ROMAIN KASONGO AMSINI

WATER RESOURCES MANAGEMENT USING SOFT COMPUTING METHODS

MASTER THESIS

NEU 2024

NEAR EAST UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DEPARTMENT OF CIVIL ENGINEERING

WATER RESOURCES MANAGEMENT USING SOFT COMPUTING METHODS

M.Sc. THESIS

ROMAIN KASONGO AMSINI

Nicosia

February, 2024

NEAR EAST UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
DEPARTMENT OF CIVIL ENGINEERING

WATER RESOURCES MANAGEMENT USING SOFT COMPUTING METHODS

M.Sc. THESIS

ROMAIN KASONGO AMSINI

SUPERVISOR
Prof. Dr. GÖZEN ELKIRAN

Nicosia
February, 2024

Approval

We certify that we have read the thesis submitted by ROMAIN KASONGO AMSINI titled **"Water Resources Management Using Soft Computing Methods"** and that in our combined opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Educational Sciences.

Examining Committee	Name-Surname	Signature
Head of the Committee:	Prof. Dr. Gözen Elkiran	
Committee Member:	Assoc. Prof. Dr Fidan Aslanova	
Committee Member:	Asst. Prof. Dr. Ikenna Uwanuakwa	
Supervisor:	Prof. Dr. Gözen Elkiran	

Approved by the Head of the Department

21/02/2024

Prof. Dr. Kabir Sadeghi
Title, Name-Surname
Head of Department

Approved by the Institute of Graduate Studies

…../…../20…

Prof. Dr. Kemal Hüsnü Can Başer
Head of the Institute

NEAR EAST UNIVERSITY INSTITUTE OF GRADUATE STUDIES 1988

Declaration

I hereby declare that all information, documents, analysis and results in this thesis have been collected and presented according to the academic rules and ethical guidelines of Institute of Graduate Studies, Near East University. I also declare that as required by these rules and conduct, I have fully cited and referenced information and data that are not original to this study.

ROMAIN KASONGO AMSINI

13/02/2024

Acknowledgments

To all Lord, all honor, I would like to give all the glory and honor to the undisputed master of times and circumstances my God and my king for wisdom, intelligence and light.

To my father Kamani Kasongo Maurice, to my mother Gisele Nawej Isolo, to my brother and sisters Mike Tshawila Koj, Plamedie Nyange Kasongo, Vancia Isolo Karaj, Amelie Kamani Kasongo. My sympathy, my love, my concern is deep for them.

This work began under the direction of Prof. Dr. Gözen Elkiran. May he find here the expression of my sincere thanks for having agreed to direct me to carry out this work. For his availability, his advice, his guidance, his fruitful criticism. He spared no effort to get me to my destination.

To all my family here in Cyprus and my base: Peniel Beya, Dan Beya, Alliance Beya, Samy Mwanza, Joseph Mwenze, Ben Tavele, Anthony Kisula, Jonathan Bukasa, Eunice Tavele, Celia Tchaly, Shekinah Mulumba, Tracy Mukengere, Consolate Kayobe, Mariette Lusagila, Victoire Kinkonde, Magloire Kapita.

I thank all the heroes in the shadows, friends and acquaintances: Joyce Nguz, El Diglo Kassata, Ruth Sam Tshiyaz, Carmel Nkosi, Owin Ntumba, Jonathan Sumbu, Elie Kalumba, Isaac Katambo, Henoc Kalamabay, Simon Nyembwe.

I cannot forget the precious help of my brothers in the struggle and survivors: Laura Kinianga, Nephtalie Samba, Virginie Feza, Peniel Kasongo, Hervé Yav, Phares Basaole, Sam Chite, Fidel Nyembo, Nathan Malamba, Paul Kalenga, Hagion Kabeya.

All my highest consideration to all the teachers and all those who contributed to my training.

To all those who, directly or indirectly, have supported me and contributed to the completion of this work, may they accept with these words the expression of my consideration.

Romain Kasongo Amsini

Dedication

To my Dad for showing me the power of work and education,
to my Mum for showing me the power of love and prayer

Abstract

Water resources management using soft computing methods

Romain Kasongo Amsini

Prof. Dr. Gözen Elkiran

MSc., Department of Civil Engineering

January 2024, 81 pages

Nowadays, with massive population growth, urban societal challenges, modernization and conflicts linked to natural resources, we more than need the good management of natural resources. This work presents a tool and a management guide in the field of hydraulic and water engineering. The purpose of this research is to assess the capacity of models built on the basis of soft computing technique, notably ANN, ANFIS and MLR, to investigate and study the management of water resources, particularly for the city of Kinshasa in the Democratic Republic of the Congo. Scenarios such as the environmental impact based on the climate change, as well as the aspect of economic and social development were taken into account for the construction of this research, 5 parameters were evaluated to select the predictors and inputs of the models constructed in particular (the annual population, the average annual temperature, the annual precipitation, the GDP as well as the agricultural land). The selected data covers a period of 60 years from 1961 to 2021. The three modes were implemented following the MATLAB software and it was noted during the presentation and the discussion of the output that the models based on artificial intelligence ANN and ANFIS were more satisfactory than MLR with the following performances for ANN 99|% of R-squared and 0.000019 of RMSE, ANFIS 99% and 0.000054 of RMSE while the results of MLR were 91% and 0.08670

Key Words: Soft Computing, Water Management, Artificial Neural Network, Adaptive Neuro-Fuzzy Inference System, Multiple Linear Regression

Özet

Esnek hesaplama yöntemlerini kullanarak su kaynakları yönetimi

Romain Kasongo Amsini

Prof. Dr. Gözen Elkıran

Yüksek Lisans, İnşaat Mühendisliği Bölümü

Ocak 2024, 81 Sayfa

Günümüzde, fazla nüfus artışı, kentsel toplumsal zorluklar, modernleşme ve doğal kaynaklarla bağlantılı çatışmalar nedeniyle, doğal kaynakların iyi yönetimine fazlasıyla ihtiyacımız var. Bu çalışma, hidrolik ve su mühendisliği alanında bir araç ve yönetim kılavuzu sunmaktadır. Bu araştırmanın amacı, özellikle Kongo Demokratik Cumhuriyetin Kinşasa şehri için su kaynaklarının yönetimini araştırmak ve incelemek amacıyla, yumuşak hesaplama tekniği temelinde oluşturulan modellerin, özellikle ANN, ANFIS ve MLR'nin, kapasitesini değerlendirmektir. Bu araştırmanın oluşturulmasında iklim değişikliğine bağlı çevresel etki, ekonomik ve sosyal kalkınma boyutu gibi senaryolar dikkate alınmış, özellikle oluşturulan modellerin tahmin edicilerini ve girdilerini seçmek için 5 parametre değerlendirilmiştir (yıllık nüfus, ortalama yıllık sıcaklık, yıllık yağış, GSYİH ve tarım arazileri). Seçilen veriler 1961'den 2021'e kadar 60 yıllık bir dönemi kapsamaktadır. Üç model, MATLAB yazılımı kullanarak uygulanmıştır ve sunum ve çıktıların tartışılması sırasında, YSA ve ANFIS'e dayalı modellerin, diğer modelden daha tatmin edici olduğu not edilmiştir. YSA; R2 =0,99 ve KOKH=0,000019, ANFIS; 0,99 ve KOKH=0,000054, ÇLR; R2 =0,91 ve KOKH=0,08670.

Anahtar Kelimeler: Esnek Hesaplama, Su Yönetimi, Yapay Sinir Ağı, Uyarlanabilir Nöro-Bulanık Çıkarım Sistemi, Çoklu Doğrusal Regresyon

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AI	**Artificial Intelligence**
ANFIS	**Adaptive Neuro-Fuzzy Inference System**
ANN	**Artificial Neural Network**
BN	**Bayesian Network**
BP	**Back Propagation**
CC	**Correlation Coefficient**
DC	**Determination Coefficient**
DRC	**Democratic Republic of the Congo**
FFNN	**Feed Foward Neural Network**
FIS	**Fuzzy Inference System**
MAE	**Mean Absolute Error**
MAPE	**Mean Absolute Percentage Error**
MSE	**Mean Square Error**
MLR	**Multiple Linear Regression**
R^2	**Determination Coefficient**
REGIDESO	**Regie de Distribution des Eaux**
RMSE	**Root Mean Square Error**
SVM	**Support Vector Machine**
SWAT	**Soil and Water Assessment Tool**

UNESCO	**United Nation Educational, Scientific and Cultural Organization**
UNICEF	**United Nations International Children Emergence Funds**
UNDP	**United Nation Development Programs**
UNEP	**United Nation Environment Programs**
USAID	**United States Agency for International Development**
WHO	**World Health Organization**
WRF	**Weather Research and Forecasting**

CHAPTER I

INTRODUCTION

1.1. Overview

It goes without saying that water and air are undeniably the necessary elements for human survival, despite this, a significant part of world populace lacks to access adequate system of water supply. Water consumption and management is defined today as a physiological, social, economic, political and cultural issue. (Prabhata K. Swamee, Ashok K. Sharma, 2008).

The use of water by humans, animals and plants is universal. Without the latter, there would be no life on earth. Every living thing needs water to exist. Water resource management is a structure that enables us to make possible and guarantee the supply of drinking water to a residential area (city, village, etc.) or industrial. Different issues are related to water resource management, including health, economic, architectural and other that are linked to the quality and delivery of water. Another challenge is to limit the water waste and poor management which are both a source of waste of resources, money and even an opening for the spread of certain water-borne diseases. Generally, such a work is extended over hundreds of kilometres, over years and is the result of a long process and processing storage using numerous devices.

For several years, numerous projects have been carried out in the Democratic Republic of Congo in the field of both urban and rural hydraulics. Indeed, to stem epidemics and illnesses linked to improper consumption of water in different remote environments, Solidarités International has been implementing sustainable, adapted solutions in urban-rural contexts. This is already happening in certain regions of the country such as North and South Kivu. (Solidarités International, 2015). Small gravity drinking water supply networks represent a solution preferred for these environments.

In urban areas, only more complex interventions and longer (5 to 10 years), can radically change the situation for those city dwellers. This is the case of Beni and Kalemie. It is in partnership with REGIDESO that some global institutions are working extensively to rehabilitate water supply systems dating mostly from the colonial era and these do not only partially cover the water needs of residents

1.2. Statement of the problem

Despite the DRC's immense resources, mainly fresh water resources, the country faces a major challenge in the water sector. The DRC is currently facing an acute crisis in the supply of drinking water. In fact, no more than 26% of the citizens have the ability to use and drink clean water. (UNEP, 2011).

In addition to the complexity of the social and rural situation in the country, the political will and international aid have brought positive dynamism to the water sector.

Knowing that the country faces challenges in water resources management and people are suffering from the rarity of pure and clean water from water scarcity and water management, it is therefore necessary to seek for better tools for predicting and managing the resource allocated to the country. In many countries especially arid countries and semi-arid the waste of the water resource can create cultural, social, political crises and tension among communities. Water resource management has a remarkable effect on the social and the emergence of the country. Different studies have been conducted and shown that water resources management can be done through different models especially Soft Computing based models (Janusz A. Starzyk, 2015).

Unfortunately, in the Democratic Republic of Congo traditional linear models based on rough estimation, assumption, linear approximation are still used. Different models have been built through the past twenty years that cover the prediction of water consumption and water resources allocation applying soft computing methods such as artificial neural networks model, support vector machine model and multiple linear regression. (Msiza I. S., et al., 2008). It is therefore urgent and necessary to raise the question of Soft Computing based models and its usage in the domain of Water Engineering and Hydraulics in the Democratic Republic of Congo.

1.3. Purpose of the Study

The purpose of this research is to thrive the capacity of Soft Computing methods such as ANNs, ANFIS and the conventional MLR in the field of water engineering and hydraulics management, to compare the results of the mentioned methods, to select the best model for water resource management for Kinshasa Capital City considering different combination of data and predictors parameters. To set up, initiate, a Soft Computing technique that can be used as a template of combined models or single models in order to increase the efficiency of the prediction performance in the field of hydraulics and water engineering.

1.4. Objectives of the Study

- To evaluate the sensitivity of data and perform a preprocessing for the determination the most dominant parameters.
- To regulate and evaluate the execution of water resource management using ANN
- To calculate the performance of water resource management using ANFIS.
- To determine the achievement of water resource management using MLR
- To develop and compare the results of the three Soft Computing based methods in modelling the performance of water resource management.
- To establish, set up and apply the three Soft Computing techniques for improving the overall effectiveness of the prediction performance in water resources engineering.

1.5. Hypothesis

Regarding the literature review and current research, the following are made as the hypothesis for this research:

- ❖ The three models selected ANN, ANFIS and MLR will be adequate for the modelling of water resources management in Kinshasa Capital City.
- ❖ Correlation Coefficient method will conduct a good sensitivity analysis for the selection of predictors and parameters.
- ❖ Due to the weather pattern and temperature on the regions, the projection of the climate in the zone will be taken into account for our study.
- ❖ Two major scenarios (socio economic development which refers to population growth, the agricultural land, the gross domestic product and water demand; the climate change impact which refers to mean average temperature and monthly precipitation in the area).

1.6. Implication of the Study

- The significance of this research investigation is the fact that it will be the first research conducted to assess the water resource management using Soft Computing methods in Kinshasa.
- This investigation will be used as a base and pilar for more relevant work and studies in the area of water resource management.

- The data, models built and made accessible here will be useful to other reaches for further studies.
- The study will show the significance of the application of ANN, ANFIS and MLR in dealing with water resource management issues.

1.7. Limits of the Study

This research is limited to the modelling of water resource management using three Soft Computing models ANN, ANFIS, MLR models for state and province of Kinshasa. A total of parameters 5 and 3 predictors were generated.

1.8. Arrangement of the Study

The present thesis will be conducted as the following:

1. The first chapter that covers the overview related to the work, the statement of the problem, the purpose of the research, the objectives of this study, the hypothesis considered for the sake of this research, the novelty or the significance of this study and finally the limitations of this investigation.
2. The second chapter presents the earlier research conducted, the findings about the research. It is presented as the literature review.
3. The third chapter gives the methodology used to reach to goals of the research the formulations applied.
4. The fourth chapter provides the outcomes, findings and discussion
5. The fifth chapter is the conclusion and recommendations for future works.

CHAPTER II

LITERATURE REVIEW

2.1. Previous study in DRC

UNEP (2011) conducted a research and published a report on water issues that the Democratic Republic of the Congo is still encountering so far, the challenges and opportunities were enlightened to strengthen the report. The paper comprises a study of the environmental problems that DRC is facing and propose solutions and recommendations. The research gave an emphasis on the main difficulties in the water sector, mostly the drinking water supply. Different results and recommendations were made. To build up a system for water policy and regulations for the whole country, strategies were given and statutory regulations to maintain the development in the sector of urban and rural water supply network. To come up with an innovative comprehensive national method which applies water statistical data system for the Democratic Republic of the Congo. To invest in the community and create a self-government system, community-based management of microscale water supply network. To carry out a capacity-building program for non-centralized institutions in the domain of water management. To construct and implement watershed-based source protection plans. To enhance national capacity and bring out different regulation for water supply interventions in rural areas. To draw a field of monitoring program that will make sure the application of drinking water standards by wash actors in humanitarian rapid response. To design and implement renewable energy pilot projects for conventional water utilities and autonomous, community-based water supply system. To create ecological sanitation (Ecosan) pilot projects in strategic urban micro-catchments zones and finally implement pilot projects to introduce rainwater harvesting technologies at household and community levels.

SOLIDARITES INTERNATIONALES (2015) built a project on the design and the sizing of a water supply network in different areas, the network could fit and cover the needs of 250 000 habitants in term of water demand. The project strengthens the capacities of local actors means guaranteeing users, beyond our humanitarian intervention phase, the continuity of drinking water service. Autonomy, sustainability and viability of the water service represent a fundamental issue for the health of populations and are, as such, the basis of our intervention.

USAID (2017) conducted a technical report as part of the better management of natural resources as well as its development and support program for nations. The record focuses on the vastness of the country's resources as well as the management policy and water policy. The profile of colonial cities and those subject to the greatest climate impacts were taken into account. The report provides a plan and assessment of both the ground water level and surface water level resources and how they are used effectively to address major challenges in the sectors. It also proposes solutions for the construction of certain national and international programs that can take care of vulnerable cities and those with increased water shortages.

In 2017, World Group Bank carried out an investigation in the DRC to show the influence of social class on access to water, in most underdeveloped countries in the DRC also the accessibility to adequate water supply network system is often a function of social class. The poorer the population, the more difficult it is to access drinking water. the diagnostic carried out covers the hygiene and sanitation of water. The report shows a satisfactory but not total evolution compared to the years 2005. The conclusion is such that the DRC is a country with extraordinarily rich wealth, especially in terms of manual resources, but the quality of services is poor due to poor management of growth. demographic but also the bad authorities working to provide water to the population, in fact there are only a few water supply systems that are operational until now

Didier Bompangue et al. (2020) presented the first community-based application of the cluster grid response technique, which was used to select cholera case clusters using a grid approach during the Kinshasa cholera outbreak. Interventions included home disinfection, promoting good hygiene, treating and storing household water safely, and providing emergency water. In order to determine a seasonal pattern of the outbreak both before and after-action strategies were put into place, they also conducted a preliminary starting community trial study. Using epidemic curves and maps, cholera monitoring information from the Ministry of Health were examined to evaluate the dynamics of the outbreak in the space and the time. A number of 1712 probable cholera cases were registered in Kinshasa between January 2017 and November 2018. The health zones that were most impacted at this time were Kingabwa, Limeté, Kokolo, Kintambo, and Binza Météo. The cholera case numbers recorded weekly in Binza Météo commune, Kintambo town, and Limete residential place decreased by a mean of 57% only after two weeks and 86% after an observation of four weeks after the response strategy was put into place. Four weeks after the outbreak's height, the cumulative weekly case count in Kinshasa Province fell by 71%.

Romain Kasongo et al. (2020) worked on the design and the sizing of a water supply network system in JOLI SITE. The methodology adopted was based, firstly, on the construction of a database cartographic data of the study. Secondly, EPANET software was selected to analyse the performance and identified failures of the water supply network system. The research emerged with an on-demand power system optimized which ensures the good management of water distribution and which meets the needs in perimeter water.

The simulation results of the network in real working condition were found to be satisfactory. It presents technology not only adapted to the needs operation and maintenance, but which also meets the requirements of standards adopted in this area.

Romain Kasongo et al. (2023) examined and evaluated the water consumption of Lubumbashi in DRC applying artificial intelligence methods. The study showed the use of artificial intelligence while investigating water demand prediction in southeast region of Congo. Three machine learning models ANN, ANFIS, MLR have been developed to forecast water demand. For each model, the socioeconomic parameters and climate change impact were taken into account. The training of each model was performed using a 10-fold CV. The three architectures were evaluated using the MSE, the RMSE, the MAE, and R^2 score metrics. The outcomes demonstrated that the ANN model outperformed all other models considered.

2.2. Theoretical Framework

Jaehak Jeong et al. (2010) developed an integration architecture model for rainfall-runoff based on a sub-houlrly time frame, the implementation was performed using a watershed model. The SWAT 2005 model has been granted the ability to simulate flow on a sub-hourly time frame. The investigation of the base flow was conducted using hourly time periods and then dispensed evenly to every timeframe step, the model built was based on sub-hourly components in SWAT and enable the investigation of the runoff/infiltration first of all, the second step was the investigation of the overland flow routing, reservoir/pond/wetland routing were investigated as well, and channel routing at any sub-daily timeframe. Although the selected area for the case of study was spread on a 1.94 km^2, the watershed demonstrates a considerable amelioration in the model parameter results, principally in the evaluation and the estimation of high-level flows in comparison to the quotidian SWAT method, the disparity in the computing timeframe among the surface runoff and base flow may be a considered as a weakness for this method. Modelers need to use a different approach when simulating sub-daily hydrological processes. The SWAT analysis revealed that the model operational timeframe step has a significant impact on how sensitive the design parameters are to the design results. According to a sensitivity analysis presented in this paper, groundwater flow characteristics become more important as time intervals grow longer, and SWAT parameters linked to channel routing become more important as time intervals get shorter, down to 15 min. The calibration procedure was quite effective in this investigation thanks to the utilization of both automatic and manual calibrations.

Elana Pacchin et al. (2017) published a work on a short period of time water demand prediction model using a moving window that was based on the observed data. A new model was proposed for short period of time consumption in a water supply network system. The model comes up with a prediction for the 24 hours following hours using a pair of coefficients, whose coefficient values are updated for each step of the process, based on demand data. An application of the implementation model was set up to fit to real-time life situation and a comparison with different results for another short period of time prediction model based on the same data have shown that the constructed model has an efficient predictive capacity.

In China, using a new, comprehensive and evolutionary model, Jiping et al. (2018) assessed the adjustability of water resources management system. The suggested model was used to study the compliance improvement of the subsystems, coordinated development status between the existing subsystem, and adaptive amelioration of the WRS. The findings demonstrated that subsystems with a poor adaptability were often influencing the WRS's adaptability within the research area. It was determined that the lack of adequate coordination between each subsystem was the primary factor in the limited adaptive evolution of the subsystems.

Jiping et al. (2019) established a model based on short time period trend construction to assess the possible hydrological statistics from hydro meteorological time period. To tern, address the issues of lacking efficient data while capturing information and building a time series data, the research paper set up a short time period tendency architecture design by applying the assessment of time periods scheme and rules, the tendency of scheme patterns and rules, and the tendency of confidence and support in order to deal with the lack of relevant statistics, captured data and patterns and laws hidden in time series data in traditional research based on time series data. Using the suggested trend structure model for a given period of time, potential information and variation laws for both rainfall and evaporation time periods parameters were thoroughly investigated. The findings demonstrate that rainfall and evaporation have a short-term continually growing or reducing trend in the research area, and that their changing tendencies are largely consistent.

Elahe Kalashak (2021) conducted a research on the estimation of water demand prediction using machine learning methods. He used machine learning based techniques to estimate water demand prediction in sustainable and developed smart city using an hour time frame. Seven models were built to conduct the research and were compared to come up with the best result.

The first model used was SVM that showed the necessity of going closer to the time of consumption. The adaBoost regressor outcome demonstrated that for the range of data examined, the result would be desirable even if the research is moved back in the time as much as possible. The other model used was Ridge Regressor (RR) it gave a better outcome for a long-time duration for past data. The RF algorithm tend to give better result for large range of data. The KNN regressor shows the most proper for a long-range data set. XGBoost algorithm works better and efficiently for long time period series. A focal point was made on the LSTM model because of the time of simulation and long computation.

2.1.1. Concept I: Artificial Neural Network

Muleta and Nicklow (2005) illustrated amply the applications and enormous versatile role that artificial intelligence methods and environmental assessments (EAs) could play in resolving genuine, challenging issues with water resources and environmental systems. Natural selection is the foundation for GAs, EPs, and SPEA, all of which have been combined for a variety of uses and goals. An artificial neural network (ANN), a technology modeled after the functioning mechanisms of the human brain, has proven effective in addressing the computing demand issue. It has been possible to apply a novel training strategy that leverages the advantages of gradient-based and EA-based search techniques to other systems. The control of other NPS contaminants might potentially be incorporated into the computational models that are showcased here. However, it was discovered that the models required a lot of computing power. An ANN-based simulation model that mimics and produces the necessary SWAT outputs was created because the influence of the CPU demand on the usefulness of the computational tools was a source of concern. The blend of the BP and EP algorithms was performed to train the ANN architecture. The developed ANN architecture reduced CPU time by 84% when it replaced SWAT, demonstrating the effectiveness and efficiency of the training procedure.

2.1.2. Concept II: Adaptive Neurro Fuzzy Inference System

Lihua Xiong et al. (2001) created the combination forecasts; they integrated the simulation data from the several models. The study explains and introduces the Takagi-Sugeno fuzzy system as a way of gathering the simulation performance of five different experimental rainfall-runoff models for flood prediction research. A number of eleven catchments using the simple average model (8SAM), the weighted average model (8WAM), and the neural network method (8NNM). When the forecast simulation effectiveness of the Takagi-Sugeno affiliation technique is compared to the other three combination methods, it can be shown that, in terms of improving the accuracy of flood forecasting, the Takagi-Sugeno model is just as effective as the WAM and the NNM. It is advised to employ the Takagi-Sugeno approach as the combination system for flood forecasting due to its effectiveness and simplicity.

In 2004, P.C Nayak et al. investigated a neuro-fuzzy method for modeling hydrological short period of time. The potential of a neuro-fuzzy computing techniques was developed using an ANFIS for modelling hydrological short period of time of a river flow basin. The observed ANFIS model is performed a traditional ARMA models that simplifies the building process, it was noticed that the model was able of persevering the numercial properties in the short time period.

Mohammad Z. Kermani et al. (2007) developed an ANFIS for hydrological time series prediction. The proposed model was used to predict daily behavior of a river flow. Two types of architecture were considered regarding the type of data, the inputs and the variables pattern. For each network respectively 2, 3, 4, 6, and 8 memberships functions were executed. The outcome of the implementation was compared to the observed data; the output shows that a regular structure for the network were not convenient. A combination of the auto regression and the neurro-fuzzy network shows better results for the research.

Nourani et al. (2011) conducted a research based on two hybrid artificial intelligence techniques for the modelling of a rainfall-runoff process. A combination of ANN, SARIMAX and ANFIS were used to construct a hybrid model, the model was able to capture both nonlinear and seasonal parameters of the runoff short period data. Those hybrid multi-variable methods showed a great enhancement in rainfall-runoff modeling. By using the hybrid model a nonlinear Kernel structure was built to simulate nonlinear behavior of the phenomena. The ANN and ANFIS models could monitor only the autoregressive scheme of the time period.

D. P. Vijayalaksmi et al. (2015) studied water demand forecasting system using ANFIS. The proposed methodology relied on six ANFIS models that were developed by considering different number of memberships functions and time periods. The research has revealed that the tested models are subjected to an improvement while the number of inputs parameters increased. The results show and manifests that the ANFIS approach is a workable tool for modelling water demand.

2.1.1. Concept III: Multi Linear Regression

Abdullhai et al. (2019) uses MLR combined with two other AI-based techniques for forecasting ET0 evapotranspiration, including single-step and multi-step forecasts. The research uses 12 input parameters which were processed sequentially with numerous time lags (up to 12 months) in order to enhance the precision of the forecast models. This being done, the Markovian strength of the settings features was checked to ensure that the seasonal nature of the process was well covered.

The outcome of this study shows that if the temperature combines with the other parameters such as TD, RH, PR, and SP have weak Markov chains, the other parameters respectively Ep and RS have strong Markov chains. This suggests that for Ep and RS, the sequence of previous events determines the ET0 values for the upcoming step, while for TD, RH, PR, Umin, Umax, Umean and SP, the ET0 values for the next step depends only on previous events.

2.2. Related research

Ishmael S. Msiza et al. (2008) investigated water demand prediction in South Africa region using two models respectively artificial neural network and support vector machine. The approach adopted was conducted into two parallel experiments the first one for the ANN and the second one for the SVM. The architecture chosen for the ANN models were respectively the multi-layer perception while the SVM relied on different Kernel functions. The performance criteria used from each experiment were respectively the validation error and the accuracy of the validation dataset target values. The outcome of the two methods were therefore compared to come up with the best and suitable result for the research.

Afshin Shabani et al. (2017) successfully used SWAT model and CE-QUAL-W2 model to evaluate the variation in both water quality and quantity combined with the arising water levels in Devils Lake. The result demonstrates the great potential of using and combining the SWAT method and CE-QUAL-W2 model for studying watershed and reservoir arrangement practices. A buoy monitoring water quality was deployed in Devils Lake and incorporated near-real-time data into the models to improve its performance for both simulation of water quality and quantity in Devis lake watershed.

Alireza Deriane (2017) used a WEAP software and created a simulation model to investigate the Urmia Lake basin's water resources system. With an NSE of 0.97, a calibration process was suggested and executed utilizing historical data covering the years 1966–2012. Further investigation revealed that the hydrology of Lake Urmia is significantly impacted by various water withdrawals, primarily from wells, Qanats, and river diversions. The second significant issue contributing to the Lake's current critical state is the operating reservoirs and the demand sites that they serve. It was discovered that in 2012, the LWL would have increased by 5.2 meters, or 22,325 mcm of lake storage, if other withdrawals had not occurred.

Tao Tian and Huifeend Xue (2017) conducted research on BNN and BP neural network algorithm, MAE and MAPE of the predicted value is 0.70% and 0.46%, respectively, which demonstrates that parameters selected for this study was found to be effective. The findings demonstrate that, in certain circumstances, Bayesian regularization energy can remove redundant information from the network architecture and lower the network's complexity, meeting the training samples' fitting accuracy requirements. The proper degree and linkage weight are compared once the two models have been trained. The BNN method outperforms the BP neural network method in terms of accuracy by 0.24%, demonstrating the effectiveness of Bayesian regularization. The technique can significantly increase the network's capacity for generalization. The Guangdong province's annual water usage can be predicted using the model as a guide. This offers precise data for overall water management and assistance in encouraging the wise use of water resources in Guangdong.

Li et al. 2017 developed for hilly river in the northern part of India a model which was based on the simulation of over a ten years' time frame, the method was used to test a two-way combined implementation of the Weather Research and Forecasting, Model and the WRF-Hydro hydrological modeling extension package in its offline version. A triple nest is used, with the hydrological components located at 300 m and the atmospheric model grids located at 3 km in the innermost domain. A quantitative evaluation of two microphysical parameterization (MP) schemes is conducted to determine the ways in which MP affects hydrological responses and orographic-related precipitation differently.

Bentolhoda Asl-Rousta et al. (2018) studied several hydrological models for Sirwan River Basin in Iran, twelve distinct SWAT method were used and compared together for the same bassin. Regarding the calibration strategies, the set for the gauged stations, predictors parameters and the calibration process for the twelve models were different. Eight MSC including AICc, CAIC, AICu, AIC, SIC, HIC, SICc, HICc were used for the calibration and the validation of the outperformed model. The results showed a significance between the ranking outputs and two examined scenarios.

Evangelos Rozos (2019). Discussed the benefits of using an automatic learning algorithm to manage an irrigation system network. For the purpose of creating a network neuronal à action directe (FFN), the output of a network programmation flow model (NFP) that simulates and optimizes the operation of an aquatic supply system was used. Appropriately, the PFN's penetrating functions have been selected to reflect operational policies with varying degrees of risk acceptance. The utilisation of long-term synthetic data has been employed in the programming of flux networks to reliably capture the risk associated with each political operation and to ensure a lengthy period of training for the FFN.

Mngereza Miraji et al. (2019) conducted a research on the impacts of water demand and its implications for future surface water resources management in Tazania's Wamu Ruvu Basin. The overall estimation of the future increase in water use primarily concentrated on the agricultural and domestics sector. The authors relied on the arithmetic and mathematical model to assess the water demand and evaluate the implication of different factors such as population growth and economic development. Economic Growth (EG) and Demand Side Management (DSM) scenarios combine with current trend (CT) of Tanzania were used to conduct to research.

Runzi et al. (2020) developed a patter for urban water quality, the patter was found to significantly influence stream E, TP, and NO3 –N in the amount of E. The study case was the Coli in the Texas Gulf Region, with the connections between them changing depending on the time of year and place taking into account the following parameters: Patch Cohesion Index (COHESION), Largest Patch Index (LPI). In developed areas, the Splitting Index (SPLIT) and Landscape Division Index (DIVISION) proved to be the most effective metrics for measuring urban growth patterns in relation to stream water quality. Between each other and the developed areas' Edge Density (ED) and Juxtaposition Index (IJI) were crucial for particular pollutants and times of year. The impact of the urban development pattern on the quality of stream water was greater in wet seasons as opposed to arid ones.

In Ghana's Tono irrigation dam, Edward Naabil et al (2020) evaluated the possible reduction in rainfall and its effect on water resources, especially the Climate Research and Forecasting (WRF) model downscaled climate parameters of Representative Concentration Pathways (RCPs) 4.5 and 8.5 in reference to historical data (1990-2010), and the ECHAM6 model projected the future climate data. A two-domain arrangement was employed, consisting of an outside domain covering the West African zone and an inner domain centered on the Tono basin at a resolution of 5 km. The evaluation was conducted using the simulated temperature and precipitation's yearly average, the relative percentage change, and geographic seasonal variation. The findings demonstrate that both scenarios disagree on the change signal for precipitation. RCP 4.5 shows a rise in yearly precipitation of +7%, whereas RCP 8.5 shows a fall in precipitation of -9.6%. In terms of temperature, both predictions concur that it will rise. These findings indicate that climate variation will have an impact on streamflow in the future.

There are signs that the flows may decrease, which would also lower the dam levels and have an impact on irrigation operations. Therefore, this study informs the operators of the Tono irrigation dam about the steps that should be taken to ensure its sustainability.

Zhang et al investigated the repercussion of increasing hydrological model based on the description of the land atmosphere and the interactions in complex local terrain in 2020. Heihe River Basin was considered as the case study, The searcher used the WRF methods with the WRF-Hydro package for the extension of the model, for this purpose. The impact of lateral water flow on land-atmosphere reactions is assessed using a joint atmospheric water budget investigation, a regional precipitation recycling assessment, and a fully three-dimensional atmospheric moisture tracing method by comparing simulation methods outcomes with and without coupling for the period going from 2008 up to 2010. Along with simulating the temperature at the near surface and the precipitation fluctuation similarly to the WRF method, the connected modeling system WRF-Hydro also shows that it can replicate daily streamflow. Due to the lateral terrestrial water flow description in the fully coupled mode, the redistribution of the infiltration excess in the mountainous area raises the soil moisture content in the root zone, increases evapotranspiration and terrestrial water storage, and reduces total runoff. As a result of the near-surface wetting and cooling, variations in precipitation are caused, which in turn alters the regional water vapor transports and water vapor content. Overall, the recycling rate is increased by the fully coupled simulation, suggesting that the regional climate in the research area is influenced by lateral terrestrial water movement.

Baolin Xue et al. (2021) suggested a hydrological method and a hydrodynamic method for water quality management based on SWAT model and HEC-RAS model. The two models were combined together to implement the simulation of the spacio-temporal patters for water quantity and water quality. The models used an “Output-Input” pattern, where the runoff from the SWAT method was adopted as the input parameters of the HEC-RAS method in the simulation.

Quing Shuang and Rui Ting Zhao (2021) presented a study that assesses water demand in China, particularly in the Beijing-Tianjin-Hebei region. This study spans 15 years and covers a period ranging from 2004-2019. Eleven machine learning models were built to predict water consumption based on economic, demographic and resource parameters availability. The models were built using 10-CV. The four best models were evaluated on the test data by the MSE, MAE and R2 score measurements. The results of the GBDT model show better performance compared to other models and the performances were evaluated on the basis of three metrics MAE, MSE and R^2 score. Performances of all the models taken into account, obtaining the lowest error rates and R2 values of 99.9578% for EPS and 99.9999% for IPS, appropriately. Statistical models performed better than machine learning models, according to a comparison of model performance. When compared to individual models, ensemble models outperformed machine learning models.

Following that, three additional Chinese regions with distinct economies, topographies, and climates were used to validate the GBDT model. Every forecast was more than 80% accurate. This illustrates how reliable the GBDT model is. The same dataset other locations' water demand can be predicted using the explanatory factors. Data from testing and training were also modified to forecast water consumption for the ensuing two years.

Mengyan Zhu et al. (2022) conducted a study on the review of certain machine learning methods for the evaluation of water quality, different applications such as surface water, ground water and drinking water were taken into account for the construction of the models. The predictors considered for each of the parameters as well as each of the models changed according to the analysis needs of the model. A total of 8 different machine learning models were necessary for the evaluation of performance including ANN, SVM, SOM, XGBoost, RF, PCA, Soft Sensor, DT. Complex models ranging from seawater harvesting, pollution of natural resources to the management of water resources and the composition of the latter have been built. Two metrics R score and RMSE were retained for the evaluation of performances. The goal here was not to compare the different methods but to make an overall evaluation using different parameters to shed light on the contribution and application of the machine learning in water quality assessment

Table 2.1. Explanatory variables found in the literature review

Explanatory variables		*Unit*	*Lu et al 2020*	*Li et al 2017*	*Zao and Chen 2014*	*Tian and Xue 2017*	*Zhang et al 2013*	*Sun et al 2017*
Economy	***GDP***	*Dollars USD*		✓		✓	✓	✓
	GDP per Capita	*Dollars USD per person*			✓		✓	✓
Socio-Economic	***Year-end population***	*Million Hab*	✓	✓			✓	✓
Water Demand	***Agricultural water demand***	*Billion m^3*				✓	✓	✓
	Irrigated area	*Thousand Hectare*				✓	✓	✓
Resources availability	***Water resources available***	*Billion m^3*				✓		
	Annual Precipitation	*mm*				✓		

CHAPTER III

MATERIALS AND METHODS

3.1. Study area and location

The case study chosen to conduct our research is the city province of Kinshasa. Kinshasa is the capital city but also the biggest city of the Democratic Republic of Congo. A historical, social and cultural emblem, it represents a moral value and a patriotic emblem for many Congolese people. For several years, Kinshasa proved to be the site and stronghold of fishing villages and traders along the Congo River, alongside its position astride the various waterways. It is also considered today as one of the megacities in the world with the fastest growing capacity in term of population but also in term of expansion. Its population is estimated at no less than 26 million inhabitants, it sits on the peak of Africa in terms of demographics of the continent because it is the most populous city in the DRC but also the densest metropolis on the African continent. It is considered the third biggest metropolitan zone in Africa, the leading social and economic, but also political and cultural middle point of the DRC. Kinshasa is home to many big constructions, such as manufacturing, industries, telecommunications, banking development site and entertainment for tourist.

Its surface area and geographical extent extends over some 9,965 km^2. It is spread out on the southern side of the Malebo Pool, creating a vast form of a crescent on low, flat ground at an mean altitude of about 300 meters. Its geolocatable position is between the following coordinates: latitudes 4° and 5° and eastern longitudes 15° and 16°32', Kinshasa shares its limits with the province of Mai-Ndombe, the province of Kwilu and the province of Kwango to the east; the Congo River demarcates its western and northern perimeters, forming a natural limit with the Republic of Congo; to the south lies the central province of Kongo. Across the river is Brazzaville, the smaller capital of the neighbouring Republic of Congo, constituting the second closest pair of capitals in the world, although it is separated by a four-kilometre-wide stretch with no bridge on the Congo River.

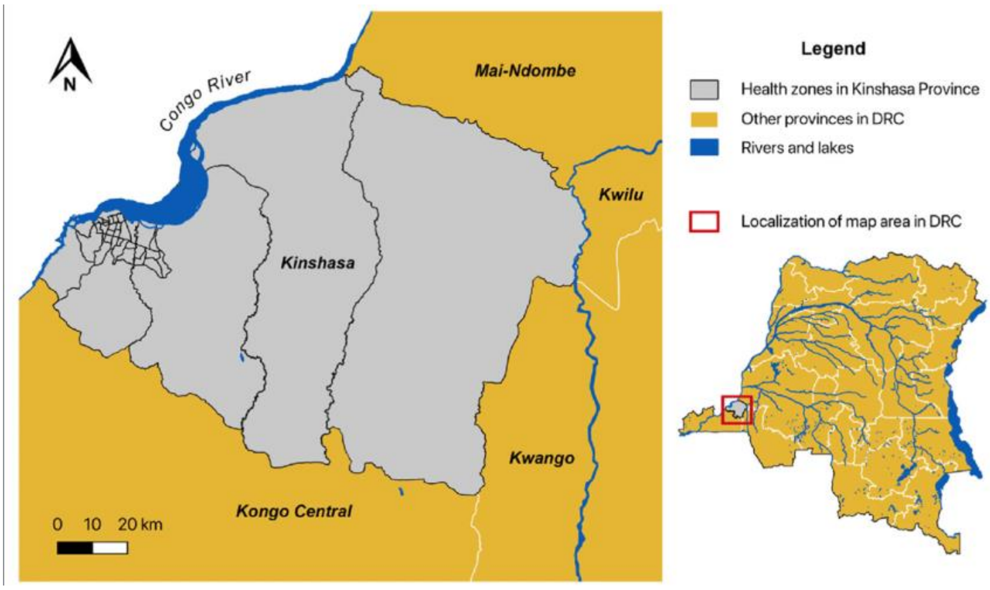

Figure 3.1. Map of Kinshasa (Dider Bompangue et al, 2020)

The second figure represents the case of the city Kinshasa in another form with different access routes but also an administrative view.

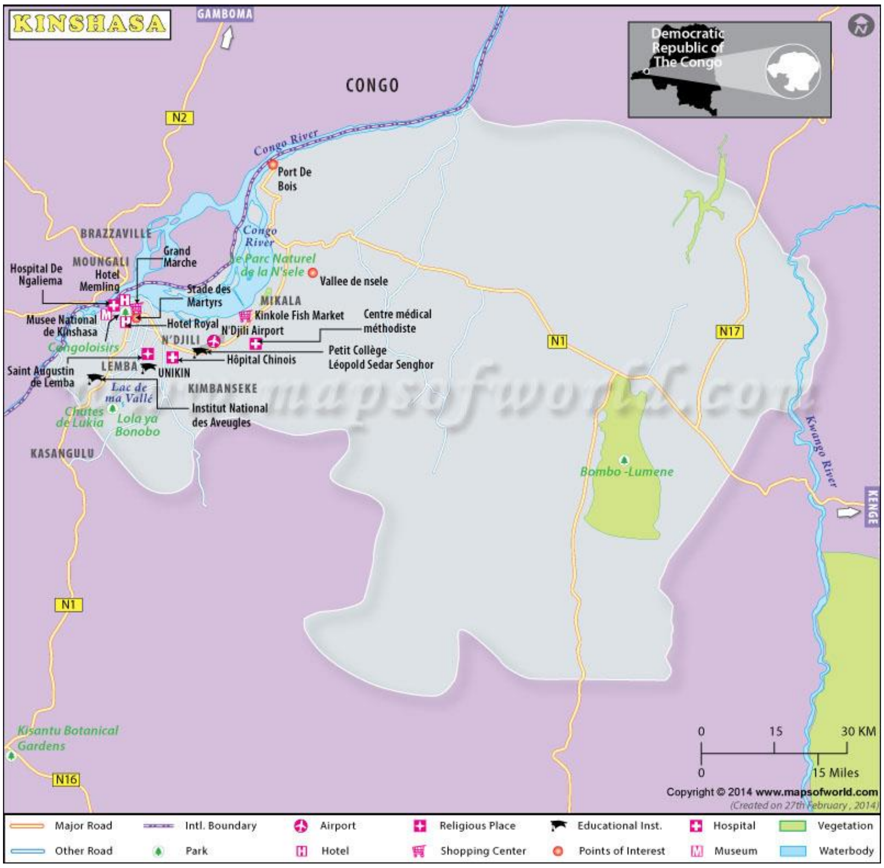

Figure 3.2. Map fo the Study area (Map of the World, 2023)

The DRC represents wealth on an ecological level but also on a hydraulic level. the following figure represents an overall situation of the DRC in relation to its resources over its entire extent.

Figure 3.3. DRC RainForest Map (Britannica, 2023)

3.2. Data collection and Pre-processing

The range of data selected for this research spans 60 years covering a period from 1961-2021. The parameters chosen for this thesis are annual parameters. The data collected was recorded and obtained through certain national institutions (the annual report of the Ministry of Planning and Development but also the Ministry of Energy and Hydraulic Resources as well) but also

from certain online databases such as the World Bank group organization https://www.worldbank.org/en/home which to provide us with updated weather data for the city province of Kinshasa. A total number of 5 parameters were retained for this study.
The following table represents the main explanatory variables for the thesis.

Table 3.1. Description of Predictors Data

Parameters	***Units***	***Min***	***Max***	***Mean***	***Std Deviation***
Mean Temperature	*Celsius*	*23.68*	*24.74*	*24.19*	*0.26*
Mean Precipitation	*millimetre*	*1414.1*	*1528.74*	*1485.60*	*33.57*
Gross Domestic Product	*Dollars*	*1250.98*	*2882.8*	*2001.3*	*447.7*
Population	*Hab*	*1841820*	*26681825*	*9726192.5*	*7198193.8*
Agricultural land	*Thousand Hectare*	*250500*	*337300*	*267967.86*	*25855.86*

The first predictors parameters considered for the selection of data refers to the climatic impact which the world as well as the City of Kinshasa is facing. The climate data contains two important parameters: the mean annual temperature and the mean annual precipitation (Tian and Xue, 2017).
The second scenario considered for the selection of predictors parameters is the socio-economic and development impact because of the growing expansion of the city. The named scenario brings out the following parameters: the annual increase in the population but also the gross domestic product of the DRC). (Lu et al. 2020).

The very last scenario considered for the construction of this research focuses on the available natural resources but also the water supply of the study area the parameter considered for this scenario is the irrigated land (Sun e al. 2017).

3.3. Data Pre-processing and Analysis

3.3.1. Data Normalization

Data preprocessing and treatment is a common, appropriate and above all very significant means of processing information. For the construction of the model and the modelling of the latter, the data are rescaled into the interval of 0 to 1. The process used to rescale the data is called normalization process, see the following equation which represents it. (Abdullahi and Elkiran, 2017).

$$z_i' = \frac{z_i - \min(z)}{\max(z) - \min(z)} \tag{3.1}$$

Where, the normalized data, the observed data, the maximum value of data, and the minimum value of data are represented by z'_i, z_i, z_{min}, and z_{max}.

The data collected afterwards was then taken to another process to eliminate noise as well as the relationship between certain data.

The normalized data were then split into three sets, notably 70% of the data for training, 15% for validation and 15% for testing. With a total of 60 instances considering yearly parameters and covering a period of 1961-2021 respectively.

3.3.2. Performance Criteria

In relation to the methodology chosen and based on the available literature, for the evaluation of statistical performances, for the research conducted for this thesis two different types of standard statistics were taken into account because of their redundancy in several works (Nourani et al. 2020). The Root Mean Square Error (RMSE) and the Determination Coefficient R-squared were applied.

The data representation for the test consists of the rest of random samples, which represent a 30% range of the overall data set from the years 1961 to 2021. These measurements were considered because they have shown good and satisfactory results while used in recent research studies for predicting water consumption and demand (Kasongo and Elkiran, 2023). For each approach for the selected methods, the following equations to measure the deviation between the actual and simulation-predicted values were used:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=0}^{n-1}(r_i - \hat{r}_i)^2} \tag{3.2}$$

r_i, $\hat{r}_i$ present the observed data values and the anticipated data values respectively and n is the number of observations. The Root Mean Square Error (RMSE), obtained by squaring the error between the predicted values and the observed values, it is used to measure the accuracy of the projected values compared to the observed. The root mean square error (RMSE) is given between 0 to infinity ($0 \geq RMSE < \infty$). The effectiveness of the performance is proportional to the decrease in the root mean square error (RMSE), when the latter approaches zero, the model shows better results in its performance.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{3.3}$$

R^2 is the measure of the goodness of the fit. It evaluates the proportion of variance that the explanatory components of a model can explain. $x_i, \hat{x}_i, \bar{x}$ represents the observed data values, the predicted data values and the mean of the observed data values, n is the total number of observations made. When R^2 approaches 1 this shows the quality of the model's fit, the more the evaluation tends towards less infinity the more the model decreases in its quality ($-\infty < R^2 \leq 1$). (Nourani et al., 2020).

3.3.3. Model Validation

To conduct this research and to write this thesis, an n-fold cross-validation approach was used to investigate the validation of the different architecture models constructed. With the n-fold method, the entire data set is randomly split into a number of subsamples that are equal.

These sub-samples are considered to be of number n. The model is then trained with the subsamples of number n minus 1 (n-1) and the rest of the subsample n is used for validation. The process is repeated in a manner continuous by taking n-1 and n sub-samples being different for the training and validation of each model built. The final results are then obtained by taking the average of the results which we will multiply by n. The advantage of this n-fold validation method is that for the construction and validation of the model, the different observations are used once (Nourani et al. 2019).

For this case study and this thesis, the random n-fold division of all the data was carried out for n=4 (4-fold) subsamples. The ¾ of the subsample were used for the construction of the model and the remaining ¼ subsample were taken into account for model validation. 4 consecutive times, the process was repeated with different ¾ and ¼ subsamples for the construction of the model and its validation.

3.3.3. Parameters Description

a. Precipitation (mm)

Because of its position to the south of the equator, the climate of the provincial city of Kinshasa is relatively warm and humid over several months, the consequence is that the dry seasons are generally very short while the hot and rainy seasons are much longer. Because of these parameters, one of the reasons why the precipitation was chosen as the first predictand for the construction of this work and its realization.

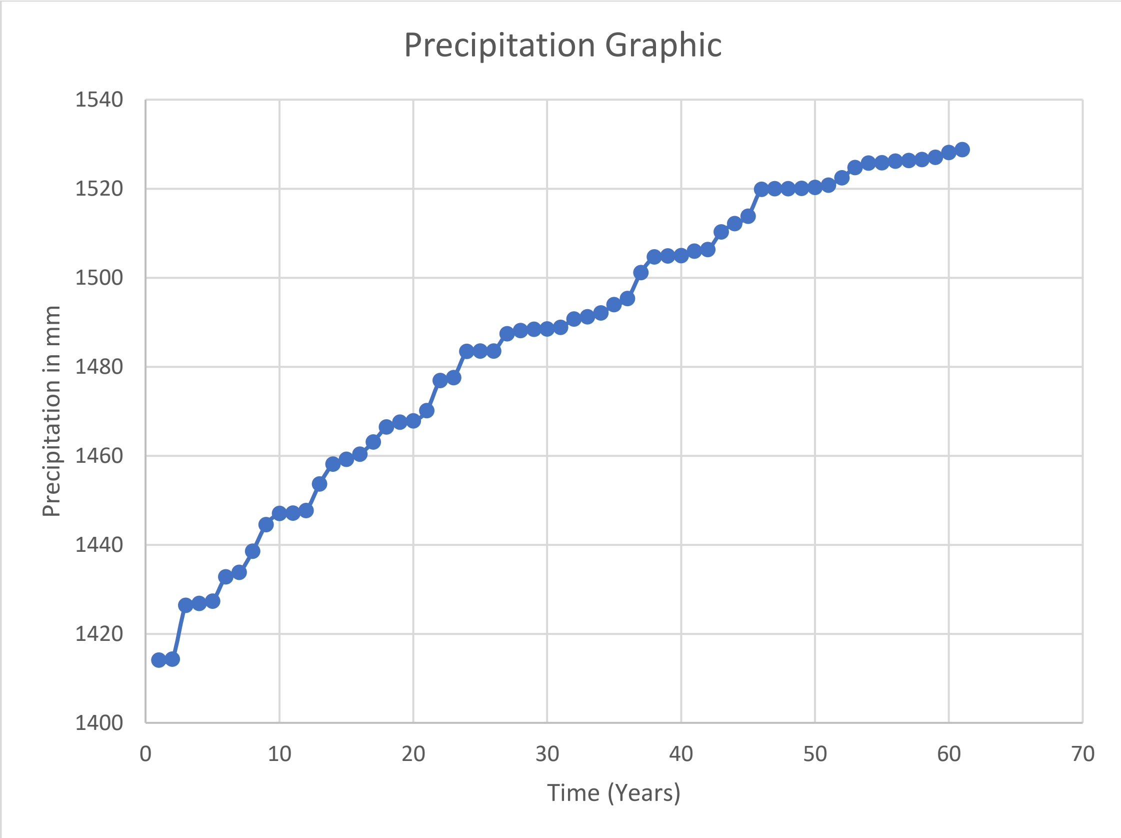

Figure 3.4. Precipitation Parameter Graphic

b. Temperature (Celsius degree)

The second parameter taken into account for the construction and the realization of the following research is the temperature. In all regions near the equator, in the city of Kinshasa also the annual climate is relatively stable and varies between 24-25 degrees Celsius. With a relatively low growth rate of approximately 0.05 degrees per decade. (World Bank Group, 2023).

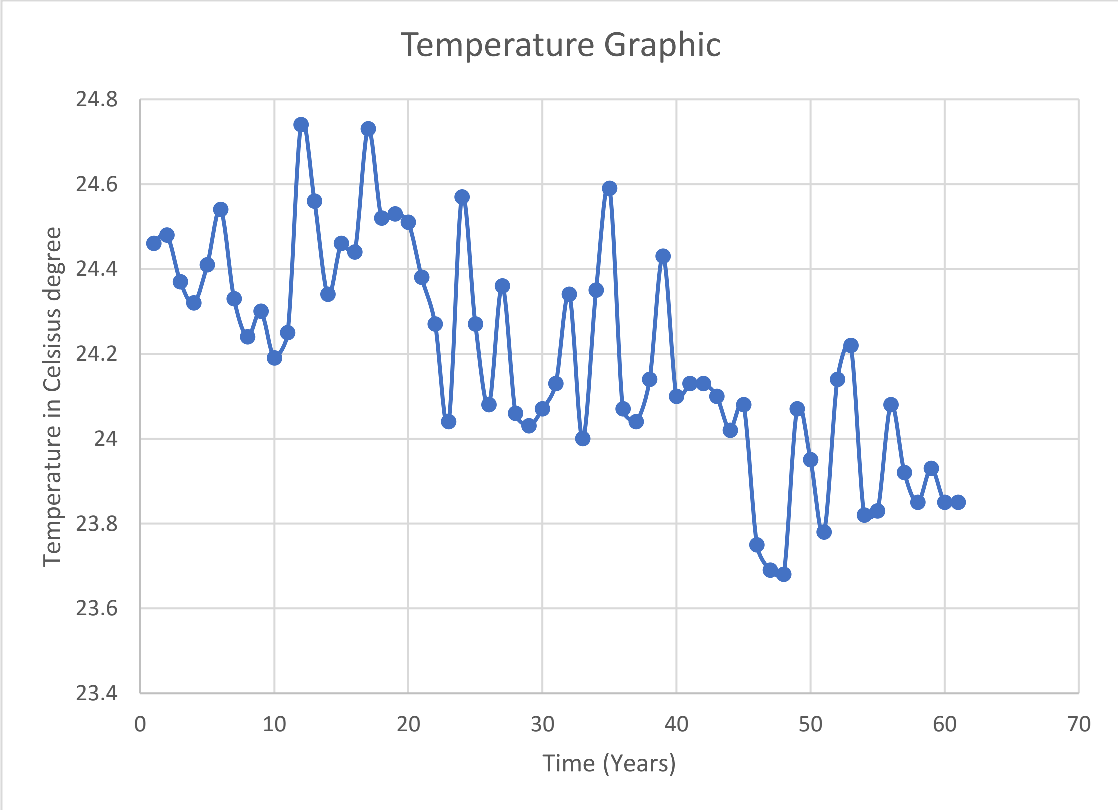

Figure 3.5. Temperature Parameter Graphic

c. Agricultural Land (Thousand Hectare)

The third parameter chosen for conducting this research is irrigated land. Because of the geographical location of Kinshasa, its equatorial climate but also its annual temperature which is generally warm from its immensity of arable land but also from the wealth that it offers in terms of its diversity in terms of resources on its ground and its basement. Sun et al. (2017) said in their research that the latter constitutes an important value for the conduct of investigation in the filed on water resources management.

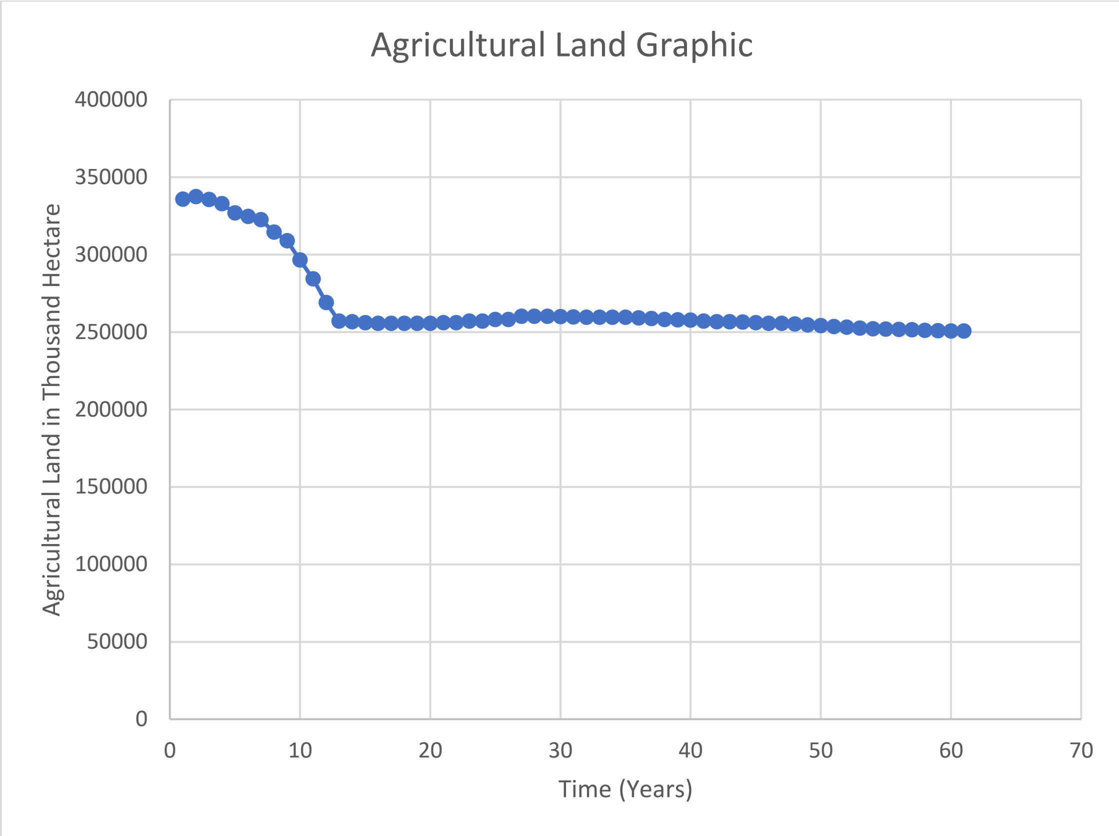

Figure 3.6. Agricultural Land Parameter Graphic

d. Gross Domestic Product per capita (Dollars)

Despite the various global challenges and the troubled political situation, Kinshasa, being the capital of the DRC, faces a rather flexible but also stable economy compared to other cities. The GDP Gross domestic product is estimated as the quantification of the market worth of all the final goods and services that cab be produced and released by a country over a given period of time. The GDP is a monetary computation.

Because of this importance of the GPD the choice was made on this parameter also for the selection and construction of the model.

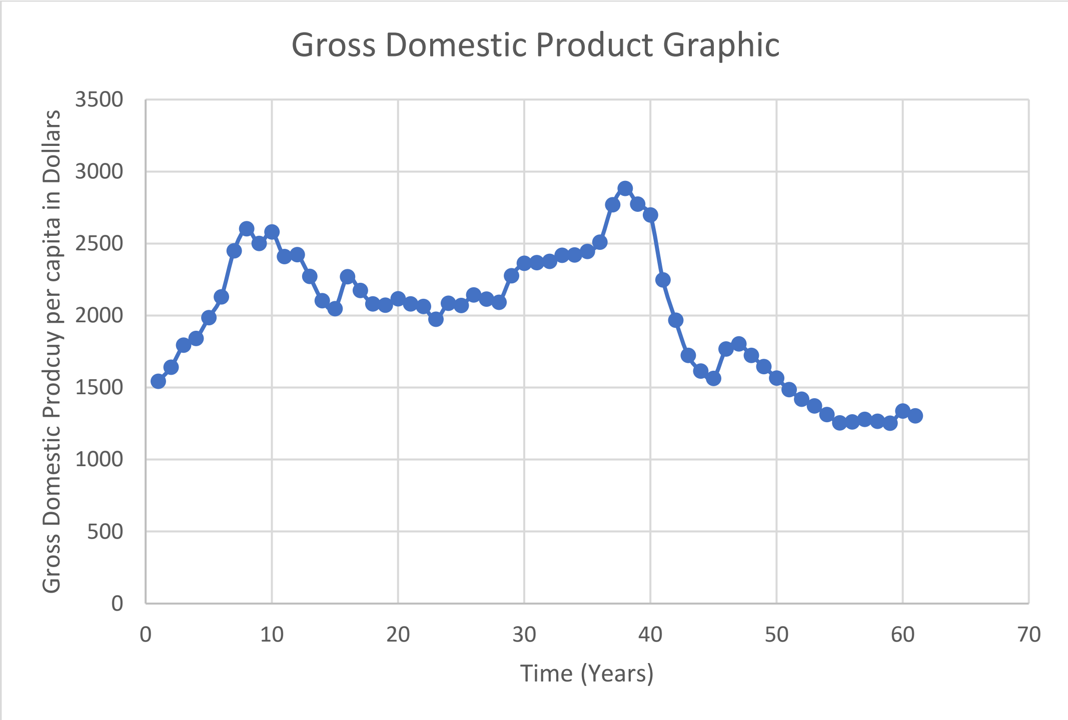

Figure 3.7. GDP Parameter Graphic

e. Population (millions of hab)

The last parameter leading to the relationship of this work is the annual population of the city province of Kinshasa. Kinshasa has the biggest population in Central Africa and is considered one of the fastest growing megacities in Africa.

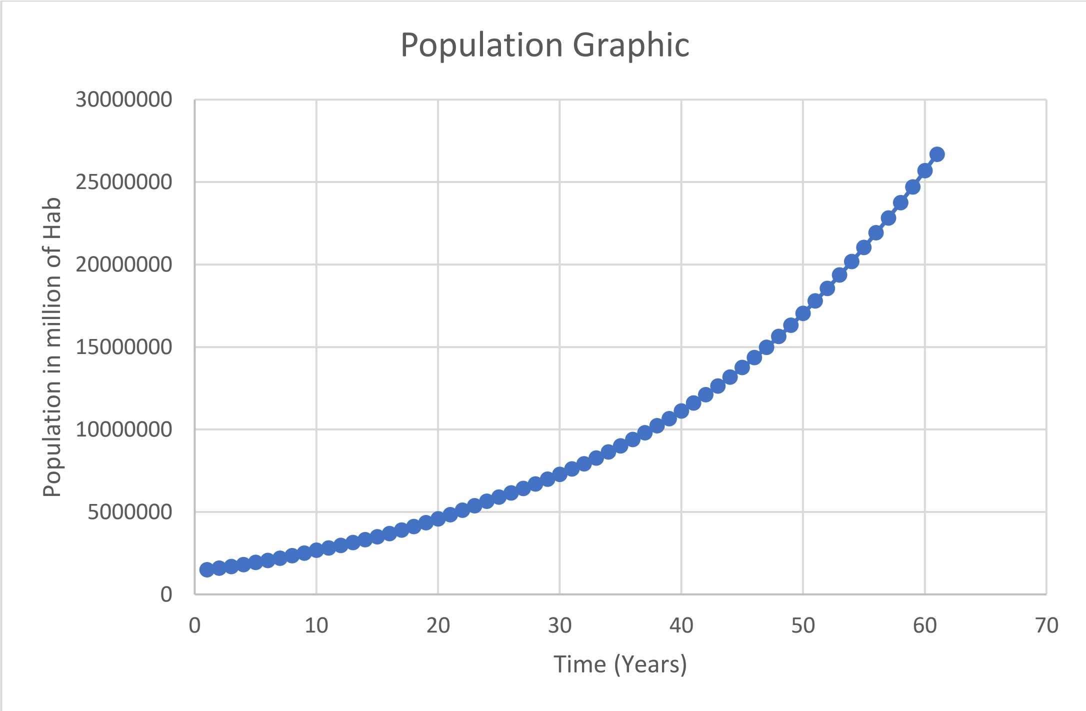

Figure 3.8. Population Parameter Graphic

3.4. Proposed Methodology

For the scenario taken into account for the writing of this thesis and the conduct of this research, the Soft Computing based models approach was used (ANFIS, ANN, MLR) to evaluate the management of water resources in the city province of Kinshasa. This was done following the next steps:

a. Input Data Selection and extraction of dominant predictors

After the selection of data, the latter were processed according to the relationship of the coefficients to eliminate the noise, evaluate the sensitivity of the data but also emerge with the most dominant parameters for the modelling but also the evaluation of performances.
Nourani et al. 2019 used the method of correlation coefficient (CC) to assess the relationship of data and determine the most influent predictor in term of linear data and non-linear data.

b. Select of single model

The construction of the different ANN, ANFIS and MLR models was carried out throughout this stage by considering the dominant predictors determined by the relation of the characteristic method coefficients extraction methods highlighted during the first step.

In modelling a method is applied to determine the relevance of predictors and predictands. To balance and assess the performance of AI based (ANN and ANFIS) models, MLR was preferred to other methods because of its ease of use but also its effectiveness in previous research. Given the scarcity of monthly time series data for the selected area of this study, the work was done with the model for which all data were reduced to the annual scale.

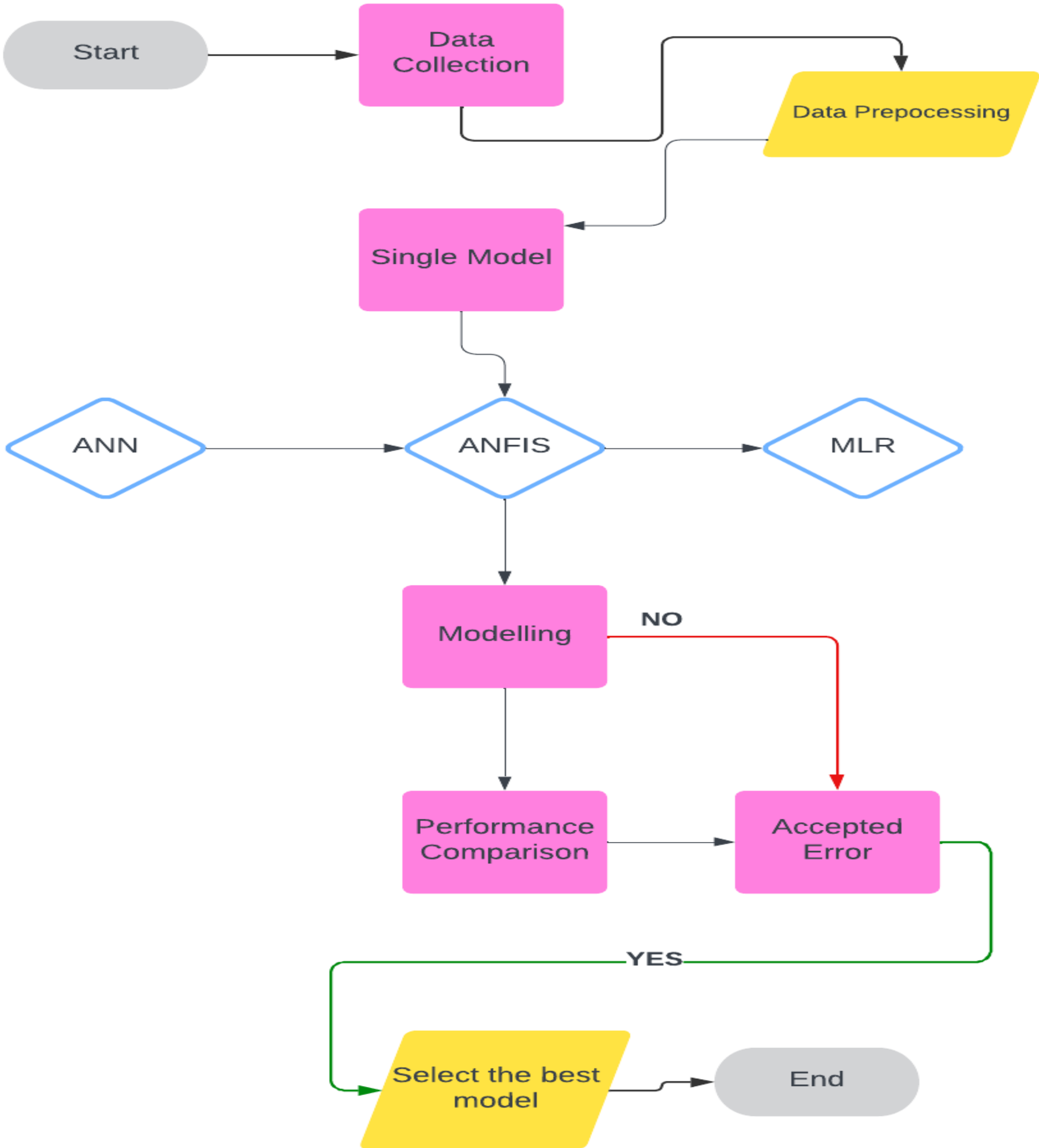

Figure 3.9. Schematic diagram for the proposed methodology

c. Performance comparison and selection of the best model

The performance of the different models in particular ANN, ANFIS and MLR were evaluated for the further selection of the model with the best performance and the least error to adopt it as for this case study.

3.5. Methodology for the scenario (ANN, ANFIS, MLR, CC)

3.5.1. Artificial Neural Network

ANN is a frequently used and widely used model and forecasting tool in the research field. Particularly in hydrology and in the field of water resources management. Researchers often use back propagation network (BP) models for artificial neural networks (ANN). Nourani et al (2011) demonstrated that every engineering problem can be predicted and simulated with the help of a three-layer back-propagation network model.

ANN is a model made up of neurons that are connected each to other by a certain number of simple process elements. The general characteristics of this model is very attractive because it includes a general capacity of learning algorithm, treatment of nonlinearity issues, noise tolerance etc, it has been shown in recent decades that ANN can handle many engineering issue. (Hornik et al., 1989).

Among all the neural networks methods spread all over the world, the most used is the FFNN model, which is fit out with a Back-Propagation algorithm.

Sahoo et al., (2015) investigated the on the training of neural network using Levenberg-Marquardt due to its ability to perform the convergence with rapidity. It is a hybrid technique which an optimum convergence of solution, it combined Gauss-Newton and deepest descent together.

The tangent sigmoid (Tansig) was considered as the function used for the transfer of the hidden layer and the output layer.

Among the different techniques spread throughout the world in the domain of artificial intelligence, the ANN is a model with convincing performances and has demonstrated its ability to process noisy, non-linear and dynamic even spatial data, in particular when the physical relationships under -underlying are not well comprehended. In artificial intelligence-based investigations, it is more frequent for all variables to have the same value when selecting and considering variables. Nourani et al. (2020), highlights this and shows it in their research.

The following diagram shows the building of an ANN model:

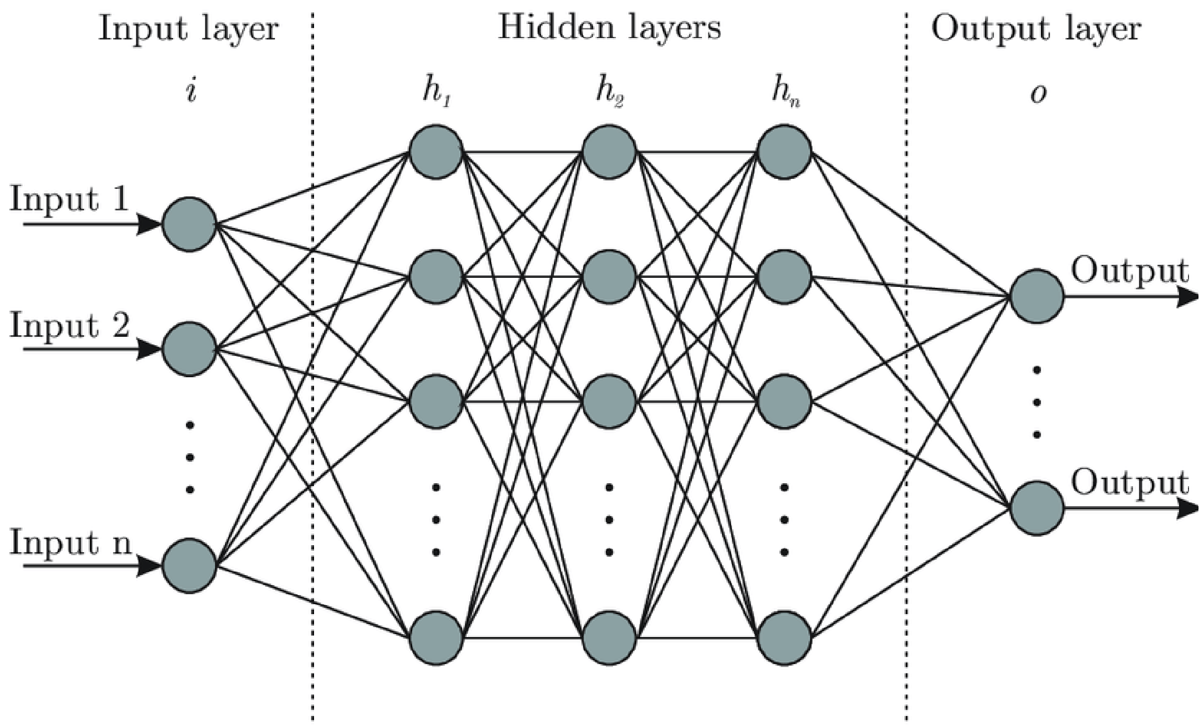

Figure 3.10. Composition of an ANNs model (Facundo Bre et al, 2018)

3.5.2. Adaptive Neuro Fuzzy Inference Systems

For the many AI-based research and studies all over the world, it is asked for researcher to put each variable at the same ground and give to them the same value of consideration. In the research domain of neural networks, he word ‘‘neurofuzzy’’ refers to the fuzzy logic modelling that called for a special learning algorithms to the fuzzy inference system (FIS). Jang presented in 1993 Adaptive Neuro-Fuzzy Inference System (ANIFS) it was considered as a unique and revisionary technique for the construction of the Fuzzy Neuro System. The model proposed and employed the training design and procedure of the neural network.

ANFIS is a Fuzzy Surgeno model incorporated to the adaptive systems framework to help with training, flexibility and strong matches. ANFIS implementation training becomes, with the support of this structure, less influenced by expert knowledge on and more systematic.

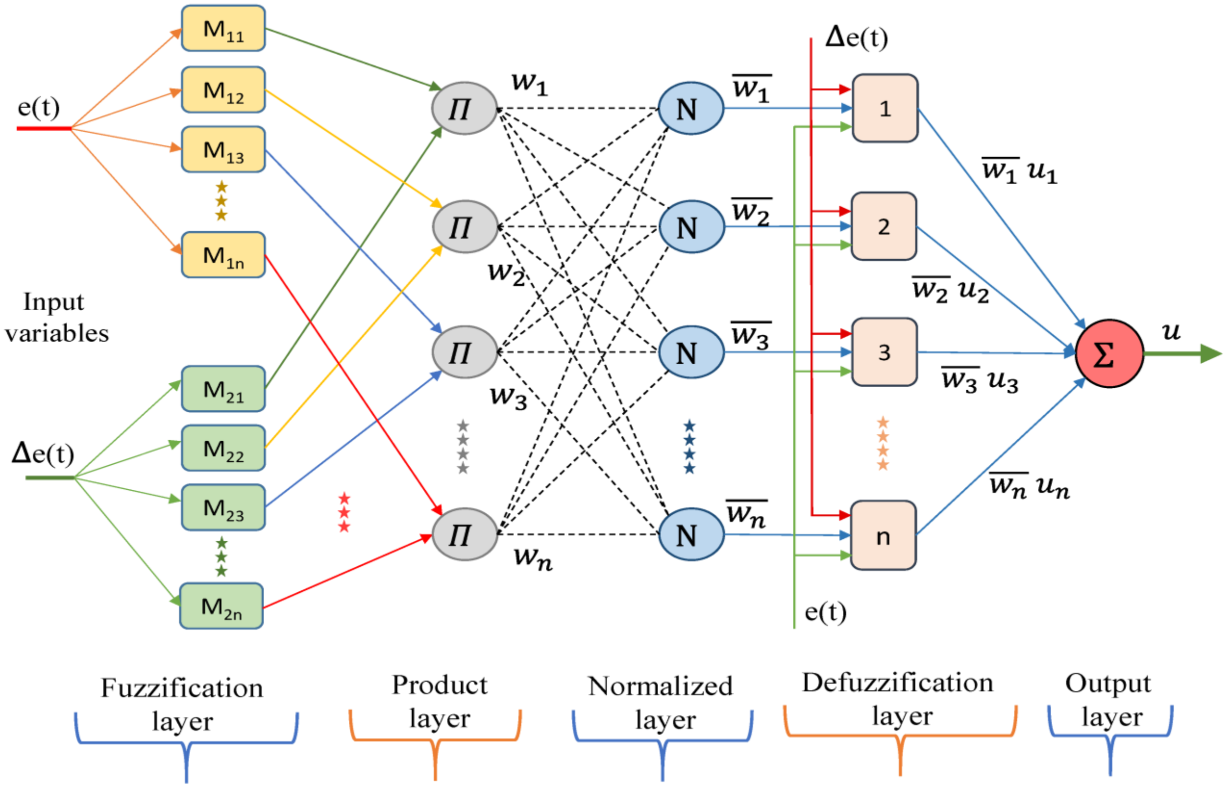

Figure 3.11. ANFIS model architecture (Ahmed F. Bendary et al, 2021)

The equations (3.4) and (3.5) represent and formulate the Fuzzy if-then rules, n is the representation of all the number of rules. Notice that for the fuzzification layer the membership function are represented by: M_{1i} and M_{2i}. For n_{th} rule all the linear components corresponding to the mentioned rule are expressed by An, Bn, and Dn.

$$D_n = if\ M_{1i}(e)\ and\ M_{2i}(\Delta e)\ \text{ then,} \tag{3.4}$$

$$f = a_n e(t) + b_n \Delta e(t) + d_n \tag{3.5}$$

For each node in the layer the representation of the adaptive function is then created by the following mathematical expression:

$$M_{1i} = \frac{1}{1 + [\frac{x - c_i}{p_i}]^{q_i}} \tag{3.6}$$

Noticed that for the second layer, it is the representation of the inference product layer, where fuzzy rules show how each node are strongly called with A. p_i, q_i and c_i represent respectively the setting parameters.

$$w_i = M_{1i}(e) * M_{2i}(\Delta e) \tag{3.7}$$

The following equation is the expression of the computed firing strength coming from the previous equation, and the expression is normalized.

$$\overline{Wi} = \frac{w_i}{\sum_i w_i} \tag{3.8}$$

After receiving the normalized results from layer 3 which is the normalized layer, layer 4 describes the defuzzification layer. It can be observed that each node in the mentioned layer indicates an adaptive mode. The equation (3.9) shows the representation of the defuzzification layer.

$$\overline{Wi}U = \overline{Wi}(a_i e + b_i \Delta e + d_i) \tag{3.9}$$

For gathering the outputs results in layer 5, the inward signals had to be computed and summed up. The expression U represents the selected control signal and (a, b, d) are the parameters set for the selected function.

$$\sum_i \overline{wi}U = \frac{\sum_i w_i U}{\sum_i w_i} \tag{3.10}$$

3.5.3. Multi Linear Regression

Multiple linear regression is a method that dates back to 1908 an refers to Pearson’s use, it is a tried-and-true method that is used to predict the variance in an interval dependency, it is typically based in the linear combination of independent variables and those variables can be presented in different ways (dummy, dichotomous or interval variables. The main role of this method of analysis is to seek out and find the link between a certain number of independent variables, predictors on a dependent or criteria variable.

MLR has been used widely over the past century to technique characterize quantitatively the linear relationship of two or more parameters (independent predictor parameters) and a predictand (dependent parameter). Nourani et al. presented typically in 2019 the relationship between the n and y as the predictors and dependent parameters.

The MLR can be expressed by the following:

$$y = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + b \tag{3.11}$$

The coefficient of the regression analysis is respectably a_1, a_2..., a_n, the describes how much the dependent variable y variate in response to a unit change in the corresponding independent parameters. When all the independent parameters are zero, the regression line will intercept the y-axis at a constant value, b which represents the number of the dependent variable y.

The following graphic represents the Multiple Linear regression scheme:

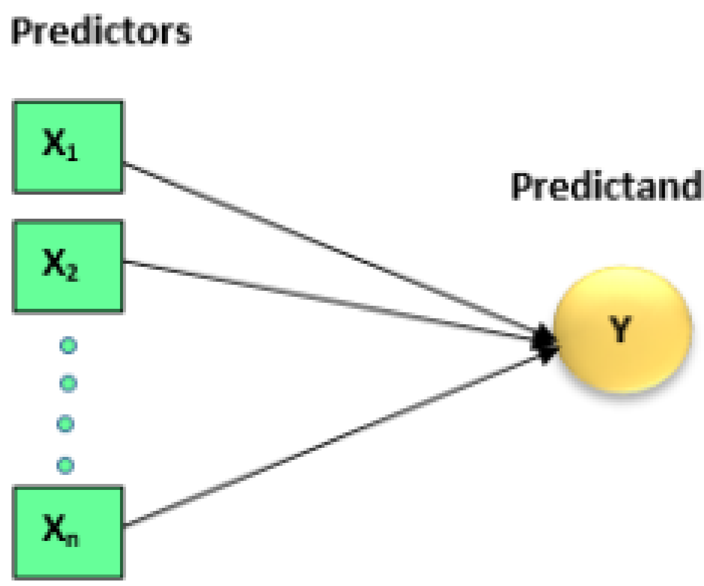

Figure 3.12. MLR model representation (Ogodor Elvis et al, 2020)

3.5.4. Correlation Coefficient CC

The correlation coefficient method is known as a metric method used to apprehend the linear relationship of variables (predictands and predictors). The values of CC range between [-1, +1]. When the values tend toward +1, it denotes the positive relationship between the data conversely when the data tend to -1 it denotes of a negative relationship between the data. Zero means the data have no relationship.
The following equation represents a CC relationship based on 2 variables:

$$CC = \frac{\sum(Z-\bar{Z}) - \sum(S-\bar{S})}{\sqrt{\sum(Z-\bar{Z})^2 \sum(S-\bar{S})^2}} \tag{3.12}$$

Where $Z,\ \bar{Z}$ represents the predictor value and the mean predictor value, $S,\ \bar{S}$ represent the predictand value and the mean predictand value, respectively.

CHAPTER IV

Findings and Discussion

The results of this research were presented in three stages following the methodological approach adopted for this research. The results will be presented for different models adopted some of the results will be given in form of figure, some others in form of tables and other will be shown as graphs according to the process followed.

Firstly, sensitivity analysis was done to determine the correlation of each variable and the dependence of these on each other, but also to determine the effect of each of them on water demand as well as the management of natural resources.

Secondly, the results of each model will be presented in various tables to present the different parameters adopted to evaluate the performance of each of them.

Thirdly, a comparison was made to show which of the methods was productive and which model showed the best performance.

The results for this research were presented as follows:

4.1. Sensitivity Analysis results

When modelling, the selection of the most dominated parameters is extremely important but also in the field of water resources management because of the range of data but also the volume of these data it is important to select the parameters to avoid noise in evaluation of model performance. Nourani et al. (2019) shows in their research that to achieve optimum performance in data driven, the selection of the dominant parameters and predictors must be evaluated according to GCM.

Because of this, a total number of 5 parameters were taken into account for the city province of Kinshasa, these data were generated by WBG and were submitted to the CC.

The following table presents the results of CC:

Table 4.1. Results of Correlation Coefficient Investigation

	Agricultural land Thousand Hectare	*Year-end Population*	*Gross Domestic product*	*Mean Average Temperature*	*Yearly Mean Precipitation*
Agricultural land Thousand Hectare	*1*				
Year-end Population	*0.85134295*	*1*			
Gross Domestic product	*0.14032181*	*0.34142591*	*1*		
Mean Average Temperature	*0.36133989*	*0.67100445*	*0.45041202*	*1*	
Yearly Mean Precipitation	*-0.80145877*	*-0.9911284*	*-0.4254877*	*-0.71815664*	*1*

As revealed in the previous table during sensitivity analysis, for the case of Kinshasa shows that for the different parameters taken into account the most influence parameters are Mean precipitation and the agricultural land respectively.

Simply, the reason these parameters are influential is because of their close connection with the other parameters.

When the average annual precipitation of an area experiences an increase, the irrigated area also experiences an increase because of the land which has abundant water and this also allows the inhabitants of the area to carry out their agricultural activity. The precipitation of a zone is also a function of the climate of the latter and therefore of the temperature of the zone which explains the negative influence between the two parameters, but also because of the population which grows according to the service in precipitation of an area which explains why people prefer to live in less arid areas because of the availability of water.

For the agricultural land or the irrigated area, when the value of the latter increases for an area, two parameters will also increase for this area, it is the Gross Domestic Product but also the population because the culture of the land influences not only the economy of an area therefore has its GDP index but also its demographic growth. Because of the influence of these two parameters in the recent studies (Zhao and Chen, 2014) but also in the study carried out for this research, the results show that these parameters are clearly linked to the others and for the good conduct of this project we will reject them, we are not going to consider them during the modelling as well as the management of the project to have more significant results.

The following table shows the representation of the data according to their influence, rank over the predictor.

Table 4.2. Results for the study area across CC

Water Resources Management			
Kinshasa	*Predictor*	*Rank*	*Validation*
	Mean Annual Precipitation	*1*	*Failed*
	Agricultural Land	*2*	*Failed*
	Yerly Population	*3*	*Pass*
	Mean Average Temperutre	*4*	*Pass*
	Gross Domestic Product	*5*	*Pass*

4.2 Results of the models (ANN, ANFIS, MLR)

The following section presents the results of water resources management modelling using the three methods (ANN, ANFIS and MLR). The results of the previous section shown (in tables 4.1. and 4.2.) reveal that only three parameters were taken into account for the modelling and evaluation of the performance of the models and we can be considered as the input parameters for our study while the surrounding parameter is the water demand of the zone. The validation method taken in code is n-fold methods. The data range was subdivided into 4-fold taking into account 3/4 for the calibration and 1/4 for the validation in other words practice 70% for the sorting of the model and 30% for the validation of this last.

4.2.1. Results for ANN

For ANN several modelling algorithms were taken into account to ensure and emerge with the most effective and the model with the best performance. The algorithm used for the training of the FFNN is Levenberg-Marquardt with a single hidden layer and a number of neurons which vary for each of the tests. In accordance with the various FEM tests, for the relaxed area of Kinshasa the satisfactory models with the best performance are respectively that of 7, 8, 10 and finally 6 respectively. To consider only one model, that of 7 turned out to be the best model based on its performance in term of the RMSE and the determination coefficient R-squared.

Table 4.3. ANN parameters for the model

Type of Network used	*FFNN*
Learning algorithm	*LM*
Training function	*TrainLM*
Number of input layers neurons	*3*
Number of hidden layer neurons	*6*
Number of Output layer neurons	*1*
Transfer function type	*TANSIG*
Maximum epochs	*1000*
Numbers of iterations	*70*

Type of Network used	*FFNN*
Learning algorithm	*LM*
Training function	*TrainLM*
Number of input layers neurons	*3*
Number of hidden layer neurons	*7*
Number of Output layer neurons	*1*
Transfer function type	*TANSIG*

Maximum epochs	*1000*
Numbers of iterations	*506*

Type of Network used	*FFNN*
Learning algorithm	*LM*
Training function	*TrainLM*
Number of input layers neurons	*3*
Number of hidden layer neurons	*8*
Number of Output layer neurons	*1*
Transfer function type	*TANSIG*
Maximum epochs	*1000*
Numbers of iterations	*478*

Type of Network used	*FFNN*
Learning algorithm	*LM*
Training function	*TrainLM*
Number of input layers neurons	*3*
Number of hidden layer neurons	*9*
Number of Output layer neurons	*1*
Transfer function type	*TANSIG*
Maximum epochs	*1000*
Numbers of iterations	*777*

Type of Network used	*FFNN*
Learning algorithm	*LM*
Training function	*TrainLM*
Number of input layers neurons	*3*

Number of hidden layer neurons	*10*
Number of Output layer neurons	*1*
Transfer function type	*TANSIG*
Maximum epochs	*1000*
Numbers of iterations	*504*

As represented in the follow figure, the best architecture built for the ANN model using FFNN training, LM algorithm and TANSIG as the transferring membership function, it has been noticed that the best and optimal architecture is a (3-7-1) which simply means that the architecture had 3 inputs parameters- 7 hidden layers and 1 output

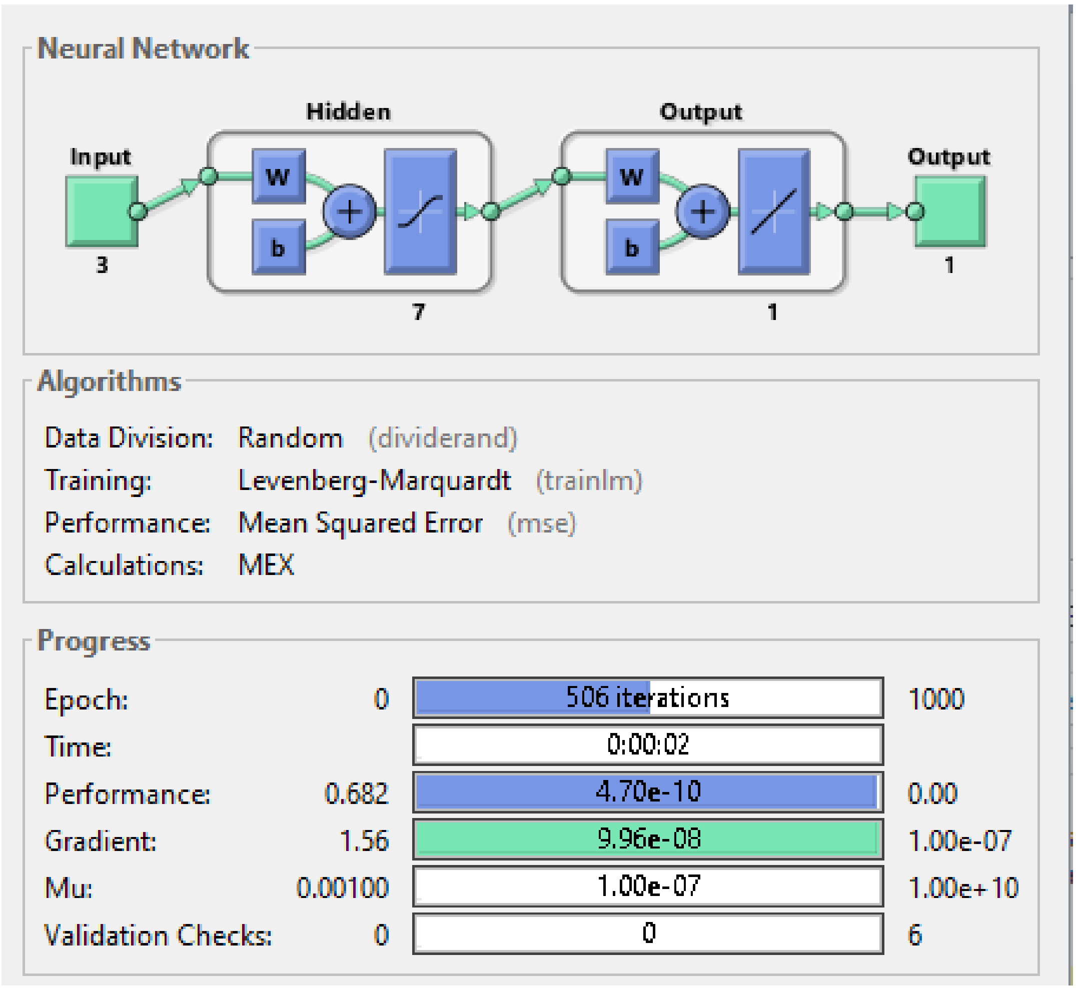

Figure 4.1. Opitimal ANN model built

4.2.2. Results for ANFIS

The ANFIS model implemented for this research study used a Surgeno Fuzzy Inference Algorithm, the structure formulation for the model was the TA and the learning algorithm used was a Hybrid model, for each of the above setting the tolerance calibration was set to zero and the generated FIS chosen was GP. For the set parameters different membership function type were trained to come up with the best results. As described in the following table the membership function chose for the research were respectively the Triangular membership function, the Trapezoidal membership function, the Bell membership function and the Gaussian membership function.

Table 4.4. ANFIS parameters for the model.

Learning algorithm	*Hybrid*
Fuzzy Inference Algorithm	*Surgero*
Structure formulation	*TA*
Membership function type	*Triangular*
Generated FIS	*GP*
Number of membership functions	*3*
Output type	*Constant*
Error tolerance	*0*
Epochs	*9*

Learning algorithm	*Hybrid*
Fuzzy Inference Algorithm	*Surgero*
Structure formulation	*TA*
Membership function type	*Trapezoidal*
Generated FIS	*GP*
Number of membership functions	*3*
Output type	*Constant*

Error tolerance	*0*
Epochs	*9*

Learning algorithm	*Hybrid*
Fuzzy Inference Algorithm	*Surgero*
Structure formulation	*TA*
Membership function type	*Gaussian*
Generated FIS	*GP*
Number of membership functions	*3*
Output type	*Constant*
Error tolerance	*0*
Epochs	*9*

Learning algorithm	*Hybrid*
Fuzzy Inference Algorithm	*Surgero*
Structure formulation	*TA*
Membership function type	*Bell*
Generated FIS	*GP*
Number of membership functions	*3*
Output type	*Constant*
Error tolerance	*0*
Epochs	*9*

TA and GP represent the trial error and grid partitioning respectively.

The training revealed that for the selected membership function implemented for the ANFIS model, the Triangular membership function showed the best result with the less RMSE and the best determination coefficient among the selected ones. The following image presents the structure of the ANFIS model

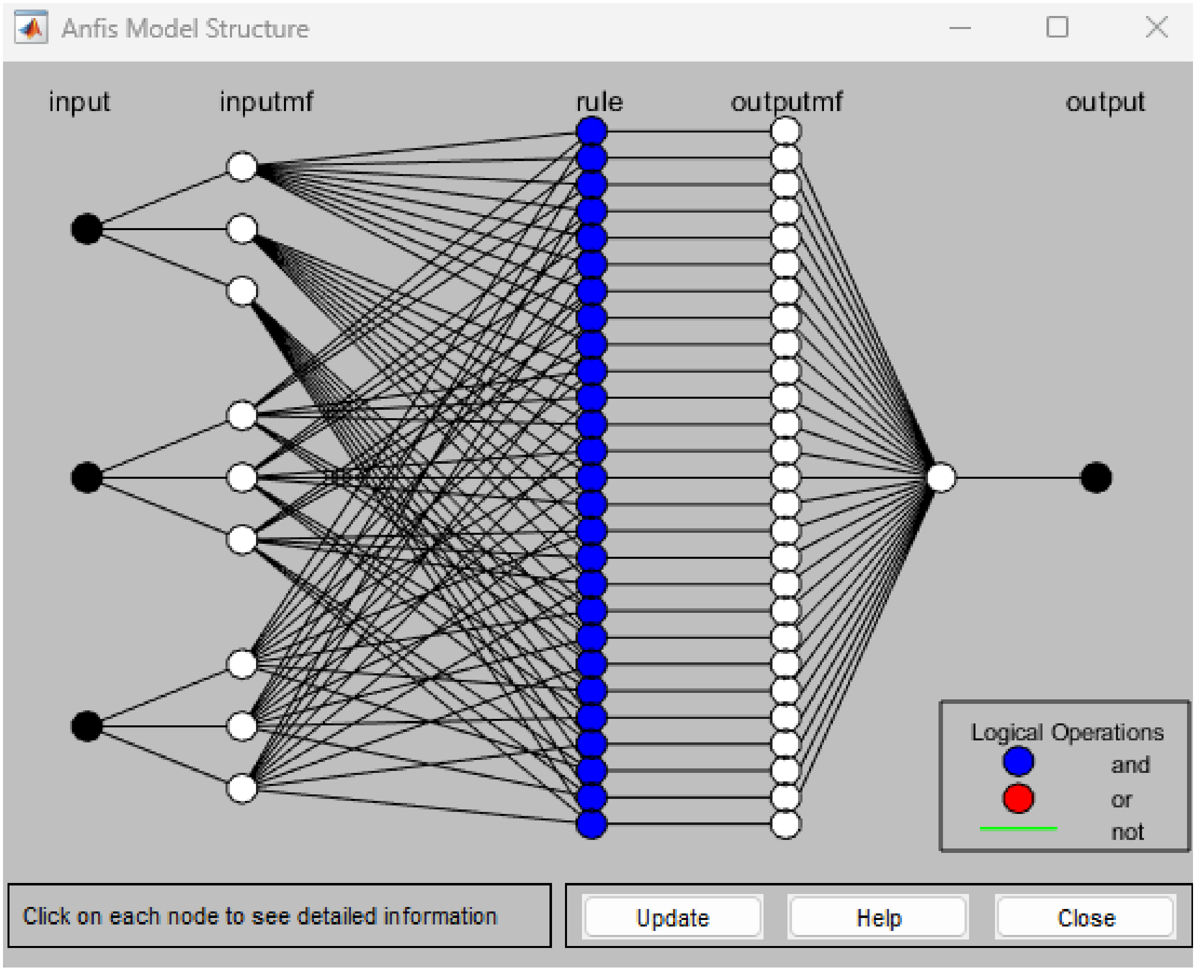

Figure 4.2. Optimal ANFIS model built

Through the review of the existing literature for the construction of this work, to evaluate the performances of ANFIS and ANN, MLR was performed to evaluate the performances of two models but also to study the correlation between the parameters which constituted the inputs but also the one that constitutes the output.

4.2.3. Results for MLR

The MLR model constructed investigated the relation between the predicted value and the observed value. Different Trees algorithms were assessed to come up with the best algorithm and architecture for this research study. The MLR structure used respectively the Fine Tree, Medium Tree and Coarse Tree algorithm different metrics were given and the RMSE and R-squared were retained for the evaluation of the permanence model. The feature selection for the construction of the MLR and the Principal Component Analysis PCA were drawn in accordance with the methodology proposed through the literature review and researches conducted. (Kasongo and Elkiran, 2023).

The following picture shows the amplified model with the optimum results given by MLR model.

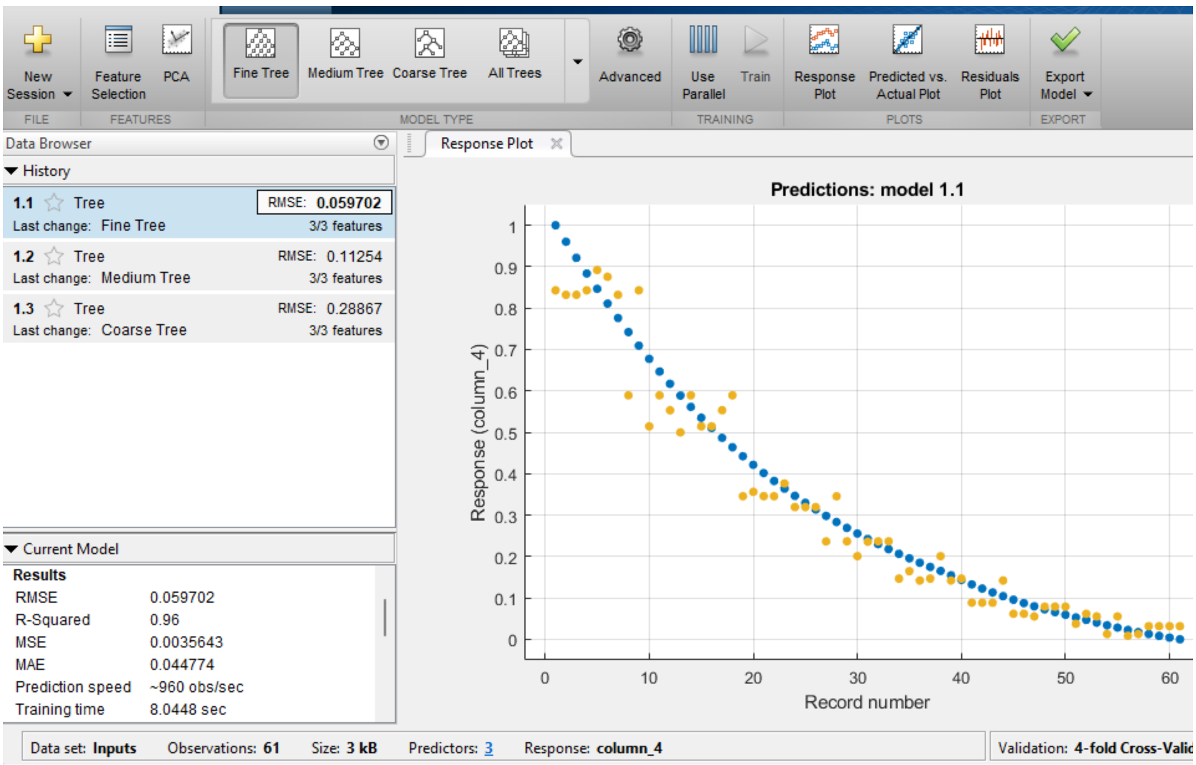

Figure 4.3. MLR model built

4.4. Discussions

As we can see in the following table, the performance of the models was evaluated taking into account two metrics, R-square and RMSE. For each of the models, 3 parameters were retained as input parameters after having carried out a sensitivity analysis of the data and retained the parameters decided and not having a strong link and correlation between them. The table 4.5. presents the training results but also the validation to emerge with better light on the performance of the constructed models, the table specifies for the case of study of Kinshasa Capital city, the models selected and for each model it clearly explained that the parameters selected through the sensitivity analysis were adequate for the construction of the three models. The selected metrics shows the differences in term of RMSE and R-squared for the model selected according to the input parameters. The goodness of the results are based mainly on the ability of each model to fit the requirements and the needs set through the methodology and fit the hypothesis that were made while preparing the work but also in the ability of each model to give us good fit of the results with the least mean error between the data for the selected time period data but also for the parameters chose knowing that the data taken into account were drawn on the annual time series instated of the monthly time series, the instance of all the model were therefore optimized using the normalization and coefficient corelation. Knowing that the implementation and modelling using MATLAB software and AI-based model is process that need attention and ability of the modeler while treating the data to avoid mistakes and errors because a simple difference in the data can lead to great change in the outputs results. Consequently, naming the above reasons and referring to the methodology used the results presents for this work according to the range of data and parameters selected can be trustable and subjected to different discussions and comments, the results of some algorithm implementation and model simulation can be optimized such as MLR, the instance selected for the validation and the calibration were insufficient to create a more precise model.

The results presented in table 4.5. are only for the best structure constructed in the implementation of the three models (ANN, ANFSI and MLR).

As seen in the table the RMSE is unitless because the inputs data is normalized and the determination coefficient is given in percentage of fit.

Table 4.5. Results for water resources management

				Training		*Validation*	
Station	*Screenin Method*	*Model*	*Inputs*	R^2	$RMSE^*$	R^2	$RMSE^*$
Kinshasa	*CC*	*ANFIS*	*Gross Domestic Product, Mean Average Temperature, Year-End Population*	*99*	*0.000054*	*92*	*0.000058*
	CC	*ANN*	*Gross Domestic Product, Mean Average Temperature, Year-End Population*	*99*	*0.000019*	*99*	*0.012080*
	CC	*MLR*	*Gross Domestic Product, Mean Average Temperature, Year-End Population*	*91*	*0.080670*	*54*	*0.020799*

The results of the modelling in MATLAB will be discussed for both the training and the validation process according to each model built for the construction of this research. With respect to the sensitivity analysis outperformed we can notice that for Kinshasa region the AI-based model has shown the best results compared to the MLR model. For the training process ANN gave the best fitting results with a determination coefficient of 99% a RMSE of 0.000019 and ANFIS followed with a determination coefficient of 99% and a RMSE of 0.000054. Lastly, we have the MLR with respectively 91% and 0.080670 for the determination coefficient and the RMSE. The significance and meaning of these results are because of the ability of AI-based model both ANN and ANFIS to learn efficiently, reproduce quickly and predict with a high degree of accuracy the response in the domain of hydro informatics and hydraulics and water resources engineering. (Kasongo and Elkiran, 2023) conducted similar research and the result presented in that work highlighted the accuracy of AI-based model compared to MLR in the domain of water resources management and hydraulics engineering.

Ogodor and Elkiran noticed in 2020 in their research that for different station in Cyprus using the same inputs parameters the AI-based model satisfied the needs of the results and gave best performance compared to MLR. It goes without saying that the results outperformed by the model especially ANN and ANFIS for Kinshasa station in DRC will similarly give high precisions and accuracy.

It has been noticed that for the validation process the results dropped significantly for MLR with the determination coefficient of 54% and a RMSE of 0.020799, while ANN gave a determination coefficient of 99% and a decreased has been noticed for the RMSE 0.012080 respectively and the ANFIS model showed a decreased in term of the determination coefficient and the RMSE with the respectively value of 92% and 0.000058. The reasons of that decreased is because the set of data selected randomly for the validation had only 30% of the overall dataset, the significance on AI-based does not only rely on the quality of data but only the quantity, as the quantity of data dropped down for the validation process the performances for the three models dropped as well especially in term of the error measured by the RMSE regarding the actual data versus the predicted.

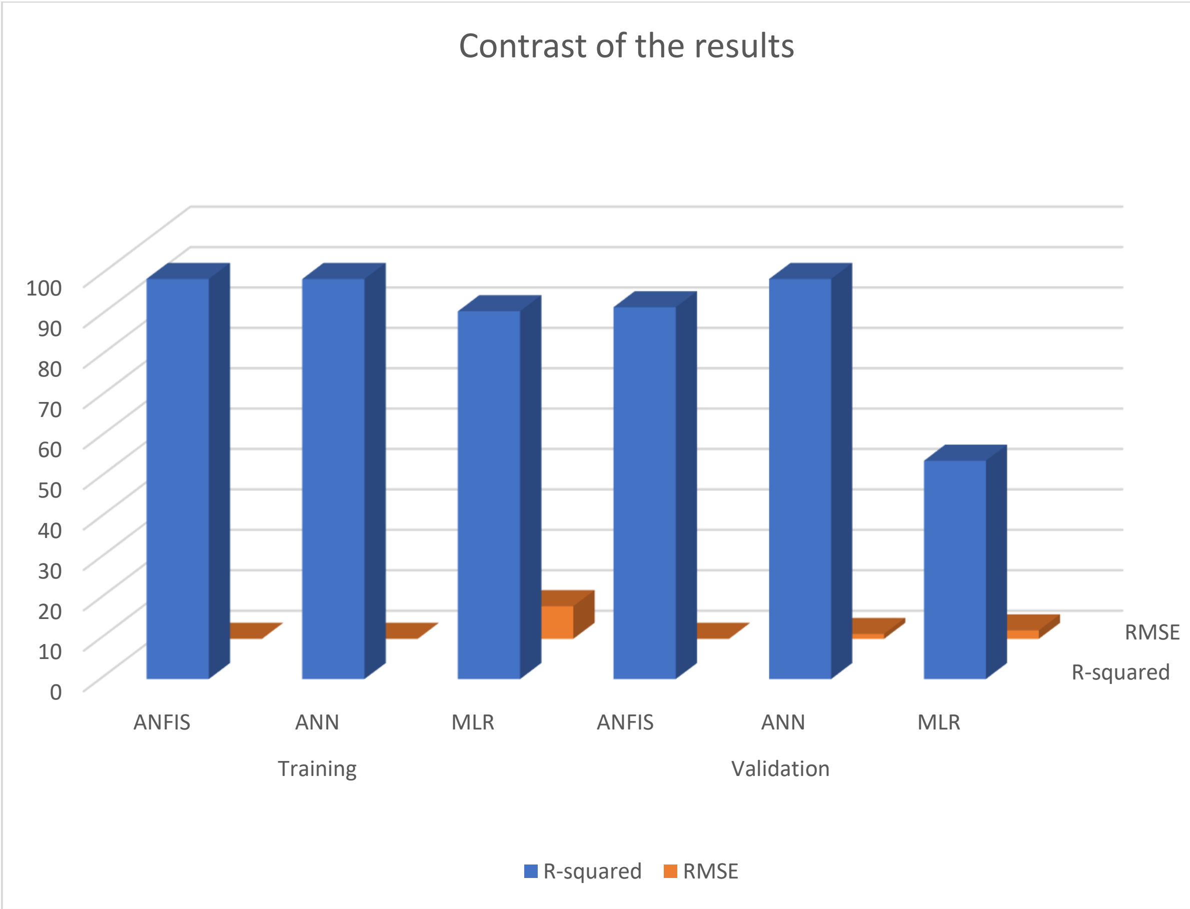

Figure 4.4. Contrast of the modelling results

The Figure 4.4. represents a contrast in terms of performances for the model built in the training process and the validation process. As shown in the diagram ANN outperformed the best metrics results in the first step which is the training of the model and the second one which is the validation. ANFIS is the second model to give satisfactory results with a decreased of 8% in term of the R-squared for the validation process then MLR presents a dropped of 45% compared to ANN and 38% compared to ANFIS. As seen in the literature review and previous work conducted in the domain of water resources management and hydraulic engineering, AI-based model outperformed the best results compared to the conventional MLR methods because of their success in the work they are strongly recommended for the application in Water resources management and water engineering.

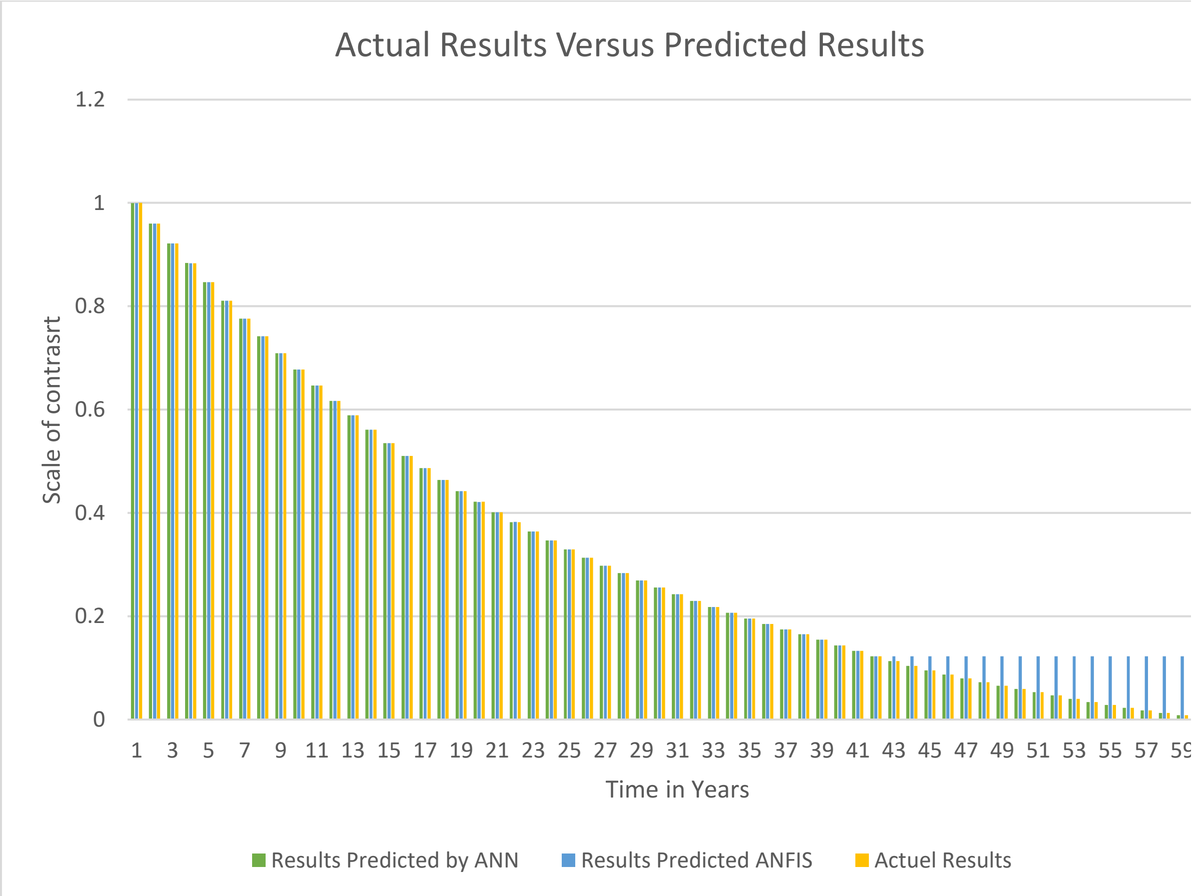

Figure 4.5. Actual results versus Predicted results

As shown in Figure 4.5. the current results and the predicted results do not show great results as well as a big difference because the performances obtained by the AI-based models are well above expectations, the ANN model shows the best results and can perhaps be used for projections as well as future research in the field and for the Kinshasa area as well as other study areas in the country.

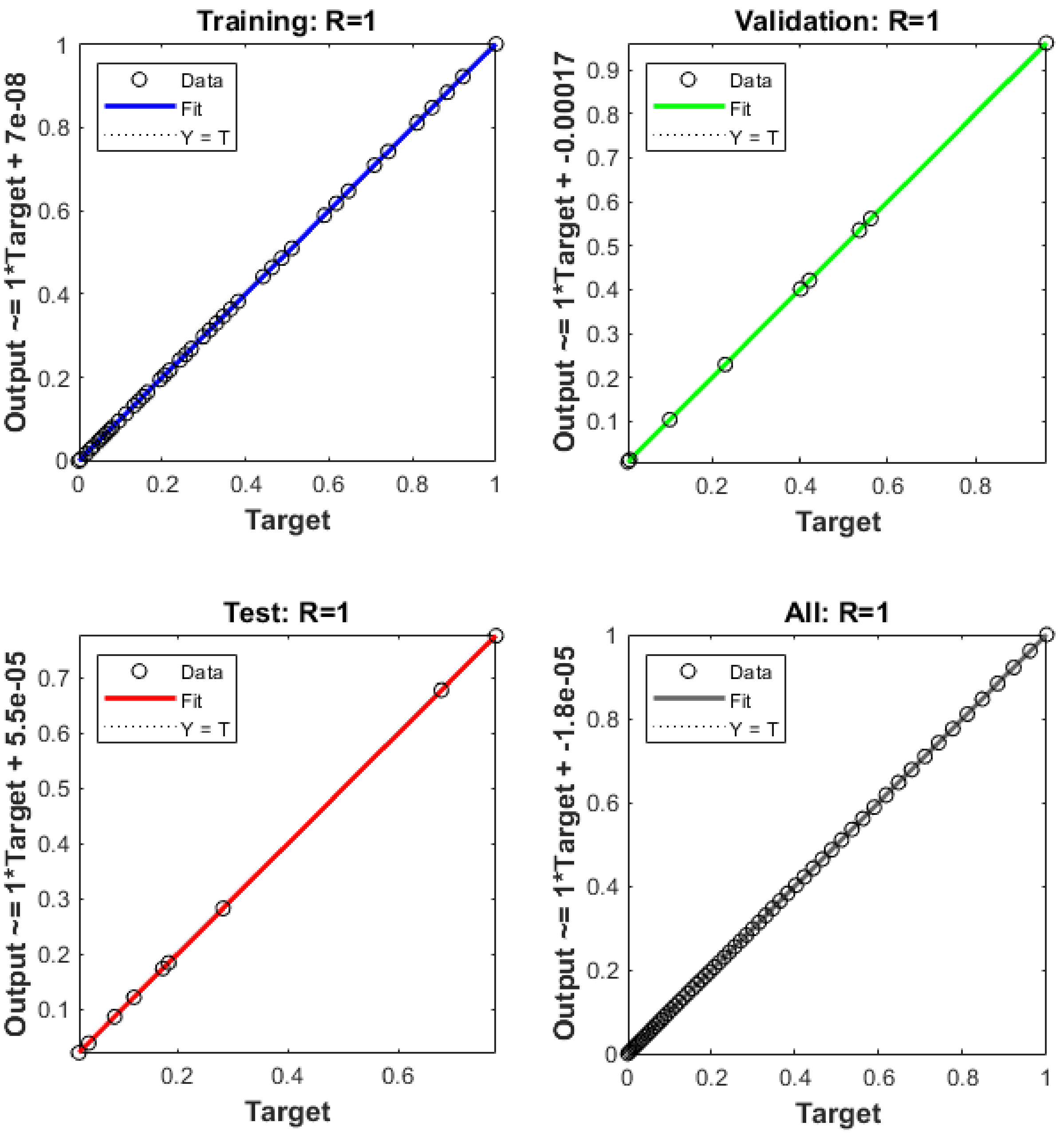

Figure 4.6. Result for the best model selected

CHAPTER V

Conclusion and Recommendations

5.1. Conclusion

For this research paper, Kinshasa City in Democratic Republic of the Congo was used as a case of study for water resources management especially water demand management, the model was developed and evaluated covering a period of time from 1961-2021. For the mentioned study area, a total of 5 parameters from WBG based on the data available in DRC and the resources provided by the ministry of plan and development were used to build up the model for the evaluation of water resources and its management through three Soft Computing Methods ANFIS, ANN and MLR models. By taking into account the different parameters that can influence water demand as well as the management of natural and water resources in an area, certain variables seem to be more important and more effective than others for this case study. To ensure the effectiveness of this research and obtain more satisfactory results, a sensitivity analysis was performed using Correlation Coefficient extraction methods to determine the relationship of data, linearity or nonlinearity between the predictors and the predictands the parameters that were related to others were dropped and only 3 parameters gave and satisfactory results and pass the test. These 3 parameters were used as inputs for the construction of the different models.

The investigation carried out under the sensitivity analysis shows that the selection of parameters plays a role in obtaining performance in the field of water resources management. The more the data are linked together, the less the model has good data performance, especially for models based on the linearity of the data. Based on the correlation coefficient, the GDP, the Mean Annual Temperature and the year-end population give the selection design for the model. The discussion of the results demonstrated that the ANN model in terms of its performance turned out to be the best model with 99% for R-squared and 0.000019 for RMSE, the ANFIS model came second with 99% and 0.000054 as determination coefficient and RMSE and the last model turned out to be MLR with 91% and 0.080670 these results undoubtedly show the precision with which AI-based models can implement and model water resources management and explain the major interest that researchers focus on this area.

5.2. Recommendations

Through the different results obtained in this work, there are several observations that have been made and which show the limits of this work but also present a list of recommendations for the improvement of this type of study and constitute an opportunity for those who are interested in launching into the same theme and tackling this subject:

1. For this study 5 parameters were taken into account (Mean average temperature, population, precipitation, agricultural land and gross domestic product) by considering different scenarios (socio economic development and climate impact change) combined together. Certain references clearly show that other parameters can be taken into account to improve the effectiveness of this work. (Quing Shuang and Rui Ting Zhao, 2021).
2. The Democratic Republic of Congo and particularly Kinshasa face challenges both on an ecological level with the different climatic challenges but also concerns in terms of economic development, for others subject to crises, conflicts and wars it is also advisable to take into account causes several factors for the improvement of this work and this particularly focuses on the scenario which most impacts the study area such as the human factor for crisis zones.
3. The work recently carried out in the country in the field of water shows a certain interest in Soft Computing Methods. It is therefore advisable to focus on other methods such as Support Vector Machine SVM, AdaBoost, Decision Tree (DT), Random Forest (RF) etc.
4. In this thesis, annual data were used for model construction because some parameters could not be reduced to the monthly scale for the selected data range. In accordance with the technical literature as well as various research, monthly data can be used such as precipitation, average temperature and others to obtain different results and make a comparison between the two approaches.
5. To obtain more promising results, this study can be carried out subsequently taking into account different configurations such as the laws and regulations of the country in the field of water but also the management policy and also the configuration of certain water works, water storage such as reservoirs and others, river polices and dam construction for the resources available.

This work is limited to the city province of Kinshasa in the Democratic Republic of Congo, there are several areas in this country which are under the influence of an increased lack of access to drinking water as well as certain areas which are cause several problems in the management of hydraulic resources because of the different reasons mentioned above in the work. It is therefore important and recommended for other regions also to make predictions through Soft Computing Methods for the best management of water resources.

REFERENCES

Afshin Shabani, Xiaodong Zhang, Mike Ell. (2017). Modelling Water Quantity and Sulfate Concentration in the Devis Lake Watershed Using Coupled SWAT and CE-QUAL-W2. *Journal of American Water Resources Association (JAWRA)*. https://doi.org/10.1111/1752-1688.12535

Ala T. Mohsin, Gözen Elkiran. (2018). Multi-Region Artificial Intelligence Based Ensemble Modelling of Daily Global Solar Radiation. Near East University.

Alireza B. Dariane, Leila Eamen. (2017). Finding the Causes and Evaluation Their Impacts on Urmia Lake Crisis Using a Comprehensive Water Resources Simulation Model. *Journal of Hydraulic Structures.*

Baolin Xue, Hanwen Zhang, Yuntao Wang, Zhongxin Tan, Yi Zhu. (2021). Modelling Water Quantity and Quality for A Typical Agricultural Plain Basin of Northern China by A Coupled Model. Science of the Total Environment. https://doi.org/10.1016/j.scitotenv.2021.148139

Bentolhoda Asl-Rousta, S. Jamshid Mousavi, Majid Ehtiat and Mehdi Ahmadi. (2018). SWAT-Based Hydrological Modelling Using Model Selection Criteria. Water resources management. https://doi.org/10.1007/s11269-018-1925-5

Chen, W.C., Chang, N.B., Shieh, W.K., 2001. Advanced hybrid fuzzy-neural controller for industrial wastewater treatment. Journal of Environmental engineering 127 (11), 1048–1059.

D. P. Vijayalaskimi, K. S. Jinesh Badu. (2015). Water Supply System Demand Forecasting Using Adaptive Neuro-Fuzzy Inference System. *International conference on water resources, coastal and ocean engineering* (ICWRCOE 2015).

Edward Naabil, Benjamin Lantei Lamptey, Kouakou Kouadio, Thompson Annor. (2020). Climate Change Impact on Water Resources; The Case of Tono Irrigation Dam in Ghana. Hydrology, 8(4), 69-78. https://doi.org/10.11648/j.hyd.20200804.12

Elahe Kalashak. (2021). Prediction of Water Consumption Using Machine Learning. Ostfold University College.

Elana Pacchin, Stephano Alvini, Marco Franchini. (2017). A Short-Term Water Demand Forecasting Model Using A Motivated Window On Previously-Observed Data. *Journal of Water 2017, 9, 172*. doi:10.3390/w9030172.

Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G. (2007). The soil and water assessment tool: historical development, applications, and future research directions. *Trans.ASABE 50,1211-1250.*

Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. Neural Networks, 2(5), 359-366.

Ishmael S. Msiza, Fulufhelo V. Nelwamondo, Tshilindzi Marwala. (2008). Water Demand Prediction Using Artificial Neural Networks and Support Vector Regression**.** *Journal of computers, vol. 3, no. 11.*

Jaehak Jeong, Narayanan Kannan, Jeff Arnold, Roger Glick, Leila Gosselink, Raghavan Srinnivasan. (2010). Development and intregration of a sub-hourly rainfall-runoff moeling capability within a watershed model. *Water Resour Manage (2010) 24:4505–4527*. DOI 10.1007/s11269-010-9670-4.

Janusz A. Starzyk. (2015). Water Resource Planning and Management Using Motivated Machine Learning. School of Electrical Engineering and Computer Science, Ohio University, Athens, Ohio 45701, USA.

Jiping Yao, Guoqiang Wang, Weina Xue, Zhipeng Yao, Baolin Xue. (2018). Assessing the adaptability of water resources system in Shandong Province, Chine, using a novel comprehensive Co-evolution model. *Water Resources Management (2019) 33:657–675*. https://doi.org/10.1007/s11269-018-2129-8.

Jiping Yao, Puze Wang, Guoqiang Wang, Sangam Shrestha, Baolin Xue, Wenchao Sun. (2019). Establishing a time series trend structure model to mine potential hydrological information from hydrometerological time series data. *Science of total environment.* https://doi.org/10.1016/j.scitotenv.2019.134227.

Lihua Xionga, Asaad Y. Shamseldin, Kieran M. O'Connor. (2001). A non-linear Combination of the forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system Journal of Hydrology 245 (2001) 196-217.

Li M., Finlayson B., Webber M., Barnett J., Webber S., Rogers S., Chen Z., Wei T., Chen, J., Wu X., et al. (2017). Estimating urban water demand under conditions of rapid growth: The case of Shanghai. Reg. Environ. Chang. 2017, 17, 1153–161.

Li L., D.J. Gochis, S. Sobolowski, MD.S. Mesquita. (2017). Evaluating the present annual water budget of a Himalayan headwater river basin using a high-resolution atmosphere-hydrology model, J. Geophys. Res. Atmos., 122, 4786–4807, doi:10.1002/2016JD026279.

Lu H, Matthews J, Han S. (2020). A hybrid model for monthly water demand prediction: A case study of Austin, Texas. AWWA Wat Sci. 2020;e1175. https://doi.org/10.1002/aws2.1175

Mengyan Zhu, Jiawei Wang, Xia Yang, Yu Zhang, Linyu Zhang, Honqiang Ren, Bin Wu, Lin Ye. (2022). A Review Of The Application Of Machine Learning In Water Quality Evaluation. Eco-Environment & Health 1 (2022) 107–116. https://doi.org/10.1016/j.eehl.2022.06.001

Mngereza Miraji, Jie Lu, Chunmiao Zheng. (2019). The impacts of Water demand and its implication for future surface water resource management: the case of Tanzania's Wami Ruvvu Bassin. Water 2019, 11,1280 https://doi.org/10.3390/w11061280.

Mohammad Zounemat-Kermani, Mohammad Teshnehlab. (2007). Using adaptive neurro-fuzzy inference system for hydrological time series prediction. Journal of Applied Soft Computing 8 (2008) 928-936.

Muleta, M. K. & Nicklow, J. W. (2005) Decision Support for Watershed

Management Using Evolutionary Algorithms. *Water Resour. Plng. And Mgmt. vol. 131, no. 1, pp. 35-44.*

Msiza I. S., Nelwamondo, F. V. & Marwala, T. (2008). Water Demand Prediction Using Artificial Neural Networks and Support Vector Regression, Journal of Computers, vol. 3, no. 11.

Nourani Vahid, Ozgur Kisi, Mehdi Komasi. (2011). Two hybrid artificial intelligence appoaches for modelling rainfall runoff process. Journal of Hydrology 402(2011) 41-59. doi:10.1016/j.jhydrol.2011.03.002.

Nourani, V., Khanghah, T. R., and Sayyadi, M. (2013). Application of the Artificial Neural Network to monitor the quality of treated water. International Journal of Management and Information Technology, 3(1), 39-45.

Nourani, V. (2017). An Emotional ANN (EANN) approach to modelling rainfall-runoff process. Journal of Hydrology, 544, 267-277.

Nourani V., Elkiran G., & Abdullahi J. (2019). Multi-Station Artificial Intelligence Based Ensemble Modeling Of Reference Evapotranspiration Using Pan Evaporation Measurements. Journal of Hydrology, 577123958.

Nourani V., Razzaghzadeh Z., Baghanam A. H., & Molajou A. (2019). ANN- Based Statistical Downscaling Of Climatic Parameters Using Decision Tree Predictor Screening Method. Theoretical and Applied Climatology, 137(3-4), 1729-1746.

Nourani V., Elkiran G., & Abdullahi, J. (2020). Multi-step ahead modeling of reference evapotranspiration using a multi-model approach. Journal of Hydrology, 581, 124434.

Ogodor Elvis, Gözen Elkiran. (2020). Impact Of Climate Change On Hydro Climatological Parameters In North Cyprus: Application Of Artificial Intelligence Based Statistical Downscaling Models. Near East University.

Parisouj, P.; Mohebzadeh, H.; Lee, T. (2020). Employing Machine Learning Algorithms for Streamflow Prediction: A Case Study of Four River

Basins with Different Climatic Zones in the United States. Water Resour. Manag. 2020, 34, 4113– 4131.

Pesantez, J.; Berglund, E.Z.; Kaza, N. Smart meters data for modeling and forecasting water demand at the user-level. *Environ. Model. Softw. 2020, 125, 104633.*

PNUE. (2011) Problématique de l'Eau en République Démocratique Congo. Janvier,2011 : pp27-55. (Report).

P. C. Nayak, K. P. Sudheer, D. M. Rangan, K. S. Ramasastri. (2004). A Neurro-Fuzzy Computing Technique for Modelling Hydrological Times Series. Journal of Hydrology 291(2004)52-66. doi:10.1016/j.jhydrol.2003.12.010.

Quing Shuang, Rui Ting Zhao. (2021). Water Demand Prediction Using Machine Learning Methods: A Case Study of the Beijing–Tianjin–Hebei Region in China *Water* 2021, *13*(3), 310; https://doi.org/10.3390/w13030310

Romain Kasongo, Bovic Kilundu. (2020). Design and Sizing of a Water Supply Network: Case of Sudy JOLI SITE. University Nouveaux Horizons.

Romain Kasongo, Gözen Elkiran. (2023). Evaluation of Water Demand of a City Applying Machine Learning Methods. *Journal of Optoelectronincs Laser, volume 42*. doi:10050086.2022.12.06v.

Rozos E., Butler D., Makropoulos C. (2016). An Integrated System Dynamics-Cellular Automata Model For Distributed Water infrastructure Planning. *Water Sci. Tech. Water Supply 2016, 16, 1519–1527.*

Runzi Wang, Jun-Hyun Kim, Ming-Han Li. (2020). Predicting Stream Water Quality Under Different Urban Devolopment Pattern Scenarios With Am Interpretable Machine Learning Approch. https://doi.org/10.1016/j.scitotenv.2020.144057.

R.Z. Wang, J.H. Kim, M.H. Li. (2021). Predicting stream water quality under Different urban development pattern scenarios with an interpretable Machine learning approach, Sci. Total Environ. 761, 144057, https://doi.org/10.1016/j.scitotenv.2020.144057.

Sani I. Abba, Gözen Elkiran. (2019). Wastewater Treatment Plant Performance Analysis Using Artificial Intelligence- An Ensemble Approach. Near East University.

Sahoo, G. B., Ray, C., & Wade, H. F. (2005). Pesticide prediction in ground water in North Carolina domestic wells using artificial neural networks. Ecological Modelling, 183(1), 29- 46.

SOLIDARITES INTERNATIONAL. (2015) Expertise-SI-Conception-et-réalisation-de-reseauxdadduction-deau-potable-2015. www.soladirites.org (Report).

Solgi, A., Zarei, H., Nourani, V., and Bahmani, R. (2017). A new approach to flow simulation using hybrid models. Applied Water Science, 7(7), 3691-3706.

Solgi, A., Pourhaghi, A., Bahmani, R., and Zarei, H. (2017). Improving SVR and ANFIS performance using wavelet transform and PCA algorithm for modeling and predicting biochemical oxygen demand (BOD). Ecohydrology and Hydrobiology, 17(2), 164-175.

Sun, Y. Liu, N. Shang, J.; Zhang, J. (2017). Sustainable utilization of water resources in China: A system dynamics model. J. Clean. Prod, 142, 613 625.

Sun M, Li Z, Yao C, Liu Z, Wang J, Hou A, Zhang K, Huo W, Liu M. (2020). Evaluation of Flood Prediction Capability of the WRF-Hydro Model Based on Multiple Forcing Scenarios. *Water*. 2020; 12(3):874. https://doi.org/10.3390/w12030874.

Tao Tian, Huifeend Xue. (2017). Prediction Of Annual Water Consumption In Guangdong Province Based On Bayesian Neural Network. *IOP Conf. Series: Earth and Environmental Science 69, 012032.* doi:10.1088/17551315/69/1/012032.

Tang, Z.Q., Liu, G., Xu, W.N., Xia, Z.Y., Xiao, H. (2012). Water Demand Forecasting in Hubei Province with BP Neural Network Model. Adv. Mater. Res. 2012, 599, 701–704.

T.M. Tung, Z.M. Yaseen Tiyasha. (2020). A survey on river water quality modelling using artificial intelligence models: 2000-2020. *J. Hydrol. 585*. https://doi.org/10.1016/j.jhydrol.2020.124670.

Tian T., Xue H. (2017). Prediction of annual water consumption in Guangdong Province based on Bayesian neural network. IOP Conf. Series Earth Environ. Sci. 2017, 69, 12032.

Tumer, A. E., and Edebali, S. (2015). An Artificial Neural Network Model for Wastewater Treatment Plant of Konya. International Journal of Intelligent Systems and Applications in Engineering, 3(4).

UNESCO (2015). Water Supply, Sanitation and Health. Report available at: https://unesdoc.unesco.org/ark:/48223/pf0000231823

USAID. (2017). Water and Development Country Plan for the Democratic Republic of Congo (DRC).; 2017. https://pdf.usaid.gov/pdf_docs/ PBAAH782.pdf.

Villarin, M.C., Rodriguez-Galiano, V.F. (2019). Machine Learning for Modeling Water Demand. J. Water Resour. Plan. Manag. 2019, 145, 04019017.

Wang, H.; Wang, W.; Cui, Z.; Zhou, X.; Zhao, J.; Ya, L. (2010) A new dynamic firefly algorithm for demand estimation of water resources.Inf. Sci. 2018, 438, 95–106.

Wilson, P., & Mantooth, H. A. (2013). Model-based engineering for complex electronic systems. Newnes.

World Health Organization (WHO); United Nation Children Fund (UNICEF) (2012). Progress on drinking water and sanitation WHO, Geneva and UNICEF, New York, pp 122–125.

Zhang, Z.G.; Shao, Y.S.; Xu, Z.X. Prediction of urban water demand on the basis of Engel's coefficient and Hoffmann index: Case studies in Beijing and Jinan, China. Water Sci. Technol, 62, 410–418.

Zhang Z., Arnault J., Wagner S., Laux P., Kunstmann H.G. (2019). Impact of lateral terrestrial water flow on land-atmosphere interactions in the Heihe River Basin in China: Fully coupled modeling and precipitation recycling analysis. Journal of Geophysical Research: Atmospheres, 124, 8401–8423. https://doi.org/10.1029/2018JD030174

Zhang Q., Diao Y., Dong J. (2013). Regional Water Demand Prediction and Analysis based on CobbDouglas Model. Water Resource Management. 2013, 27, 3103–3113.

Zhao, Y.; Chen, X.J. (2014). Prediction Model on Urban Residential Water Based on Resilient BP Learning Algorithm. Appl. Mech. Mater. 2014, 543, 4086 4089.

APPENDICES

Appendix A

Similarity Check Report

thesis

ORIGINALITY REPORT

10%	8%	4%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Zhenyu Zhang, Joël Arnault, Sven Wagner, Patrick Laux, Harald Kunstmann. "Impact of Lateral Terrestrial Water Flow on Land-Atmosphere Interactions in the Heihe River Basin in China: Fully Coupled Modeling and Precipitation Recycling Analysis", Journal of Geophysical Research: Atmospheres, 2019 Publication	1%
2	docs.neu.edu.tr Internet Source	1%
3	postconflict.unep.ch Internet Source	1%
4	repository.maseno.ac.ke Internet Source	1%
5	neu.edu.tr Internet Source	1%
6	www.researchgate.net Internet Source	1%
7	idr.nitk.ac.in Internet Source	1%

8	Jaehak Jeong, Narayanan Kannan, Jeff Arnold, Roger Glick, Leila Gosselink, Raghavan Srinivasan. "Development and Integration of Sub-hourly Rainfall–Runoff Modeling Capability Within a Watershed Model", Water Resources Management, 2010 Publication	<1 %
9	elibrary.buse.ac.zw:8080 Internet Source	<1 %
10	ugspace.ug.edu.gh Internet Source	<1 %
11	hopedrc.org Internet Source	<1 %
12	ir.mu.ac.ke:8080 Internet Source	<1 %
13	etd.hu.edu.et Internet Source	<1 %
14	erepository.uonbi.ac.ke Internet Source	<1 %
15	rcastoragev2.blob.core.windows.net Internet Source	<1 %
16	ikp.nilebasin.org Internet Source	<1 %
17	Fausto Favero, Lieke Hüsken, Jochen Hinkel, Heleen Vreugdenhi, Umberto Pernice,	<1 %

Michael Sedlmeier. "D3.1 Framework for developing funding and finance arrangements for coastal restoration", Pensoft Publishers, 2023
Publication

18	Hanoi National University of Education Publication	<1 %
19	bnrc.springeropen.com Internet Source	<1 %
20	eprints.undip.ac.id Internet Source	<1 %
21	eprints.utm.my Internet Source	<1 %
22	www2.mdpi.com Internet Source	<1 %
23	ourspace.uregina.ca Internet Source	<1 %
24	stacks.cdc.gov Internet Source	<1 %
25	bioone.org Internet Source	<1 %
26	bjbas.springeropen.com Internet Source	<1 %
27	dspace.cbe.ac.tz:8080 Internet Source	<1 %

28	en.wikipedia.org Internet Source	<1 %
29	eprints.kfupm.edu.sa Internet Source	<1 %
30	www.dissertations.mak.ac.ug Internet Source	<1 %
31	www.hindawi.com Internet Source	<1 %
32	www.mdpi.com Internet Source	<1 %
33	www.slideshare.net Internet Source	<1 %
34	Bentolhoda Asl-Rousta, S. Jamshid Mousavi. "A TOPSIS-Based Multicriteria Approach to the Calibration of a Basin-Scale SWAT Hydrological Model", Water Resources Management, 2018 Publication	<1 %
35	Vahid Nourani, Gozen Elkiran, Jazuli Abdullahi. "Multi-step ahead modeling of reference evapotranspiration using a multi-model approach", Journal of Hydrology, 2020 Publication	<1 %

Prof. Dr. Gözen Elkıran

Romain Kasongo Amsini

Appendix B

Ethical Approval Letter

November 20, 2023

Nicosia

ETHICS EVALUATION

Dear Romain Kasongo Amsini

Your application titled **Water Resources Management Using Soft Computing Methods** has been evaluated by me (instead of the Scientific Research Ethics Committee) and granted approval. You can start your research on the conditions that will abide by the information provided in your application.

This evaluation has been done by me because you have not used a questionary and there is no need for data collection from the people, and your work will be based on calculations and the applications of soft computing methods using a software.

Sincerely

Prof. Dr. Gozen Elkiran

Civil Engineering Department

Faculty of Civil and Environmental Engineering

Near East University, Near East Boulevard, ZIP: 99138, Nicosia/TRNC, Mersin 10 – Turkey