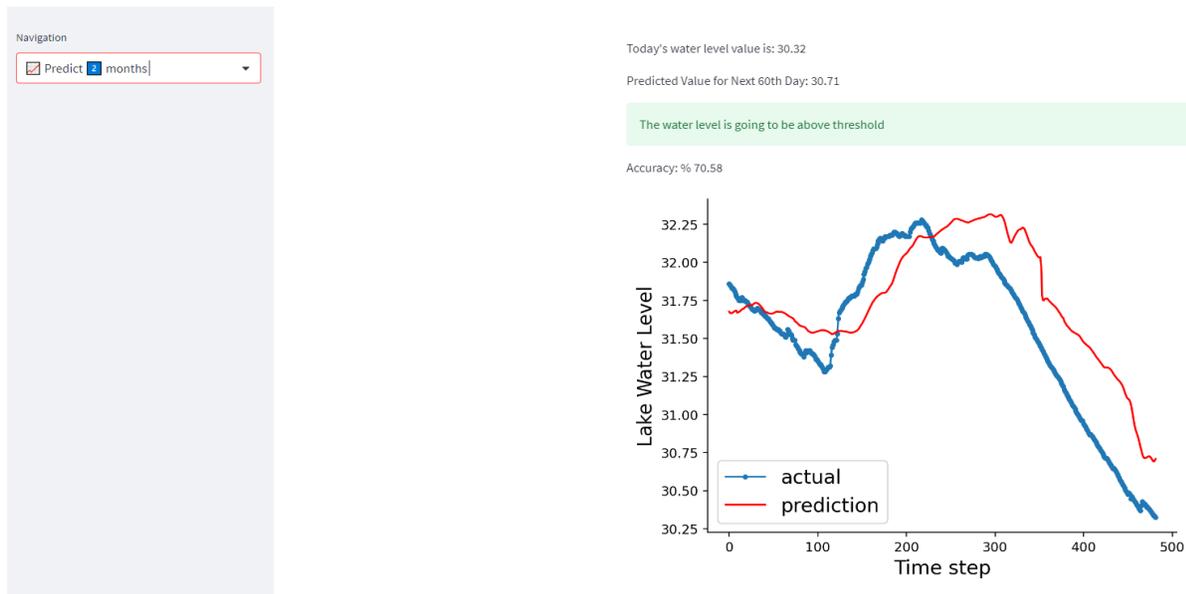


Figure 48: Actual versus predicted values graph

c) *Warning Mechanism*

The warning mechanism that is intended to be incorporated within the decision support tool is among the most crucial aspects to be established during the conceptual design phase. But when the system is consistently fed online (real-time) data, the warning system will be able to function effectively. This warning mechanism can only regulate data that is manually entered into the system because there isn't an online monitoring station in the Sapanca Lake area at this time.

Since there is no online data entry into the system, this control is carried out when the analysis data is manually entered. Control can, however, become continuous with the installation of online stations in the lake and the system's integration of online data-producing stations.

After the algorithm is incorporated and the results are presented, the water level for the future period is stated in the interface, comparing against today's LWL value. If the water level drops below the threshold of 30 m, which affects water quality, the application warns water managers to take action for the lake; otherwise, the tool prints, 'The water level is going to be above the threshold'.

CHAPTER 7

7. DISCUSSION

Based on the experimental result of this case study that applies an ANN and RNN-based DL algorithm for LWL prediction, it is possible to forecast the next 120 days with a smaller RMSE (0.3838 m), a reasonable Naïve Benchmark comparison value (58.00%), and significant Diebold Mariano test results ($p < 0.05$). However, compared with other models, the prediction result based on the LSTM proposed in this thesis is optimal for the next 60 days of LWL forecasting, with a smaller RMSE (0.1762 m), the highest Naïve Benchmark comparison value (78.55%), and a significant Diebold Mariano test p -value (< 0.003). The prediction performance of the investigated ANN and RNN algorithms aligns with previous research based on the RMSE and the Naïve Benchmark. In Donghae City, Korea, Yoon et al. (2011) used ANN and SVM to forecast the level of groundwater in the nearshore aquifer for a pair of wells having RMSE levels of 0.13 m and 0.136 m, respectively. In addition, Piasecki et al. (2018) found several RMSE results for four different samples of dataset which are BIOR2.4, DB2, COIF2 and BIOR2.2. The authors found RMSE results as 1.872 m, 1.968 m, 2.132 m, and 2.154 m, respectively. However, these studies lack a comparison of the proposed algorithms with the baseline models and other algorithms from DL. Therefore, the performance of the models cannot be evaluated for predicting water levels. The algorithms are also not evaluated against basic benchmark methods such as Naïve Benchmark, which raises the question of whether it is necessary to create fancy DL algorithms for LWL prediction. Thus, this thesis could be a milestone for further water level studies that attempt to develop every single DL algorithm available in the field of data science.

Hrnjica and Bonacci (2019) discovered that on datasets with a specific number of attributes and a one-month time span, the LSTM and RNN algorithms outperformed the conventional ANN algorithms. Additionally, they discovered that the LSTM and feed-forward neural network models outperformed conventional time series predictions using ARIMA and other related methods. Lee et al. (2020) demonstrated that the model developed by LSTM outperforms standard models in reproducing the critical metrics of the initial temporal domain along with the variation and association structure of the larger time scale. The improved representation of long-term variability is critical for water managers, as they rely on this data to plan and manage future water resources. In the future, the performance improvement over the Naïve Benchmark can be tested with other novel models, such as attention-based algorithms or other derivatives. However, the recent attempt to use an attention-based algorithm did not perform better than a recurrent network (Zhu et al., 2023).

The main hypothesis of the present thesis is confirmed by the fact that RNN-based algorithms achieve better predictive performance than LWL when using long-term daily

data from a decade and improve predictive accuracy for 60-day forecasts. The trends of observations and model predictions in Figures 19 through 27 suggest that the potential performance of RNN algorithms can also be extended beyond 120-day forecasts by incorporating more data into the models. Because gradient explosion caused by intervals of time during backpropagation in the process of learning for model networks, RNN-based DL methods are used. The LSTM algorithm has demonstrated its ability to learn from sequential data and has been a successful model in the past. According to LeCun et al. (2015), it can effectively learn from sequences of different durations and grasp dependence over time. The GRU algorithm is computationally more efficient than LSTM because the structure is simpler and more straightforward. Although the structure of Stacked LSTM and Bidirectional LSTM contributes to the model's ability to learn higher-level temporal representations, it can lead to degradation problems due to the low convergence rate of the LSTM layers. Another important point to consider is that the ANN algorithm performs better among all algorithms in short time periods thanks to its ability not to take long sequence periods into account with its lack of gated structure. To confirm our results, Zhu et al. (2023) studied 69 lakes in Poland for 30-day water level prediction and concluded that the DL models performed similarly to conventional ML models in terms of predictive performance. The results of the LSTM algorithm between its variants, namely the Stacked LSTM and the Bidirectional LSTM, in the present study show that there is no significant difference in predicting less than 30 days ahead. The LSTM algorithm requires long observation datasets and the selection and optimization of hyperparameters, learning rate, and number of epochs to achieve correct prediction results (Ozdemir et al., 2023). For example, Morovati et al. (2021) reported better prediction performance of LSTM when using daily recorded data over 20 years. The results obtained for LSTM in this thesis are consistent with these findings. The findings also show that the LSTM algorithm well reflects the fluctuation trend of the real LWL value. This is due to the use of gated structure in the LSTM model, so the LSTM algorithm is good at extracting short-term temporal correlations. However, due to the cyclic periods of water level variations, the performance increase drops when it reaches the next LWL cycle after 60 days. The best performance of the LSTM compared to the other models is also due to the successful optimization of hyperparameters in the LSTM network.

Another important aspect is that although the prevailing opinion suggests using all available DL algorithms to find the algorithm that performs best according to the RMSE or MAPE results, the results of the algorithms do not seem to differ significantly with respect to the Diebold-Mariano test. Therefore, in order to suggest a better-performing algorithm, the statistical difference must be shown in addition to the RMSE or MAPE results (Van der Heijden et al., 2021), and in some cases, the ANN and gated RNN derivatives, as indicated in the Results section, do not appear to have statistical significance and can be used interchangeably.

The fluctuations of LWL are associated with meteorological processes and anthropogenic activities, which lead to a nonlinear and complex system. In this context, the thesis has several limitations due to its nature. One of the limitations is that the results depend on the geographical location. The experiment is to be conducted at Lake Sapanca in the

northeastern Marmara region of Turkey. This location has characteristics of both the Black Sea and Mediterranean climates. Therefore, the results may change in regions with different climate characteristics. Another limitation of the thesis is that the dataset produced by the Turkish Meteorological Service contains several missing data points for selected parameters. Although it is possible to interpolate missing data, the results with interpolated data rows are limited. The results could change with a dataset containing all records for a longer period without missing data. Therefore, it is further suggested to apply other appropriate preprocessing methods to improve the predictive performance of the RNN DL models with different time horizons. Another limitation is the univariate analysis. The thesis focused only on multivariate analysis since future water level values have more than one dimension and other features need to be included. (Young et al., 2015; Mpallas et al., 2011). In the future, the LWL results can be combined with climate scenarios to see the far-future effects of current events. The research can be used in the future to optimize the economic worth of the electrical energy generated by dams by determining the frequency and amount of reservoir releases. Moreover, the LWL prediction could be practiced by using GIS methods with satellite dataset. The performance difference between time series prediction and prediction with image data could be compared with Naïve Method Benchmark. Lastly, in the availability of more features, the researchers can conduct sensitivity analysis and uncertainty analysis to eliminate some of the features to prevent possible overfitting issues.

Several well-known nutrient inputs and relatively less known meteorological parameters, together with hydrological disturbances, cause excessive growth of cyanobacteria in freshwater ecosystems, which degrade water quality with their toxins. Extreme heat waves are becoming more common as global and regional warming continues and are expected to become the norm in future scenarios. Microcystin concentration correlated positively with temperature variables (max, min, and mean, $p < 0.01$), including evaporation and light intensity ($p < 0.05$), and not significantly with precipitation (negative correlation), which is directly related to LWL (Figure 32). Significant correlations between meteorological parameters and microcystin concentrations in freshwater bodies have been reported previously (Novais et al., 2023). Light intensity in the Metalimnion zone leads to greater development of cyanobacteria and the presence of large amounts of microcystins, posing potential problems for the use of water resources (Boscaini et al., 2017). Since freshwater lakes are used as drinking water sources, proper water and algae management is necessary to ensure a clean and safe water supply. The use of tap water is restricted when large amounts of algae are found in water reservoirs because various water treatment problems can occur, such as clogged treatment systems, bad odor and color in the water, and regulated toxic substances such as microcystin. Predicting the correlation of algal blooms with easily measured meteorological or hydrological parameters in advance and taking rapid response actions to algal growth can minimize damage and ensure uninterrupted production of purified water.

CHAPTER 8

8. CONCLUSION

Compared to the traditional approaches, the gated RNN-based algorithms provide more accurate estimates of reservoir level changes. The gated RNN algorithms correctly adapt to changing input conditions, such as adjustments in water demand policy during reservoir operation. The fact that the gated RNN structure accounts for the nonlinear dynamics of the problem throughout the data set can be used to explain why gated RNNs perform better than conventional approaches in predicting reservoir levels. With respect to the RMSE, the results demonstrated here show the ability of the models used to understand the nonlinear behavior of LWLs.

Water resource management can greatly benefit from the development of a gated RNN algorithm for a given area. Estimates of periodic water reservoir levels are important for water supply planning, hydropower calculations, and flood management studies, among others. In addition, global climate change is leading to an increase in microcystin levels in freshwater lakes. To prevent this effect, this thesis proposes a DL-based prediction of future LWLs. The results show that RNN algorithms are a valuable alternative technique for predicting RWLs in different prediction periods.

The modeling results support the following findings:

- 1- The results of the algorithms can be compared, and although there could be different but similar results, the algorithms can be used interchangeably.

- 2- Overall, the GRU algorithm performs better than other gated RNN algorithms because it has a lower RMSE. However, it does not perform better in all time periods, so the algorithm needs to be replaced by another one to achieve better results for more future LWL prediction cases.

- 3- Gated RNN-based algorithms appear to have higher RMSE results as the prediction horizon increases, indicating poorer performance in lower prediction time periods. A more accurate comparison is possible using the Naïve Benchmark, and the percentage increase could provide a healthier result for comparing algorithm results with different prediction time periods. Although the prediction may differ from the actual values as the time period increases, the performance increase is much higher compared to the Naïve Method benchmark, making it more attractive for use in LWL prediction cases.

In addition, this thesis also examined the relationship between global warming and microcystin levels in freshwater lakes and demonstrated a clear relationship with meteorological data. However, more research is needed in this area to close the gap between LWL predictions with different geographical locations using same available features, algal growth, and microcystin levels.

Overall, the prediction results suggest that the proposed RNN algorithms can be successfully used to predict the future state of LWL for drinking water resource management leading to the achievement of sustainability under changing climatic conditions. The prediction result comparison between different time horizons and climate characteristics could be handled with additional benchmark methods such as Naïve Method and Diebold Mariano test.

REFERENCES

- Abbaspour, M., Javid, A. H., Mirbagheri, S. A., Givi, F. A., & Moghimi, P. (2012). Investigation of lake drying attributed to climate change. *International Journal of Environmental Science and Technology*, 9(2), 257-266. <https://doi.org/10.1007/s13762-012-0031-0>
- Adhikary, S. K., Muttill, N., & Yilmaz, A. G. (2018). Improving streamflow forecast using optimal rain gauge network-based input to artificial neural network models. *Hydrology Research*, 49(5), 1559-1577. <https://doi.org/10.2166/nh.2017.108>
- Ahmad, S. and S. P. Simonovic. (2006). An intelligent decision support system for management of floods. *Water Resources Management* 20(3): 391–410. <https://doi.org/10.1007/s11269-006-0326-3>
- Ahn, S. R., Jeong, J. H., & Kim, S. J. (2016). Assessing drought threats to agricultural water supplies under climate change by combining the SWAT and MODSIM models for the Geum River basin, South Korea. *Hydrological Sciences Journal*, 61(15), 2740-2753. <https://doi.org/10.1080/02626667.2015.1112905>
- Akcaalan, R., Young, F. M., Metcalf, J. S., Morrison, L. F., Albay, M., & Codd, G. A. (2006). Microcystin analysis in single filaments of *Planktothrix* spp. in laboratory cultures and environmental blooms. *Water research*, 40(8), 1583-1590. <https://doi.org/10.1016/j.watres.2006.02.020>
- Akkoyunlu, A., & Akiner, M. E. (2012). Pollution evaluation in streams using water quality indices: A case study from Turkey's Sapanca Lake Basin. *Ecological Indicators*, 18, 501-511. <https://doi.org/10.1016/j.ecolind.2011.12.0188>

- Albay, M., Akcaalan, R., Tufekci, H., Metcalf, J. S., Beattie, K. A., & Codd, G. A. (2003). Depth profiles of cyanobacterial hepatotoxins (microcystins) in three Turkish freshwater lakes. *Hydrobiologia*, 505, 89-95. <https://doi.org/10.1023/B:HYDR.0000007297.29998.5f>
- Altunkaynak, A. (2007). Forecasting surface water level fluctuations of Lake Van by artificial neural networks. *Water resources management*, 21(2), 399-408. <https://doi.org/10.1007/s11269-006-9022-6>
- Anneville, O., Gammeter, S., & Straile, D. (2005). Phosphorus decrease and climate variability: mediators of synchrony in phytoplankton changes among European peri-alpine lakes. *Freshwater Biology*, 50(10), 1731-1746. <https://doi.org/10.1111/j.1365-2427.2005.01429.x>
- Anupa, A., Sugathadasa, R., Herath, O., & Thibbotuwawa, A. (2021). Artificial neural network based demand forecasting integrated with federal funds rate. *Applied Computer Science*, 17(4), 34-44. <https://doi.org/10.23743/acs-2021-27>
- Ashaary, N. A., Wan Ishak, W. H., & Ku-Mahamud, K. R. (2015). Neural network application in the change of reservoir water level stage forecasting. *Indian Journal of Science and Technology*, 8(13), 1-6. <http://doi.org/10.17485/ijst/2015/v8i13/70634>
- Ashagre, H. M., Hatiye, S. D., Goshime, D. W. (2023). A pragmatic approach to the combined effect of climate change and water abstraction on Lake Ziway water balance, Ethiopia. *Journal of Water and Climate Change*, 14(1), 83-103. <https://doi.org/10.2166/wcc.2022.201>
- Atlas, L., Cole, R., Connor, J., El-Sharkawi, M., Marks, R., Muthusamy, Y., & Barnard, E. (1989). Performance comparisons between backpropagation networks and classification trees on three real-world applications. *Advances in neural information processing systems*, 2.
- Aykulu, G., M. Albay, R. Akcaalan, H. Tüfekçi and Y. Aktan. (2006). Species composition, abundance and seasonality of phytoplankton in a moderately deep Turkish Lake, *Nova Hedwigia*, Beiheft, 130, 325- 338.
- Azad, A. S., Sokkalingam, R., Daud, H., Adhikary, S. K., Khurshid, H., Mazlan, S. N. A., & Rabbani, M. B. A. (2022). Water level prediction through hybrid SARIMA and ANN models based on time series analysis: Red hills reservoir case study. *Sustainability*, 14(3), 1843. <https://doi.org/10.3390/su14031843>

- Azevedo, S. M. F. O., Carmichael W. W., Jochimsen E. M., Rinehart K. L., Lau S., Shaw G. R., Eaglesham G. K. (2002). Human intoxication by microcystins during renal dialysis treatment in Caruaru-Brazil. *Toxicology 181*, 441-446.
[https://doi.org/10.1016/S0300-483X\(02\)00491-2](https://doi.org/10.1016/S0300-483X(02)00491-2)
- Backer, L. C., Manassaram-Baptiste, D., LePrell, R., & Bolton, B. (2015). Cyanobacteria and algae blooms: review of health and environmental data from the harmful algal bloom-related illness surveillance system (HABISS) 2007–2011. *Toxins*, 7(4), 1048-1064. <https://doi.org/10.3390/toxins7041048>
- Bernard, C., Ballot, A., Thomazeau, S., Maloufi, S., Furey, A., Mankiewicz Boczek, J., PawlikSkowrońska B., Capelli C. & Salmaso, N. (2017). Cyanobacteria associated with the production of cyanotoxins. In *Handbook on cyanobacterial monitoring and cyanotoxin analysis* (pp. 501-525). Wiley.
- Bertone, E., O'Halloran, K., Stewart, R. A., & de Oliveira, G. F. (2017). Medium-term storage volume prediction for optimum reservoir management: A hybrid data-driven approach. *Journal of cleaner production*, 154, 353-365.
<https://doi.org/10.1016/j.jclepro.2017.04.003>
- Bláha, L., Babica, P., & Maršálek, B. (2009). Toxins produced in cyanobacterial water blooms–toxicity and risks. *Interdisciplinary toxicology*, 2(2), 36.
<https://doi.org/10.2478/v10102-009-0006-2>
- Bonakdari, H., Ebtehaj, I., Samui, P., & Gharabaghi, B. (2019). Lake water-level fluctuations forecasting using minimax probability machine regression, relevance vector machine, Gaussian process regression, and extreme learning machine. *Water Resources Management*, 33(11), 3965-3984.
<https://doi.org/10.1007/s11269-019-02346-0>
- Bond, N. R., Lake, P. S., & Arthington, A. H. (2008). The impacts of drought on freshwater ecosystems: an Australian perspective. *Hydrobiologia*, 600, 3-16.
<https://doi.org/10.1007/s10750-008-9326-z>
- Boscaini, A., Brescancin, F., Cerasino, L., Fedrigotti, C., Anna, F. E., Salmaso, N. (2017). Vertical and horizontal distribution of the microcystin producer *Planktothrix rubescens* (Cyanobacteria) in a small perialpine reservoir. *Advances in Oceanography and Limnology*, 8(2), 208-221.
<https://doi.org/10.4081/aiol.2017.7134>
- Bourdeau, M., qiang Zhai, X., Nefzaoui, E., Guo, X., & Chatellier, P. (2019). Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustainable Cities and Society*, 48, 101533.
<https://doi.org/10.1016/j.scs.2019.101533>

- Buyukyildiz, M., Tezel, G., & Yilmaz, V. (2014). Estimation of the change in lake water level by artificial intelligence methods. *Water resources management*, 28, 4747-4763. <https://doi.org/10.1007/s11269-014-0773-1>
- Cai, Z., Jin, T., Li, C., Ofterdinger, U., Zhang, S., Ding, A., & Li, J. (2016). Is China's fifth-largest inland lake to dry-up? Incorporated hydrological and satellite-based methods for forecasting Hulun lake water levels. *Advances in Water Resources*, 94, 185-199. <https://doi.org/10.1016/j.advwatres.2016.05.010>
- Carmichael, W. W., Azevedo, S. M., An, J. S., Molica, R. J., Jochimsen, E. M., Lau, S., Rinehart K. L., Shaw G. R. & Eaglesham, G. K. (2001). Human fatalities from cyanobacteria: chemical and biological evidence for cyanotoxins. *Environmental health perspectives*, 109(7), 663-668. <https://doi.org/10.1289/ehp.01109663>
- Castillo-Botón, C., Casillas-Pérez, D., Casanova-Mateo, C., Moreno-Saavedra, L. M., Morales-Díaz, B., Sanz-Justo, J., Gutiérrez, P.A. & Salcedo-Sanz, S. (2020). Analysis and prediction of dammed water level in a hydropower reservoir using machine learning and persistence-based techniques. *Water*, 12(6), 1528. <https://doi.org/10.3390/w12061528>
- Chang, F. J., & Chang, Y. T. (2006). Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Advances in water resources*, 29(1), 1-10. <https://doi.org/10.1016/j.advwatres.2005.04.015>
- Chen, Y., Fan, R., Yang, X., Wang, J., & Latif, A. (2018). Extraction of urban water bodies from high-resolution remote-sensing imagery using deep learning. *Water*, 10(5), 585. <https://doi.org/10.3390/w10050585>
- Chen, S., Johnson, F., Drummond, C., & Glamore, W. (2020). A new method to improve the accuracy of remotely sensed data for wetland water balance estimates. *Journal of Hydrology: Regional Studies*, 29, 100689. <https://doi.org/10.1016/j.ejrh.2020.100689>
- Contesse, L., Donoso, P., & Prina, J. (2003). The value of information in a long-term hydro-Thermal electrical planning model. *International Transactions in Operational Research*, 10(1), 89-100. <https://doi.org/10.1111/1475-3995.00395>
- Costa Nogueira Jr, A., de Sousa Almeida, J. L., Auger, G., & Watson, C. D. (2021). Reduced order modeling of dynamical systems using artificial neural networks applied to water circulation. *arXiv e-prints*, arXiv-2103. <https://doi.org/10.48550/arXiv.2103.10931>
- Croley, T. E. (2006). Using climate predictions in Great Lakes hydrologic forecasts. *Climate variations, climate change, and water resources engineering*, 166-187.

- Damova, M., Stoyanov, E., Kopchev, M., Kuzmanova, K., & Natali, S. (2020, June). Linked Data Approach to Water Resources Management of Hydropower Reservoirs. In *IOP Conference Series: Earth and Environmental Science* (Vol. 509, No. 1, p. 012011). IOP Publishing.
<https://doi.org/10.1088/1755-1315/509/1/012011>
- Davis, P. A., & Walsby, A. E. (2002). Comparison of measured growth rates with those calculated from rates of photosynthesis in *Planktothrix* spp. isolated from Blelham Tarn, English Lake District. *New Phytologist*, *156*(2), 225-239.
<https://doi.org/10.1046/j.1469-8137.2002.00495.x>
- Dinka, M. O. (2020). Estimation of groundwater contribution to Lake Basaka in different hydrologic years using conceptual netgroundwater flux model. *Journal of Hydrology: Regional Studies*, *30*, 100696.
<https://doi.org/10.1016/j.ejrh.2020.100696>
- Dokulil, M. T., & Teubner, K. (2000). Cyanobacterial dominance in lakes. *Hydrobiologia*, *438*, 1-12. <https://doi.org/10.1023/A:1004155810302>
- Dokulil, M. T., & Teubner, K. (2012). Deep living *Planktothrix rubescens* modulated by environmental constraints and climate forcing. *Phytoplankton responses to human impacts at different scales*, 29-46. https://doi.org/10.1007/978-94-007-5790-5_4
- Duru, U. (2017). Shoreline change assessment using multi-temporal satellite images: a case study of Lake Sapanca, NW Turkey. *Environmental monitoring and assessment*, *189*(8), 1-14. <https://doi.org/10.1007/s10661-017-6112-2>
- Ebtehaj, I., Bonakdari, H., & Gharabaghi, B. (2019). A reliable linear method for modeling lake level fluctuations. *Journal of Hydrology*, *570*, 236-250.
<https://doi.org/10.1016/j.jhydrol.2019.01.010>
- Ehteram, M., Ferdowsi, A., Faramarzpour, M., Al-Janabi, A. M. S., Al-Ansari, N., Bokde, N. D., & Yaseen, Z. M. (2021). Hybridization of artificial intelligence models with nature inspired optimization algorithms for lake water level prediction and uncertainty analysis. *Alexandria Engineering Journal*, *60*(2), 2193-2208.
<https://doi.org/10.1016/j.aej.2020.12.034>
- Faber, B. A., & Stedinger, J. R. (2001). Reservoir optimization using sampling SDP with ensemble streamflow prediction (ESP) forecasts. *Journal of Hydrology*, *249*(1-4), 113-133. [https://doi.org/10.1016/S0022-1694\(01\)00419-X](https://doi.org/10.1016/S0022-1694(01)00419-X)
- Fang, T., & Lahdelma, R. (2016). Evaluation of a multiple linear regression model and SARIMA model in forecasting heat demand for district heating system. *Applied energy*, *179*, 544-552. <https://doi.org/10.1016/j.apenergy.2016.06.133>

Fowler, J., Cohen, L., & Jarvis, P. (2013). *Practical statistics for field biology*. John Wiley & Sons.

Fry, L. M., Apps, D., & Gronewold, A. D. (2020). Operational seasonal water supply and water level forecasting for the Laurentian Great Lakes. *Journal of Water Resources Planning and Management*, 146(9), 04020072.
[https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0001214](https://doi.org/10.1061/(ASCE)WR.1943-5452.0001214)

García Molinos, J., Viana, M., Brennan, M., & Donohue, I. (2015). Importance of long-term cycles for predicting water level dynamics in natural lakes. *Plos one*, 10(3), e0119253. <https://doi.org/10.1371/journal.pone.0119253>

Ghorbani, M. A., Khatibi, R., Aytak, A., Makarynsky, O., & Shiri, J. (2010). Sea water level forecasting using genetic programming and comparing the performance with artificial neural networks. *Computers & Geosciences*, 36(5), 620-627.
<https://doi.org/10.1016/j.cageo.2009.09.014>

Gillies, R. R., Chung, O. Y., Wang, S. Y. S., DeRose, R. J., & Sun, Y. (2015). Added value from 576 years of tree-ring records in the prediction of the Great Salt Lake level. *Journal of Hydrology*, 529, 962-968.
<https://doi.org/10.1016/j.jhydrol.2015.08.058>

Guevara-Ochoa, C., Medina-Sierra, A., Vives, L. (2020). Spatio-temporal effect of climate change on water balance and interactions between groundwater and surface water in plains. *Science of the Total Environment*, 722, 137886.
<https://doi.org/10.1016/j.scitotenv.2020.137886>

Guinaldo, T., Munier, S., Le Moigne, P., Boone, A., Decharme, B., Choulga, M., & Leroux, D. J. (2021). Parametrization of a lake water dynamics model MLake in the ISBA-CTRIP land surface system (SURFEX v8. 1). *Geoscientific Model Development*, 14(3), 1309-1344. <https://doi.org/10.5194/gmd-14-1309-2021>

Guyennon, N., Salerno, F., Rossi, D., Rainaldi, M., Calizza, E., & Romano, E. (2021). Climate change and water abstraction impacts on the long-term variability of water levels in Lake Bracciano (Central Italy): A Random Forest approach. *Journal of Hydrology: Regional Studies*, 37, 100880.
<https://doi.org/10.1016/j.ejrh.2021.100880>

Haddout, S., Jamali, A., Rhazi, M., & Aghfir, M. (2018). Finite volume coastal ocean model for water-level fluctuation due to climate change in Aguelmam Sidi Ali Lake (Middle Atlas, Morocco). In *Annales de Limnologie-International Journal of Limnology*, 54, 5. EDP Sciences. <https://doi.org/10.1051/limn/2017033>

- Haq, M. M., Seidou, O., Mohammadian, A., & Khalidou, B. A. (2021). Effect of rating curve hysteresis on flood extent simulation with a 2D hydrodynamic model: A case study of the Inner Niger Delta, Mali, West Africa. *Journal of African Earth Sciences*, 178, 104187. <https://doi.org/10.1016/j.jafrearsci.2021.104187>
- Harke, M. J., Steffen, M. M., Gobler, C. J., Otten, T. G., Wilhelm, S. W., Wood, S. A., & Paerl, H. W. (2016). A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp. *Harmful algae*, 54, 4-20. <https://doi.org/10.1016/j.hal.2015.12.007>
- Hirsch, P. E., Schillinger, S., Weigt, H., & Burkhardt-Holm, P. (2014). A hydro-economic model for water level fluctuations: combining limnology with economics for sustainable development of hydropower. *PloS one*, 9(12), e114889. <https://doi.org/10.1371/journal.pone.0114889>
- Hrnjica, B. & Bonacci, O. (2019). Lake level prediction using feed forward and recurrent neural networks. *Water Resources Management*, 33, 2471– 2484. <https://doi.org/10.1007/s11269-019-02255-2>
- Hu, T., Mao, J., Zhang, P., Xu, D., Chen, W., & Dai, H. (2018) (a). Hydrological utilization of satellite precipitation estimates in a data-scarce lake region. *Water Supply*, 18(5), 1581-1589. <https://doi.org/10.2166/ws.2017.223>
- Hu, T., Mao, J., Pan, S., Dai, L., Zhang, P., Xu, D., & Dai, H. (2018) (b). Water level management of lakes connected to regulated rivers: An integrated modeling and analytical methodology. *Journal of Hydrology*, 562, 796-808. <https://doi.org/10.1016/j.jhydrol.2018.05.038>
- Hussain, D., Khan, A. A., Hassan, S. N. U., Naqvi, S. A. A., & Jamil, A. (2021). A time series assessment of terrestrial water storage and its relationship with hydro-meteorological factors in Gilgit-Baltistan region using GRACE observation and GLDAS-Noah model. *SN Applied Sciences*, 3(5), 1-11. <https://doi.org/10.1007/s42452-021-04525-4>
- Ishak, W. H. W., Ku-Mahamud, K. R., & Norwawi, N. M. (2011). Intelligent decision support model based on neural network to support reservoir water release decision. In *International Conference on Software Engineering and Computer Systems* (pp. 365-379). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-22170-5_32
- Jaafar, O., Toriman, M. E. H., Idris, M. H., Mastura, S. A. S., Juahir, H. H., Aziz, N. A. A., Kamarudin, K. A., & Jamil, N. R. (2010). Study of water level-discharge relationship using artificial neural network (ANN) in Sungai Gumum, Tasik Chini Pahang Malaysia. *Research Journal of Applied Sciences*, 5(1), 20-26.

- Jacquet, S., Briand, J. F., Leboulanger, C., Avois-Jacquet, C., Oberhaus, L., Tassin, B., Vinçon-Leite, B., Paolini, G., Druart, J.C., Anneville, O. & Humbert, J. F. (2005). The proliferation of the toxic cyanobacterium *Planktothrix rubescens* following restoration of the largest natural French lake (Lac du Bourget). *Harmful algae*, 4(4), 651-672. <https://doi.org/10.1016/j.hal.2003.12.006>
- Jahani, E., Mousavi, S. J., Zadeh, N. A., & Kim, J. H. (2016). Assessing the role of foresight on future streamflows in storage-yield-reliability analysis of surface water reservoirs. *Procedia Engineering*, 154, 1163-1168. <https://doi.org/10.1016/j.proeng.2016.07.530>
- Jiang, F., Dong, Z., Wang, Z. A., Zhu, Y., Liu, M., Luo, Y., & Zhang, T. (2021). Flood forecasting using an improved NARX network based on wavelet analysis coupled with uncertainty analysis by Monte Carlo simulations: a case study of Taihu Basin, China. *Journal of Water and Climate Change*, 12(6), 2674-2696. <https://doi.org/10.2166/wcc.2021.019>
- Joehnk, K. D., Huisman, J. E. F., Sharples, J., Sommeijer, B. E. N., Visser, P. M., & Stroom, J. M. (2008). Summer heatwaves promote blooms of harmful cyanobacteria. *Global change biology*, 14(3), 495-512. <https://doi.org/10.1111/j.1365-2486.2007.01510.x>
- Karamouz, M., & Mousavi, S. J. (2003). Uncertainty Based Operation of Large Scale Reservoir Systems: Dez and Karoun Experience. *JAWRA Journal of the American Water Resources Association*, 39(4), 961-975. <https://doi.org/10.1111/j.1752-1688.2003.tb04419.x>
- Kenda, K., Peternelj, J., Mellios, N., Kofinas, D., Čerin, M., & Rožanec, J. (2020). Usage of statistical modeling techniques in surface and groundwater level prediction. *Journal of Water Supply: Research and Technology-AQUA*, 69(3), 248-265. <https://doi.org/10.2166/aqua.2020.143>
- Khandelwal, I., Adhikari, R., & Verma, G. (2015). Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition. *Procedia Computer Science*, 48, 173-179. <https://doi.org/10.1016/j.procs.2015.04.167>
- Kim, Y. O., & Palmer, R. N. (1997). Value of seasonal flow forecasts in Bayesian stochastic programming. *Journal of Water Resources Planning and Management*, 123(6), 327-335. [https://doi.org/10.1061/\(ASCE\)0733-9496\(1997\)123:6\(327\)](https://doi.org/10.1061/(ASCE)0733-9496(1997)123:6(327))

- Kim, Y. O., Eum, H. I., Lee, E. G., & Ko, I. H. (2007). Optimizing operational policies of a Korean multireservoir system using sampling stochastic dynamic programming with ensemble streamflow prediction. *Journal of Water Resources Planning and Management*, 133(1), 4-14. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2007\)133:1\(4\)](https://doi.org/10.1061/(ASCE)0733-9496(2007)133:1(4))
- Kisi, O., Shiri, J., & Nikoofar, B. (2012). Forecasting daily lake levels using artificial intelligence approaches. *Computers & Geosciences*, 41, 169-180. <https://doi.org/10.1016/j.cageo.2011.08.027>
- Koch, J., Gotfredsen, J., Schneider, R., Troldborg, L., Stisen, S., & Henriksen, H. J. (2021). High resolution water table modelling of the shallow groundwater using a knowledge-guided gradient boosting decision tree model. *Frontiers in Water*, 3, 81. <https://doi.org/10.3389/frwa.2021.701726>
- Lai, X. J., Jiang, J. H., & Huang, Q. (2011). Effect of Three Gorge Reservoir on the water regime of the Dongting Lake during important regulation periods. *Resources and Environment in the Yangtze Basin*, 20(2), 167-172.
- Latif, S. D., Ahmed, A. N., Sherif, M., Sefelnasr, A., & El-Shafie, A. (2021). Reservoir water balance simulation model utilizing machine learning algorithm. *Alexandria Engineering Journal*, 60(1), 1365-1378. <https://doi.org/10.1016/j.aej.2020.10.057>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- Lee, T., Shin, J.Y., Kim, J.S. & Singh, V.P. (2020) Stochastic simulation on reproducing long-term memory of hydroclimatological variables using deep learning model. *Journal of Hydrology*, 582, 124540. <https://doi.org/10.1016/j.jhydrol.2019.124540>
- Li, C., Sun, B., Jia, K., Zhang, S., Li, W., Shi, X., Cordovil C. M.d.S., & Pereira, L. S. (2013). Multi-band remote sensing based retrieval model and 3D analysis of water depth in Hulun Lake, China. *Mathematical and Computer Modelling*, 58(3-4), 771-781. <https://doi.org/10.1016/j.mcm.2012.12.027>
- Li, Y., Liao, J., Guo, H., Liu, Z., & Shen, G. (2014). Patterns and potential drivers of dramatic changes in Tibetan lakes, 1972–2010. *PloS one*, 9(11), e111890. <https://doi.org/10.1371/journal.pone.0111890>
- Li, B., Yang, G., Wan, R., Dai, X., & Zhang, Y. (2016) (a). Comparison of random forests and other statistical methods for the prediction of lake water level: a case study of the Poyang Lake in China. *Hydrology Research*, 47(S1), 69-83. <https://doi.org/10.2166/nh.2016.264>

- Li, Z., Zhu, D., Chen, Y., Fang, X., Liu, Z., & Ma, W. (2016) (b). Simulating and understanding effects of water level fluctuations on thermal regimes in Miyun Reservoir. *Hydrological Sciences Journal*, 61(5), 952-969. <https://doi.org/10.1080/02626667.2014.983517>
- Liang, C., Li, H., Lei, M., & Du, Q. (2018). Dongting lake water level forecast and its relationship with the three gorges dam based on a long short-term memory network. *Water*, 10(10), 1389. <https://doi.org/10.3390/w10101389>
- Liboriussen, L., Landkildehus, F., Meerhoff, M., Bramm, M. E., Søndergaard, M., Christoffersen, K., Richardson, K., Søndergaard M., Lauridsen T. L., & Jeppesen, E. (2005). Global warming: Design of a flow-through shallow lake mesocosm climate experiment. *Limnology and Oceanography: Methods*, 3(1), 1-9. <https://doi.org/10.4319/lom.2005.3.1>
- Lin, P., Yang, Z. L., Cai, X., & David, C. H. (2015). Development and evaluation of a physically-based lake level model for water resource management: A case study for Lake Buchanan, Texas. *Journal of Hydrology: Regional Studies*, 4, 661-674. <https://doi.org/10.1016/j.ejrh.2015.08.005>
- Liu, J., Huang, C. H., & Zeng, G. Z. (2017). Comparative analysis of year-end water level determining methods for cascade carryover storage reservoirs. In *IOP Conference Series: Earth and Environmental Science*, 82(1), 012060. IOP Publishing. <https://doi.org/10.1088/1755-1315/82/1/012060>
- Liu, Z., Zhou, P., Chen, X., & Guan, Y. (2015). A multivariate conditional model for streamflow prediction and spatial precipitation refinement. *Journal of Geophysical Research: Atmospheres*, 120(19), 10-116. <https://doi.org/10.1002/2015JD023787>
- Livingstone, D. M., & Lotter, A. F. (1998). The relationship between air and water temperatures in lakes of the Swiss Plateau: a case study with palaeolimnological implications. *Journal of paleolimnology*, 19, 181-198. <https://doi.org/10.1023/A:1007904817619>
- Lofgren, B. M., & Rouhana, J. (2016). Physically plausible methods for projecting changes in Great Lakes water levels under climate change scenarios. *Journal of Hydrometeorology*, 17(8), 2209-2223. <https://doi.org/10.1175/JHM-D-15-0220.1>
- Loucks, D. P. (1992). Water resource systems models: their role in planning. *Journal of Water Resources Planning and Management*, 118(3), 214-223. [https://doi.org/10.1061/\(ASCE\)0733-9496\(1992\)118:3\(214\)](https://doi.org/10.1061/(ASCE)0733-9496(1992)118:3(214))

- Loucks, D. P. (1995). Developing and implementing Decision Support Systems: A Critique and a Challenge. *JAWRA Journal of the American Water Resources Association*, 31(4), 571-582. <https://doi.org/10.1111/j.1752-1688.1995.tb03384.x>
- Lukman, Q. A., Ruslan, F. A., & Adnan, R. (2016). 5 Hours ahead of time flood water level prediction modelling using NNARX technique: Case study terengganu. In *2016 7th IEEE Control and System Graduate Research Colloquium (ICSGRC)* (pp. 104-108). IEEE. <https://doi.org/10.1109/ICSGRC.2016.7813310>
- Lv, N., Liang, X., Chen, C., Zhou, Y., Li, J., Wei, H., & Wang, H. (2020). A long Short-Term memory cyclic model with mutual information for hydrology forecasting: A Case study in the xixian basin. *Advances in Water Resources*, 141, 103622. <https://doi.org/10.1016/j.advwatres.2020.103622>
- Magyar, N., Hatvani, I. G., Arató, M., Trásy, B., Blaschke, A. P., & Kovács, J. (2021). A New Approach in Determining the Decadal Common Trends in the Groundwater Table of the Watershed of Lake “Neusiedlersee”. *Water*, 13(3), 290. <https://doi.org/10.3390/w13030290>
- M Dawam, S. R., & Ku-Mahamud, K. R. (2019). Reservoir water level forecasting using normalization and multiple regression. *Indonesian Journal of Electrical Engineering and Computer Science*, 14(1), 443-449. <https://doi.org/10.11591/ijeecs.v14.i1.pp443-449>
- Mengistu, D., Bewket, W., Dosio, A., Panitz, H. J. (2021). Climate change impacts on water resources in the upper blue Nile (Abay) river basin, Ethiopia. *Journal of Hydrology*, 592, 125614. <https://doi.org/10.1016/j.jhydrol.2020.125614>
- Merel, S., Villarín, M. C., Chung, K., & Snyder, S. (2013). Spatial and thematic distribution of research on cyanotoxins. *Toxicon*, 76, 118-131. <https://doi.org/10.1016/j.toxicon.2013.09.008>
- Metcalf, R. A., Chang, C., & Smakhtin, V. (2005). Tools to support the implementation of environmentally sustainable flow regimes at Ontario's waterpower facilities. *Canadian water resources journal*, 30(2), 97-110. <https://doi.org/10.4296/cwrj3002097>
- Mislan, M., Gaffar, A. F. O., Haviluddin, H., & Puspitasari, N. (2018). Water Level Prediction of Lake Cascade Mahakam Using Adaptive Neural Network Backpropagation (ANNBP). IOP Conf. Series: Earth and Environmental Science 144. <https://doi.org/10.1088/1755-1315/144/1/012009>

- Mpallas, L., Tzimopoulos, C., & Evangelides, C. (2011). Comparison between neural networks and adaptive neuro-fuzzy inference system in modeling lake Kerkini water level fluctuation lake management using artificial intelligence. *Journal of Environmental Science and Technology*, 4(4), 366-376.
- Mohammadi, B., Guan, Y., Aghelpour, P., Emamgholizadeh, S., Pillco Zolá, R., & Zhang, D. (2020). Simulation of Titicaca Lake water level fluctuations using hybrid machine learning technique integrated with Grey Wolf Optimizer Algorithm. *Water*, 12(11), 3015. <https://doi.org/10.3390/w12113015>
- Montroull, N. B., Saurral, R. I., Camilloni, I. A., Grimson, R., & Vasquez, P. (2013). Assessment of climate change on the future water levels of the Iberá wetlands, Argentina, during the twenty-first century. *International journal of river basin management*, 11(4), 401-410. <https://doi.org/10.1080/15715124.2013.819807>
- Mooij, W. M., Hülsmann, S., De Senerpont Domis, L. N., Nolet, B. A., Bodelier, P. L., Boers, P. C., Pires, L. M. D., Gons, H. J., Ibelings, B. W., Noordhuis, R., Portielje R., Wolfstein, K., & Lammens, E. H. (2005). The impact of climate change on lakes in the Netherlands: a review. *Aquatic Ecology*, 39, 381-400. <https://doi.org/10.1007/s10452-005-9008-0>
- Morgan, H., Hocking, A., & Henderson, S. (2019). Simplified method to predict final void water levels. In *Proceedings of the 13th International Conference on Mine Closure* (pp. 1337-1352). Australian Centre for Geomechanics. https://doi.org/10.36487/ACG_rep/1915_105_Morgan
- Morovati, K., Nakhaei, P., Tian, F., Tudaji, M., & Hou, S. (2021). A Machine learning framework to predict reverse flow and water level: A case study of Tonle Sap Lake. *Journal of Hydrology*, 603, 127168. <https://doi.org/10.1016/j.jhydrol.2021.127168>
- Mtilatila, L., Bronstert, A., Bürger, G., & Vormoor, K. (2020). Meteorological and hydrological drought assessment in Lake Malawi and Shire River basins (1970–2013). *Hydrological Sciences Journal*, 65(16), 2750-2764. <https://doi.org/10.1080/02626667.2020.1837384>
- Myakisheva, N., Gaidukova, E., Shanochkin, S. V., & Batmazova, A. A. (2021). Seasonal and annual probabilistic forecasting of water levels in large lakes (case study of the Ladoga Lake). *International Letters of Natural Sciences*, 82. <https://doi.org/10.18052/www.scipress.com/ILNS.82.13>
- Naselli-Flores, L., Barone, R., Chorus, I., & Kurmayer, R. (2007). Toxic cyanobacterial blooms in reservoirs under a semiarid Mediterranean climate: the magnification of a problem. *Environmental Toxicology: an International Journal*, 22(4), 399-404. <https://doi.org/10.1002/tox.20268>

- Nhu, V. H., Shahabi, H., Nohani, E., Shirzadi, A., Al-Ansari, N., Bahrami, S., Miraki S., Geertsema M., & Nguyen, H. (2020). Daily water level prediction of Zrebar Lake (Iran): A comparison between M5P, random forest, random tree and reduced error pruning trees algorithms. *ISPRS International Journal of Geo-Information*, 9(8), 479. <https://doi.org/10.3390/ijgi9080479>
- Nõges, P., Mischke, U., Laugaste, R., & Solimini, A. G. (2010). Analysis of changes over 44 years in the phytoplankton of Lake Võrtsjärv (Estonia): the effect of nutrients, climate and the investigator on phytoplankton-based water quality indices. *Hydrobiologia*, 646, 33-48. <https://doi.org/10.1007/s10750-010-0178-y>
- Nourani, V., Tootoonchi, R., & Andaryani, S. (2021). Investigation of climate, land cover and lake level pattern changes and interactions using remotely sensed data and wavelet analysis. *Ecological Informatics*, 64, 101330. <https://doi.org/10.1016/j.ecoinf.2021.101330>
- Nouri, H., Ildoromi, A., Sepehri, M., & Artimani, M. (2019). Comparing three main methods of artificial intelligence in flood estimation in Yalphan catchment. *Geography and Environmental Planning*, 29(4), 35-50. <https://doi.org/10.22108/GEP.2018.98036.0>
- Novais, M. H., Penha, A. M., Catarino, A., Martins, I., Fialho, S., Lima, A., & Palma, P. (2023). The usefulness of ecotoxicological tools to improve the assessment of water bodies in a climate change reality. *Science of The Total Environment*, 166392. <https://doi.org/10.1016/j.scitotenv.2023.166392>
- Obringer, R., & Nateghi, R. (2018). Predicting urban reservoir levels using statistical learning techniques. *Scientific reports*, 8(1), 5164. <https://doi.org/10.1038/s41598-018-23509-w>
- Odongo, V. O., van Oel, P. R., van der Tol, C., Su, Z. (2019). Impact of land use and land cover transitions and climate on evapotranspiration in the Lake Naivasha Basin, Kenya. *Science of the total environment*, 682, 19-30. <https://doi.org/10.1016/j.scitotenv.2019.04.062>
- Ouni, H., Sousa, M. C., Ribeiro, A. S., Pinheiro, J., M'Barek, N. B., Tarhouni, J., Tlatli-Hariga N., & Dias, J. M. (2020). Numerical modeling of hydrodynamic circulation in Ichkeul Lake-Tunisia. *Energy Reports*, 6, 208-213. <https://doi.org/10.1016/j.egyr.2019.08.044>
- Ozdemir, S., Yaqub, M., & Yildirim, S. O. (2023). A systematic literature review on Lake water level prediction models. *Environmental Modelling & Software*, 105684. <https://doi.org/10.1016/j.envsoft.2023.105684>

- Özer, T., Erkaya, İ. A., Udoh, A. U., Akbulut, A., Yildiz, K., & Şen, B. (2012). New records for the freshwater algae of Turkey (Tigris Basin). *Turkish Journal of Botany*, 36(6), 747-760. <https://doi.org/10.3906/bot-1108-16>
- Paerl, H. W., & Otten, T. G. (2013). Harmful cyanobacterial blooms: causes, consequences, and controls. *Microbial ecology*, 65, 995-1010. <https://doi.org/10.1007/s00248-012-0159-y>
- Paerl, H. W., & Huisman, J. (2009). Climate change: a catalyst for global expansion of harmful cyanobacterial blooms. *Environmental microbiology reports*, 1(1), 27-37. <https://doi.org/10.1111/j.1758-2229.2008.00004.x>
- Páliz Larrea, P., Zapata Ríos, X., & Campozano Parra, L. (2021). Application of Neural Network Models and ANFIS for Water Level Forecasting of the Salve Faccha Dam in the Andean Zone in Northern Ecuador. *Water*, 13(15), 2011. <https://doi.org/10.3390/w13152011>
- Paul, N., & Elango, L. (2018). Predicting future water supply-demand gap with a new reservoir, desalination plant and waste water reuse by water evaluation and planning model for Chennai megacity, India. *Groundwater for Sustainable Development*, 7, 8-19. <https://doi.org/10.1016/j.gsd.2018.02.005>
- Paul, S., Ooppelstrup, J., Thunvik, R., Magero, J. M., Ddumba Walakira, D., & Cvetkovic, V. (2019). Bathymetry development and flow analyses using two-dimensional numerical modeling approach for Lake Victoria. *Fluids*, 4(4), 182. <https://doi.org/10.3390/fluids4040182>
- Paynter, S., & Nachabe, M. (2011). Use of generalized extreme value covariates to improve estimation of trends and return frequencies for lake levels. *Journal of Hydroinformatics*, 13(1), 13-24. <https://doi.org/10.2166/hydro.2010.077>
- Person, M., Roy, P., Wright, H., Gutowski Jr, W., Ito, E., Winter, T., Rosenberry, D., & Cohen, D. (2007). Hydrologic response of the Crow Wing Watershed, Minnesota, to mid-Holocene climate change. *Geological Society of America Bulletin*, 119(3-4), 363-376. <https://doi.org/10.1130/B26003.1>
- Phan, T.T.H., Nguyen, X. H. (2020). Combining statistical machine learning models with ARIMA for water level forecasting: The case of the Red river. *Advances in Water Resources*, 142, 103656. <https://doi.org/10.1016/j.advwatres.2020.103656>
- Piasecki, A., Jurasz, J., & Adamowski, J. F. (2018). Forecasting surface water-level fluctuations of a small glacial lake in Poland using a wavelet-based artificial intelligence method. *Acta Geophysica*, 66(5), 1093-1107. <https://doi.org/10.1007/s11600-018-0183-5>

- Piasecki, A., Jurasz, J., & Skowron, R. (2017). Forecasting surface water level fluctuations of Lake Serwy (Northeastern Poland) by artificial neural networks and multiple linear regression. *Journal of Environmental Engineering and Landscape Management*, 25(4), 379-388. <https://doi.org/10.3846/16486897.2017.1303498>
- Ricko, M., Carton, J. A., & Birkett, C. (2011). Climatic effects on lake basins. Part I: modeling tropical lake levels. *Journal of Climate*, 24(12), 2983-2999. <https://doi.org/10.1175/2010JCLI3602.1>
- Rani, D., & Moreira, M. M. (2010). Simulation–optimization modeling: a survey and potential application in reservoir systems operation. *Water resources management*, 24, 1107-1138. <https://doi.org/10.1007/s11269-009-9488-0>
- Rodríguez-Rodríguez, M., Green, A. J., López, R., & Martos-Rosillo, S. (2012). Changes in water level, land use, and hydrological budget in a semi-permanent playa lake, Southwest Spain. *Environmental Monitoring and Assessment*, 184(2), 797-810. <https://doi.org/10.1007/s10661-011-2002-1>
- Sapitang, M., M Ridwan, W., Faizal Kushiar, K., Najah Ahmed, A., & El-Shafie, A. (2020). Machine Learning Application in Reservoir Water Level Forecasting for Sustainable Hydropower Generation Strategy. *Sustainability*, 12(15), 6121. <https://doi.org/10.3390/su12156121>
- Schulz, S., Darehshouri, S., Hassanzadeh, E., Tajrishy, M., Schüth, C. (2020). Climate change or irrigated agriculture–what drives the water level decline of Lake Urmia. *Scientific Reports*, 10(1), 236. <https://doi.org/10.1038/s41598-019-57150-y>
- Segrera, S., R. Ponce-Hernandez, and J. Arcia. (2003). Evolution of decision support system architectures: Applications for land planning and management in Cuba. *Journal of Computer Software and Technology* 3(1): 40–46.
- Sethia, A., & Raut, P. (2019). Application of LSTM, GRU and ICA for stock price prediction. In *Information and Communication Technology for Intelligent Systems: Proceedings of ICTIS 2018, Volume 2* (pp. 479-487). Springer Singapore. https://doi.org/10.1007/978-981-13-1747-7_46
- Shanker, M., Hu, M. Y., & Hung, M. S. (1996). Effect of data standardization on neural network training. *Omega*, 24(4), 385-397. [https://doi.org/10.1016/0305-0483\(96\)00010-2](https://doi.org/10.1016/0305-0483(96)00010-2)

- Singh, D. K., Singh, N. (2019). Drying Urban lakes: A consequence of climate change, urbanization or other anthropogenic causes? An insight from northern India. *Lakes & Reservoirs: Research & Management*, 24(2), 115-126. <https://doi.org/10.1111/lre.12262>
- Soylu Pekpostalci, D., Tur, R., Danandeh Mehr, A., Vazifekhah Ghaffari, M. A., Dąbrowska, D., & Nourani, V. (2023). Drought monitoring and forecasting across Turkey: A contemporary review. *Sustainability*, 15(7), 6080. <https://doi.org/10.3390/su15076080>
- Stedinger, J. R., Sule, B. F., & Loucks, D. P. (1984). Stochastic dynamic programming models for reservoir operation optimization. *Water resources research*, 20(11), 1499-1505. <https://doi.org/10.1029/WR020i011p01499>
- Su, D., Hu, X., Wen, L., Lyu, S., Gao, X., Zhao, L., Li, Z., Du, J., & Kirillin, G. (2019). Numerical study on the response of the largest lake in China to climate change. *Hydrology and Earth System Sciences*, 23(4), 2093-2109. <https://doi.org/10.5194/hess-23-2093-2019>
- Talebizadeh, M., & Moridnejad, A. (2011). Uncertainty analysis for the forecast of lake level fluctuations using ensembles of ANN and ANFIS models. *Expert Systems with Applications*, 38(4), 4126-4135. <https://doi.org/10.1016/j.eswa.2010.09.075>
- Talsma, J., Werkman, W., & van Loenen, A. (2016). Anticipatory real-time management in the Lake IJssel: implementation and practical application. *Procedia Engineering*, 154, 49-57. <https://doi.org/10.1016/j.proeng.2016.07.418>
- Taminskas, J., Petrosius, R., Simanauskiene, R., Satkunas, J., & Linkeviciene, R. (2013). Prediction of change in wetland habitats by groundwater: case study in Northeast Lithuania. *Estonian journal of earth sciences*, 62(2), 57. <https://doi.org/10.3176/earth.2013.06>
- Tejada-Guibert, J. A., Johnson, S. A., & Stedinger, J. R. (1995). The value of hydrologic information in stochastic dynamic programming models of a multireservoir system. *Water resources research*, 31(10), 2571-2579. <https://doi.org/10.1029/95WR02172>
- Teubner, K., Tolotti, M., Greisberger, S., Morscheid, H., Dokulil, M. T., & Kucklantz, V. (2006). Steady state of phytoplankton and implications for climatic changes in a deep pre-alpine lake: epilimnetic versus metalimnetic assemblages. *Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen*, 29(3), 1688-1692. <https://doi.org/10.1080/03680770.2005.11902973>

- Tsao, H. H., Leu, Y. G., Chou, L. F., & Tsao, C. Y. (2021). A method of multi-stage reservoir water level forecasting systems: A case study of Techí hydropower in Taiwan. *Energies*, 14(12), 3461. <https://doi.org/10.3390/en14123461>
- Turon, C., J. Comas, J. Alemany, U. Cortés, and M. Poch. (2007). Environmental decision support systems: A new approach to support the operation and maintenance of horizontal subsurface flow constructed wetlands. *Ecological Engineering* 30(4): 362–372. <https://doi.org/10.1016/j.ecoleng.2007.04.012>
- Üneş, F., Demirci, M., & Kişi, Ö. (2015). Prediction of millers ferry dam reservoir level in USA using artificial neural network. *Periodica Polytechnica Civil Engineering*, 59(3), 309-318. <https://doi.org/10.3311/PPci.7379>
- Üneş, F., Demirci, M., Taşar, B., Kaya, Y. Z., & Varçin, H. (2019). Estimating dam reservoir level fluctuations using data-driven techniques, *Polish Journal of Environmental Studies*, 28(5), 3451-3462. <https://doi.org/10.15244/pjoes/93923>
- Valizadeh, N., El-Shafie, A., Mirzaei, M., Galavi, H., Mukhlisin, M., & Jaafar, O. (2014). Accuracy enhancement for forecasting water levels of reservoirs and river streams using a multiple-input-pattern fuzzification approach. *The Scientific World Journal*, 2014. <https://doi.org/10.1155/2014/432976>
- Van der Heijden, T., Lago, J., Palensky, P., & Abraham, E. (2021). Electricity price forecasting in European Day Ahead Markets: A greedy consideration of market integration. *IEEE Access*, 9, 119954-119966.
- Viccione, G., Guarnaccia, C., Mancini, S., & Quartieri, J. (2020). On the use of ARIMA models for short-term water tank levels forecasting. *Water Supply*, 20(3), 787-799. <https://doi.org/10.2166/ws.2019.190>
- Voulanas, D., Theodossiou, N., & Hatzigiannakis, E. (2021). Assessment of potential hydrological climate change impacts in the Kastoria basin (Western Macedonia, Greece) using EURO-CORDEX regional climate models, *Global NEST Journal*, 23(1), 35-46. <https://doi.org/10.30955/gnj.003444>
- Walsby, A. E., Schanz, F., & Schmid, M. (2006). The Burgundy-blood phenomenon: a model of buoyancy change explains autumnal waterblooms by *Planktothrix rubescens* in Lake Zürich. *New phytologist*, 169(1), 109-122. <https://doi.org/10.1111/j.1469-8137.2005.01567.x>
- Wang, M., Dai, L., Dai, H., Mao, J., & Liang, L. (2017). Support vector regression based model for predicting water level of Dongting Lake. *Journal of Drain Irrigation Machine Engineering*, 35(11), 954-961.

- Wang, Q., & Wang, S. (2020). Machine learning-based water level prediction in Lake Erie. *Water*, 12(10), 2654. <https://doi.org/10.3390/w12102654>
- Wang, Y., Yue, J., Liu, S., & Wang, L. (2018). Copula entropy coupled with wavelet neural network model for hydrological prediction. In *IOP Conference Series: Earth and Environmental Science*, 113(1), 012160. IOP Publishing. <https://doi.org/10.1088/1755-1315/113/1/012160>
- Watkins, D. W., and D. C. McKinney. (1995). Recent developments associated with decision support systems in water resources. *Review of Geophysics* 33(2): 941–948. <https://doi.org/10.1029/95RG00179>
- Westrick, J. A., Szlag, D. C., Southwell, B. J., & Sinclair, J. (2010). A review of cyanobacteria and cyanotoxins removal/inactivation in drinking water treatment. *Analytical and bioanalytical chemistry*, 397, 1705-1714. <https://doi.org/10.1007/s00216-010-3709-5>
- Woolway, R. I., Jennings, E., Shatwell, T., Golub, M., Pierson, D. C., Maberly, S. C. (2021). Lake heatwaves under climate change. *Nature*, 589(7842), 402-407. <https://doi.org/10.1038/s41586-020-03119-1>
- World Bank. (2021). Population growth (annual %) – Turkiye. Retrieved on March 8, 2023, from <https://data.worldbank.org/indicator/SP.POP.GROW?locations=TR>
- Wu, N., Luo, Y., Liu, T., & Huang, Z. (2014). Experimental study on the effect of the Three Gorges Project on water level in Lake Poyang. *Journal of Lake Science*, 26, 522-528.
- Xiang, Z., & Demir, I. (2020). Distributed long-term hourly streamflow predictions using deep learning—A case study for State of Iowa. *Environmental Modelling & Software*, 131, 104761. <https://doi.org/10.1016/j.envsoft.2020.104761>
- Yadav, B., & Eliza, K. (2017). A hybrid wavelet-support vector machine model for prediction of lake water level fluctuations using hydro-meteorological data. *Measurement*, 103, 294-301. <https://doi.org/10.1016/j.measurement.2017.03.003>
- Yeşilköy, S., & Şaylan, L. (2022). Spatial and temporal drought projections of northwestern Turkey. *Theoretical and Applied Climatology*, 149(1-2), 1-14. <https://doi.org/10.1007/s00704-022-04029-0>
- Yoon, H., Jun, S. C., Hyun, Y., Bae, G. O., & Lee, K. K. (2011). A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of hydrology*, 396(1-2), 128-138. <https://doi.org/10.1016/j.jhydrol.2010.11.002>

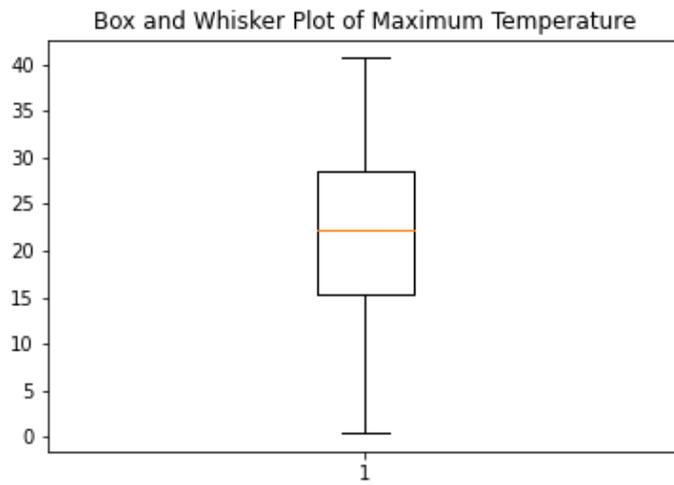
- Young, C. C., Liu, W. C., & Hsieh, W. L. (2015). Predicting the water level fluctuation in an alpine lake using physically based, artificial neural network, and time series forecasting models. *Mathematical Problems in Engineering*, 2015. <https://doi.org/10.1155/2015/708204>
- Yu, Z., Lei, G., Jiang, Z., & Liu, F. (2017). ARIMA modelling and forecasting of water level in the middle reach of the Yangtze River. In *2017 4th International Conference on Transportation Information and Safety (ICTIS)* (pp. 172-177). IEEE. <https://doi.org/10.1109/ICTIS.20178047762>
- Yuan, Z., Liu, J., Liu, Y., Zhang, Q., Li, Y., & Li, Z. (2022). A two-stage modelling method for multi-station daily water level prediction. *Environmental Modelling & Software*, 156, 105468. <https://doi.org/10.1016/j.envsoft.2022.105468>
- Zambelli, M., Siqueira, T. G., Cicogna, M., & Soares, S. (2006). Deterministic versus stochastic models for long term hydrothermal scheduling. In *2006 IEEE Power Engineering Society General Meeting* (pp. 7). IEEE, New York.
- Zappa, M., Bernhard, L., Spirig, C., Pfaundler, M., Stahl, K., Kruse, S., Seidl, I., & Stähli, M. (2014). A prototype platform for water resources monitoring and early recognition of critical droughts in Switzerland. *Proceedings of the International Association of Hydrological Sciences*, 364, 492-498. <https://doi.org/10.5194/piahs-364-492-2014>
- Zhang, X., Liu, P., Zhao, Y., Deng, C., Li, Z., & Xiong, M. (2018). Error correction-based forecasting of reservoir water levels: Improving accuracy over multiple lead times. *Environmental Modelling & Software*, 104, 27-39. <https://doi.org/10.1016/j.envsoft.2018.02.017>
- Zheng, F., Maier, H. R., Wu, W., Dandy, G. C., Gupta, H. V., & Zhang, T. (2018). On lack of robustness in hydrological model development due to absence of guidelines for selecting calibration and evaluation data: Demonstration for data-driven models. *Water Resources Research*, 54(2), 1013-1030. <https://doi.org/10.1002/2017WR021470>
- Zhu, S., Ji, Q., Ptak, M., Sojka, M., Keramatfar, A., Chau, K. W., & Band, S. S. (2023). Daily water-level forecasting for multiple polish lakes using multiple data-driven models. *The Geographical Journal*, 189(2), 357-369. <https://doi.org/10.1111/geoj.12488>
- Zhu, W., Yan, J., & Jia, S. (2017). Monitoring recent fluctuations of the southern Pool of Lake Chad using multiple remote sensing data: Implications for water balance analysis. *Remote Sensing*, 9(10), 1032. <https://doi.org/10.3390/rs9101032>

Zounemat-Kermani, M., Batelaan, O., Fadaee, M., & Hinkelmann, R. (2021). Ensemble machine learning paradigms in hydrology: A review. *Journal of Hydrology*, 598, 126266. <https://doi.org/10.1016/j.jhydrol.2021.126266>

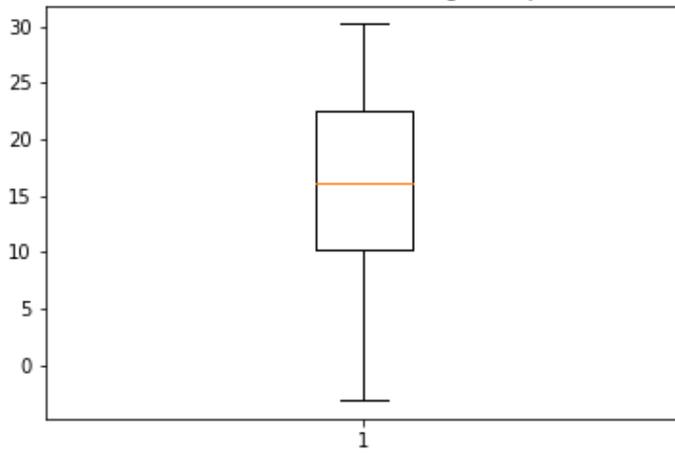
APPENDICES

APPENDIX A

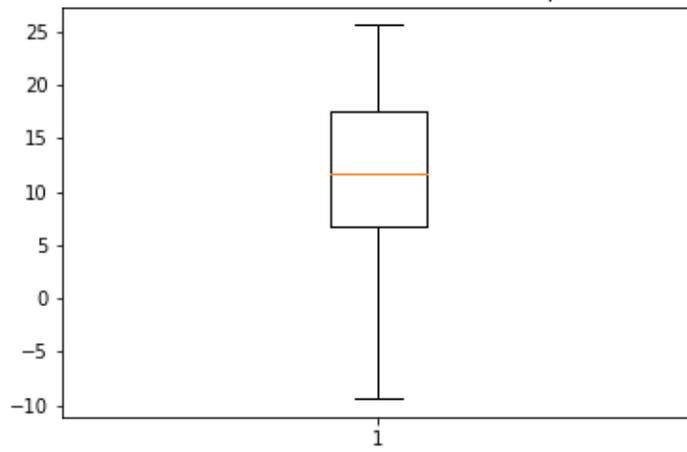
BOX AND WHISKER PLOTS



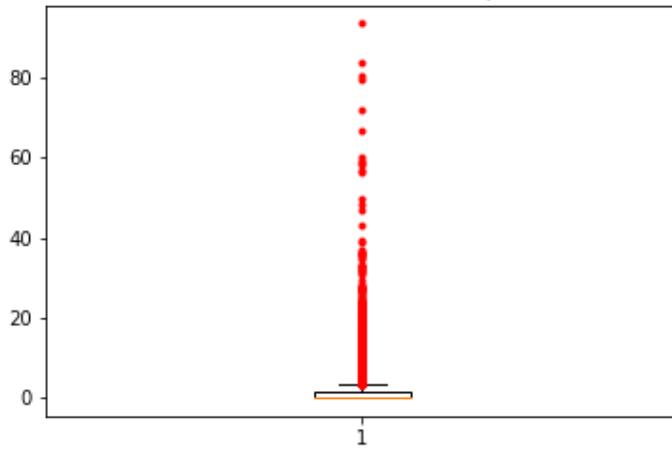
Box and Whisker Plot of Average Temperature



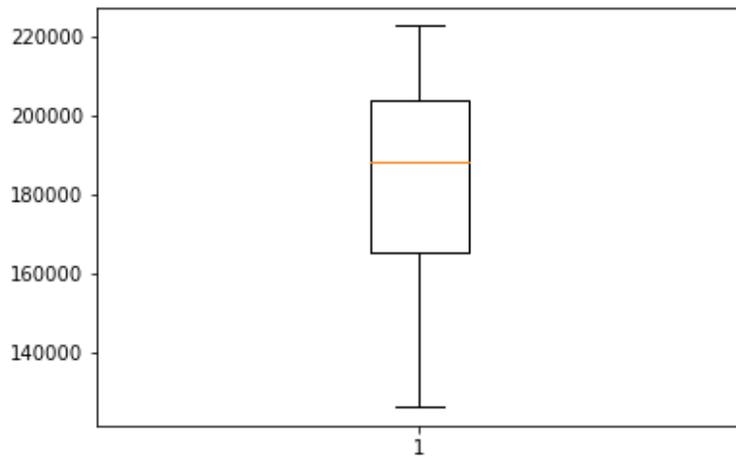
Box and Whisker Plot of Minimum Temperature



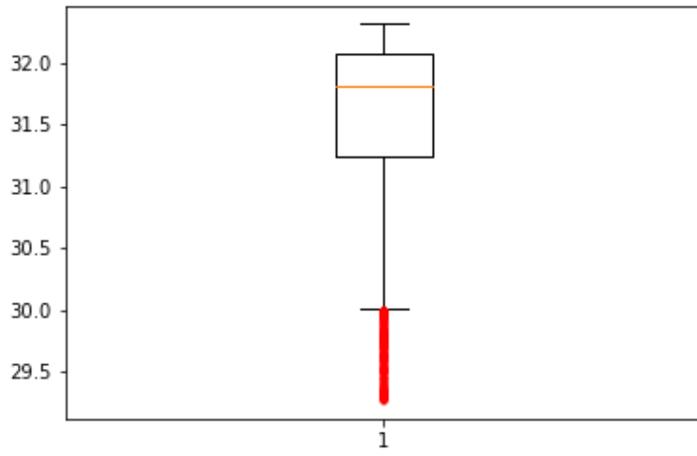
Box and Whisker Plot of Precipitation



Box and Whisker Plot of Withdrawal



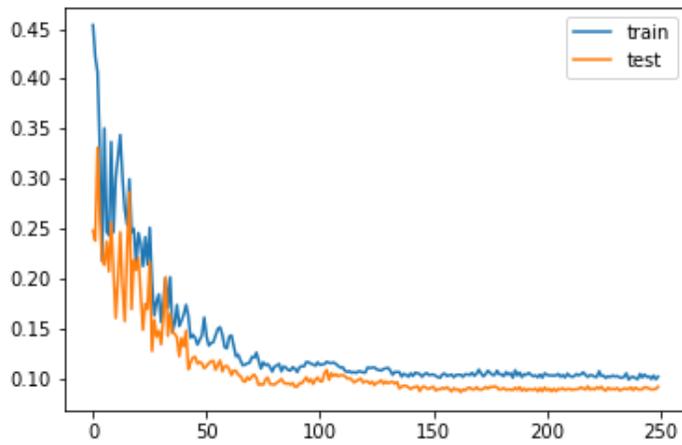
Box and Whisker Plot of Lake Water Level



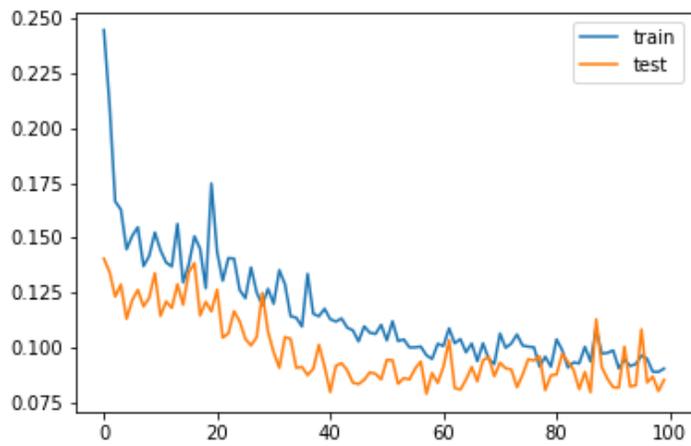
APPENDIX B

LOSS FUNCTION GRAPHS

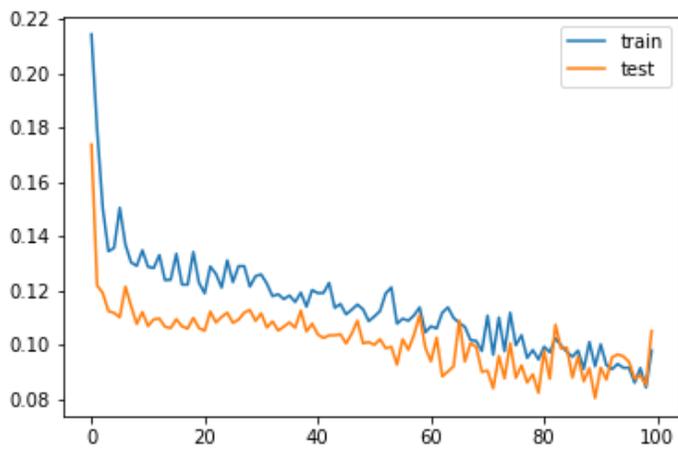
ANN



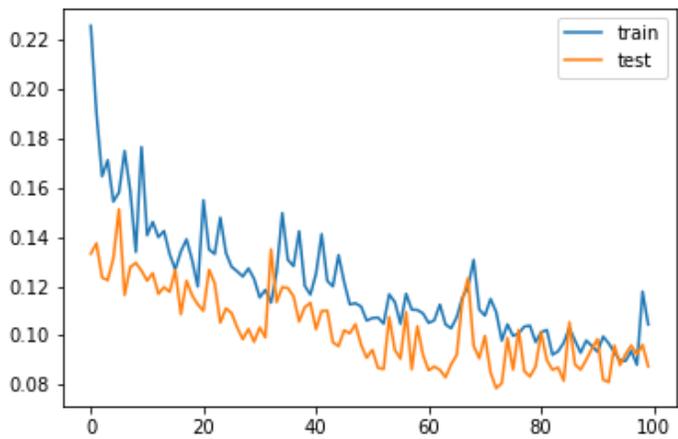
LSTM



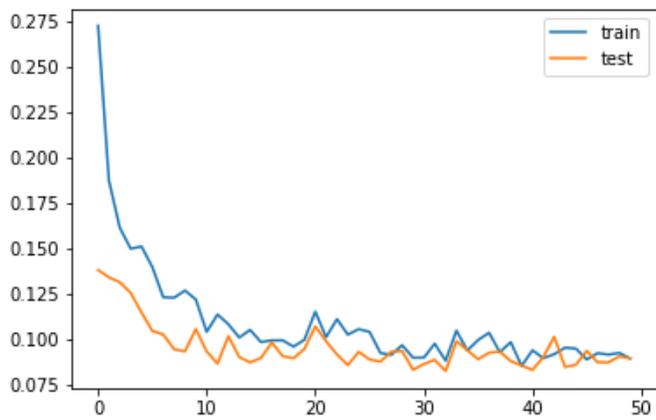
GRU



Stacked LSTM



Bidirectional LSTM



CURRICULUM VITAE

BIOGRAPHY

Serkan Özdemir is a data scientist in primarily areas of deep learning, machine learning, time series prediction, and lake water levels. He graduated from Middle East Technical University at the department of Business Administration. He received his master's degree from Boğaziçi University on the thesis of Country of Origin determination on digital transformation products. He currently works as a Research Assistant in METU focusing on several projects.

AREAS OF INTEREST

Data Science, Deep Learning, Machine Learning, Time Series Prediction, Lake Water Levels, Water Quality, Artificial Intelligence, Country of Origin, Digital Transformation, Industry 4.0.

EDUCATION

Degree	Institution	Year of Graduation
PhD.	Middle East Technical University	2023
MS	Boğaziçi University	2019
BS	Middle East Technical University	2017

WORK EXPERIENCE

Year	Place	Enrollment
2019-Present	Middle East Technical University, Graduate School of Informatics	Research Assistant
2021-Present	Clemson University-Online	Deep Learning Researcher
2023-2023	Delft University of Technology	Internship
2018-2018	HSBC	Internship
2017-2017	EG Consulting	Business Analyst
2016-2016	LNL Technology	Internship

TEACHING ASSISTANTSHIPS

IS100 Introduction to Information Technologies and Applications

IS501 Introduction to Information Systems

IS538 Contemporary Socio-technical Perspectives in Information Systems

IS539 Information Systems in Organizational Design and Applied Systems Thinking

IS540 Information Technology Acceptance in Organizations

PUBLICATIONS

SCI Indexed Journals

Ozdemir, S., Ozkan Yildirim, S. (2023). Prediction of Water Level in Lakes by RNN-Based Deep Learning Algorithms to Preserve Sustainability in Changing Climate and Relationship to Microcystin. *Sustainability*, 15, 16008.

Ozdemir, S., Yaqub, M., & Ozkan Yildirim, S. (2023). A systematic literature review on lake water level prediction models. *Environmental Modelling & Software*, 105684.

Ozdemir, S., Wynn, M., & Metin, B. (2022). Cybersecurity and Country of Origin: Towards a New Framework for Assessing Digital Product Domesticity. *Sustainability*, 15(1), 87.

Ozdemir, S., Şimşek, A., **Ozdemir, S.,** & Dede, C. (2022). Investigation of poultry slaughterhouse waste stream to produce bio-fuel for internal utilization. *Renewable Energy*, 190, 274-282.

Ozdemir, S., **Ozdemir, S.,** Ozer, H., & Yetilmezsoy, K. (2021). A techno-sustainable bio-waste management strategy for closing chickpea yield gap. *Waste Management*, 119, 356-364.

Ozdemir, S., **Ozdemir, S.,** & Yetilmezsoy, K. (2019). Agro-economic and ecological assessment of poultry abattoir sludge as bio-nutrient source for walnut plantation in low-fertility soil. *Environmental Progress & Sustainable Energy*, 38(6), 13225.

International Conference

Ozdemir S., Mutluturk M. E., Kor B., Metin B. (2019, October). Country of Origin Criteria for Digitalization with National IT Products. In *6th International Management Information Systems Conference 2019*.

Other Journal

Ozdemir S. (2019). Endüstri 4.0 Tarım. *Sakarya Ticaret Borsası*, 62, 6-7.

FOREIGN LANGUAGES

Native Turkish, Advanced English, Elementary Dutch, Elementary German