DEVELOPMENT OF A SUSTAINABLE EVOLUTIONARY-INSPIRED ARTIFICIAL INTELLIGENT SYSTEM FOR MUNICIPAL WATER DEMAND MODELLING

UNIVERSITY OF KWAZULU - NATAL ™

INYUVESI YAKWAZULU-NATALI

By

Oluwaseun Kunle OYEBODE

Student Number: 214585625

MARCH 2020

DEVELOPMENT OF A SUSTAINABLE EVOLUTIONARY-INSPIRED ARTIFICIAL INTELLIGENT SYSTEM FOR MUNICIPAL WATER DEMAND MODELLING

UNIVERSITY OF KWAZULU - NATAL ™

INYUVESI YAKWAZULU-NATALI

By

Oluwaseun Kunle OYEBODE

Student Number: 214585625

Submitted in fulfillment of the academic requirements for the degree of Doctor of Philosophy in Civil Engineering, School of Engineering, University of KwaZulu-Natal (Howard College Campus), Durban 4000, South Africa

PREFACE

The information presented in this thesis is an original work by the candidate. Unless otherwise indicated in the chapters that some works have been done partly by the local or international researcher collaborators, this work was carried out in the Department of Civil Engineering, School of Engineering, College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Howard College Campus, Durban, South Africa from January 2015 to December 2018 and has not otherwise been submitted in any form for any degree or diploma to any other University in the world. Apart from research collaborators, all the assistance obtained from others has been duly acknowledged in the acknowledgement section.

PhD Candidate:

____

Oluwaseun Kunle Oyebode

Supervisors:

____

Professor Akshay Kumar Saha

____

Professor Albert Modi

DECLARATION 1 – PLAGIARISM

I, **Oluwaseun Kunle Oyebode**, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
   a. Their words have been rewritten but the general information attributed to them has been referenced
   b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed

____

DECLARATION 2 – PUBLICATIONS

Details of contribution to publications that form part and/or include research presented in this thesis (include publications in preparation, submitted, *in press* and published and give details of the contributions of each author to the experimental work and writing of each publication).

PUBLICATIONS FROM THIS THESIS

1. **Oyebode, O.,** Babatunde, D. E., Monyei, C. G. & Babatunde, O.M. (2019). Water demand modelling using evolutionary computation techniques: Integrating water equity and justice for realization of the sustainable development goals. Heliyon 5 (11): e02796. https://doi.org/10.1016/j.heliyon.2019.e02796
2. **Oyebode, O.** & Stretch, D. (2019). Neural network modeling of hydrological systems: A review of implementation techniques. Natural Resource Modeling 32 (1): e12189. https://dx.doi.org/10.1111/nrm.12189
3. **Oyebode, O.** (2019). Evolutionary modelling of municipal water demand with multiple feature selection techniques. Journal of Water Supply: Research and Technology–AQUA 68 (4): 264-281. https://doi.org/10.2166/aqua.2019.145
4. **Oyebode, O.** & Ighravwe, D. E. (2019). Urban water demand forecasting: A comparative evaluation of conventional and soft computing techniques. Resources 8 (3): 156. https://doi.org/10.3390/resources8030156
5. **Oyebode, O.,** Monyei, C. G. & Stretch, D. (2018). A sustainability framework for integrating equity and justice in evolutionary-based water demand modelling. Presented at the 2018 College of Agriculture, Engineering, and Sciences Postgraduate Research and Innovation Symposium, University of KwaZulu-Natal, Durban, South Africa. Awarded Best Poster Presentation: PhD category.
6. **Oyebode, O.,** Monyei, C. G. & Stretch, D. (2019). Harnessing the potentials of evolutionary computation via the integration of rights-based distributive principles in water demand management. In: Proceedings of the 2019 European International Conference on Transforming Urban Systems (EICTUS–2019), Strasbourg, France, 26-28 June 2019; p. 36.

DEDICATION

To **God Almighty**

To my Late Father, Pa. **Johnson O. Oyebode**

To my mother, Mrs. **Victoria O. Oyebode**

To my darling wife, **Olajumoke** and our handsome sons, **David** and **Joshua**

ACKNOWLEDEGEMENTS

Who am I, that you are mindful of me, O Lord? Words are not enough to appreciate your wondrous works and acts in my life and, more especially, during the course of this PhD program. It can only be you, Lord – the Source of Wisdom and Fountain of Knowledge. I return all glory and adoration to your name.

I would like to express my sincere appreciation to the following individuals and institutions that formed an integral part of the successful completion of this research:

Professor Akshay Kumar Saha and Professor Albert Modi, for supervising and ensuring the smooth completion of this study.

Professor Derek Stretch, for his guidance and support in the course of this study.

My beautiful and supportive wife, Olajumoke Oyebode, thank you for your indefatigable moral support, endurance and unfailing commitment to my academic pursuit, and for the confidence you have in me. Your academic inclination, perseverance, diligence, hard work and multitasking ability, no doubt contributed to the success of this study. To my cute sons – David and Joshua Oyebode, thank you for always being a source of joy to me. Your arrival during the course of this study invigorated my drive towards the completion of this program. You guys are my best gifts of all time.

My wonderful parents, Mr Johnson Oyebode (Late) and Mrs Victoria Oyebode, for the outstanding upbringing and solid academic foundation you bestowed. I am forever grateful to you. I also appreciate my sister, Dr Olubukola Oyebode, and my brother, Kolawole Oyebode, for their unrelenting, enthusiastic, moral, inspirational and spiritual support. Thank you all for believing in me.

I cannot fail to mention the endearing support of my in-laws, Lt. Col (Rtd.) and Chief Mrs Daramola. Many thanks for your encouragement and prayers.

Dr Chukwuka Monyei, an amiable brother and brilliant scholar. I am grateful for your assistance and guidance, especially in addressing the policy and sustainability aspects of this study. Your input to this thesis is invaluable.

Ms Nombuso Dlamini – Postgraduate Administrator (School of Engineering), and Ms Ausie Luthuli – College Academic Administrative Officer (Higher Degrees), for facilitating the external examination process of this thesis.

Mr and Mrs Olubayo Babatunde, for your unalloyed love, encouragement, inspiration and noble acts of selflessness. I will forever appreciate you.

To my mentors, Professor Olufunso Olorunsogo and Engr. Greg Veerasamy, for their inspirational and fatherly advice, which span across all facets of life. May the Lord reward you abundantly.

My pastors, Pastor and Mrs Adejimi and their families, for the prayers, spiritual guidance, godly counsels and moral support you gave during the course of the work. May the Lord enlarge your ministries in Jesus Name. I also thank the members of the Redeemed Christian Church of God (RCCG), Chapel of Praise, Durban, for their prayers, godly counsels and moral support. I treasure you all.

City of Ekurhuleni, Department of Water Affairs and the South African Weather Service, for providing me with the requisite data for carrying out this research.

The entire staff of Royal HaskoningDHV, for their support and motivation throughout the program.

My sincere appreciation goes to friends and families who have encouraged and supported this work in one way or the other – the Oyewoles, the Adegbites, the Olokobas, Professor Temitayo Ewemoje and family, Mr Deji Audu and family, Engr. Nathaniel Sawyerr, Engr. Oluwatoyin Gbadeyan, A/Prof. Olufunke Oladipupo, Dr Oluwatosin Olofintoye, Dr Akin Ikudayisi, Dr Desmond Ighravwe, Dr and Dr (Mrs) Akinrinde, Mr Lindokuhle Magagula, Mr Ntala Lelimo, Mr Adesoji Jaiyeola, Mr Bukunmi Fatoyinbo and my friends at Tongaat Hulett Developments. Thank you all for being there for me.

Finally, to all who have contributed to the success of this research, but who due to space could not be explicitly mentioned, I say a BIG thank you. May God bless you all, Amen.

Oluwaseun OYEBODE

ABSTRACT

This study presents the development of a differential evolution (DE)-inspired artificial neural network (ANN) that incorporates climate and socioeconomic information for a more accurate and reliable water demand forecasting. The study also addresses the limitations of ANN. Multiple feature selection techniques were employed to identify the minimal subset of features for optimal learning. The performance of the feature selection techniques was validated and compared to a baseline scenario comprising a full set of data covering potential casual variables including weather, socio-economic and historical water consumption data. The performance of the models was evaluated based on accuracy. Results show that all the feature selection techniques outperformed the baseline scenario. More importantly, the subset of features obtained from the Pearson correlation technique produced the most superior model in terms of model accuracy. Findings from the study suggests that inclusion of weather and socioeconomic variables in water demand modelling could enhance the accuracy of forecasts and cater for the impacts of climate and socioeconomic variations in water demand planning and management.

The performance of the optimal DE-inspired model was thereafter compared to those developed via conventionally-used multiple linear regression and standard time series technique – exponential smoothing as well as other prominent soft computing techniques, namely support vector machines (SVM) and conjugate-gradient (CG)-trained multilayer perceptron (MLP). Results show that the DE-inspired ANN model was superior to the four other techniques for both the baseline scenario and optimal subset of features. DE showcased robustness in fine-tuning algorithm parameter values thereby producing good performance in terms of the solution efficiency and quality. Generally, this study demonstrates that water demand models can account for the impacts of weather and socioeconomic variations by incorporating explanatory variables based on weather and socioeconomic factors. This study also suggests that the synergetic use of feature selection techniques, DE algorithm and an early stopping criterion could be used in addressing the limitations of ANN and developing an improved and more reliable water demand forecasting model.

This work goes further to propose for a novel and more comprehensive integrated water demand and management modelling framework (IWDMMF) that is capable of syncing conventional evolutionary computation techniques and social aspects of society. The methodologies, principles and techniques behind this study fosters sustainable development and thus could be adopted in planning and management of water resources.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMI	Average Mutual Information
ANN	Artificial Neural Networks
APSO	Adaptive Particle Swarm Optimization
ARIMA	Autoregressive Integrated Moving Average
BME	Bayesian Maximum Entropy
BP	Back Propagation
CG	Conjugate Gradient
CoCT	City of Cape Town
CR	Crossover Rate
DE	Differential Evolution
DFS	Demand Forecasting System
EA	Evolutionary Algorithm
EANN	Evolutionary Artificial Neural Networks
EC	Evolutionary Computation
EKF	Extended Kalman Filter
EP	Evolutionary Programming
ES	Evolution Strategy
ES_m	Exponential Smoothing
F	Mutation Rate
GA	Genetic Algorithm
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GDP	Gross Domestic Product
GDS	Growth and Development Strategy
GEP	Gene Expression Programming
GNP	Gross National Product

GP	Genetic Programming
GRNN	Generalized Regression Neural Network
GVA	Gross Value Added
HDI	Human Development Index
HH	Number of Household Connections
HRWS	Human Right to Water and Sanitation
HWD	Historical Water Demand
I	Income-based Variables
IDP	Integrated Development Plan
IPART	Independent Pricing and Regulatory Tribunal of New South Wales
IWDMMF	Integrated Water Demand and Management Modelling Framework
MAPE	Mean Absolute Percent Error
MDGs	Millennium Development Goals
MLP	Multilayer Perceptron
MLR	Multiple Linear Regression
NP	Population Size
NSE	Nash-Sutcliffe Efficiency Index
NSGA-II	Non-dominated Sorting Genetic Algorithm II
O	Other Variables
P	Population
PC	Principal Components
QoL	Quality of Life
R	Rainfall
R^2	Coefficient of Determination
RBF	Radial Basis Function
RBFNN	Radial Basis Function Neural Network

RH	Relative Humidity
RMSE	Root Mean Square Error
SDGs	Sustainable Development Goals
Stats SA	Statistics South Africa
SVM	Support Vector Machines
T_{max}	Maximum Temperature
T_{min}	Minimum Temperature
UN	United Nations
UNDESA	United Nations Department of Economic and Social Affairs
UNDP	United Nations Development Programme
UNEP	United Nations Environment Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
USCB	United States Census Bureau
UWOT	Urban Water Optioneering Tool
W	Weather-based Variables
WC	Water Consumption
WDN	Water Distribution Network
WWAP	United Nations World Water Assessment Programme

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Water scarcity has been a global issue in recent times. Global use of available freshwater has grown at roughly twice the rate of global population for the past century (Figure 1-1). With a rapidly growing global population, demand for water is expected to increase by nearly one-third by 2050 (WWAP, 2018). This threatening situation is further being exacerbated by rising water demands due to unanticipated factors such as climate variation, rapid urbanization, changing consumption patterns and socioeconomic transitions among others (Olofintoye, 2015). The need for an integrated and sustainable approach to water resources management across the globe is increasingly more imperative and of urgent attention to water managers and decision-makers.

Water distribution networks (WDNs) are designed to satisfy consumers' requirements in the short-, medium- and long-terms. One of the key factors in planning and management of WDNs is the satisfaction of consumer demand, which presumes providing adequate water with acceptable quality and at a reasonable pressure (Cabral, 2014). Moreover, the United Nations (UN), in a resolution adopted at its General Assembly on 28 July 2010, gave recognition to "the Human Right to Water and Sanitation" (HRWS) as a human right; stating with respect to water that, the human right to water entitles everyone to sufficient, safe, acceptable, physically accessible and affordable water for personal and domestic uses" (UN, 2010). The Resolution further acknowledges HRWS as an integral component in realizing all human rights. The adoption of HRWS allows for its recognition in international law through human rights treaties, declarations and other standards. These developments therefore elucidate the importance of assessing and projecting the evolution of water demands, as accurately as possible, to ensure appropriate future service levels.

Water demand forecasting is a critical component of effective water resource planning and management as it assists in determining the timing and capacity of developing new water resources (Shang et al., 2017). For instance, decisions on water infrastructure investments are critically dependent on the profile of future water demands. In addition, water demand forecasting allows for modelling of future conditions which assists in facilitating appropriate management options that balances water demand and supply (Mohamed and Al-Mualla, 2010). Water demand forecasting therefore enables the formulation of appropriate management policies to ensure continuity of service with the lowest possible cost.

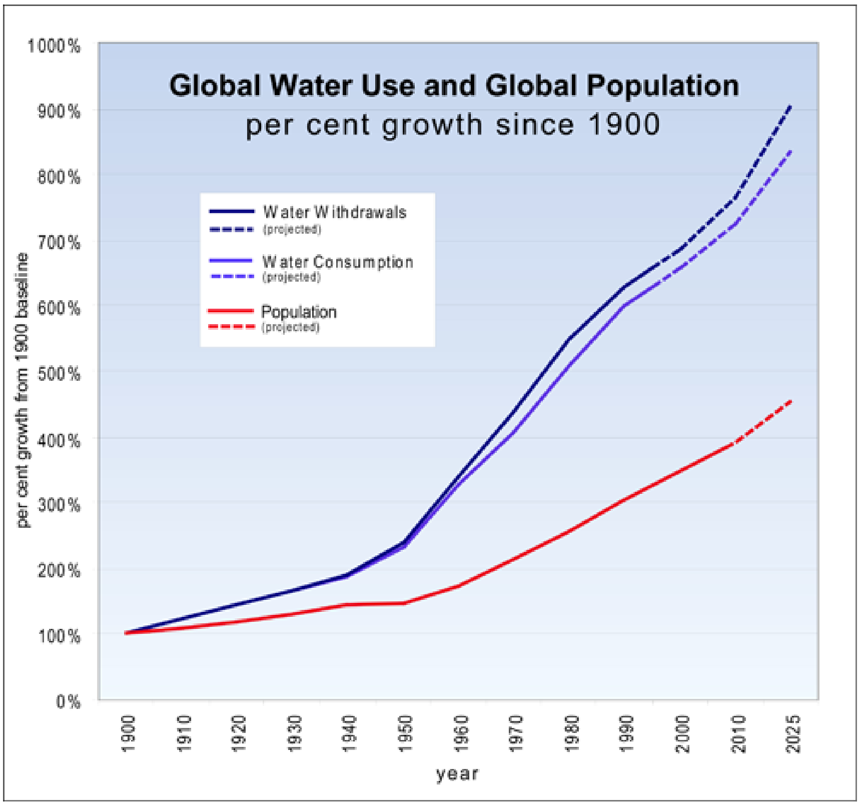


Figure 1-1: Rate of growth in global freshwater withdrawal and consumption. Source: UNEP (2012)

Water utilities, researchers, governments and other related stakeholders have been making concerted efforts towards ensuring that cities meet the water demand of their residents. However, in view of the complex and dynamic interactions among contributing factors and their intimate ties to urban hydrological processes, estimation of water demand remains a complex yet imperative task (House-Peters and Chang, 2011). Over the years, water utilities and consultants have generally employed a “fixture-unit” method which considers fixture unit demands, facility types, and socioeconomic factors in forecasting water demand for infrastructure planning and design (Blokker et al., 2010; Blokker et al., 2012; Buchberger and Wu, 1995). However, to compensate for several uncertainties associated with demand, this approach involves inclusion of large safety factors which usually overestimate the actual water demand by as much as 100% (Shabani et al., 2016). The overestimation of the actual water demand, in turn, translates to an over-designed system with high operation and maintenance costs (e.g. pumping, pipelines, etc.) as well as high price of water as resultant effects. The over-designed systems may also result in negative environmental impacts in areas located downstream of the system. Furthermore, the conventional “fixture-unit” approach (classically based on the assumption of collinearity), does not usually account for nonlinearities which are often associated with the contributing factors such as population, stand area, stand value, tourism, etc. (House-Peters and Chang, 2011; Shabani et al., 2017). The conventional approach has also been reported in the literature to be characterized by a lack of a climate variation perspective to water demand forecasting (UNESCO, 2016). These drawbacks, coupled with the complexity of water demand analysis, have necessitated the search for more robust and reliable water demand forecasting methods that can assist in designing more environmentally sustainable systems and in managing available water resources more efficiently (Shabani et al., 2016).

1.2 WATER DEMAND FORECASTING APPROACHES

Over the last decade, researchers have focused on improving water demand forecasting methodologies. They have largely focused on understanding the different factors that influence water consumption, improving forecasting methods and reducing forecasting uncertainty (Cabral, 2014).

Explanatory variables that have been considered in water demand forecasting can be categorized into three: (i) weather variables (e.g., levels of snow and rainfall, temperature, evaporation, wind speed, relative humidity, etc.) (Coomes et al., 2010; Dos Santos and Pereira Filho, 2014; Yousefi et al., 2017); (ii) socioeconomic variables (e.g., income-level, consumption patterns, tourism, population and water price, etc.) (Arbués et al., 2003; Qu et al., 2010; Shaw, 2007); and (iii) consumption variables which often include historical profiles of consumption, varying from previous hours to years (Nasseri et al., 2011; Shabani et al., 2018; Walker et al., 2015). Research has however shown that limited studies have considered weather variations in developing water demand forecasting models, thereby making water availability prone to uncertainties due to meteorological factors (Shabani et al., 2016). Furthermore, the selection of explanatory variables is dependent on data availability and quality; necessitating the adoption of a robust modelling approach in water demand forecasting. The robust modelling approach must ensure that the impacts of each explanatory variable is well-analyzed and adequately captured.

Table 1-1 presents a summary of the forecasting techniques that have been applied in water demand forecasting. According to Cabral (2014), these techniques are classified as subjective, extrapolation, regression, soft computing and other methods.

The selection of the most suitable forecasting technique involves the consideration of several factors including but not limited to expert's knowledge, intended use of forecast results, size and other characteristics of the utility and its service area, etc. (Billings and Jones, 2008). None of the techniques mentioned in Table 1-1 is yet to be considered as universally applicable to any water supply system or referentially relative to other methods (Kozłowski et al., 2018). However, researchers have recently argued that soft computing techniques, especially evolutionary computation techniques, have the potential of being universally embraced. This is due to their ability to exploit the tolerance for imprecision and uncertainty to achieve tractability, robustness, and low-solution cost", while concurrently taking into cognizance nonlinearities inherent in contributing variables (Ghalehkhondabi et al., 2017; Maier et al., 2019; Maier et al., 2014; Reed et al., 2013). Considering the drawbacks of other techniques and the inherent properties of soft computing techniques, more effort is now being tailored towards exploiting the potential of soft computing techniques. Reed et al. (2013) suggest the need to further explore the potentials of evolutionary computation techniques as this approach could accelerate the discovery of water resource planning innovations required for enabling a sustainable future.

This study thus hypothesizes that evolutionary-inspired soft computing techniques could offer new high-end solutions required to forecast water demand. This would assist water resource managers in offsetting the rising challenges to water security from population growth, socioeconomic and weather variations among other factors.

Table 1-1: Overview of water demand forecasting approaches. Adapted from: Cabral (2014)

Approach	Description	Forecasting technique	Research studies	Drawbacks
Subjective	Range from informed opinions of utility management to highly structured Delphi and scenario construction methods	Expert opinion	Billings and Jones (2008)	Based on several assumptions about the future which may be untrue; Prone to sentiments.
		Delphi methods	Billings and Jones (2008)	
Extrapolation (Time series models)	Involves making statistical forecast by using historical trends of only the variable forecasted. Based on the assumption that recent and historical trends will continue.	Moving average	Kozłowski et al. (2018); Mun (2006); ;	Produces large errors if discontinuities occur within the predicted period.
		Exponential smoothing	Billings and Jones (2008); Mun (2006); Donkor et al. (2012)	
		Autoregressive Integrated Moving Average (ARIMA)	Donkor et al. (2012)	
		Generalized Autoregressive Conditional Heteroskedasticity (GARCH)	Caiado (2009)	
Regression	Uses a set of drivers or explanatory variables to explain why a target variable has changed historically and to forecast its future values using time series and cross-sectional data.	Simple/ Multiple linear regression	Babel et al. (2007); Caiado (2009)	Considers linear relationship among variables and water demand
Soft computing methods	Acquires knowledge via a self-learning process to define a relationship (whether linear or nonlinear) between explanatory and target variables thereby describing the behaviour of the process being modelled and capable of forecasting future values.	Artificial intelligent methods (e.g., artificial neural networks, support vector machines, evolutionary algorithms, fuzzy inference systems, etc.)	Bennett et al. (2013); Firat et al. (2009); Ghalehkhondabi et al. (2017); Shabani et al. (2018)	Requires high level of expertise for model development and interpretations.
Other methods	As applicable	Bayesian Maximum Entropy (BME); Hybrids	Fagiani et al. (2015); House-Peters and Chang (2011); Nasseri et al. (2011); Oshima and Kosuda (1998)	May comprise many intricate parts

1.3 PROBLEM STATEMENT

Despite the recent advances in the application of soft computing tools, research suggests that some of the techniques have not been exhaustively applied to water demand prediction. Ghalehkhondabi et al. (2017) and Oyebode et al. (2019) in their comprehensive reviews of the application of soft computing methods in water demand prediction, identify several areas that are yet to be explored which should be considered for future research. These areas include:

i. Application of hybrid and ensemble models and algorithms.
ii. Investigating the potential of emerging artificial intelligence and metaheuristic techniques (for example, evolutionary algorithms and deep learning).
iii. Shifting focus from short term to medium- and long-term forecasting to enhance the accuracy and reliability of medium- and long-term planning and management decisions.
iv. Incorporating economic factors into medium- to long-term forecasting.
v. Dealing with noisy data associated with feedforward ANN models.

Furthermore, Shabani et al. 2016 recommends that greater attention be given to the impacts of varying weather conditions on medium- to long-term water demand predictions for effective water resource management.

The above suggestions serve as a motivation for this study. To this end, this study explores the potential of a hybrid soft computing technique which entails coupling an artificial intelligence technique – artificial neural network (ANN) to an evolutionary computation technique – differential evolution (DE) algorithm, in forecasting water demand. Monthly forecast models were developed by training a feedforward ANN using a DE algorithm. Although the development of hybrid soft computing models is gradually increasing within the science and engineering domains, including water resources, there has been no report of the techniques used in this study in the existing literature. Specifically, the application of DE, despite its application in some water resources studies (Abdul-Kader, 2009; Olofintoye et al., 2016; Oyebode, 2014) is yet to be used as a training algorithm for feedforward ANNs in water demand forecasting. This study therefore pioneers the application of DE in a training multilayer feedforward ANN in water demand forecasting. Furthermore, this study incorporates weather and socioeconomic variables as model inputs to assess their impacts on water demand forecasts. The performance of the ANN-DE was thereafter evaluated against widely used soft computing techniques (an ANN trained using a conjugate gradient algorithm and a support vector machine) in terms of forecast accuracy and model complexity.

This study also examines the role of EC techniques in enabling sustainable development with the United Nations Sustainable Development Goals (SDGs) in perspective. The UN SDGs (Figure 1-2), with specific reference to goals 6 and 10, elucidates the imperativeness of universal and equitable access to clean and affordable water. Goal 6 relates to availability and sustainable management of water and sanitation for all, while goal 10 promotes reduction in inequality within and among countries. This implies that in managing water resources vis-à-vis its allocation and safeguarding, balance must be achieved in ensuring equitable distribution of water resources to everyone and among competing needs. However, research has shown that engineers and water managers often focus on the technical, environmental and economic aspects of water demand and supply when conceptualizing and implementing water allocation, optimization and conservation strategies, leaving the social aspects to social scientist to address (Marques et al., 2018; Marques et al., 2015; Menapace et al., 2018; Zeng et al., 2012). This is evident in many of the existing optimization models, including EC optimization models, which lack a social perspective to water demand or allocation analysis. Social scientists on the other hand, often focus majorly on the social aspects of water such as quality of life, consumer satisfaction or perception, poverty, consumption patterns or consumer behavior, legal and political constructs, productivity, etc., but place less emphasis on other important aspects (Chomba et al., 2017; Johnson et al., 2016; Roa Garcia, 2014). Sustainability however centers around technical, environmental, economic and social aspects of life. There is therefore a need for a framework that is capable of syncing existing optimization models with social aspects of society to foster the realization of the UN SDGs.

Figure 1-2: The 2030 United Nations Sustainable Development Goals (SDGs)

1.4 STUDY OBJECTIVES

This research aims to develop an intelligent model that will seamlessly utilize weather and socioeconomic factors for municipal water demand prediction.

Specific aims of the study are as follow:

i. To conduct an extensive review of the extent to which evolutionary-inspired artificial intelligent models have been employed in water demand modelling.
ii. To identify and analyze the factors that affect municipal water demand.
iii. To develop an intelligent system model for municipal water demand prediction.
iv. To evaluate the performance of the intelligent model.
v. To propose a framework for sustainable allocation of water resources.

1.5 SIGNIFICANCE OF STUDY

The methodologies, techniques and models developed in this study could be adopted in improving the accuracy of water demand forecasts, and serve as an alternative to the conventional method which is often characterized by overestimations and several uncertainties. The study could help municipalities, water utilities and other stakeholders in planning for adaptive water resource allocation at different scales. It could also enhance the future operation and management of water resources facilities for both the immediate community and the global society.

The development of the novel framework that integrates equity and justice in EC optimization models fosters the realization of the UN SDGs. The novel framework could therefore assist in simultaneously addressing technical, environmental, economic and social concerns relating to water allocation and conservation within communities, countries and continents around the world.

This study therefore provides the pathway to strategically and accurately plan for the implementation, operation and management of water resources and associated infrastructure.

1.6 SCOPE AND LIMITATION OF THE STUDY

This study is limited to the application of ANN and DE in forecasting water demand at a municipal level. An ANN model is developed for forecasting monthly water demand while DE is applied in training and optimizing the network architecture of the model resulting in an improved water demand forecasting model. The choice of DE is due to its numerous advantages reported in the literature. This study is also limited by data availability. Finding lengthier data samples for model development is a challenging process in data analytics, especially in developing countries like South Africa. Revenue (i.e. billed) water consumption is utilized in this study, hence, the impacts of non-revenue water is not considered.

1.7 THESIS OUTLINE

This thesis presents manuscripts that were prepared, compiled or published during the course of the research work. This thesis is organized into five chapters.

The work starts with a general introduction in chapter 1. It provides a general background to current issues with respect to water demand modelling, identifies research problems and presents a brief review of modelling techniques employed in the subject area. The statement of the problem, study objectives, significance and limitations of the study are also presented.

Chapter 2 is organized into two sections (Sections A and B). Section A presents a comprehensive review of EC techniques to establish and classify the ways wherein they have been employed in water demand modelling and identifies important research challenges and future directions. Section B proceeds to investigate the inadequacies of conventional EC techniques in influencing water demand management policies and presents a novel and more comprehensive integrated water demand and management modelling framework (IWDMMF) that is capable of syncing conventional EC techniques and social aspects of society.

Chapter 3 presents the methods used in developing an artificial intelligent model for water demand forecasting. The integration of DE into the model and the parameters used is discussed. The impacts of climate, socioeconomic and consumption information on the model were investigated and multiple feature selection techniques were used to identify the minimal subset of features for optimal learning. The performance of the models as per each feature selection technique was compared to a baseline scenario comprising a full set of data covering potential casual variables. The ability of DE in solving real world water demand forecasting problems was investigated in terms of model complexity and accuracy.

In Chapter 4, the performance of the DE-inspired ANN model developed in the preceding chapter is evaluated comparatively with four prominent modelling techniques - a conventional multiple linear regression (MLR) model, a standard time series technique – exponential smoothing (ES_m), an ANN trained using a conjugate gradient algorithm (ANN-CG) and a support vector machine (SVM). The techniques were thereafter tested across two of the scenarios developed in chapter 3. The superior models were identified based on their underlying performance using standard performance evaluation criteria and their results discussed extensively.

Chapter 5 presents a general summary and conclusion based on the results of the previous chapters. It also gives suggestions and recommendations for future research.

This thesis represents a compilation of manuscripts where each chapter is an individual entity, hence, some repetitions between chapters are unavoidable.

CHAPTER 2

WATER DEMAND MODELLING USING EVOLUTIONARY COMPUTATION TECHNIQUES: INTEGRATING WATER EQUITY AND JUSTICE FOR REALIZATION OF THE SUSTAINABLE DEVELOPMENT GOALS

Oluwaseun Oyebode[1]*, Damilola E. Babatunde[2], Chukwuka G. Monyei[3] and Olubayo M. Babatunde[4]

[1]Centre for Research in Environmental, Coastal and Hydrological Engineering (CRECHE), Department of Civil Engineering, University of KwaZulu-Natal, Durban 4014, South Africa

[2]Department of Chemical Engineering, Covenant University, Sango Ota, Nigeria

[3]Big Data Enterprise and Artificial Intelligence Laboratory (Big-DEAL), University of the West of England, Bristol, United Kingdom

[4]Department of Electrical and Electronics Engineering, University of Lagos, Akoka, Yaba, Lagos, Nigeria

*Corresponding author: oluwaseun.oyebode@gmail.com; Tel: +27 (0) 84 807 3576

2.1 OVERVIEW

The purpose of this review is to establish and classify the diverse ways in which evolutionary computation (EC) techniques have been employed in water demand modelling and to identify important research challenges and future directions. This review also investigates the potentials of conventional EC techniques in influencing water demand management policies beyond an advisory role while recommending strategies for their use by policy-makers with the sustainable development goals (SDGs) in perspective. This review ultimately proposes a novel integrated water demand and management modelling framework (IWDMMF) that enables water policy-makers to assess the wider impact of water demand management decisions through the principles of egalitarianism, utilitarianism, libertarianism and sufficientarianism. This is necessary to ensure that water policy decisions incorporate equity and justice.

Keywords: Artificial intelligence; evolutionary computation; sustainable development goals; water demand; water equity; water justice

2.2 SECTION A: APPLICATION OF EVOLUTIONARY COMPUTATTION TECHNIQUES IN WATER DEMAND MODELLING

2.2.1 Introduction

Over the past several decades, ever-growing demands for freshwater resources have increased the risks of severe water stress in many parts of the world (Figure 2-1). According to the 2015 United Nations (UN) World Water Development Report, the world is projected to face a 40 per cent deficit in water supply in 2030, unless the international community intensely improves water supply management (UNESCO, 2015). This figure is expected to increase to 55 per cent by 2050, under a business-as-usual scenario (UNESCO, 2015). The management of available water resources is therefore important to many decision-makers in the public and private sectors, with concerted efforts being made towards ensuring that cities meet their water demands in the future. However, factors such as increasing population, socioeconomic growth, water leakages, excessive water withdrawals and evolving climate conditions remain intimately tied to urban hydrological processes, thereby making the estimation of water demand a complex task (House-Peters and Chang, 2011). Increasing water demand makes a restoration of the balance between demand and limited supplies necessary to avoid severe global water crisis, and attaining the United Nations' 2050 vision – "achieving a water secure world, where every person has access to adequate quantities of water of an acceptable quality and from sustainable sources, to meet their basic needs and sustain their well-being and development" (UNESCO, 2015).

The planning and management of water resources as well as design and operation of water infrastructure remains critical to the provision of water supply services, and forms the basis for water demand forecasting (Oyebode et al., 2014a). Decisions on water-related investments are critically dependent on how future water demands are to be forecasted (Almutaz et al., 2013). Water demand forecasting is therefore of strategic importance, especially in regions with limited water supplies where the role of demand management policy becomes increasingly significant. Over the years, the conventional approach employed by water utilities and consultants in planning and designing water treatment, supply and distribution systems has been the "fixture-unit" method which considers the sum of fixture unit demands, facility types, and socioeconomic factors to determine the peak demand. However, to compensate for several uncertainties associated with demand, this approach involves inclusion of large safety or peak factors which usually overestimates the actual water demand by as much as 100% (Shabani et al., 2016), with resulting high operation and maintenance costs and high prices for water. Furthermore, the conventional approach (classically based on the assumption of collinearity), does not usually account for nonlinearities which may be inherent in the contributing factors (House-Peters and Chang, 2011; Shabani et al., 2017). Another deficiency in the application of the conventional approach is the lack of a climate change perspective in the water demand planning phase. Research has shown that, for each degree of global warming, nearly 7% of the global population will be exposed to a decrease of renewable water resources of at least 20% (UNESCO, 2016). The exclusion of the impacts of climate change in conventional approaches to water demand forecasting may consequently deprive water managers of the opportunity to put in place effective early warning systems and implement adaptive interventions to variations in water availability and extreme water-related events.

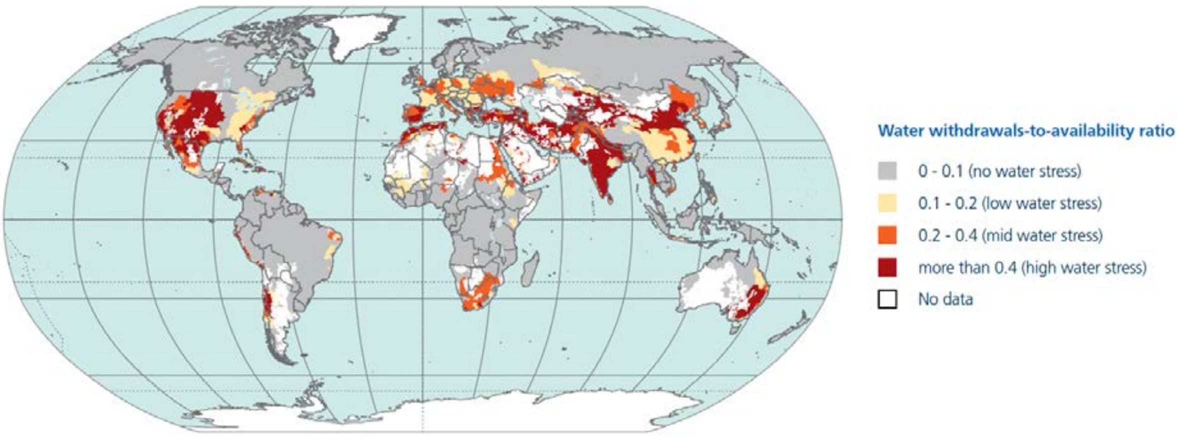


Figure 2-1: Annual average water stress based on withdrawals-to-availability ratio (1981-2010). Source: (UNESCO, 2016)

The complexity of water demand analysis has necessitated a search for more sophisticated tools for accurate water demand prediction. More recently, researchers have explored soft computing techniques to develop models to achieve more accurate water demand forecasts. Soft computing is “a collection of methodologies that aim to exploit the tolerance for imprecision and uncertainty to achieve tractability, robustness, and low-solution cost", while also taking into cognizance nonlinearities inherent in contributing variables (Ghalehkhondabi et al., 2017). Examples of robust soft computing techniques that have found application in water resources include, but are not limited to, artificial neural networks (ANN), fuzzy and neuro-fuzzy methods, support vector machines (SVMs), and more recently, evolutionary computation (EC) techniques. These soft computing techniques and many more have been reported to have achieved varying degrees of successes in diverse water resource applications, including streamflow forecasting (Kisi and Cigizoglu, 2007; Oyebode et al., 2014a), reservoir inflow prediction (Oyebode and Adeyemo, 2014), water quality modelling (Chang et al., 2015; Dragoi et al., 2011), wastewater treatment (Enitan et al., 2014) and sediment yield modelling (Ch et al., 2013; Guven and Kişi, 2011). These techniques have also been hybridized to allow for complementary modelling; resulting in improved performance (Adeyemo et al., 2018; Bhagwat and Maity, 2013; Londhe and Narkhede, 2017).

Previous studies have reported the application of soft computing techniques to water demand forecasting (Firat et al., 2009; Shabani et al., 2017; Tabesh and Dini, 2009; Varahrami, 2010). However, research suggests that despite the recent advances in soft computing in water resources, some tools have not been exhaustively applied to water demand forecasting (Ghalehkhondabi et al., 2017). These tools include recently developed artificial intelligence and metaheuristic techniques like evolutionary computation, deep learning, simulated annealing, ant colony optimization and particle swarm optimization. Ghalehkhondabi et al. (2017)'s finding therefore supports the United Nations' call for exploitation of new data sources, improved models and more powerful data analysis methods for implementation of adaptive management strategies to foster effective response to varying and uncertain conditions (UNESCO, 2015).

This paper aims to contribute to the literature that reviews the general application of soft computing techniques in water resources (Maier and Dandy, 2000; Maier et al., 2014; Oyebode et al., 2014b; Reed et al., 2013) and specifically to water demand modelling (Donkor et al., 2012; Ghalehkhondabi et al., 2017; House-Peters and Chang, 2011; Shabani et al., 2016). A provisional search of scholarly databases however returned no significant review with specific focus on the application of EC techniques in water demand forecasting. Moreover, no attempt has been made to review the extent to which EC techniques have been used to address the UN SDGs within the context of water demand modelling. This review is therefore pertinent in that its specific emphasis is on the application of EC techniques to water demand modelling and how it can be positioned to implement water policy decisions based on equity and justice, and foster the realization of SDGs. The aims are to (a) establish and classify the diverse ways in which EC techniques have been employed in water demand modelling; (b) identify important research challenges and future directions; (c) recommend implementation strategies for the adoption by policy-makers with water equity and justice and SDGs in perspective.

2.2.2 Water demand forecast variables and determinants

One of the initial steps in carrying out any modelling study is to collect data of explanatory variables that may possibly influence the system to be modelled (Oyebode, 2014). The development of models for accurate water demand forecasting therefore requires identification of explanatory variables that directly and indirectly influence water demand. The identification of explanatory variables forms the basis upon which final input parameters will be selected during model development.

This section briefly discusses key explanatory variables that may influence water demand and justifies the need for their consideration in the development of water demand forecasting models. The key explanatory variables discussed in this paper include weather-based variables (e.g., rainfall, temperature, evaporation, wind speed, relative humidity, etc.), population growth, income-level and water price.

Other factors that may influence water demand include those based on water conservations and demand management initiatives such as period of restricted use, source substitution, adoption of water sensitive urban designs, land use, consumer education and end use consumption profiling, and non-revenue water (losses from leakages and other sources).

2.2.2.1 Weather-based variables

Weather plays a huge role in water availability as both short-and long-term factors can significantly affect water consumption rates. Short-term factors are typically daily weather variables like rainfall, temperature, wind speed and humidity, while long-term factors which occur as a result of climate change, is likely to impact annual average temperature, rainfall and evapotranspiration (White et al., 2003). Short-term factors often have an immediate effect on water demand. For instance, extremely hot or dry weather will increase the rate of evaporation, thereby heightening water consumption rates for drinking, irrigation and recreational (swimming) purposes, and ultimately result in a surge in water demand. Climate change however can induce changes in seasonal runoff regimes and inter-annual runoff variability which can significantly impact water availability.

According to Coomes et al. (2010), weather is hysteretic, dynamic and state-dependent in nature, and has been proven to have non-linear effects on water consumption. Research has also shown that climate change impacts, like the changes in the occurrence of droughts, can have significant influence on economic growth as evident in Figure 2-2. It is therefore important to take into consideration climatic or weather factors when developing a water demand forecasting model to account for the impacts of short-and long-term weather variability, and potentially improve the performance of the model. It is however important to note that inclusion of weather inputs may increase computational demand and possibly introduce new risks associated with data quality or handling (Bakker et al., 2014). Consequently, consideration should be given to performance enhancement as well as computational demand and data quality risks during the process of model development.

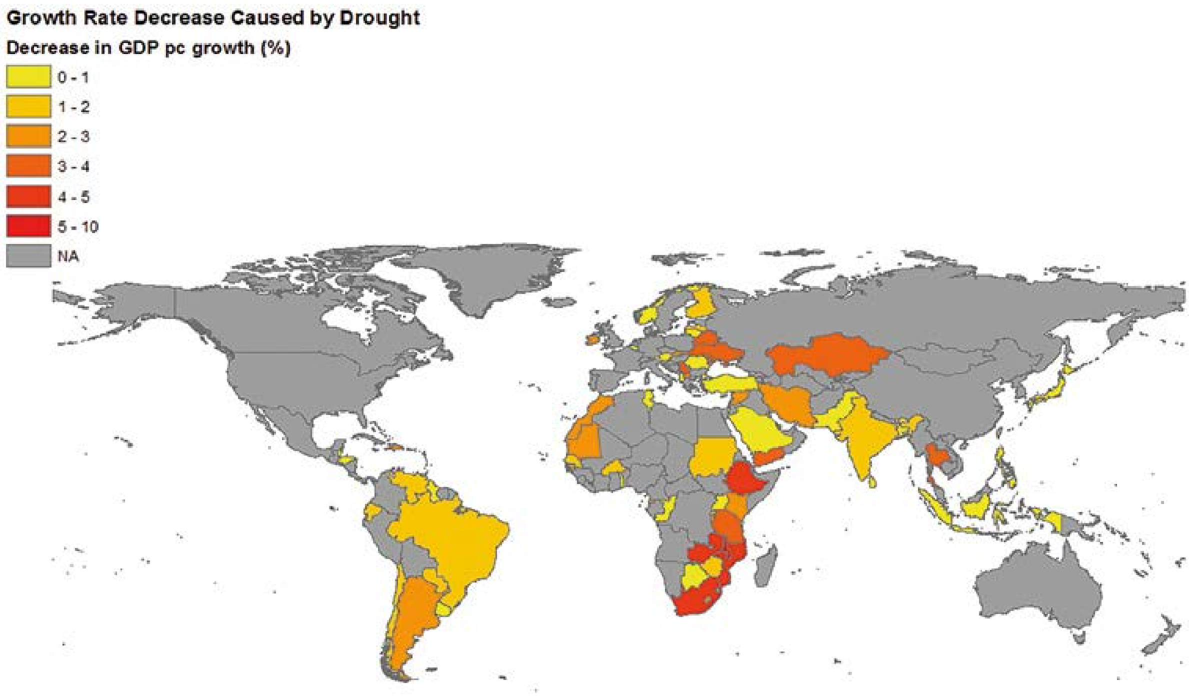


Figure 2-2: Countries with the largest reduction in growth due to drought. Source: Sadoff et al. (2015)

2.2.2.2 Population growth

Population growth, usually due to urbanization, migration and industrialization, is characterized by a nonlinear relationship, and considered as one of the key factors that impact water demand (UNESCO, 2015). The rate of demand for water has reportedly doubled the rate of population growth, with global population growth estimated at about 80 million people per year, and projected to reach 9.1 billion by 2050 (UNDESA, 2013; USCB, 2012). While some areas are experiencing significant increases in population, others are experiencing significant decline and depopulation (Fox et al., 2009). Thus, population growth is a key component to consider when planning and designing water supply and distribution systems. An underestimation in water demand will make the system inadequate for the intended purpose; similarly, an overestimation will result to high capital and operation costs. It is therefore imperative that population-based information be considered when developing water demand forecasting models to ensure that the impacts of population growth are taken into account. Variables that have been used to depict population growth of a particular area in water demand forecasting include historical population, occupancy rate of dwellings, number of households or household connections and lot size.

2.2.2.3 Income levels

Water demand of high-income areas is generally higher than those of low-income areas considering that large-sized lots that typically require more water are often associated with more affluent households or societies (Almutaz et al., 2013; CSIR, 2019; Jacobs et al., 2004). Research has shown that the average yearly water consumption for high-income households can be up to 40 percent higher than the average consumption for low-income households (IPART, 2004). This can be explained by the fact that high-income households tend to have extra features connected to higher water consumption such as swimming pools, gardens and additional water-based appliances. However, it is important to note that considerable variability in consumption exists within income classes, and a substantial proportion of high water users are low-income households (Husselmann and Van Zyl, 2006; IPART, 2004; Jacobs and Haarhoff, 2007). Additionally, higher income households may be less responsive to water price, as it represents a smaller portion of household income, hence they may be more disposed to use high volumes of water. Conversely, low-income household tend to be more populated than high-income households, and thus may consume higher volumes of water. White et al. (2003) suggested that, although consumer income levels have some influence on water consumption patterns, there is no automatic link between increasing income (longitudinal) and rising water consumption. The authors further argued that high-income households tend to purchase newer and more efficient appliances which may result in a reduction in water use as they replace older units. Husselmann and Van Zyl (2006) investigated the independent effects of both stand size and income on water demand, using stand value as a surrogate for income. Their research found a definite trend of increasing water demand with increases in both stand size and stand value. The authors submitted that stand size is a good measure for defining an annual average daily demand range. These varying perspectives indicate that there is not yet a consensus in the manner with which income level affects demand. These arguments elucidate the need to further investigate the impacts of income level on water demand and to ensure that the impacts, if any, are accounted for appropriately. This can be achieved by considering income-based variables as potential model inputs during model development. Variables that have been used to represent income level of consumers in water demand forecasting studies include but are not limited to per capita gross domestic product (GDP), gross national product (GNP), per capita gross provincial product (GPP), human development index (HDI), stand size and stand value as well as hotel occupancy rate (Babel et al., 2007; Babel and Shinde, 2011; Firat et al., 2009; Husselmann and Van Zyl, 2006; Shabani et al., 2016).

2.2.2.4 Water price

Water price has been reported in the literature as the major tool for controlling demand (Arbués et al., 2003). This is based on the premise that consumers are likely to reduce their water use (e.g., for irrigating lawns and gardens, car washing, and swimming pools) if water prices are substantially increased, and that leaks, due to poor or faulty plumbing, that might be ignored under low prices would be repaired under high prices. A low water price could thus promote wastage. According to Coomes et al. (2010), economic theory predicts that residential water demand will be more inelastic towards price, as there are no alternatives for water in its basic household uses. However, given the viability of on-site supplementary household water sources such as groundwater abstraction, rainwater harvesting and greywater reuse in providing non-potable/second class water to residential consumers (Nel et al., 2017), it can be said that, even though an alternative product cannot replace water, supply of water from piped municipal system is replaceable. Some researchers have reported that low price elasticity abounds in residential areas, limiting the efficacy and worthiness of using water price as a conservation instrument (Gaudin, 2006; Gaudin et al., 2001). Coomes et al. (2010) however submits that there is a high likelihood that a significant change in price would affect consumption, especially if sound pricing structures are established to incentivize water conservation.

Non-payment of water bills is a huge problem, especially in developing countries like South Africa, where it seems to have become an established ‘norm’. Non-payment often depicts a demonstration of consumer dissatisfaction with current water services, affordability (i.e. inability to pay, especially in low-income households), an “entitlement culture” or a “culture of non-payment” (Fjeldstad, 2004; Vásquez, 2015). Considering that researchers have differing opinions about water being a commodity versus a human right (Bakker, 2007; Kornfeld, 2012), the non-payment of water bills add a new dynamic to the debate on the impact of price on water demand as the price of water becomes irrelevant if consumers do not pay for the water. For example, it is argued that the socioeconomic status of many low-income households makes them unable rather than unwilling to pay, hence the need for free basic services to the poorer segments of the population and/or a lowering of the rates (Fjeldstad, 2004). From this perspective, water price may not provide a true reflection of actual water demand.

2.2.3 Application of EC techniques for WD forecasting

Evolutionary computation techniques (also referred to as "evolutionary algorithms") belong to a class of solution methods referred to as metaheuristics that are inspired by observations of natural phenomena for a robust exploration and exploitation of a solution space, while integrating variables of structured randomness to find near-optimal solutions (Maier et al., 2014). EC techniques are a unique set of search methods inspired by the principle of biological evolution; yielding outcomes that are based on a collective learning process from a population of possible solutions. Bi et al. (2016) itemized the advantages of EC techniques over traditional deterministic approaches to include (i) superior potential in exploring the entire search space, resulting to higher possibility of achieving near-optimal solutions; (ii) ease of integration with any simulation model; and (iii) greater degree of adaptability in solving complex multi-objective problems that are typical of those concerned with water resources.

EC techniques have become increasingly popular in the field of water resources and have found application in water demand forecasting. Applications and development of EC techniques in water demand forecasting can be categorized into two parts – (i) predictive modelling; and (ii) optimization modelling. In predictive modelling, EC techniques are either used directly in developing water demand forecasting models (Nasseri et al., 2011; Yousefi et al., 2017) or as optimization algorithms in intelligent models such as artificial neural network (Perea et al., 2015; Varahrami, 2010). In solving optimization problems, EC techniques are being widely applied for optimizing model parameters of other modelling techniques (Bárdossy et al., 2009; Di Nardo et al., 2015) and in estimating the coefficients of functions (Ehteram et al., 2017; Qu et al., 2010). An overview of EC techniques that have found application in water demand forecasting is presented in the next section. These techniques are categorized based on their mode of application as found in the literature.

2.2.3.1 Predictive modelling

The main objective of predictive modelling is to directly estimate a response (output) from a defined set of explanatory variables (input), or to indirectly drive the choice of decision rules (Steyerberg, 2008). This typically results to a generalized functional relationship as presented below:

$$\boldsymbol{S} = (\boldsymbol{D^p})$$

where $\boldsymbol{D^p}$ is a p-dimensional input vector consisting of explanatory variables $\boldsymbol{d_1}$, $\boldsymbol{d_i}$,… $\boldsymbol{d_p}$, and $\boldsymbol{S}$ is the output variable. In water demand forecasting, values of $\boldsymbol{d_i}$ may include water demand values with different time lags and the value $\boldsymbol{S}$ is typically the water demand in the succeeding period (Wang et al., 2009).

The development of predictive models require the application of a series of processes which comprise data collection, data pre-processing,, input data selection, data splitting, determination of model type and model architecture, model training and testing as well as model performance evaluation (Maier and Dandy, 2000; Oyebode, 2014). These processes are ultimately targeted at achieving optimal model predictive accuracy and reliability.

(a) Direct application of EC techniques in predictive modelling

Genetic Programming (GP)

Genetic programming (GP), developed by Koza (1994) is an EC technique and population-based search founded on the principle of natural selection (survival of the fittest). GP is a member of the evolutionary algorithm (EA) family and an extension of genetic algorithm (GA) – an evolutionary-based optimization technique which seeks to arrive at the global optimum of a function. Unlike GA which operates on a set of binary strings, GP genetically breeds a population of candidate solutions (computer programs), via genetic operations like reproduction, crossover and mutation, to evolve solutions or models that give the best representation of a system (Oyebode and Adeyemo, 2014). The model structure and coefficients are simultaneously determined by optimizing a population of computer programs based on a fitness function which accounts for how well a given computational task is solved.

The main steps in the implementation of the GP algorithm, following Sette and Boullart (2001), are summarized below and illustrated in Figure 2-3:

- Step 1: Generation of a random population of candidate solutions.
- Step 2: Fitness evaluation of all candidate solutions in the population. Moreover, if the fitness function is reached, the algorithm is terminated and the computer program with the highest fitness is selected as the ultimate result.
- Step 3: Replacement of the current population by a new population via a probabilistic application of genetic operators (reproduction, crossover and mutation).
- Step 4: Return to step 2.

Details on the working principles and applications of GP in the water resources modelling domain can be found in Wang et al. (2009) and Oyebode and Adeyemo (2014). GP has been applied successfully especially in symbolic regression where the main objective is to define a functional relationship between explanatory and target variables (Guven and Kişi, 2011; Londhe and Charhate, 2010; Oyebode, 2014). Furthermore, the ability of GP to “offer a good compromise between accuracy and complexity” makes it suitable for solving complex real-world problems such as predictive modelling (Chadalawada et al., 2017).

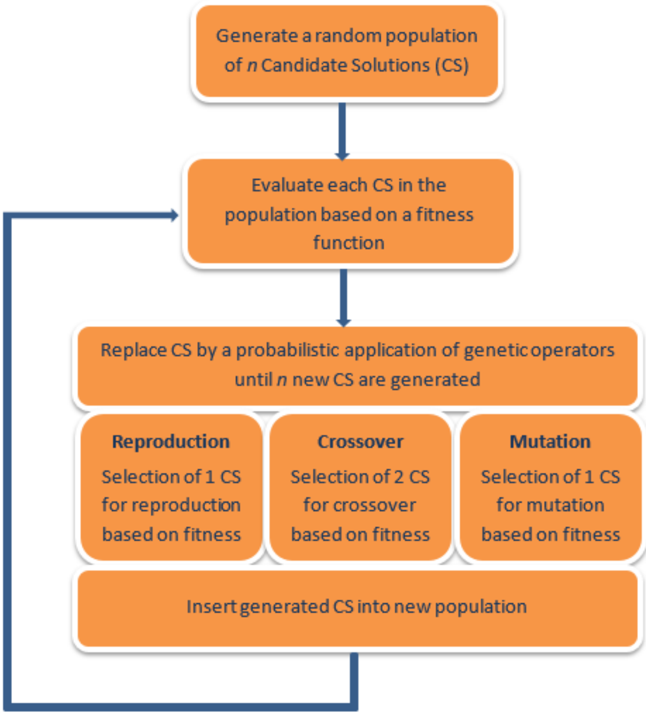


Figure 2-3: Steps in GP implementation process. Source: Sette and Boullart (2001)

GP has been successfully applied and validated in the field of water demand forecasting. Wu and Yan (2010) investigated the ability of GP to evolve daily water demand forecast models for a district water system located in a large demand monitoring zone in the UK. The daily forecast models were developed using various combinations of weather variables and historical water consumption data. Results showed that models produced were physically interpretable, easy to implement and provided a good representation of the complex and nonlinear input-output relationship between the variables utilized.

In form of a hybrid model, Nasseri et al. (2011) coupled an extended Kalman filter (EKF) and GP for dynamic monthly water demand forecasting monthly for Tehran, Iran. A time series modelling approach was adopted in the study with lags of previous water consumption considered as initial inputs. The EKF was employed to deduce latent variables based on results obtained from initial GP simulations for forecasting purposes. The proposed hybrid model was found capable of replicating, with precision, the pattern of water demand for the city, thus, providing a means for reducing the risks associated with online or dynamic water demand forecasting.

Fagiani et al. (2015) also applied a combination of an EKF and GP to forecast domestic water and natural gas consumptions on an hourly basis using heterogeneous data comprising different resource types. The study investigated the performance of the hybrid model at multiple time resolutions and the mutual correlation between the resource types. Results for water consumption prediction showed high forecast accuracies across different time resolutions. The authors however noted that improved model performance can be achieved if the sample size is increased and additional explanatory variables considered.

Generally, results of studies focused on the application of GP in the field of predictive modelling including water demand forecasting have confirmed its aptitude for problem-solving in terms of generation of solutions that are accurate, transparent and structured in representing complex and nonlinear real-world processes. However, GP may encounter challenges in finding constants, as there is a tendency to create more intricate functions as the forecast horizon increases (Giustolisi and Savic, 2006).

Gene Expression Programming (GEP)

Gene expression programming (GEP), advanced by Ferreira (2001) is a variant of GA and GP, with similarities in terms of initialization of populations of candidate solutions, selection of based on fitness, and application of genetic operators. GEP however differs from GA and GP in the manner with which it evolves a new generation of candidate solutions. In GAs, the candidate solutions are in a form of linear strings of fixed length referred to as “chromosomes”, while in GP, the candidate solutions are nonlinear tree-like structures of different sizes and configurations. In GEP, however, the individuals are encoded as linear threads of constant length chromosomes, enabling the genetic operators function at chromosome level, thereby resulting to a remarkably simplified method for achieving genetic diversity (Martí et al., 2013). This distinctive, multi-genic characteristic of GEP in turn allows for the evolution of more robust programs composed of multiple subprograms (Ferreira, 2006). GEP therefore exerts superiority over the GP algorithm in 100-10,000 times (Karimi et al., 2016). A schematic representation of the GEP algorithm is presented in Figure 2-4. A detailed explanation on the implementation of GEP can be found in Ferreira (2001).

GEP has been reported to be effective in finding explicit formulations of relationships governing different physical phenomena, making it useful for validating popular physical relationships, in knowledge mining and for improving conventional science- and engineering-based theoretical frameworks (Guven and Aytek, 2009; Martí et al., 2013).

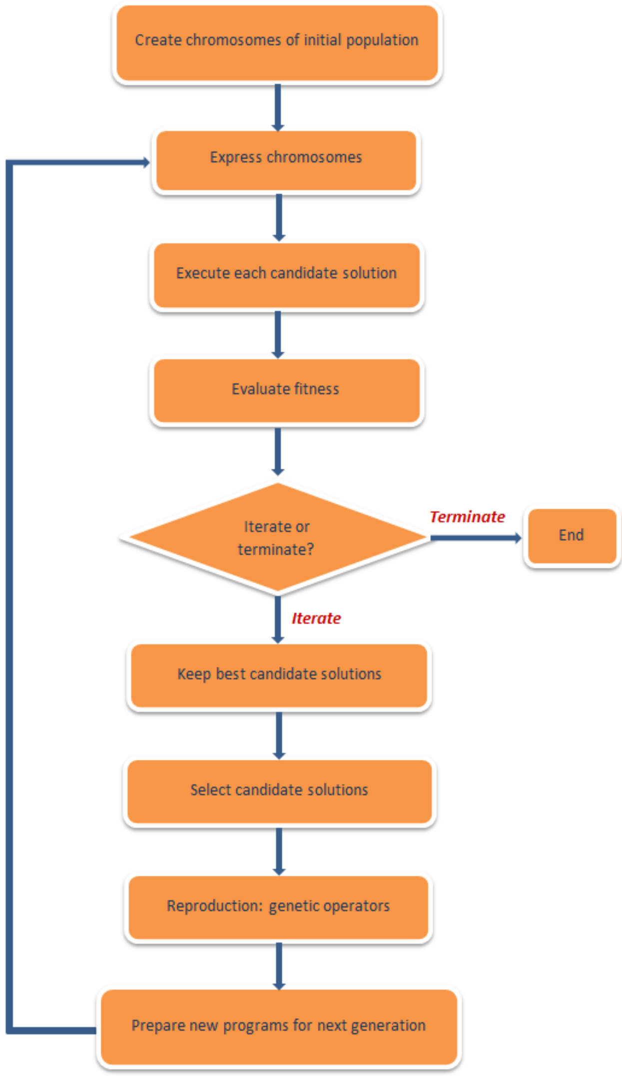


Figure 2-4: Steps in GEP implementation process. Source: Martí et al. (2013)

Only a few studies related to the application of GEP in water demand forecasting exist in the literature. This can be attributed to the fact that it is a relatively new EC technique. Yousefi et al. (2017) coupled GEP with wavelet transform in developing models for long term forecasting of water demand in the City of Kelowna, British Columbia, Canada. The performance of the GEP models was complemented with a three-levelled wavelet transform comprising two transfer functions. The input vector space of the model was populated with different combinations of variables including temperature, precipitation and water demand. Average mutual information (AMI) was employed to determine the optimal number of lags for each input variable. Results showed that GEP models can be highly sensitive to wavelet decomposition if all combinations of suitable lag times are carefully selected. The authors recommended GEP as one of the emerging techniques in water demand forecasting that should be giving more attention. This suggestion is supported by findings from earlier studies conducted by Wu and Yan (2010) and Shabani et al. (2016) wherein GEP models were found to be effective for constructing short- and medium-term water demand forecasting models; with an average forecast accuracy of above 90% reported by Wu and Yan (2010).

In a recent study, Shabani et al. (2018) proposed a hybrid model which comprises a GEP-supervised and K-means clustering (unsupervised) learning process for short-term water demand forecasting. The hybrid model was verified using hourly water demand data for the City of Milan. The unsupervised module of the hybrid model was applied to organize daily water consumption in six distinct clusters to account for seasonality and recurring patterns while GEP was employed to evolve explicit water demand forecast models for each of the clusters. AMI was used for determining the most suitable lags of the water demand time-series that will serve as model inputs. Results show that the hybrid model produced accurate forecasts across the six distinct clusters, with the 1-hour lead time models considerably outclassing models based on other sampling frequencies. This study further confirms the "ease of integration" attribute of EC techniques, and that techniques like GEP could be coupled with unsupervised learning algorithms to improve the forecast accuracy of water demand models.

(b) EC techniques as optimization algorithms in intelligent models

Intelligent models are models that apply the basic working principles of the human nervous system in decision-making. These models consist of a large pool of processing units (referred to as neurons) which receive, process and send information to each other over a large number of weighted connections. They are therefore referred to as artificial neural network (ANN). ANN operates in a similar fashion to the human brain as experiential knowledge is acquired through a search process aimed at determining an optimal set of weights for the connections and threshold values (biases) for the neurons (Elshorbagy et al., 2010). Each individual neuron computes an output, based on the weighted sum of all its inputs, according to a nonlinear function called the activation or transfer function (Kalteh et al., 2008). Finding the optimal weight values within the network is considered as key to having a well-trained ANN. A learning algorithm is usually employed to supervise an iterative adjustment of the connection weights thereby minimizing the error measure between the network output and target outputs. A network with an output ($\boldsymbol{y^P}$), inputs ($\boldsymbol{z_k}$), $\boldsymbol{k = 1, \ldots K}$, weights ($\boldsymbol{w, v}$) and number of hidden neurons ($\boldsymbol{J}$) can be represented using the following expression:

$$\boldsymbol{y^P = v_0 + \sum_{j=1}^{J} v_j f\left(w_{j0} + \sum_{k=1}^{K} w_{jk}\, z_k\right)}$$

ANN is, thus, an approximation function mapping inputs to outputs, thereby developing learning, adaptive and generalization features. These features make ANN suitable for solving a wide variety of problems relating to input and output variables in complex systems such as water demand forecasting. ANN can be classified as single, bilayer and multilayers according to the number of layers, and as feed-forward, recurrent and self-organizing according to the direction of information flow and processing (Govindaraju, 2000). A detailed review of procedural steps, implementation techniques and applications of ANN can be found in Oyebode and Stretch (2019). The authors provided useful insights into how the performance of ANN can be improved and potential areas of application that are yet to be explored in water resources.

Over the past decades, ANN has been widely applied in different domains and made remarkable developments. Despite its prominence, it is widely accepted that ANN is prone to certain problems which include difficulty in network training or learning process, over-parameterization and poor generalization (Ding et al., 2013; Oyebode, 2014). Although a number of optimization methods based on gradient descent (e.g., back-propagation, Levenberg-Marquardt and conjugate descent algorithms) have been used in ANN model development, these methods have been reported to be susceptible to being trapped in local optima and could also generate negative values, especially if the error surface is fairly rugged (Kisi and Cigizoglu, 2007; Maier et al., 2010).

To solve the aforementioned ANN-inherent problems, EC techniques are now being integrated with ANN. The combined use of ANN and EC techniques has given rise a new category of ANN referred to as evolutionary ANNs (EANNs) (also known as neuro-genetic models). EC techniques have been principally applied in EANNs for evolution of connection weights, model architectures and learning rules (Ding et al., 2013; Yao, 1993). Studies have shown that EANNs do not only showcase better learning capabilities than other ANNs, but also exhibit robustness in the design and implementation of ANNs, thereby enhancing overall model performance (Ding et al., 2013; Kim et al., 1999).

The successful applications of EANNs in water demand forecasting have been reported in a number of studies. EC techniques that have found application in ANN optimization include GA and differential evolution (DE). The following sections provide a brief description of each technique followed by their applications in ANN development with water demand forecasting in context.

Genetic algorithms

GA is the most widely used global optimization technique (Nicklow et al., 2009), and is based on the rules of evolution and natural selection. It initializes with a preliminary population of candidate solutions, distinguishing each as a chromosome, and thereafter computes and grades each solution based on a fitness function. GA subsequently performs three genetic operations namely selection, mutation and crossover to create a new population of offspring solutions which may be superior than their parents (Behera et al., 2015). GA thus employs this systematic approach to achieve continuous improvement of individual solutions, and ultimately evolve an optimal or near optimal solution after successive iterations.

A schematic representation of the fundamental methodologies for implementing GA is shown in Figure 2-5. The GA methodology is regarded as the framework upon which several other single and multi-objective EC techniques were developed. Successful applications of GAs in water resources include vital areas like design and operation of water distribution systems, urban drainage and sewer systems, water supply and wastewater treatment applications, hydrologic and fluvial systems, and groundwater systems design. A review of the extensive applications of GA in water resources can be found in the literature (Nicklow et al., 2009).

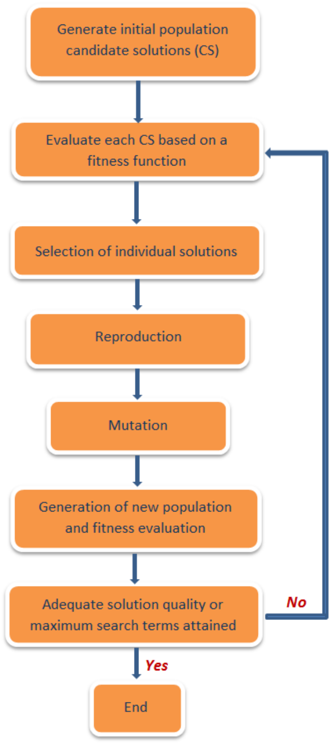


Figure 2-5: Fundamental methodologies for implementation of GA. Source: Nicklow et al. (2009)

GA has been widely employed in optimizing the performance of ANNs in water demand forecasting. Kim et al. (2001) employed the neuro-genetic approach based on GA to develop ANN models for daily water demand forecasting for the City of Seoul, South Korea. Nine modelling scenarios were implemented using different combinations of explanatory variables and their associated lags. The predictive performance of the GA-trained ANNs was compared to those trained using a back-propagation algorithm for the nine scenarios. Results showed that the neuro-genetic approach performed better than the back-propagated ANNs in all nine models; suggesting GA as an effective and reliable technique for training ANNs in the water demand forecasting domain. Varahrami (2010) applied GA to ANN forecasting of short-term water demand, comparing its performance with that of a conventional back-propagated ANN. It was found that the ANN-GA model consistently outperformed the back-propagated ANN, showing higher precision on unseen data sets.

In an irrigation water demand forecasting study conducted by Perea et al. (2015), GA was applied to determine the optimal architecture of an ANN model. Twelve ANNs were trained with different gradient-based optimization techniques and applied to predict water demand one-day ahead. A GA-based multi-objective algorithm – Non-dominated Sorting Genetic Algorithm II (NSGA II) was employed to optimize the model architecture of the ANNs in terms of computational speed and forecast accuracy. The ANNs were tested with actual data recorded in the water distribution network of a real irrigation district in Spain. Results showed that the GA was capable of evolving an ANN model with optimal sets of architecture parameters while simultaneously maximizing model predictive accuracy. The authors argued that the appropriate method for achieving an optimal generalization in ANNs is to employ GA in determining the optimal network architecture. This study therefore elucidates the crucial role of EANNs in agricultural water management and in guiding the development of nexus-sensitive policies, considering the inextricable link between water and food security.

By coupling a GA to a modified adaptive particle swarm optimization (APSO) algorithm, Mohammadi et al. (2014) proposed a new hybrid evolutionary algorithm to simultaneously determine the architecture and network parameters of radial basis function neural networks (RBFNNs). The GA was applied to optimize the model architecture (input variables and hidden layer neurons) of the RBFNN while the APSO supervised the learning process thereby determining the network parameters which include centers, width and weights of the RBFNN. The performance of the hybrid algorithm was initially analyzed comparatively with several benchmark time series modelling and algorithms. The proposed hybrid model was subsequently extended to forecast emergency supplies, including water demand, after earthquakes in Iran. Simulation results indicated that proposed GA-APSO algorithm demonstrated better forecast accuracy with computational efficiency. Findings from this study imply that the performance of ANNs can be improved through learning methods that simultaneously adjust the entire set of model parameters. More importantly, these findings demonstrate the applicability of EANNs to disaster management thereby ensuring water security and strengthening the resilience of vulnerable communities, and thus directly contributing to sustainable development. This further suggests that the adoption of EC techniques offer great potential in terms of improving individual and institutional capacity for achieving a post-2015 development agenda of the UN (UNESCO, 2015), and in reducing the impacts of water-related disaster risks.

Other water demand-based studies wherein GA has been used for ANN development include forecasting of irrigation water demand (Pulido-Calvo and Gutierrez-Estrada, 2009); regional water demand (Papageorgiou et al., 2016); and domestic water demand (Rangel et al., 2017; Walker et al., 2015).

Differential evolution (DE)

DE, proposed by Storn and Price (1997), is an EC-based optimization technique which evolves candidate solutions in a similar manner as the GA. DE however differs from GA in the manner with which the mutation operation is executed. In DE, mutation precedes crossover. The mutation operation entails generating a mutated population by adding the weighted difference between two random candidate solutions (Zheng et al., 2012). Crossover operation is thereafter introduced to combine the mutated population with a target population to evolve a trial or experimental population (Sedki and Ouazar, 2012). Parameters that influence the operation of the DE algorithm are: the population size (NP), the mutation scale factor (F), and the crossover constant (CR) (Behera et al., 2015). Research has however shown that the performance of DE is significantly governed by F and CR (Zheng et al., 2012). The implementation steps for the DE algorithm are presented in Figure 2-6.

DE has received significant attention due to its consistency in converging towards the global optimum solution, and has been used in solving several complex optimization problems (Qian et al., 2009; Zheng et al., 2012), including water resources applications (Kişi, 2010; Suribabu, 2010; Tanyimboh and Seyoum, 2016; Vasan and Simonovic, 2010).

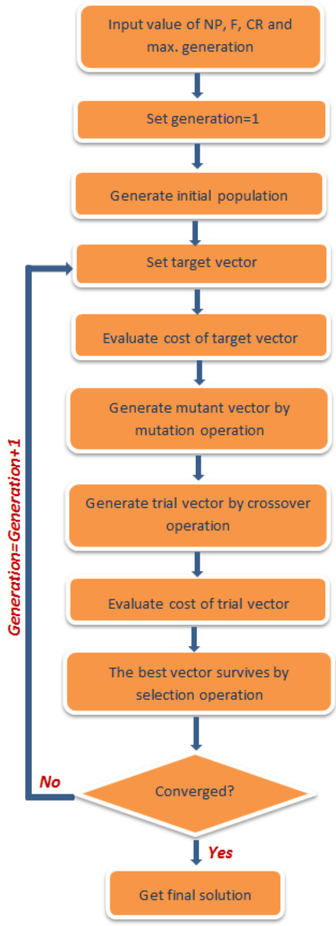


Figure 2-6: Steps for implementing the DE algorithm. Source: Zheng et al. (2012)

Unlike its GA counterpart, the DE algorithm is yet to be widely applied in water demand forecasting. Within the scope of the authors' knowledge, the only application of a DE-trained ANN in water demand forecasting was reported in a study conducted by Qu et al. (2010). The study entailed the use of DE to optimize a generalized regression neural network (GRNN) to forecast annual industrial, agricultural and domestic water demands in Yellow River Basin in China. DE was specifically employed to optimize the value of the smoothing parameter which is known for influencing the prediction performance of GRNNs. The key explanatory variables considered as input parameters include industrial output, agricultural output, irrigation quota as well as urban, rural and livestock population. The DE-GRNN model was used in making water demand projections for years 2010, 2020 and 2030. Results showed that the model was capable of assimilating the complex non-linear relationships between the three different water demands and their respective explanatory variables, producing reasonable and comparable forecasts with those made by BP-GA and GRNN-GA forecast models.

Although the potential of DE has not be fully exploited in the area of water demand forecasting, it has however figured prominently in other areas of water resources like river flow forecasting (Piotrowski and Napiorkowski, 2011), reservoir inflow forecasting (Oyebode and Adeyemo, 2014), reservoir optimization (Olofintoye et al., 2016), sediment yield modelling (Kişi, 2010) and optimization of water distribution networks (Suribabu, 2010; Zheng et al., 2012). The meagre application of DE in water demand forecasting studies is evidential to the assertion of Ghalehkhondabi et al. (2017) that the potential of soft computing techniques have not been fully utilized in the field of water demand forecasting. There is a therefore a need for DE to be given more attention in the development of intelligent models for water demand forecasting, to foster the realization of the full potential of modelling complex water demand processes using soft computing techniques.

Other EC techniques that have found application in other areas of water resources but not in water demand forecasting include evolution strategies (ES) (Mirghani et al., 2009) and evolutionary programming (EP) (Muleta and Nicklow, 2004). ES and EP have however found application in groundwater modelling and watershed management studies respectively.

2.2.3.2 Extent of application of EC techniques in predictive modelling of water demand

A review of existing literature on the application of EC techniques in water demand forecasting was conducted via a search on reputable academic databases, namely, Google Scholar and Scopus using keywords including “water”, “prediction” or “forecast”, “demand” or “consumption” and “evolutionary” or “evolutionary algorithm”. A summary of the extents of application of EC techniques for water demand forecasting based on the following themes: forecast technique, location, forecast periodicity and explanatory variables is presented in Table 2-1 and Figure 2-7. It can be observed that GA is the leading EC technique for the water demand forecasting. This is followed by GEP and GP in descending order. The ES and DE are the least applied techniques for water demand forecasting. With regards to periodicity of forecast, majority of the reviewed literature have performed water demand predictions based on daily usage/consumption. From Figure 2-7, it is obvious that most studies were carried out in Europe. No study in Africa was listed as at the time of this review. This is a pointer to the fact that water demand forecasting studies have not been given priority in most African countries. This is however of priority based on the UN SDGs. For example, studies based on EC techniques could assist in the planning and management of water in water-stressed countries like South Africa. Furthermore, majority of the literature reviewed employed historical water demand as the sole explanatory variable, while a few considered a combination of historical water demand and weather variables. Only a couple of studies considered the impact of socioeconomic factors like population and income levels.

Table 2-1: Application of EC techniques in water demand forecasting

SI. no.	Author & Year	Forecast technique	Location (by continent)	Forecast periodicity	Explanatory variables
1.	Kim et al. (2001)*	GA	East Asia	Daily	HWD, W
2.	Pulido-Calvo and Gutierrez-Estrada (2009)*	GA	Europe	Daily	HWD, W
3.	Qu et al. (2010)*	DE	East Asia	Annual	HWD, P
4.	Wu and Yan (2010)	GP, GEP	Europe	Daily	HWD. W
5.	Fagiani et al. (2015)	GP	North America	Hourly	HWD, W, O
6.	Varahrami (2010)*	GA	West Asia	Monthly	HWD
7.	Nasseri et al. (2011)	GP	West Asia	Monthly	HWD
8.	Mohammadi et al. (2014)*	GA	West Asia	Daily	HWD
9.	Romano and Kapelan (2014)*	ES	Europe	Hourly, Daily	HWD
10.	Perea et al. (2015)*	GA	Europe	Daily	HWD, W
11.	Walker et al. (2015)*	GA	Europe	Hourly, Daily	HWD
12.	Shabani et al. (2016)	GEP	North America	Monthly	HWD, W, I
13.	Papageorgiou et al. (2016)*	GA	Europe	Daily	HWD, W, P
14.	Yousefi et al. (2017)	GEP	North America	Monthly	HWD, W
15.	Rangel et al. (2017)*	GA	Europe	Hourly, Daily	HWD
16.	Shabani et al. (2018)	GEP	Europe	Hourly	HWD

*ANN optimization; HWD: Historical water demand; W: Weather-based; P: Population; I: Income-based; O: Others

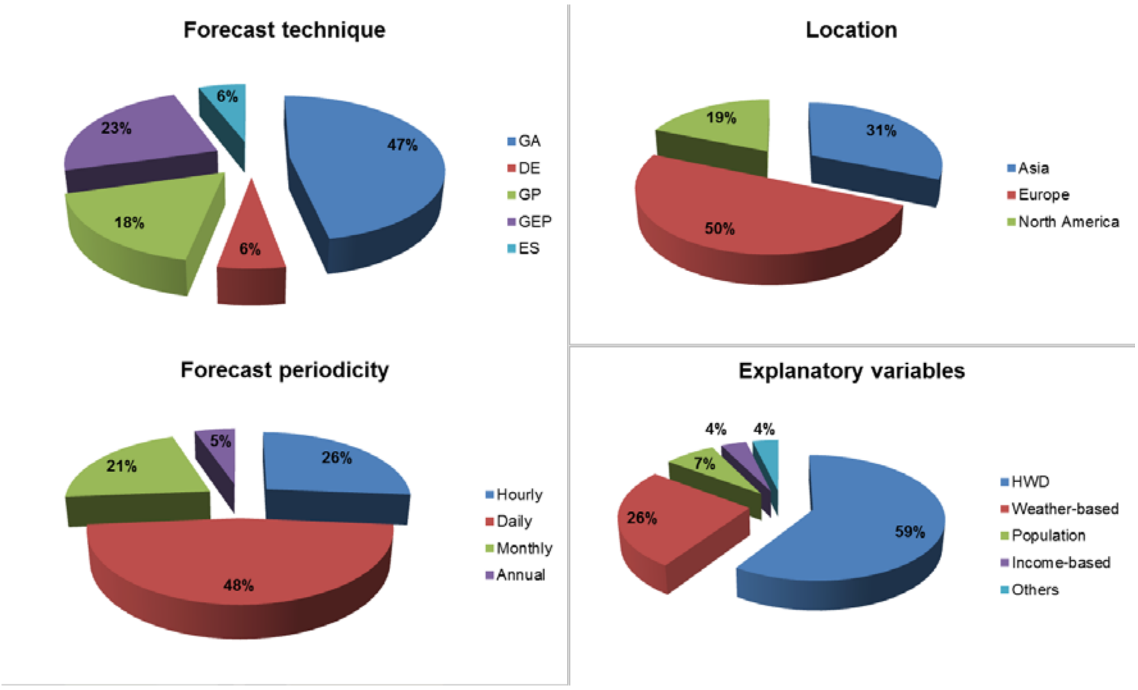


Figure 2-7: Overview of the application of EC techniques in water demand forecasting

2.2.3.3 Optimization modelling

Owing to the limitation of water resources and increasing water demand, different optimization strategies and techniques are being formulated and implemented to manage water supply, demand and consumption profiles to achieve the ultimate aim of water conservation (Tabesh and Hoomehr, 2009). Optimization of water distribution networks (WDNs) and water demand forecasting is therefore closely connected. Since this review focusses on the application of EC techniques for water demand forecasting, it is therefore essential to present a brief review of popular EC techniques used in optimizing WDNs as provided in this section. EC techniques that have been successfully applied for water resources optimization include GA, DE, ES and EP.

Savic and Walters (1997), in one of earliest applications of evolutionary optimization techniques in water resources, developed and applied a GA-based model (GANET) to a combinatorial optimization problem of a least-cost design of WDNs. The model was applied for design of a new three-loop WDN and for expansion of an existing system. The objective function adopted in the study is the minimization of the cost of the solution (i.e., overall costs of the pipes within the network), while the minimum pipe pressures were used as constraints to identify the optimal network design. Solutions produced by the GANET model was compared to those previously published in the literature. Results show that GANET produced suitable network designs without needless limitations imposed by split-pipe or presumptions of linearity. It was concluded that the GA-based technique can be effortlessly modified to benefit the design process of completely new and existing water networks. The study formed the basis for application of GA to solving least-cost design problems and has been widely implemented in several network design and optimization studies (Tanyimboh and Seyoum, 2016).

Other areas of application of GA in WDN as reported in the literature include consumption management and leakage estimation (Di Nardo et al., 2015; Tabesh and Hoomehr, 2009), optimal irrigation water allocation (Hassan-Esfahani et al., 2015; Lewis and Randall, 2017), development of IWRM decision support systems (Nouiri, 2014; Zheng et al., 2012), optimization of reservoir operations (Chang et al., 2010; Huang et al., 2002), and pumping scheduling (Atkinson et al., 2000; Van Zyl et al., 2004).

DE is increasingly becoming more popular in WDN optimization, and has been applied in solving many real-world problems. For example, Vasan and Simonovic (2010) developed a simulation-optimization model (DENET) by integrating a DE algorithm into a hydraulic simulation model - EPANET, for optimal design of WDNs. DENET was applied to benchmark two WDN problems in two distinct regions for minimization of network cost and maximization of network reliability, and its performance compared with those reported in previous researches. It was reported that DENET produced a parallel performance to those reported in earlier studies in terms percentage of convergence to global optimum, thereby offering feasible cost-effective network solutions while also maximizing network resilience. The authors submitted that the simplicity and robustness of the DE are key attributes that makes DE a promising optimization technique for design and rehabilitation for WDNs.

In a real-time reservoir optimization study, Olofintoye et al. (2016) coupled an ANN model with a novel combined Pareto multi-objective differential evolution (CPMDE) for flow forecasting and mathematical optimization of hydropower generation from the Vanderkloof reservoir, South Africa. While the ANN was used for real-time forecasting of flow into the reservoir, CPMDE was implemented to identify practicable solutions that offer optimal daily operating policies of the reservoir. Results showed that the proposed ANN-CPMDE model produced solutions that offer policies that trade-off power generation against storage depletion and reservoir storage head drop, thereby balancing short-term and long-term objectives. In conclusion, the authors argued that there is a significant degree of potentiality in the adoption of cutting-edge optimization systems like DE for provision of low-cost solution methodology, appropriate for sustainable real-time optimization of reservoir operations.

Zheng et al. (2012) proposed a self-adaptive DE (SADE) algorithm and integrated it into an EPANET network simulator for least-cost single objective WDN optimization problems. Unlike in other EC techniques wherein the termination criterion (i.e., computational requirements) for solving an optimization problem is predefined, the SADE algorithm employs a new convergence benchmark based on the coefficient of variation of the objective function values for the existing DE population of solutions. Thus, the SADE algorithm terminates when individuals in the DE find the same or extremely close final solutions, thereby preventing the challenges of computational redundancy and deficiency associated with predefining the computational requirement. The SADE algorithm also allowed for automatic tuning of the governing parameters (F and CR) via an evolution process, and in so doing, it drastically reduces the effort required to determine the optimal DE parameters. The efficiency of the proposed SADE algorithm was tested using four WDNs as case studies. Results show that the SADE algorithm produced optimal least cost solutions, with greater efficiency than other EAs; showcasing great potential in terms of percentage of best solutions found and convergence speed.

EP and ES are two other EC techniques that have been applied for optimization of WDNs. EP and ES are noticeably different from GA and DE due to their reliance on mutation as a principal genetic operator. They are therefore referred to as mutation-based EC techniques (Muleta and Nicklow, 2004). Both EP and ES typically apply a Gaussian mutation for real-valued functions, contemporary modifications however exhibit more diversity (Yang, 2009). Although EP and ES share common attributes, the major distinguishing factor in their operation is that there is no recombination or crossover operation between individuals in EP (Bäck et al., 1993; Yang, 2009).

Romano and Kapelan (2014) developed a self-learning water demand forecasting technique with the aim of promoting near real-time management of smart WDN. The novel technique was implemented in a demand forecasting system (DFS) which comprise an intelligent model, and tested using information from three District Metered Areas and a Water Supply Zone in the UK. An ES algorithm was adopted for determination of the optimal input structure and parameters of the intelligent model in the DFS. Highly accurate forecasts (with Nash–Sutcliffe efficiency >0.9), were reportedly achieved by the DFS. The authors attributed the accurate forecasts to the ability of the ES algorithm to identify the best input structure and parameter sets; utilizing less number of explanatory variables.

EC techniques have found application in the selection and optimization of demand-side management approaches which are often targeted at reducing the volume of water being drawn from a network by scaling-down end-user demand. Evolutionary optimization techniques have been used to simultaneously assess the performance of different technological options and strategies using multiple quantitative and qualitative sustainability criteria and indicators; thereby facilitating decision-making (Makropoulos et al., 2008). These technological options and strategies may include installation of water saving devices or appliances (e.g., low-flush toilets and low-flow shower heads and taps), best management practices such as rainwater harvesting, greywater reuse, as well as behavioral changes regarding water usage.

Makropoulos et al. (2008) developed a decision-support tool – Urban Water Optioneering Tool (UWOT) for planning of water cycle management for new urban developments, including sustainable option selection in water supply and demand management. A GA was embedded in UWOT to drive an optimization process which was aimed at identifying the most appropriate and feasible water-saving solutions based on technical, environmental, social and economic objectives. The UWOT modelling framework was thereafter applied to identify water-saving technological options for a new urban development in the UK. Results show that the set of technological options evolved by the GA enabled trade-offs across a series of sustainability indicators. Moreover, the GA-based solutions provided a more significant reduction in water demand compared to end-user-based optimization techniques. These results further prove the applicability of EC techniques to address potentially conflicting views and priorities via a rapid assessment of alternative what-if scenarios, and can ultimately serve as an anchor for the delivery of integrated, sustainable water management for new developments.

Generally, it can be said that EC techniques have achieved great successes in optimization and management of water supply and demand systems. Key areas of water supply and demand management wherein EC techniques have been successfully applied can be summarized to include optimization of system components during the planning and design stage, operational optimization such as pumping scheduling, real-time operations, leakage estimation, network rehabilitation and water demand analysis as well as in demand-side management.

This review shows that EC techniques are increasingly gaining recognition due to their aptitude to fully explore the search space, and greater tendency of producing optimal or near optimal solutions when dealing with complex, nonlinear, and discrete optimization problems. In addition, the ease with which they can be linked to any simulation model further gives them an edge over other optimization techniques. Considering the improved interoperability of EC techniques, we posit that they could extend their significance beyond advisory roles, and be positioned as an effective tool for developing proper standard operating procedures as recommended in the 2016 World Health Organization (WHO) report on Water Safety in Distribution Systems (WHO, 2016). EC techniques could thus be instrumental to the making of sustainable policies in the water sector. One of the key pathways wherein the application of conventional EC techniques could be enhanced is by employing them in establishing a synergy between ensuring optimality in the water network management and the wider issues of society such as poverty, economic growth (productivity) and welfare. To this end, a comprehensive framework, that is capable of incorporating EC techniques and extending their influence to assessing the impact of policy actions (based on optimized variables) on society, is thus advocated. Such framework should be capable of socializing EC techniques by ensuring that their operations harmonize optimization with justice and equity. The next section thus presents a discussion on the need for balance and equity in adopting EC techniques to address the UN SDGs.

2.3 SECTION B: EXTENDING THE CAPABILITIES OF EC TECHNIQUES IN WATER DEMAND MANAGEMENT

From the review presented in the last section, it is evident that vast research efforts have been directed towards the development, improvement and application of EC techniques in solving water demand and allocation problems for over the last two decades. Results from our survey shows that the EC techniques are robust and flexible when applied appropriately. Although EC techniques have been extensively researched, there are still vast opportunities that can be explored regarding water demand management. In many parts of the world, there are emerging issues that EC techniques must evolve to address. One of such emerging themes is the need for equitable and sustainable water resource management; including conservation and allocation (UN, 2010). This is inevitable due to scarce water resources and disparity in income class. Furthermore, emerging and multidisciplinary research is fast redefining the disciplinary confines of water resource management. Moreover, sustainability centers, not only around technical, environmental and economic circles (as often analyzed by water engineers/ modellers), but also and more importantly, the social aspects of water. As such water resource (including demand) management frameworks must be robust enough to accommodate the impacts (e.g. new complexities, system dynamics, uncertainties, nonlinearities, etc.) associated with including social factors and/or objectives in water demand modelling and management. These emerging subjects pose important challenges that motivate the need for extending the applications of EC techniques to water justice and equity. As a contribution, we propose a novel integrated water demand and management modelling framework (IWDMMF) that will enable water policy-makers to assess the wider impact of water demand management decisions through the principles of egalitarianism, utilitarianism, libertarianism and sufficientarianism. The next section presents a framework that has the potential of extending the capabilities of EC techniques in water demand modelling. This is necessary to ensure that future water demand optimization and allocation models allow for water policy decisions that incorporate equity and justice. The section starts by discussing the role of UN SDGs in improving water demand management and modelling. The SDGs are then linked to the proposed framework.

2.3.1 The United Nations Sustainable Development Goals (SDGs)

As a successor to the millennium development goals (MDGs), the Sustainable Development Goals (SDGs), otherwise known as the Global Goals according to UNDP (2018) "are a universal call to action to end poverty, protect the planet and ensure that all people enjoy peace and prosperity." To this effect, some researchers have given attention to the assessment of various components that make up the SDGs (Casini et al., 2019; Escrig-Olmedo et al., 2017; Lehner et al., 2018; Saladini et al., 2018; Siksnelyte et al., 2018). For example, Saladini et al. (2018) developed a monitoring tool based on 12 sustainable development indicators for the Mediterranean region. This tool was developed specifically for monitoring progress towards food security (SDG 2) and sustainable water management (SDG 6), and to keep track of the impact generated by projects promoted by Partnership for Research and Innovation in the Mediterranean Areas (PRIMA) in the region. The authors suggest modelling as an essential component in monitoring the progress of sustainable development. To foster the realization of the 2030 UN SDGs, Casini et al. (2019) proposed a step-by-step procedure based on the fuzzy set theory for the construction of a multidimensional index for sustainability assessment. The study was focused on agro-food sustainability, using unique composite or multidimensional indicators that allow for identification of key independent factors that determines the sustainability of a system. The framework was applied to assess the progress of 17 countries in terms of sustainable development with country scores calculated for each independent factor and an overall index defined.

Water is one of the highest priorities for healthy living and economic development, as well as a crucial factor in maintaining peace and security (Rosemeyer, 2017). SDGs 6 and 10 are thus of high importance. While goal 6 seeks to "ensure availability and sustainable management of water and sanitation for all", goal 10 promotes the reduction in inequality within and among countries. This balance thus ensures that in managing water resources vis-à-vis its allocation and safeguarding, balance must be achieved in ensuring equitable distribution of water resources to everyone and among competing needs. Such distribution must thus promote sustainable consumption (goal 12) and ensure sustainability of the ecosystem. However, a major determinant in water resource availability in recent times has been climate change which has altered the existing water management paradigm of command and control (Neal et al., 2014). Furthermore, the increasing influence of climate change on water resource availability and the need for the incorporation of justice in water resource management has exposed a gap in the adoption of most techniques adopted in forecasting water demand. For instance, the common water demand forecast variables adopted by most water demand forecasting tools include the weather-based variables (e.g., rainfall, temperature, evaporation, wind speed etc.) and social variables (population growth, income-level, water price etc.). In utilizing these variables in forecasting water demand, it is generally assumed that demand should follow historical trends with seasonal variations. To compensate for discrepancies between demand and supply, tolerance levels are usually included in forecast with the attendant problems of increased cost and wastage. In ensuring that the goals of the SDGs are achieved especially goals 10 and 16 through goal 6, there is the need to define a framework that allows for the incorporation of values into the management of water resources. An implication of the conventional application of EC techniques in water resource management is that EC techniques could extend their influence beyond ‘advisory roles’. Aside providing the utility with data for planning, EC techniques need to be positioned to ‘intervene’ or extend their influence on issues surrounding the implication of adopted water allocation strategies. Considering the domino impact of policy on lifestyle (quality of life (QoL), poverty etc.), behaviour (increase/decrease in water consumption) and the society at large (productivity), EC techniques must be able to provide beyond conventional data, other information such as the effect of an allocation strategy on [1]water poverty (water burden), non-revenue water (losses), productivity (gross Value Added, GVA), estimated revenue to the utility, pressure

[1] For this research, a household is said to experience water poverty when it expends more than 10% of its income on water and sanitation services. Water poverty thus presents itself in the forms of access (i.e. the ease of accessing sufficiently clean and quality water and sanitation services) and mobility (i.e. the ability for households to upscale water consumption easily).

profile etc. This framework must thus be able to extend the benefits of conventional EC techniques by also providing answers as to how the adopted allocation strategy guarantees egalitarianism, utilitarianism, libertarianism and sufficientarianism. In navigating the gap therefore between descriptive and prescriptive claims, there is a need to formulate a [2]realistic utopia. This is important in enabling water resource experts apply more conveniently the normative principles within a realistic context.

2.3.2 Water demand modelling and Policy discussions

This section presents policy discussions that are relevant to water demand modelling, geared towards realizing the UN SDGs. Considering our realistic utopia, emerging water crisis occasioned by climate change places further constraints on water management. We thus examine the policy implications of conventional water demand modelling and management on the socio-economic aspect of society within our realistic utopia and proffer recommendations that enhance justice in water resource management. The definition of essential terms used in this section is given in Table 2-2.

Table 2-2: Definition of essential policy terms

Theory	Meaning	Applied approach
Egalitarianism	Favours equality among living entities. Advocates the removal of inequalities among people.	Bounded by sufficientarianism and household's ability to increase water demand
Libertarianism	Emphasizes freedom, liberty, voluntary association, and respect of property rights.	Bounded by the water utility being able to provide households with services that enable them to determine how and when they intend to utilize their water allocation without impediments from the utility
Utilitarianism	The proper course of action is the one that maximises the overall "happiness". In other words, actions are right if they are useful or for the benefit of the majority.	Bounded by households being able to derive optimum utility from water allocation
Sufficientarianism	Rather than ensuring equality and all as well of as possible, the aim is to make sure that everyone has enough.	Bounded by adequate minimum access with provision for water mobility.

[2] We define for this paper a realistic utopia to be a setting in which water availability is adequate (not surplus), with sufficientarianism bounded by adequate minimum access with provision for water mobility, egalitarianism bounded by sufficientarianism and household's ability to increase water demand, utilitarianism bounded by households being able to derive optimum utility from water allocation and libertarianism bounded by the water utility being able to provide households with services that enable them determine how and when they intend to utilize their water allocation without impediments from the utility.

2.3.2.1 Policy discussion on water demand modelling and sufficientarianism

Here, we seek to explore the effect of water demand modelling on sufficientarianism. Furthermore, we seek to answer the following questions.

- What constitutes sufficient water under constrained scenarios?
- Does the adopted water allocation strategy guarantee water mobility for households?

Considering the primacy of water to life and the need to achieve sustainability between demand, supply and future use, questions have arisen over what constitutes sufficient water for survival. However, beyond immediate water access is water mobility (the ability of a person or household to increase water consumption due to an improvement in lifestyle, family size or production activity). EC techniques must thus adopt measures that ensure that water demand modelling does not unnecessarily constrain consumers and stifle productivity. The provision of water to households must thus be of sufficient quantity to facilitate normal activities (cooking, drinking, sanitation etc.) and production (small scale business). While it is generally established that wealthy households have the means to pay for water access for sundry purposes (gardening, lawn maintenance, swimming pools etc.) beyond normal uses and productivity, water allocation strategies adopted must prioritize utilization purposes that directly impact of the QoL of residents and their livelihood.

2.3.2.2 Policy discussion on water demand modelling and libertarianism

Here, we are concerned with how adopted water demand management affects the ability of water users in utilizing water resources as they deem fit. Questions to be answered here include:

- Do adopted water demand management strategies negatively impact on the prerogative (in terms of usage) of water users?
- Can adopted water demand management techniques guarantee libertarianism while also ensuring water demand-supply balance?

The water crisis currently plaguing Cape Town has led to the implementation of various water demand strategies such as fines, installation of pressure reduction devices, reduction in water allocation etc. While it can be argued that all these measures are geared towards averting a potential water crisis, there have been severe consequences as reported in (Mtembu, 2018). Furthermore, the water demand management adopted has not indicated any potential benefits for households using below the normally allocated values aside from the ‘City Water Map” (Capeetc, 2018). This is due to the fact that water rates adopted are flat within a usage band as shown in City of Cape Town – CoCT (2018). In guaranteeing libertarianism, water demand management must be able to roll over daily net surplus from consumers who are unable to utilize their daily/monthly allocation and lack on-site storage facilities. In essence, water demand management must ensure that consumers decide how and when they intend to use their allocated water resources. In planning, allowances must be made for guaranteeing storage of surplus unused water from households with a metering infrastructure that updates households in [near] real time as to how much extra water they have saved. Households are thus left with the options of either increasing consumption (up to their accumulated water reserve) or negotiating a reduction in water bill.

2.3.2.3 Policy discussion on water demand modelling and egalitarianism

Within our realistic utopia, egalitarianism is bounded by the minimum water supply quantity needed for meeting normal daily activities (cooking, drinking, sanitation etc.) and basic productivity, with additional water consumption dependent on the purchasing ability of the household. We thus seek to answer the following questions.

- Do adopted water management strategies exacerbate the water burden of poor households?
- Are poor households afforded enough opportunities in exploiting water for productive actives beyond access?

For instance, an examination of CoCT (2018) shows that water rate increment for the Step 1 users is 556% compared to 195% and 202% for Step 2 and Step 6 users. Similarly, sewage tariff increased by 484% for Step 1 users compared to 102% for Step 5 users. In addition, the adopted water and sewage tariffs which are flat within a range of usage further discriminates low end users of a usage band. With increasing poverty levels in South Africa for instance (now estimated at over 55%), this implies that a household with a monthly income of R2000 and water usage of 6 kilolitres/month will expend about 22.32% of its income on water and sewage. With additional costs of electricity, rent etc., the disposable income of households become diminished leading to consequences of reduction (if possible) of water, electricity (already shown for South Africa and Nigeria in Monyei and Adewumi (2017) and Monyei et al. (2017)) and other services. Considering arguments against the commodification of water (Smith, 2017), and the need for a sustainable pricing regime that guarantees the availability of funds to ensure proper maintenance and management of water resources and its delivery, the utilization of EC techniques in water demand modelling must incorporate variables that provide assessments on the causal effects between pricing regimes, water demand and general economic productivity. This is to provide policy makers an overview of the far-reaching implications of their policies. Also, in encouraging water mobility especially for low-end water users, incentives must be provided that assuage or minimize the effects of increased wage bills. This could be in form of a graduated and transitory increment in water rates rather than the usual step increase as observed in CoCT (2018).

2.3.2.4 Policy discussion on water demand modelling and utilitarianism

Here, we are concerned with the benefits households can derive from water allocation. We thus seek to answer the question:

- Is the allocated water able to improve the QoL of households vis-à-vis hygiene and well-being?

Acknowledging the difficulties placed on water availability as a result of climate change, there is the possibility for water demand management policies to allocate water quantities that are ‘useless’ to households due to the insufficient quantities. In modelling water demand within our realistic utopia, sufficientarianism must guide water allocation to ensure that households are able to derive the maximum utility from any allocation. This becomes necessary especially for low water users, vulnerable households (poor households) and classes (elderly, sick, children and women). Furthermore, in making water available to indigent communities, measures must be put in place to ensure that less time is spent in accessing water supplies by providing water access points close to households, away from sewage collection points and with adequate pressure.

2.3.3 The way forward

Based on the highlighted policy discussions, there is a need for an integrated water demand and management modelling framework (IWDMMF) as shown in Figure 2-8. The IWDMMF is aimed at providing a platform for the incorporation of conventional EC techniques for ensuring optimality in water management, and assessing their wider impacts on the socio-economic aspect of society in line with justice requirements. The proposed IWDMMF thus advances the conventional EC techniques by socializing conventional EC techniques and synergizing the socio-economic impacts of adopted EC techniques with national imperatives on water (access, quality, pricing etc.). Furthermore, the proposed IWDMMF must be able to provide water policy makers an opportunity in planning for adverse conditions ahead of time, investigating cost-effective mitigation strategies and optimizing cost-recovery measures (through pricing and fines) while assessing the impacts of such cost-recovery measures on water consumption, water poverty, economic productivity and general QoL. Water demand and management must thus advance beyond traditional projections to investigating impact of water allocation on the wider society.

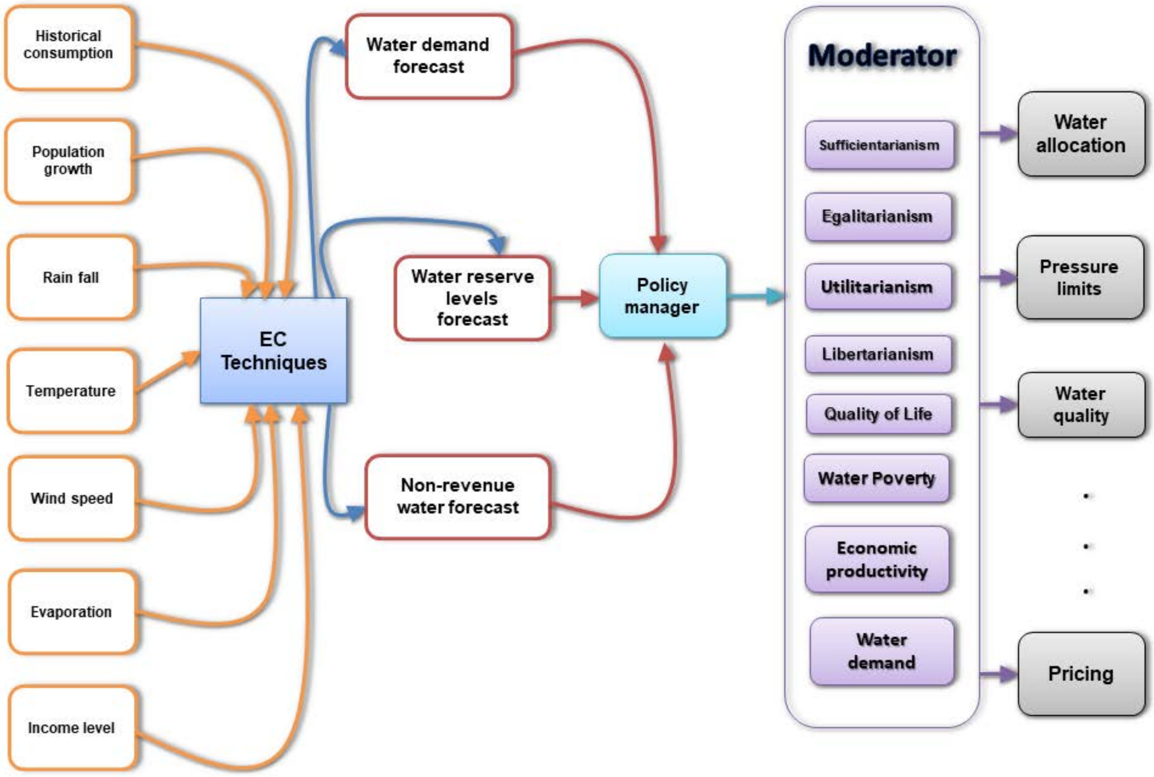


Figure 2-8: The proposed IWDMMF

2.4 CONCLUSION, RECOMMENDATIONS AND FUTURE WORK

This paper has examined the extent to which EC techniques have been applied in water demand modelling and therefore classified their application into 2 major categories namely, (i) predictive modelling, and (ii) optimization modelling. The predictive modelling category was further sub-categorized into (i) direct application in developing forecast models and (ii) indirect application which encapsulates their use as optimization engines and learning algorithms in intelligent models. In predictive modelling, an analysis of the existing literature has shown that developing countries, especially Africa, have not fully harnessed the potentials of EC techniques as the current body of knowledge lacks studies that focusses on the application of EC techniques in water demand forecasting. This may be linked largely to skills shortage and limited knowledge base in soft computing. There is therefore a need for Africa to develop policies and create platforms to build the requisite capacity in this specialized field to enable them to harness the potentials thereof. Other areas that require more attention, as identified in this review include, incorporation of weather and socioeconomic variables in forecasting studies, application of EC techniques like DE and ES in intelligent model development as well as the need to shift focus from short-term forecasting to medium- and long-term forecasting. The impacts of input variables like land use and water price on water demand should be investigated in future research especially for long-term water demand forecasting. The adoption of these recommendations will ensure that the potentials of EC techniques evolve further, thus translating from concept to demonstration and then to commercialization, and by doing so, guaranteeing their adoption in real-world water resource applications.

This study further highlights the fact that the application of EC techniques in water resource optimization and allocation could be extended by integrating wider issues of society such as poverty, economic growth (productivity), and welfare in determining the optimality of water supply and distribution networks. This will not only ensure an equitable allocation of water resources but also foster the realization of the UN SDGs. This work thus advocates for a more comprehensive framework (IWDMMF) that is capable of syncing conventional EC techniques and the social aspect of society. This is to ensure that water policy makers and administrators are able to assess the wider impact of policy decisions emanating from optimal values provided by EC techniques. Furthermore, considering the need for justice and equity in water demand management, the advocated framework (IWDMMF) offers a platform for scrutinizing policy decisions through the doctrines of egalitarianism, libertarianism, utilitarianism and sufficientarianism. This is necessary in ensuring that water demand management does not adversely affect the vulnerable and poor in the society. Future studies will focus on the application of IWDMMF in resolving real-world multi-objective water demand problems and conflicts.

2.5 RESEARCH OUTPUTS

1. **Oyebode, O.,** Babatunde, D. E., Monyei, C. G. & Babatunde, O.M. (2019). Water demand modelling using evolutionary computation techniques: Integrating water equity and justice for realization of the sustainable development goals. Heliyon 5 (11): e02796. https://doi.org/10.1016/j.heliyon.2019.e02796

2. **Oyebode, O.** & Stretch, D. (2019). Neural network modeling of hydrological systems: A review of implementation techniques. Natural Resource Modeling 32 (1), e12189. https://dx.doi.org/10.1111/nrm.12189

3. **Oyebode, O.,** Monyei, C. G. & Stretch, D. (2018). A sustainability framework for integrating equity and justice in evolutionary-based water demand modelling. Presented at the 2018 College of Agriculture, Engineering, and Sciences Postgraduate Research and Innovation Symposium, University of KwaZulu-Natal, Durban, South Africa. Awarded Best Poster Presentation: PhD category.

4. **Oyebode, O.,** Monyei, C. G. & Stretch, D. (2019). Harnessing the potentials of evolutionary computation via the integration of rights-based distributive principles in water demand management. In: Proceedings of the 2019 European International Conference on Transforming Urban Systems (EICTUS–2019), Strasbourg, France, 26-28 June 2019; p. 36.

CHAPTER 3

EVOLUTIONARY MODELLING OF MUNICIPAL WATER DEMAND WITH MULTIPLE FEATURE SELECTION TECHNIQUES

Oluwaseun Oyebode

Centre for Research in Environmental, Coastal and Hydrological Engineering (CRECHE), Department of Civil Engineering, University of KwaZulu-Natal, Durban 4014, South Africa

*Corresponding author: oluwaseun.oyebode@gmail.com; Tel: +27 (0) 84 807 3576

3.1 OVERVIEW

This paper presents the development of an artificial intelligent water demand forecasting model. The model comprises a single-hidden layer feedforward neural network trained using a differential evolution algorithm. Multiple feature selection techniques were employed to identify the minimal subset of features for optimal learning, namely Pearson correlation, information gain, symmetrical uncertainty, Relief-F attribute and principal component analysis. The performance of the feature selection techniques was compared to a baseline scenario comprising a full set of data covering potential casual variables including weather, socio-economic and historical water consumption data. The performance of the models was evaluated based on accuracy. Results show that the five feature selection techniques outperformed the baseline scenario. More importantly, the subset of features obtained from the Pearson correlation technique produced the most superior model in terms of model accuracy. Findings from the study suggest that the inclusion of weather and socio-economic variables in water demand modelling could enhance the accuracy of forecasts and cater for the impacts of climate and socioeconomic variations in water demand planning and management.

Keywords: Artificial neural network, differential evolution, feature selection, water demand forecasting

3.2 INTRODUCTION

Water demand forecasting is of crucial importance in water resource planning and management as it is a prerequisite for optimal allocation of available water resources (Qu et al., 2010). City managers and water utilities often rely on water demand forecasts to guide their decision-making on infrastructure investments as well as the scheduling and operation of water distribution systems. For example, long-term forecasts are imperative in providing new water supplies and upgrading the capacity of existing water treatment plants while short-term forecasts guide day-to-day operation of treatment plants and reservoirs to meet daily demands. Accurate water demand forecasts are therefore required for both short- and long-term infrastructure planning, operation and coordination. Moreover, the importance of water demand forecasting in realizing the sustainable development goals (SDGs) has been stressed by the United Nations (UN). The UN in its 2015 water development report called for improvement of water demand models as competing demands may lead to increasingly difficult allocation decisions and restrict the growth of sectors critical to sustainable development (UNESCO, 2015). This implies that, amidst growing demands for freshwater across the globe, sustainable development can only be achieved if competing sources of demand are well defined to enable restoration of the balance between demand and supply. Water demand forecasting therefore provides useful information for promoting a more economical use of water resources and ensuring the sustainability of water distribution systems in the short, medium and long-terms.

Accurate water demand prediction is reliant on the explanatory variables adopted in model development. Research has shown that model accuracy is a function of the impacts of each explanatory variable (Babel and Shinde, 2011; Firat et al., 2009; Toth et al., 2018). However, many city managers, consultants and water utilities still assume that water demand will evolve simply as a function of per-capita demand and a prognosis of population, although the predictive power of such approaches is deficient under changing conditions (Toth et al., 2018). Babel and Shinde (2011) argue on the need to develop improved and city specific water demand models as water demand is influenced not only by population but also by various weather and socioeconomic variables as well as government policies and strategy related factors, which are often location-based. Therefore, the need to carefully define, evaluate, understand and model the explanatory variables that directly and indirectly influence water demand is now acknowledged as crucial in obtaining accurate demand forecast.

Over the past few decades, many techniques have been used in forecasting water demand. These techniques mainly include traditional forecasting techniques such as multivariate regression and time series analysis (Babel et al., 2007), system dynamics modelling (Qi and Chang, 2011), and more recently, advanced computational intelligence techniques like expert systems or agent-based models (Ali et al., 2017), and artificial neural networks – ANN (Bennett et al., 2013). The application of ANN in water demand forecasting is becoming increasingly popular due to its superiority over traditional techniques and its ability to account for nonlinear patterns observed in real problems (Babel and Shinde, 2011; Kofinas et al., 2014). ANN is capable of learning and analyzing data attributes and thereafter, implement nonlinear approximation function without any initial assumption on the physics of the system being modelled or its data distributions (Ardabili et al., 2018). As a result, ANN is now being adopted as an alternative to the traditional methods which are limited due to their linear pre-assumption of the form of the model (Kofinas et al., 2014). A review of the capability, implementation and application of ANN in water resources modelling including water demand forecasting is available in Ghalehkhondabi et al. (2017) and Oyebode and Stretch (2019).

Water demand forecasting using ANN is characterized by some complexities. According to Kofinas et al. (2014), these complexities can be summarized as (i) inability to adequately extrapolate outside the range of primary (training) data; (ii) diminishing forecast accuracy when lagged values of target variable are used as input; and (iii) disregarding the impacts of other explanatory variables affecting water demand due to the high correlation between future water demand and its historical values. In a comprehensive review of techniques used in forecasting water demand, Oyebode et al. (2019) noted the non-inclusion of weather-based variables as inputs in most of the studies reviewed. It was further argued that, due to the non-inclusion of weather-based variables, most studies in the literature lack a climate variability perspective to water demand modelling. This jeopardizes the opportunity to put in place effective early warning systems and to implement adaptive interventions to deal with variations in water availability and the occurrence of extreme climate-linked events. The inclusion of climate-based parameters is likely to enhance the outcome of existing water demand forecasting models.

This study aims to suggest possible ways of addressing the limitations of ANN in water demand forecasting. First, this study allows for the enhancement of demand forecasting models proposed in earlier works, by integrating new explanatory variables related to climate and socioeconomic variations. The study explores the capabilities of five feature selection techniques in providing the optimal set of explanatory variables required for accurate prediction of water demand. Furthermore, the study investigates the ability of an evolutionary-based technique – differential evolution (DE) in evolving an ANN model with optimal model complexity and accuracy. Research has shown that despite the prominence and successful applications of evolutionary algorithms in water resources (Maier et al., 2014), DE is yet to be fully explored in water demand forecasting (Oyebode et al., 2019; Oyebode and Stretch, 2019). In this study, the ability of DE in training a multilayer feed-forward neural network is explored, and by doing so, developing a water demand forecasting model with optimal complexity and accuracy.

3.3 METHODOLOGY

This section presents a brief background on the modelling technique, training algorithm and feature selection techniques applied in this study.

3.3.1 Artificial neural networks

ANN is a computational intelligence technique inspired by the configuration and working principles of the human brain (Tomić et al., 2018). The ANN architecture comprises a collection of processors (neurons), typically arranged in three layers which collect, interpret, and exchange information over a framework of weighted connections (Oyebode and Stretch, 2019). ANN is popularly used for mapping an input-output relationship for a given system by combining the input information and estimating their weights. The connection weights are a product of the impact of each input on the processor, and a threshold value (known as bias) must be exceeded for a processor to be triggered. Each processor returns an output based on the weighted sum of all inputs collected and according to a nonlinear activation function. ANN thus undergoes a learning process by adjusting the weights iteratively between its processors and comparing the resulting error between actual and modelled values (Shahin et al., 2008). Given a sigmoidal activation function, the relationship between inputs and output(s) is expressed as:

$$\boldsymbol{P} = \mathbf{1}/[\mathbf{1} + \boldsymbol{e}^{\boldsymbol{s}}] \tag{1}$$

$$\boldsymbol{s} = (\boldsymbol{a}_{\mathbf{1}}\boldsymbol{w}_{\mathbf{1}} + \boldsymbol{a}_{\mathbf{2}}\boldsymbol{w}_{\mathbf{2}} + \ldots) + \boldsymbol{B} \tag{2}$$

where $\boldsymbol{P}$ is the output of each node, $\boldsymbol{a_i}$ is the input value, $\boldsymbol{w_i}$ is the weight, and $\boldsymbol{B}$ is the bias. The key objective of ANN training is to reduce the overall error $\boldsymbol{E}$ between the outputs and actual observations by adjusting the weights. The overall error, $\boldsymbol{E}$ can be mathematically expressed as (Mafi and Amirinia, 2017):

$$\boldsymbol{E} = \frac{\mathbf{1}}{\boldsymbol{m}}\sum \boldsymbol{E_m} \tag{3}$$

where $\boldsymbol{m}$ is the total number of training patterns and $\boldsymbol{E_m}$ can be expressed as:

$$\boldsymbol{E_m} = \frac{\mathbf{1}}{\mathbf{2}}\sum(\boldsymbol{O_n} - \boldsymbol{P_n})^{\mathbf{2}} \tag{4}$$

where, $\boldsymbol{O_n}$ and $\boldsymbol{P_n}$ are actual and predicted values for $\boldsymbol{n}$th output processor respectively. To be concise, details on the configuration of ANN and its implementation are not presented in this study, however they are available in the literature (Oyebode and Stretch, 2019; Shahin et al., 2008).

The study investigates the ability of ANN to forecast monthly water demand considering the nonlinear, and dynamic nature of input variables based on climate and socioeconomic factors.

3.3.2 Differential evolution training algorithm

DE is a population-based heuristic algorithm for global optimization over continuous spaces. Thus, it can find the optimal weights required for error minimization in ANNs (Ilonen et al., 2003). According to Piotrowski (2014), the classic DE algorithm evolves a population of $\boldsymbol{NP}$ individuals, $\boldsymbol{x_{i,g}} = \{\boldsymbol{x^1_{i,g}}, \ldots, \boldsymbol{x^D_{i,g}}\}, \boldsymbol{i} = \mathbf{1}, \ldots \boldsymbol{NP}$ during successive generations $\boldsymbol{g}$ to obtain the global optimum of a function $\boldsymbol{f}$ in a subset $\prod_{j=1}^{D}[L^j, U^j]$ within a decision domain $\boldsymbol{R^D}$. A preliminary location of individuals is randomly initiated from a uniform distribution expressed as:

$$x_{i,0}^{j} = L^j + rand_i^j(0,1) \cdot (U^j - L^j); \quad j = 1, \ldots, D; \quad i = 1, \ldots, NP \tag{5}$$

where $rand_i^j(0,1)$ creates an arbitrary value within the range [0, 1] for every component of each individual.

In newer generations, each parent individual$(x_{i,g})$, generates an offspring $(u_{i,g})$ using a dual-staged approach. The initial stage involves creating a donor vector $(v_{i,g})$ via mutation. In the second stage, a crossover operation is executed between the donor and parent vectors; resulting in an offspring. A parent and an offspring is subjected to a competition-based selection process (greedy selection) and only the superior proceeds to the succeeding generation.

$DE/rand/1$ mutation strategy with a scaling factor F used in implementing the classic DE can be expressed as:

$$v_{i,g} = x_{r1,g} + F \cdot (x_{r2,g} - x_{r3,g}) \tag{6}$$

where r_1, r_2, and r_3 are randomly chosen integers within the interval $[1, NP]$, such that $r_1 \neq r_2 \neq r_3 \neq i$, $x_{best,g}$ signifies the most superior individual in the present population at generation g.

A binomial crossover operation is executed on the parent and target vectors after mutation, producing an offspring $(u_{i,g})$, and consequently requiring the value of a crossover control parameter (CR) to be defined, whence:

$$u_{i,g}^{j} = \begin{cases} v_{i,g}^{j} & if\ rand_i^j(0,1) \leq CR \quad or \quad j = j_{rand,i} \\ x_{i,g}^{j} & otherwise \end{cases} \tag{7}$$

The CR values are typically defined within the [0, 1] interval. $j_{rand,i}$, is a randomly chosen integer within the [1, D] interval, as an assurance that an offspring acquires a minimum of one element from a donor vector. Ultimately, the greedy selection between the parent and the offspring is expressed by:

$$x_{i,g+1}^{j} = \begin{cases} u_{i,g} & if\ f(u_{i,g}) \le f(x_{i,g}) \\ x_{i,g} & otherwise \end{cases} \tag{8}$$

The DE algorithm typically proceeds with the exploration until a predefined number of iterations is attained. DE is employed to optimize the architecture (complexity) and network parameters of ANN models developed in this study, thus pioneering the application of DE in training multilayer feed-forward ANN models in water demand forecasting.

3.3.3 Feature selection

To develop a model with high degree of accuracy and minimal complexity, it is important to select only a small number of variables with significant predictive features. Research has shown that the inclusion of irrelevant or redundant or noisy variables could increase model complexity, reduce model interpretability, heighten computational demands, and consequently lead to non-convergence (Bowden et al., 2005). Feature selection has been widely reported as a technique that can be beneficial to learning as it seeks to identify and possibly remove all the irrelevant and redundant attributes, thereby reducing the dimensionality of the data and the size of the hypothesis space accordingly (Hall, 1999; Oyebode, 2014). Feature selection achieves this aim by finding a minimum set of variables such that the resulting probability distribution of the data classes is, to a great extent, close to the original distribution obtained using all variables (Azhagusundari and Thanamani, 2013). Feature selection, thus, enables learning algorithms to execute faster and more effectively. The techniques utilized in evaluating the worth of features (variables) used in this study are briefly described in turn.

3.3.3.1 Pearson correlation

Pearson correlation belongs to the class of "filter" feature selection techniques which are founded on data pre-processing to isolate the features $\boldsymbol{X_1}$, …, $\boldsymbol{X_p}$ that most impact the target $\boldsymbol{Y}$. Pearson Correlation provides a straightforward approach to filter features based on their correlation coefficient. The Pearson correlation coefficient between a feature $\boldsymbol{X_i}$ and the target $\boldsymbol{Y}$ is expressed as:

$$\boldsymbol{\rho_i} = \frac{\boldsymbol{cov}(\boldsymbol{X_i}, \boldsymbol{Y})}{\boldsymbol{\sigma}(\boldsymbol{X_i})\boldsymbol{\sigma_Y}} \tag{9}$$

where $\boldsymbol{cov}(\boldsymbol{X_i}, \boldsymbol{Y})$ represents the covariance, and $\boldsymbol{\sigma}$, the standard deviation (Mangal and Holm, 2018). The coefficient is typically bounded within the interval [-1, 1], and applicable to regression and numerical classification problems. The Pearson correlation thus serves as a quick criterion for ranking features according to the absolute correlation coefficient to the target.

3.3.3.2 Information gain

Information gain is a symmetric-based index used to rank features. The index computes the number of bits of information gained by an independent variable about a target variable (Karimi et al., 2013). Given the entropy is a function of impurity in a training set $\boldsymbol{S}$, an index, $\boldsymbol{IG}$, denoting additional information about $\boldsymbol{Y}$ as provided by $\boldsymbol{X}$ can be defined; representing the amount by which the entropy of $\boldsymbol{Y}$ decreases. This index is mathematically expressed as:

$$\boldsymbol{IG} = \boldsymbol{H}(\boldsymbol{Y}) - \boldsymbol{H}(\boldsymbol{Y} \backslash \boldsymbol{X}) = \boldsymbol{H}(\boldsymbol{X}) - \boldsymbol{H}(\boldsymbol{X} \backslash \boldsymbol{Y}) \tag{10}$$

Information gain is thus founded on the premise that the information gained about the target variable $\boldsymbol{Y}$ after observing an independent variable $\boldsymbol{X}$ is equal to the information gained about $\boldsymbol{X}$ after observing $\boldsymbol{Y}$. The limitation in using information gain is in its bias towards features with more values even when they are not more informative (Phyu and Oo, 2016).

3.3.3.3 Symmetrical uncertainty

Symmetric uncertainty is a feature selection system that operates based on the principle of mutual information. Symmetrical uncertainty measures the correlation, $\boldsymbol{SU}$, between the features and the target class using the following expression (Karimi et al., 2013):

$$\boldsymbol{SU} = (\boldsymbol{H}(\boldsymbol{X}) + \boldsymbol{H}(\boldsymbol{Y}) - \boldsymbol{H}(\boldsymbol{X} \backslash \boldsymbol{Y})) / (\boldsymbol{H}(\boldsymbol{X}) + \boldsymbol{H}(\boldsymbol{Y})) \tag{11}$$

where $\boldsymbol{H}(\boldsymbol{X})$ and $\boldsymbol{H}(\boldsymbol{Y})$ are the entropies according to the probability associated with each feature and class value respectively, and $\boldsymbol{H}(\boldsymbol{X}, \boldsymbol{Y})$, the mutual probabilities of all combinations of values of $\boldsymbol{X}$ and $\boldsymbol{Y}$.

3.3.3.4 Relief-F attribute

Relief-F attribute is a feature selection technique for detecting conditional dependencies between data attributes and providing an integrated assessment on the attribute estimation in regression and classification-based problems (Robnik-Šikonja and Kononenko, 2003). It seeks to draw instances at random, calculate their nearest neighbors, fine-tune a feature weighting vector, and consequently award additional weight to features that discriminate the instance from neighbors of different classes (Phyu and Oo, 2016). Mathematically, Relief-F attribute attempts to assign a weight for each feature $\boldsymbol{f}$ using a probabilistic estimate expressed as:

$$\boldsymbol{w_f} = \boldsymbol{P}(\text{different value of } \boldsymbol{f}/\text{different class}) - \boldsymbol{P}(\text{different value of } \boldsymbol{f}/\text{same class}) \quad (12)$$

3.3.3.5 Principal component analysis

Principal component analysis is a feature selection which seeks to identify linear combinations of unique explanatory variables (referred to as principal components – PC) capable of summarizing the data, with the aim of retaining maximum information during the process. Principal component analysis operates by transforming a given set of variables orthogonally such that the transformed variables are uncorrelated and independent of each other, especially if the initial variables are normally distributed (Hu et al., 2007).

Given a data set of $\boldsymbol{G}$ variables $\boldsymbol{X}$ on every $\boldsymbol{n}$ individuals, $\boldsymbol{X} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_G})$ such as water consumption-explanatory variables, the aim is to find a new set of variables $\boldsymbol{\xi} = (\boldsymbol{\xi_1}, \boldsymbol{\xi_2}, \ldots, \boldsymbol{\xi_G})$, that are linearly related to the $\boldsymbol{X}$'s but are themselves uncorrelated with a declining variance from most significant to least significant:

$$\boldsymbol{\xi_i} = \boldsymbol{\alpha_{i1} x_1} + \boldsymbol{\alpha_{i2} x_2} + \ldots + \boldsymbol{\alpha_{ij} x_j} + \ldots + \boldsymbol{\alpha_{iG} x_G} \quad (13)$$

To apply a condition that the modification is self-orthogonal, the requisite constraints are expressed as follows:

$$\sum_{i=1}^{G} \boldsymbol{\alpha_{ij} \alpha_{ik}} = \boldsymbol{0} \qquad \boldsymbol{j \neq k} \quad (14)$$

$$\sum_{i=1}^{G} \alpha_{ij}\alpha_{ik} = 1 \qquad j = k \tag{15}$$

The setup of PCs depends on the magnitude of importance. In particular, the first PC should provide the most important relation between the original variables based on the largest variance while the second PC should give the second most important relation, and is orthogonal to the first PC, etc. The variances of the succeeding PCs would be smaller if high correlation between the original variables occurs. Consequently, principal component analysis can provide guidance in reducing the number of potential explanatory variables and offer the best representation in a fewer number of transformed PCs (Hu et al., 2007).

3.4 STUDY AREA AND DATA DESCRIPTION

This research focused on the City of Ekurhuleni, a metropolitan municipality, located in the Gauteng province of South Africa – the most populous province in South Africa with a population of approximately 14.7million people (Stats-SA, 2018). The City of Ekurhuleni was established in the year 2000 from the amalgamation of two existing regional entities, namely Kyalami Metropolitan and the Eastern Gauteng Services Council, thereby agglomerating a set of relatively small and fragmented nine towns (Figure 3-1). The City of Ekurhuleni currently accounts for about 26% of Gauteng's population and plays a dominant role in the national economy, contributing 8.8% to South Africa's Gross Added Value as of 2016 (IDP, 2018). The City also has a Human Development Index (HDI) of 0.704, greater than the National value of about 0.653. However, the City is at the epicenter of migration, resulting in increased pressure on limited water resources (IDP, 2018). The Ekurhuleni area has no significant local water resource. Consequently, water is imported over a long distance**,** via bulk purchase, from the Lesotho Highlands transfer scheme and fed into the Vaal dam (IDP, 2018). Table 3-1 provides current figures relevant to Ekurhuleni Water Infrastructure.

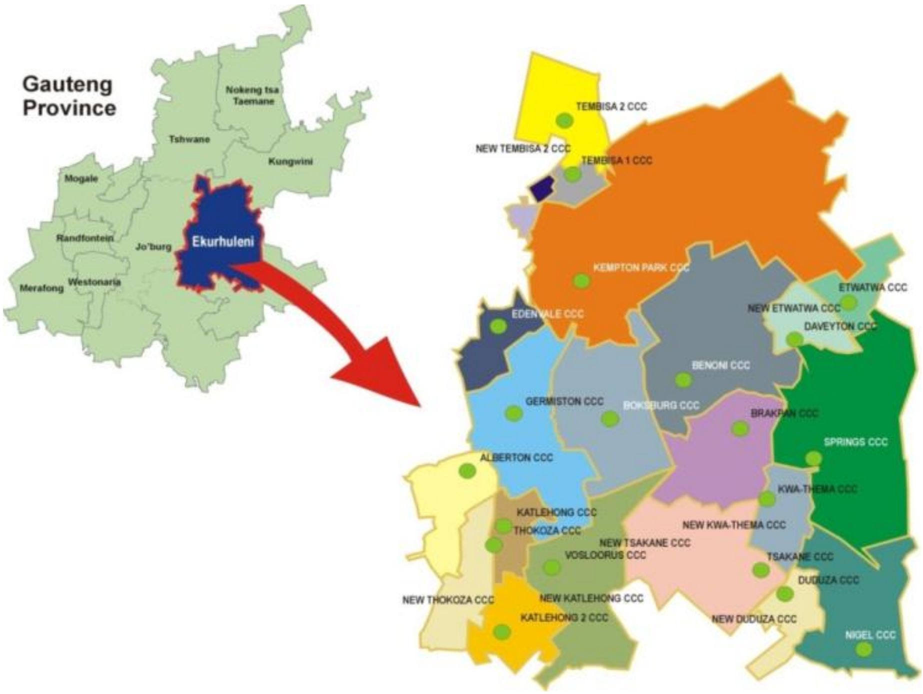


Figure 3-1: Overview of Ekurhuleni Metropolitan Municipality Service Area

The City Management seeks to ensure that Ekurhuleni transitions from being a fragmented City to being a “Delivering City” from 2012 to 2020, a “Capable City” from 2020 – 2030, and lastly a “Sustainable City” from 2030 to 2055 (IDP, 2018). To achieve these milestones, a long-term development strategy referred to as the Ekurhuleni Growth and Development Strategy 2055 (GDS 2055) has been developed to systematically analyze Ekurhuleni’s history and its development challenges, wherein it therefore outlines the desired growth and development trajectory. One of the strategic objectives and the key focus areas and/ interventions is the promotion of urban integration and continued investment in water infrastructure to ensure security of supply. This is critical to attaining the state of a “Sustainable City” and realizing the African Union Agenda 2063 and 2030 UN SDGs which includes access to clean water and sanitation, innovation and infrastructure as well as reduced inequality. Earlier works have focused on water demand analysis in the City of Ekurhuleni using stand size and land use as well as related water demand patterns in estimating future demand of residential households (Jacobs et al., 2004; Vorster et al., 1995). This study adopts an approach that considers multiple factors including those related to population, weather and socioeconomic profile of the City in developing a water demand forecast model that is geared towards sustainability of the City. This would assist in the planning and management of the City's water resources, thereby fostering the achievement of the City's objectives. Considering the real-world profile of the water infrastructure of the City of Ekurhuleni coupled with its associated challenges, this water network is thus considered as representative and relevant for use as a case of general interest. The methodologies and model development techniques applied could therefore be adopted globally.

Table 3-1: Ekurhuleni water infrastructure data

Ekurhuleni Water Infrastructure	Data
Average Water Demand (Mℓ/annum)	365 000
Water Resources/Supply	Vaal dam
Number of Reservoirs	73
Number of Towers	32
Number of Bulk Connections	186
Pipes (km)	11 448
Number of Distribution Zones	124
Population (million, 2016)	3.5
Annual Population Growth	2.51%

*Source: Gubuza (2017)

One of the initial steps in water demand forecasting is identifying explanatory variables that directly and indirectly influence water demand. The identification of explanatory variables forms the basis upon which final input parameters are selected for model development. Details on key explanatory variables to be considered in water demand forecasting is available in Oyebode et al. (2019). Based on availability, the explanatory variables considered in this study include monthly total rainfall ($\boldsymbol{R}$), monthly average minimum and maximum temperatures ($\boldsymbol{T_{min}}$ and $\boldsymbol{T_{max}}$), monthly average relative humidity ($\boldsymbol{RH}$), monthly average wind speed ($\boldsymbol{WS}$), number of household connections ($\boldsymbol{HH}$), population ($\boldsymbol{P}$), human development index ($\boldsymbol{HDI}$) and water consumption ($\boldsymbol{WC}$). Although, the weather-based variables usually employed in water demand forecasting studies are temperature and rainfall, however, considering the semi-arid characteristics of the City of Ekurhuleni (and generally, South Africa), relative humidity and wind speed were included as potential explanatory variables as they could influence outdoor water consumption (Huntra and Keener, 2017). Moreover, previous studies conducted in areas with similar characteristics have reported the inclusion of these parameters in forecasting water demand (Babel and Shinde, 2011; Firat et al., 2009; Huntra and Keener, 2017).

Monthly data records for each variable were obtained from relevant government departments (South African Weather Service, Statistics South Africa (Stats SA), and the City of Ekurhuleni) for the period August 2010 to March 2018. Water consumption data was based on total monthly billed (revenue) water consumption of water users. Weather information were supplied from a representative weather station located in the OR Tambo International Airport. The number of household connections provides an indication of the number of dwelling units served by the authority while population represents the total number of people domiciled in the City. The HDI is a measure of the City's overall achievement in its socio-economic dimensions including life expectancy, education and income levels. Yearly population and HDI data were transformed into monthly values for use in this study by linear interpolation.

The statistical properties and historical trends of the data collected in this study are presented in Table 3-2 and Figure 3-2 respectively. A new and increased trend in water consumption can be observed between mid-2015 and March 2018. As clarified by the City's water services planning manager, in mid-2015, the city management implemented one of its strategic objectives which aimed at reducing water loss within the City's water distribution network. The implementation of this strategy entailed the installation of new water meters and repair of faulty water pipes to address high water losses which were mainly due to leaks, theft and metering inaccuracies. As a result, water initially categorized as non-revenue water (i.e. real and apparent water losses) thereafter counts as revenue water.

Table 3-2: Descriptive statistics of data used in the study

Statistical parameter	$\boldsymbol{R}$ $(\boldsymbol{mm})$	$\boldsymbol{T_{min}}$ $(°\boldsymbol{C})$	$\boldsymbol{T_{max}}$ $(°\boldsymbol{C})$	$\boldsymbol{RH}$ (%)	$\boldsymbol{WS}$ $(\boldsymbol{m/s})$	$\boldsymbol{HH}$	$\boldsymbol{P}$	$\boldsymbol{HDI}$	$\boldsymbol{WC}$ $(\boldsymbol{M\ell})$
Mean	59.37	11.11	23.15	51.10	4.19	607 096	3 280 134	0.69	216 917
Maximum	210.00	16.20	29.40	75.07	5.60	698 407	3 543 077	0.71	247 135
Minimum	-	2.50	15.10	28.07	3.23	526 700	2 975 216	0.66	196 908
Standard deviation	59.42	3.65	3.38	11.56	0.58	50 953	165 046	0.01	14 524
Kurtosis coefficient	-0.30	-0.91	-0.72	-0.87	-0.64	-1.21	-1.12	0.01	-0.86
Skewness coefficient	0.81	-0.55	-0.60	0.07	0.48	0.12	-0.21	-1.15	0.70

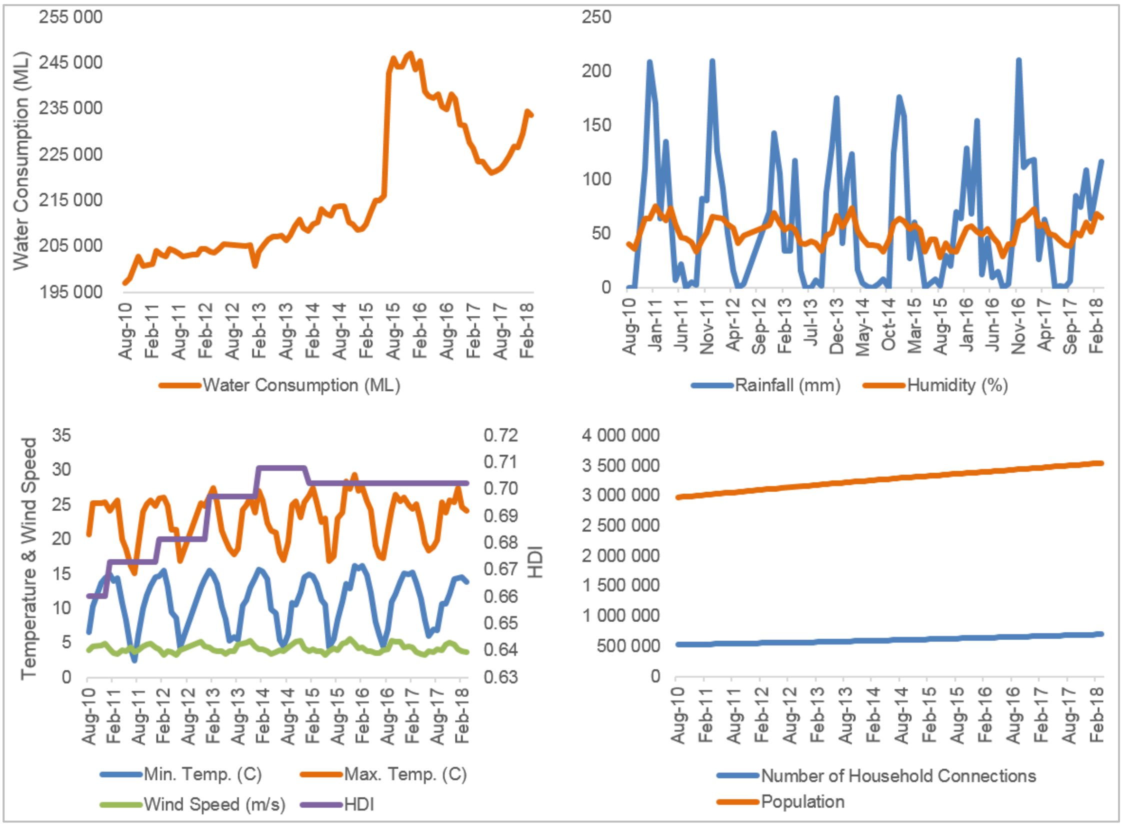


Figure 3-2: Historical trend of variables considered for model development

3.5 MODEL DEVELOPMENT

To identify feature subsets that can describe the water consumption data of the City of Ekurhuleni as good or better than the primary data set, the five feature selection techniques described above were investigated. The feature selection algorithms were implemented by means of a Ranker search method (Witten et al., 2016). The features selected by each of the techniques are presented in Table 3-3.

Table 3-3: Functional relationship between water consumption and features selected

Feature selection technique	Functional relationship of features selected	
Pearson correlation	$\boldsymbol{WC = f(HH, P, HDI, WS)}$	(16)
Information gain	$\boldsymbol{WC = f(HH, P, T_{max}, T_{min}, RH)}$	(17)
Symmetrical uncertainty	$\boldsymbol{WC = f(HH, P, T_{max}, T_{min})}$	(18)
Relief-F attribute	$\boldsymbol{WC = f(HH, P, R, T_{min}, HDI)}$	(19)
Principal component analysis	$\boldsymbol{WC = f(HH, P, HDI, RH, T_{max})}$	(20)

The functional relationship between water consumption and the original data set (i.e. all the potential explanatory variables) is expressed below. This is henceforth referred to as "baseline scenario".

$$\boldsymbol{WC} = \boldsymbol{f}(\boldsymbol{R}, \boldsymbol{T_{min}}, \boldsymbol{T_{max}}, \boldsymbol{RH}, \boldsymbol{WS}, \boldsymbol{HH}, \boldsymbol{P}, \boldsymbol{HDI}) \qquad (21)$$

The data sets were split into two subsets of similar statistical properties in line with ANN modelling standards and norms (Maier and Dandy, 2000; Modaresi et al., 2018; Oyebode and Stretch, 2019), with 70% of the data (61 instances) used for model training and the outstanding 30% (26 instances) for validation.

To investigate the performance of the feature selection techniques, a multilayer feed-forward ANN comprising three layers: one input, one hidden and one output layer was developed. The feature subsets produced by each of the feature selection techniques were used as model inputs in turn. The baseline scenario was also implemented on the ANN. The optimal architecture of the models was established by incrementally changing the number of hidden layer processors from 1 to 10 using a single stepping function. The output layer consists of only one neuron; representing the target variable – water consumption while a logistic sigmoidal-type activation function within [0, 1] interval was utilized in the hidden layer to rescale the inputs in the range [0.1, 0.9]. A linear activation function was used in the output layer to transform nonlinearities in the inputs into a linear space.

The ANN was trained using a classic DE algorithm (Storn and Price, 1997). The crossover probability, $\boldsymbol{CR}$, and mutation scale factor $\boldsymbol{F}$, were used to govern the genetic operations during the algorithm run. Following the suggestion of Montgomery and Chen (2010), $\boldsymbol{NP}$ was set at "$\boldsymbol{D}$ multiplied by 10", where $\boldsymbol{D}$ is the number of weights and biases in the selected architecture. Adopting a stepping value of 0.1, the DE algorithm was subjected to sensitivity analysis by varying $\boldsymbol{CR}$ and $\boldsymbol{F}$ incrementally within [0.5, 0.9] and [0.1, 0.5] intervals respectively. This was aimed at determining the optimal parameter settings to govern the evolution process. The algorithm was thereafter run for 1 000 generations for each of the models.

Early stopping (Raskutti et al., 2014) was integrated in the ANN models to address overfitting problems. Early stopping aims to identify the point where minimum error on the validation data set begins to rise, and then halts training to prevent overfitting. Early stopping thus ensures that model performance balances model complexity with the errors observed during training and validation.

The methodological framework developed and implemented for this study is depicted in Figure 3-3.

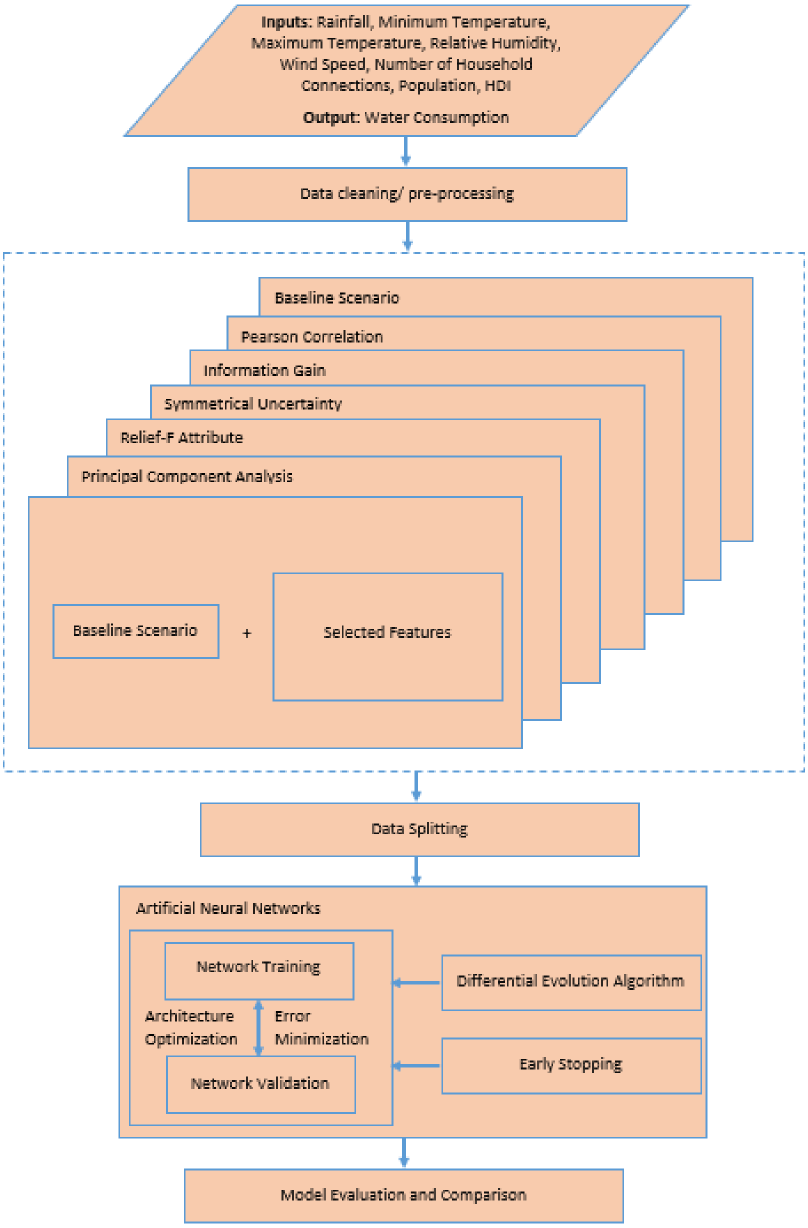


Figure 3-3: Methodological framework

3.6 MODEL EVALUATION

To evaluate the predictive capabilities of the models developed using the baseline scenario and feature selection techniques, three statistical measures were applied namely root mean-square error (RMSE), Nash–Sutcliffe efficiency index (NSE) and coefficient of determination (R^2). The RMSE is a measurement of the error variance in the model prediction, while the NSE scores the error variance within the interval [-∞; 1] (Amaranto et al., 2018). R^2 measures the degree of collinearity between observed values and predicted values, thereby defining the proportion of variance in observed values as explained by the models. Both NSE and R^2 indicate a better model as their value approaches 1. The mathematical expression for the three statistical measures is expressed as follows:

$$\boldsymbol{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(\boldsymbol{O_i} - \boldsymbol{P_i})^2}{\boldsymbol{N}}} \tag{22}$$

$$\boldsymbol{NSE} = \boldsymbol{1} - \frac{\sum_{i=1}^{N}(\boldsymbol{O_i} - \boldsymbol{P_i})^2}{\sum_{i=1}^{N}(\overline{\boldsymbol{O_i}} - \boldsymbol{O_i})^2} \tag{23}$$

$$\boldsymbol{R^2} = \left[\frac{\sum_{i=1}^{N}(\boldsymbol{O_i} - \overline{\boldsymbol{O}})(\boldsymbol{P_i} - \widetilde{\boldsymbol{P}})}{\sqrt{\sum_{i=1}^{N}(\boldsymbol{O_i} - \overline{\boldsymbol{O}})^2 \cdot \sum_{i=1}^{N}(\boldsymbol{P_i} - \widetilde{\boldsymbol{P}})^2}}\right]^2 \tag{24}$$

where $\boldsymbol{N}$ is the number of instances in the set, and $\boldsymbol{P_i}$, $\boldsymbol{O_i}$, $\widetilde{\boldsymbol{P}}$ and $\overline{\boldsymbol{O}}$ are the predicted and observed values, and their respective average values.

3.7 RESULTS AND DISCUSSION

The performance of the ANN models developed in the study were evaluated based on learning accuracy and model complexity, and the performance evaluation results presented in Table 3-4 and Table 3-5. Table 3-4 compares the performance of the ANN models in reproducing the actual water consumption at the City of Ekurhuleni, while Table 3-5 presents the optimal model architectures, optimal DE control parameters and ranks for each of the ANN models. The results show a highly competitive performance amongst the techniques employed in this study, with minimal errors (RMSEs) observed in all the ANN models. All the models were ranked based on their average performance across the three statistical measures and over the validation data sets. Overall, the ANN model developed using the Pearson correlation subset performed better than other techniques; producing the lowest error (RMSE) estimate of 4 172 Mℓ. Similarly, the Pearson correlation-based ANN model produced the highest R^2 and NSE values at 0.9233 and 0.9001 respectively.

Table 3-4: Performance of models developed from each scenario

Techniques	Statistical Parameters					
	R^2	R^2	RMSE	RMSE	NSE	NSE
	Training	Validation	Training	Validation	Training	Validation
Baseline scenario	0.9038	0.8576	4 766	5 160	0.8808	0.8472
Pearson correlation	0.8812	0.9233	5 092	4 172	0.8639	0.9001
Information gain	0.8647	0.8961	5 105	4 505	0.8632	0.8835
Symmetrical uncertainty	0.8375	0.8611	5 659	5 227	0.8319	0.8454
Relief-F attribute	0.8576	0.9075	5 372	4 178	0.8485	0.8998
Principal component analysis	0.8397	0.8943	5 720	4 528	0.8283	0.8823

The ANN model developed using the Relief-F Attribute technique produced the second-best performance, while those developed using the information gain, principal component analysis and symmetrical uncertainty came third, fourth and fifth respectively. It is interesting to note that all the ANN models developed using the five feature selection techniques converged better during validation than training, implying that the models do not suffer from the “curse of dimensionality” and overfitting which typically plagues ANN models (Adeyemo et al., 2018). This also suggests that the early stopping criterion was effective in preventing overfitting.

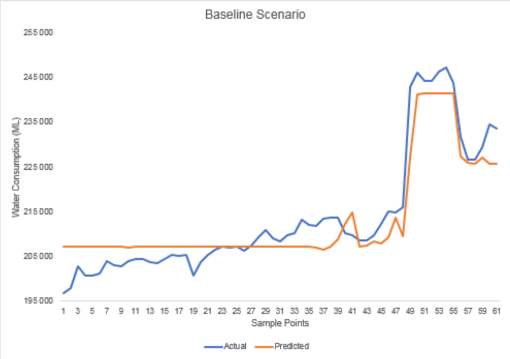
Table 3-5: Model comparison and ranks based on model architecture and forecast accuracy during validation

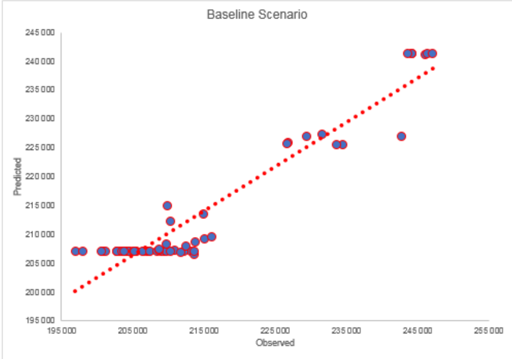
Techniques	**Model Architecture**	**Optimal DE Algorithm Control Parameters**		**Rank**			**Average: Overall Model Accuracy**
		Cr	**F**	**R^2**	**RMSE**	**NSE**	
Baseline scenario	8-2-1	0.7	0.5	6	6	6	6
Pearson correlation	4-4-1	0.8	0.3	1	1	1	1
Information gain	5-2-1	0.7	0.3	3	3	3	3
Symmetrical uncertainty	4-3-1	0.7	0.3	5	5	5	5
Relief-F attribute	5-3-1	0.7	0.3	2	2	2	2
Principal component analysis	5-2-1	0.8	0.4	4	4	4	4

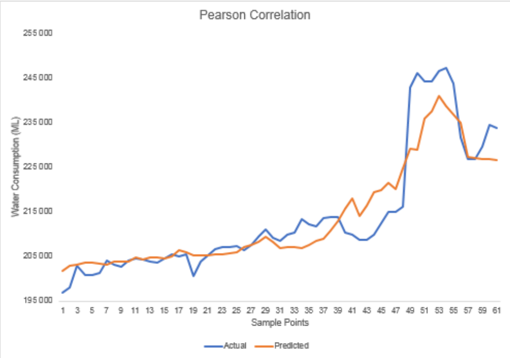
Contrastingly, the ANN developed using the baseline scenario had the worst model performance amongst the six ANN models. A slight overfit can be noticed in its training and validation results. This slight overfit could be due to overparameterization; suggesting parameter irrelevancy or redundancy in the full set of potential explanatory variables considered. The performance of the baseline scenario ANN models agrees with the argument of Phyu and Oo (2016) that, if feature selection techniques are employed, the consistency of the full set of attributes can never be higher than that of any subset of attributes. Notwithstanding, the rank of the ANN model developed from the baseline scenario, its performance could be referred to as reasonable considering its model architecture (8-2-1) which seems to have the minimal complexity.

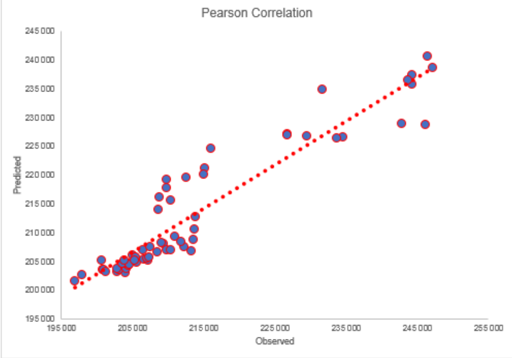
Sensitivity analysis performed on the DE control parameters show that the optimal crossover and mutation probabilities were in the [0.7, 0.8] and [0.3, 0.5] intervals respectively across the six ANN models. This suggests a high exploratory search by the DE algorithm, which is often a product of continuous productive search (Montgomery and Chen, 2010).

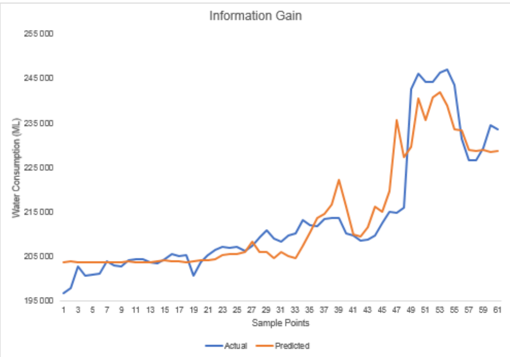
Figure 3-4 and Figure 3-5 present plots of observed and forecasted water demands for the training and validation phases respectively. The plots clearly show that all the models produced a good representation of the water demand pattern in the City of Ekurhuleni. Both the peaks and troughs including sharp spikes in the water demand pattern were reproduced by the feature selection-based models. Some constant values are however noticeable in the baseline scenario model during training; possibly due to the inclusion of irrelevant variables that have little or no significant influence on the learning process. The corresponding scatter plots also depict high accuracy and correlation as the observed and forecasted values are close to the line of equality in all the models. The best model representation and best line of fit is produced by the Pearson correlation-based model.

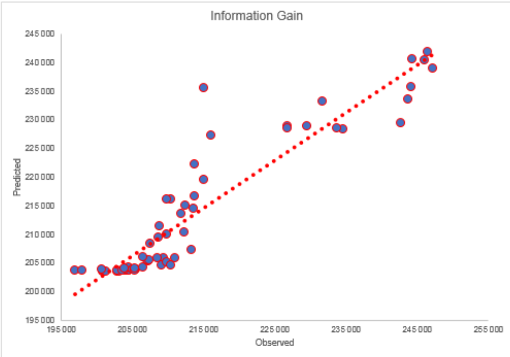

Baseline Scenario
Water Consumption (ML)
Sample Points
Actual
Predicted

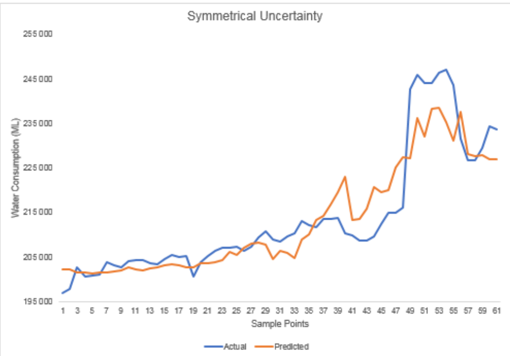

Baseline Scenario
Predicted
Observed

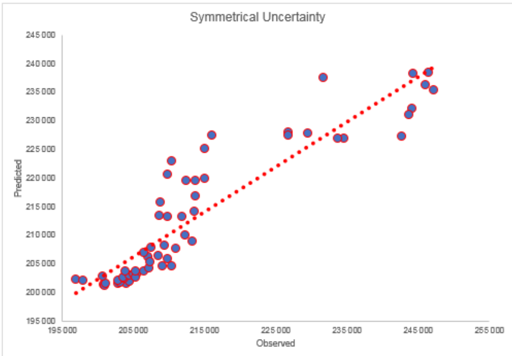

Pearson Correlation
Water Consumption (ML)
Sample Points
Actual
Predicted


Pearson Correlation
Predicted
Observed


Information Gain
Water Consumption (ML)
Sample Points
Actual
Predicted


Information Gain
Predicted
Observed


Symmetrical Uncertainty
Water Consumption (ML)
Sample Points
Actual
Predicted


Symmetrical Uncertainty
Predicted
Observed

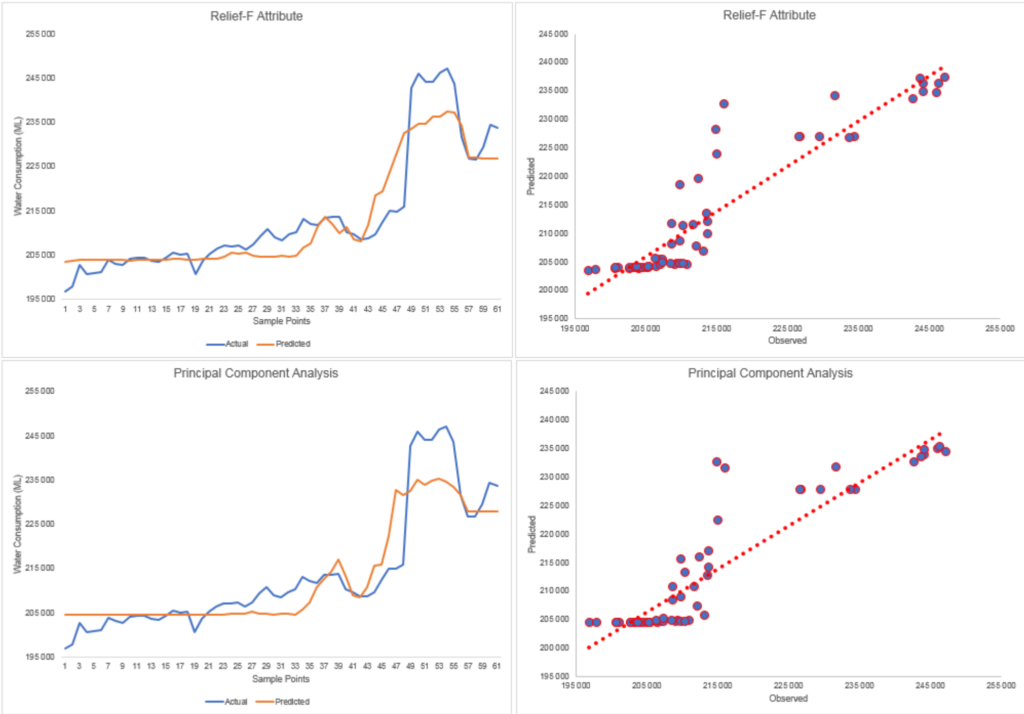
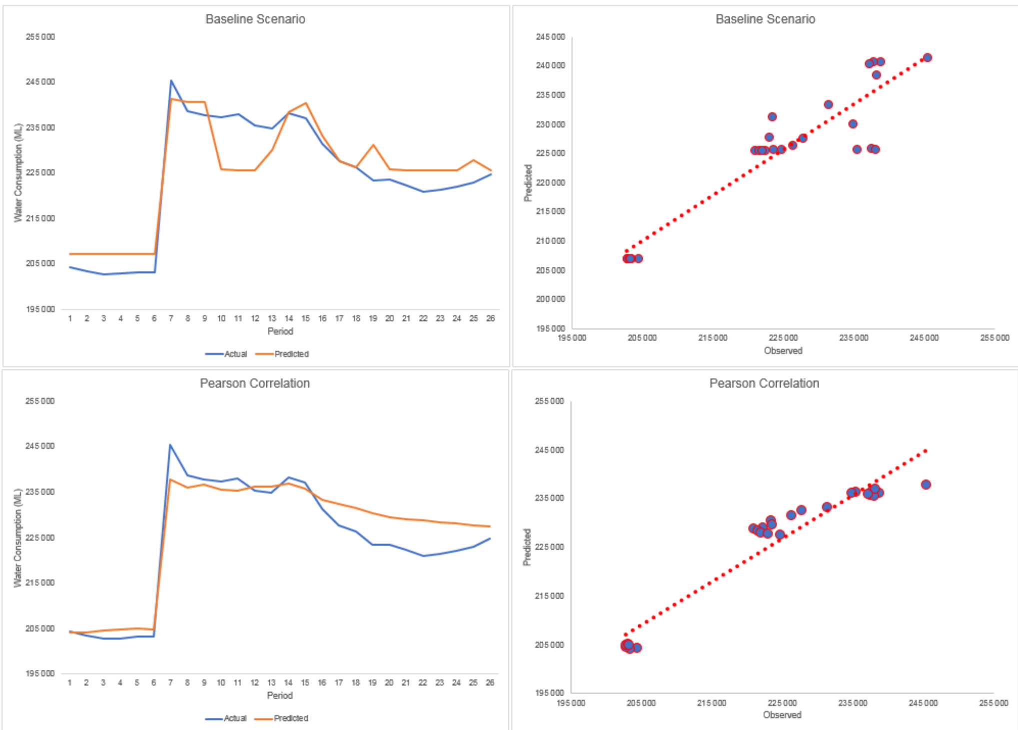


Figure 3-4: Comparison of performances of developed models during training phase



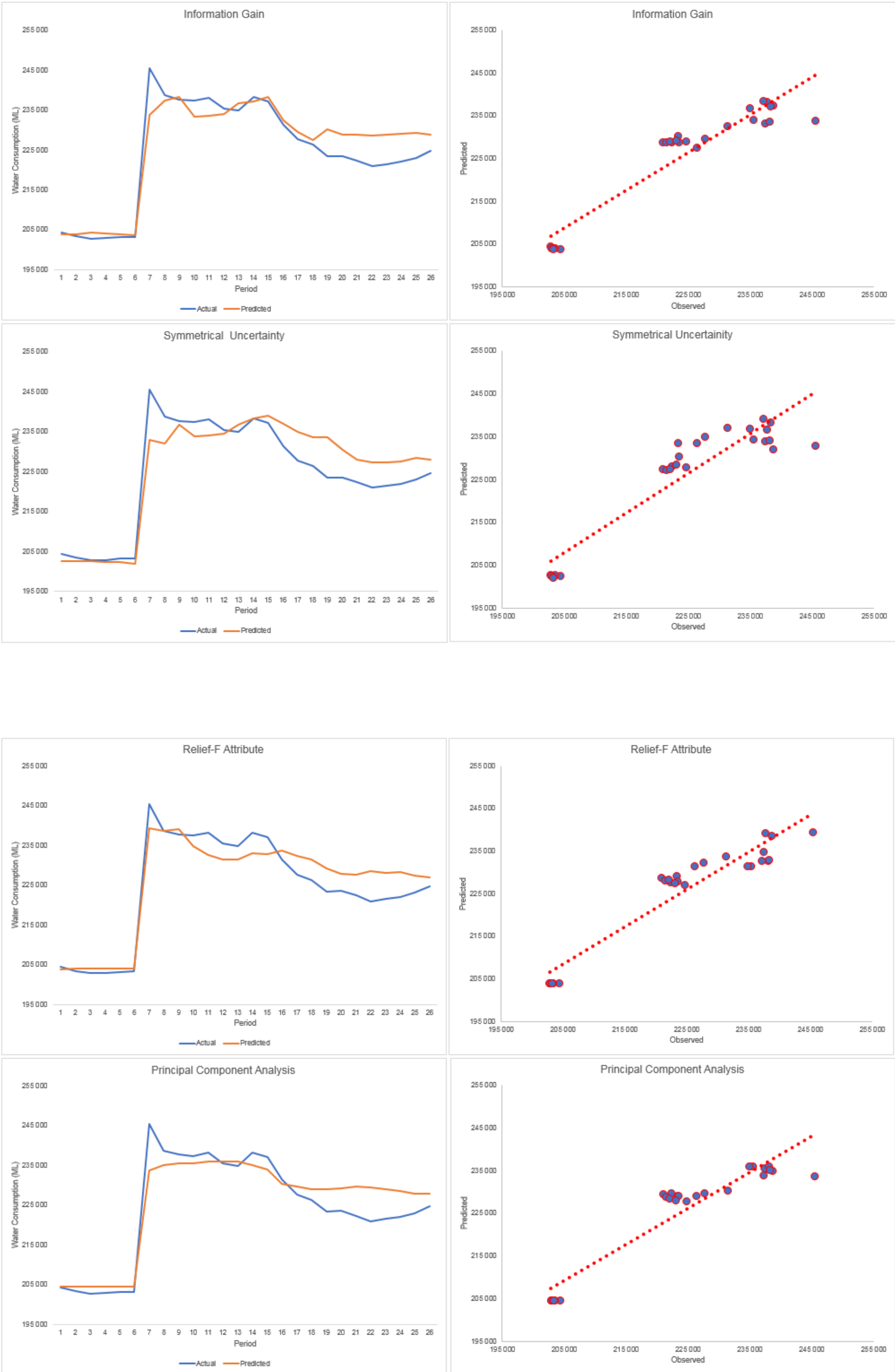


Figure 3-5: Comparison of performances of developed models during validation phase

Table 3-6 shows the contribution of each potential explanatory variable. The contribution of each variable was determined by a total count across the subsets derived from the five feature selection techniques. The results show that the number of household connections and population contributed the most to model performance; appearing in the subsets produced by the five feature selection techniques. This is followed by the minimum and maximum temperatures as well as human development index which appeared in three of the five subsets. Although, wind speed and rainfall are individually evident in only one of the subsets, their contribution is noteworthy as they, respectively, belong to the subsets that produced the most superior (Pearson correlation) and second-best (Relief-F Attribute) performances. Similarly, relative humidity appeared in the subsets of the third- and fourth-best models. These results suggest that, besides temperature, wind speed, rainfall and relative humidity have some influence on water consumption in the City of Ekurhuleni. This may be explained by the climate of the study area which is characterized by wet, windy and humid summer, resulting in higher water use. Results from this study thus agree with the findings of Babel and Shinde (2011) and Huntra and Keener (2017) which found that relative humidity, wind speed and rainfall could have some degree of influence on water consumption, especially in semi-arid regions.

Table 3-6: Contribution of explanatory variables used for model development

Feature selection techniques	R	T_{min}	T_{max}	RH	WS	HH	P	HDI
Pearson correlation					✓	✓	✓	✓
Information gain		✓	✓	✓		✓	✓	
Symmetrical uncertainty		✓	✓			✓	✓	
Relief-F attribute	✓	✓				✓	✓	✓
Principal component analysis			✓	✓		✓	✓	✓
Total count	**1**	**3**	**3**	**2**	**1**	**5**	**5**	**3**

3.8 CONCLUSIONS AND FUTURE WORK

The capability of five feature selection techniques in finding the optimal subset of features for a water demand forecasting model has been investigated in this study. The performance of the subsets generated by the five feature selection techniques were compared to that of a baseline scenario which comprise eight potential explanatory variables; totaling six scenarios. The aim was to develop an improved and reliable municipal water demand model that accounts for the impacts of weather and socioeconomic variations. Human development index – HDI was introduced for the first time in water demand forecasting as a socioeconomic variable and used alongside weather-, population- and water demand-based variables. Using a combination of evolutionary computation and artificial intelligence approach, DE-inspired ANN models were developed; one for each scenario. Results show that minimum and maximum temperatures as well as HDI were selected alongside population and number of household connections which are popularly used in water demand forecasting. Results further show that these three variables contributed significantly to the performance of three of the five models. Pearson correlation proved to be the most superior feature selection technique. DE showcased robustness in fine-tuning algorithm parameter values thereby producing good performance in terms of the solution efficiency and quality. Generally, this study demonstrates that ANN water demand models can now account for the impacts of weather and socioeconomic variations by incorporating explanatory variables based on weather and socioeconomic factors. This study also suggests that the synergetic use of feature selection techniques, DE algorithm and early stopping criterion could be used to address the limitations of ANN, thereby improving model generalization and forecast accuracy as well as providing a climate variability perspective to water demand forecasting. The methodologies, principles and techniques behind this study fosters sustainable development and thus could be adopted in planning and management of water resources. This study is limited to the use of historical water demand, weather and socioeconomic variables in predicting water demand. However, to enhance the applicability of the current ANN predictions, future research will focus on the impacts of other explanatory factors like non-revenue water, land use, recharge and run-off on the City's water demand when the information becomes available.

3.9 RESEARCH OUTPUT

1. **Oyebode, O.** (2019). Evolutionary modelling of municipal water demand with multiple feature selection techniques. Journal of Water Supply: Research and Technology–AQUA. 68 (4): 264-281. https://doi.org/10.2166/aqua.2019.145

CHAPTER 4

URBAN WATER DEMAND FORECASTING: A COMPARATIVE EVALUATION OF CONVENTIONAL AND SOFT COMPUTING TECHNIQUES

Oluwaseun Oyebode[1]* and Desmond Eseoghene Ighravwe[2]

Centre for Research in Environmental, Coastal and Hydrological Engineering (CRECHE), Department of Civil Engineering, University of KwaZulu-Natal, Durban 4014, South Africa

[2]Department of Mechanical and Biomedical Engineering, Bells University of Technology, Ota 112233, Nigeria

*Corresponding author: oluwaseun.oyebode@gmail.com; Tel: +27 (0) 84 807 3576

4.1 OVERVIEW

This paper presents a comparative evaluation of the potential of soft computing and conventional water demand forecasting techniques using the City of Ekurhuleni, South Africa as a case study. To this end, three soft computing techniques were implemented comprising two feedforward artificial neural networks (ANN) and a support vector machine. The two ANN models were trained using different algorithms namely, differential evolution (DE) and conjugate gradient (CG). Two input combination scenarios were developed in this study comprising (i) a baseline scenario of all potential explanatory variables, and (ii) an optimal subset of the baseline scenario. The performance of the three soft computing models were evaluated and compared to those obtained from conventionally-used exponential smoothing (ES_m) and multiple linear regression (MLR) techniques. Results show that, across the two scenarios, the DE-inspired ANN model was superior to the other four techniques in terms of accuracy. The results further demonstrate the robustness of evolutionary computation techniques amongst soft computing techniques.

Keywords: Artificial neural network, differential evolution, evolutionary algorithms, water demand forecasting

4.2 INTRODUCTION

The United Nations' (UN) Vision 2050 aims to ensure that water is made accessible in an appropriate quantity and quality to meet consumers' basic requirements, with healthy lifestyles and behaviors simply sustained through dependable and inexpensive water supply and sanitation services (UNESCO, 2015). To this end, in its World Water Development Report 2015, it identified the verification and transformation of data for decision-making in water resource planning and management as one of the outstanding challenges to be met in knowledge generation and policy formulation. Water demand forecasting remains one of the key implementation tools for putting in place effective planning and management of water resources, thereby making water management policies more efficient (Pulido-Calvo et al., 2007). Water demand forecasting could thus assist in addressing the knowledge management and policy formation challenges and realizing the UN 2050 Vision. However, for water demand forecasting to be effective in achieving this aim, there is a need for a shift in the manner with which water demand models are being developed. The conventional "fixture-unit" approach (typically based on multivariate regression and time-series analysis), often employed by water utilities and municipalities, has been criticized for (i) having its working principles based on the assumption of collinearity (Donkor et al., 2012; House-Peters and Chang, 2011), and (ii) having several inherent uncertainties resulting in overestimations of actual water demand as much as 100% (Shabani et al., 2016). Moreover, the ever-increasing trend in urbanization, rapid population, socioeconomic growth, climate change and their attendant threats to the sustainability of available water resources necessitate a new perspective to water demand forecasting (UNESCO, 2016). This implies that water demand models, in today's world, must be developed to better account for the dynamic and complex interactions among demographic, environmental, technological and socioeconomic characteristics of the water system. This is key to building a secure water future at both local and global scales, thereby fostering the realization of the UN objectives.

Research has shown a rapidly growing interest towards the application of soft computing techniques in water resources. This growing interest has been attributed to their ability to deliver a high level of accuracy, tractability, robustness, and cost-effective solutions to complex, ambiguous, dynamic and nonlinear real-world problems (Ghalehkhondabi et al., 2017). Examples of prominent soft computing techniques that have found application in water resources include the artificial neural networks – ANN (Adeyemo et al., 2018), support vector machines – SVM (Ch et al., 2013), adaptive neuro-fuzzy inferences systems – ANFIS (Soltani et al., 2010), systems dynamics (Dhungel and Fiedler, 2014) and evolutionary computation (EC)-based metaheuristics (Olofintoye et al., 2016). These techniques are now being explored in water demand forecasting and have been yielding some promising results (Adamowski and Karapataki, 2010; Ali et al., 2017; Ji et al., 2014; Varahrami, 2010; Vijayalaksmi and Babu, 2015; Wu and Yan, 2010; Zhai et al., 2009). In light of the successes recorded in the application of soft computing techniques, they are now being envisaged to replace or complement the conventional and/or traditional techniques (Oyebode and Stretch, 2019).

Research suggests that, despite the recent advances in the application of soft computing techniques, several areas are yet to be maximized in water demand forecasting (Ghalehkhondabi et al., 2017; Oyebode et al., 2019). Ghalehkhondabi et al. (2017), in an extensive review of the application of soft computing techniques in water demand forecasting, call for investigations into the potential of new artificial intelligence and metaheuristic techniques including deep neural nets and EC techniques, and how they can be used in optimizing model architectures. The authors also motivated for a shift from short-term to medium- and long-term forecasting as existing studies are mostly focused on short-term forecasting. In a more recent appraisal of the utilization of EC techniques in water demand modeling, Oyebode et al. (2019) found that EC techniques like differential evolution (DE) have not been adequately tested within the water demand modeling domain. In fact, findings from the review suggests that the adoption of EC techniques in water demand forecasting is yet to be embraced in developing regions, especially in Africa. Furthermore, Shabani et al. (2016) recommends the inclusion of weather and socioeconomic variables in long-term water demand forecasting to assess the impacts of evolving weather and socioeconomic conditions on water demand. Considering the importance of water demand forecasting, and the need to improve the accuracy of forecasts, more research is required to fully harness the potentials inherent in soft computing techniques.

To address the gaps in knowledge identified in the above-listed studies, this study investigates the potential of two soft computing techniques namely ANN and SVM in estimating municipal water demand. The study also seeks to investigate the impacts of training algorithms on the learning ability of feedforward ANNs. To this end, two ANN models are developed using different training algorithms – DE and conjugate gradient (CG) algorithms. The performance of the models (ANN and SVM) is thereafter compared to that of the conventionally-used multiple linear regression (MLR) and subsequently benchmarked against a standard time series technique – exponential smoothing (ES_m).

4.3 METHODOLOGY

Five different techniques have been employed in this paper for water demand forecasting in the City of Ekurhuleni, resulting in five water demand models. Two of the models are based on the working principles of artificial neural network while the other three include SVM and conventionally-used MLR and ES_m. This section presents an overview of each method.

4.3.1 Multiple linear regression

Linear regression (LR) is a popular statistical technique that has been widely applied in water demand forecasting in most water utilities and municipalities across the globe (Polebitski and Palmer, 2009; Uca et al., 2018). A regression model with two or more predictors (regressors) is referred to as a multiple linear regression (MLR) model. The generic formula for implementing a MLR model of k independent predictors can be expressed by:

$$\boldsymbol{Y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + E} \quad (1)$$

where $\boldsymbol{\beta_o}$, $\boldsymbol{\beta_1, \beta_2, \ldots, \beta_k}$ are the model parameters that must be determined, while $\boldsymbol{E, Y}$ and $\boldsymbol{X_1, X_2, \ldots, X_x}$ e are the error, target and predictor variables respectively. The values of the predictor variable ($\boldsymbol{X}$) is typically correlated with that of the target variable ($\boldsymbol{Y}$), and the performance of the model expressed as a function of the error, $\boldsymbol{E}$ (variances between actual and modeled values), correlation coefficient ($\boldsymbol{r}$) and coefficient of determination ($\boldsymbol{R^2}$). Generally, the model of best-fit obtained via a least-squares method seeks to reduce, to the smallest possible amount, the sum of squares of the difference between the observed and modeled variable (Uca et al., 2018).

4.3.2 Exponential Smoothing

ES_m is a member of the moving average forecasting methods. Forecasts produced using exponential smoothing methods are weighted averages of past observations, with the weights decaying exponentially as the observations get older (Su et al., 2018). In other words, the new observations are given relatively bigger weight in forecasting than the old observations. The prediction value is the weighted sum of observed values. The fundamental idea of smoothing model is that the trend of the time series is stable or regular, and the time series trend can be reasonably postponed, and thus, the latest historical trend will persist into the future (Hyndman et al., 2008). The forecast accuracy in the ES_m technique is mainly dependent on the value of a smoothing (α), or damping factor (1-α). Although, no formally correct procedure exists for choosing the value of α, to minimize the error influence of the smoothing factor, a trial-and-error method is often applied to select the optimal value of α (Su et al., 2018). The simplest form of ES_m is given by the formula:

$$\boldsymbol{F_{t+1} = F_t + \alpha(A_t - F_t)} \quad (2)$$

where $\boldsymbol{A_t}$ is the actual value at time $\boldsymbol{t}$; $\boldsymbol{F_t}$ is the forecast value at time $\boldsymbol{t}$; $\boldsymbol{F_{t+1}}$ is the forecast value at time $\boldsymbol{t+1}$; and $\boldsymbol{\alpha}$ is a smoothing factor, $\boldsymbol{0 \leq \alpha \leq 1}$.

4.3.3 Artificial neural network

Please refer to chapters 2 and 3 for details on the configuration and implementation of ANN.

The study investigates the performance of two feedforward ANNs, (trained using two different training algorithms), to forecast monthly water demand considering the nonlinear, and dynamic nature of input variables based on climate and socioeconomic factors. The two training algorithms investigated are (CG) and (DE) algorithms. Details on both algorithms are available in the literature (Hanke, 2017; Ilonen et al., 2003).

4.3.4 Support vector machine

SVM is a soft computing method that originates from the statistical learning theory (Vapnik, 1999). It adopts a supervised learning approach to solve regression, density estimation and classification problems. SVM initialize by defining a practical limit or boundary on the generalization error using a Structural Risk Minimization (SRM) principle (Elshorbagy et al., 2010).). It thereafter advances to search for the optimal structure of the model, using predefined model training parameters to guarantee an exclusive global minimum of the error surface. SVM thus uses a nonlinear approach to transform input domain into a higher dimensional attribute domain (Mafi and Amirinia, 2017). This approach enables SVMs to have a good performance in terms of generalization. The mapping function which is implemented by using a specified kernel may either be a linear, polynomial, sigmoidal, radial basis or hybrid function.

The architecture of SVM is similar to that of ANN in terms of its working principles. Both techniques can be denoted as two-layered networks wherein the weights are nonlinear and linear in the foremost and output layers respectively (Oyebode et al., 2014b). However, unlike ANN wherein an adaptive learning approach is adopted in optimizing all network parameters, SVM selects model parameters for the foremost layer as training input vectors. One of the advantages of SVM is that it works with smaller amount training samples and variables, and remain highly sensitive to variations in the variables (Karimi et al., 2016).

In a regression-based SVM, the training samples can be characterized as $[\boldsymbol{x_i}, \boldsymbol{y_i}]$ where $\boldsymbol{x_i} \in \boldsymbol{R^n}$ refers to the input vector, $\boldsymbol{n}$ signifies the input vector dimension, and $\boldsymbol{y_i} \in [-\mathbf{1}, \mathbf{1}]$, the target vector. The regression-based SVM employs quadratic instructions to discover optimal hyperplanes that partition the input and target classes. The quadratic instructions can be mathematically expressed as (Mafi and Amirinia, 2017):

$$\boldsymbol{min} \frac{\mathbf{1}}{\mathbf{2}} \boldsymbol{w^t w} + \boldsymbol{C} \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{\xi_i} \quad (3)$$

$$\boldsymbol{y_i}(\boldsymbol{w\varphi}(\boldsymbol{x_i}) + \boldsymbol{b}) + \boldsymbol{\xi_i} - \mathbf{1} \geq \mathbf{0} \quad (4)$$

where $\boldsymbol{\varphi}(\boldsymbol{x_i})$ transforms the input(s) into a higher dimensional attribute domain. $\boldsymbol{w}$, $\boldsymbol{b}$, $\boldsymbol{C}$ and $\boldsymbol{\xi_i} \geq \mathbf{0}$ represent the weight vector, bias, error penalty and slack variable respectively. Respectively, $\boldsymbol{C}$ and $\boldsymbol{\xi_i}$ are employed to preclude the influence of noisy information and avoid overfitting. An illustration of the SVM theory for selecting the optimal hyperplane that maximizes the limits is shown in Figure 4-1.

Equations (3) and (4) are solvable using Lagrange methods. Upon creation the optimal hyperplane, the regression function is implemented using the mathematical expression:

$$\boldsymbol{f}(\boldsymbol{y}) = \boldsymbol{sign}\left(\sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{N}} \boldsymbol{y_i c_i k}(\boldsymbol{x_i}, \boldsymbol{x_j}) + \boldsymbol{b}\right) \tag{5}$$

where $\boldsymbol{sign}()$ provides an indication of the sign; $\boldsymbol{c_i}$ signifies the Lagrange multiplier; the expression, $\boldsymbol{k}(\boldsymbol{x_i}, \boldsymbol{x_j}) = \boldsymbol{\varphi}(\boldsymbol{x_i})^{\boldsymbol{T}}(\boldsymbol{x_j})$ represents the kernel function, wherein $\boldsymbol{T}$ denotes the transpose matrix. Additional information on SVMs is available in the literature (Oyebode et al., 2014b; Vapnik, 2013).

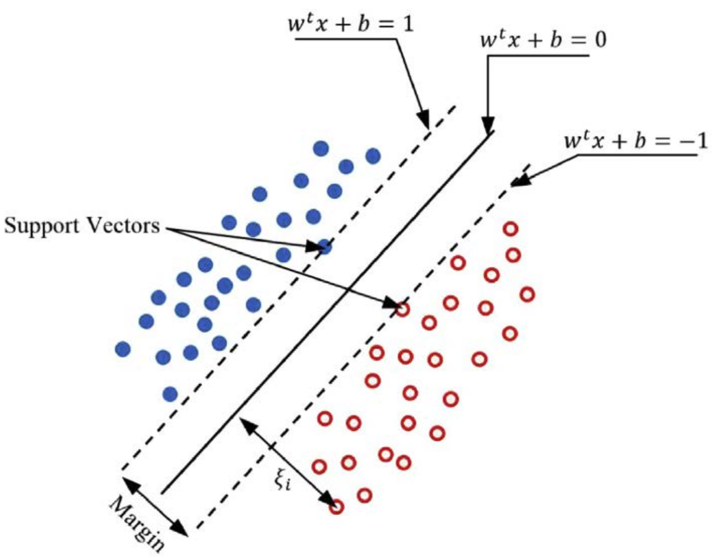


Figure 4-1: Graphical illustration of SVM working principles. Source: Mafi and Amirinia (2017)

4.4 DESCRIPTION OF STUDY AREA

Readers should refer to section 3.4 of chapter 3 for details on the description of study area and data sets employed.

4.5 MODEL DEVELOPMENT

Model development often involves subjecting potential explanatory variables to a screening process that ensures only inputs that can provide an optimal representation of the system being modeled are selected. However, some researchers believe that data-driven models including soft computation techniques, have the ability to determine important input variables, and thus introduce a large number of potential inputs to the model (Adeyemo et al., 2018). Bowden et al. (2005) argue that the inclusion of unrelated inputs may increase model complexity inhibit the learning process and consequently result in poor generalization.

To test the above assertions, this study employs two scenarios for model development. The first scenario henceforth referred to as “baseline scenario” involves the use of all potential explanatory variables (as per collected data) that could influence water consumption in the City of Ekurhuleni. The functional relationship that governs the baseline scenario can be expressed as:

$$\boldsymbol{WC} = \boldsymbol{f}(\boldsymbol{R}, \boldsymbol{T_{min}}, \boldsymbol{T_{max}}, \boldsymbol{RH}, \boldsymbol{WS}, \boldsymbol{HH}, \boldsymbol{P}, \boldsymbol{HDI}) \tag{6}$$

The second scenario, henceforth referred to as “scenario 2”, employs correlation analysis (based on Pearson correlation) to determine the dependencies between the potential explanatory variables and water consumption. The result of the correlation analysis is presented in Table 4-1. A correlation coefficient of 0.5 was adopted as cut-off point. The results show high correlation (≥0.5) between number of household connections, population, human development index and wind speed. Other potential explanatory variables produced lower correlation coefficients and were therefore discarded. The functional relationship that governs the development of scenario 2 models can be expressed as:

$$\boldsymbol{WC} = \boldsymbol{f}(\boldsymbol{HH}, \boldsymbol{P}, \boldsymbol{HDI}, \boldsymbol{WS}) \tag{7}$$

The data sets were split into two subsets of similar statistical properties in line with ANN modelling standards and norms (Maier and Dandy, 2000; Modaresi et al., 2018; Oyebode and Stretch, 2019) with 70% of the data (61 samples) used for model training and the remaining 30% (26 samples) for validation.

Table 4-1: Results of correlation analysis

Potential explanatory variables	Target Variable ($\boldsymbol{WC}$)
$\boldsymbol{R}$	-0.06
$\boldsymbol{T_{min}}$	0.07
$\boldsymbol{T_{max}}$	0.15
$\boldsymbol{RH}$	-0.23
$\boldsymbol{WS}$	0.50
$\boldsymbol{HH}$	0.79
$\boldsymbol{P}$	0.79
$\boldsymbol{HDI}$	0.59

The modeling of the water consumption in the City of Ekurhuleni was based on four modeling techniques - MLR, ES_m, ANN and SVM. As earlier stated, two ANN models were developed. The first ANN, henceforth referred to as ANN-CG, was trained using a CG algorithm while the second (ANN-DE) was trained using an EC technique – a classic DE. The training of the ANN using the classic DE algorithm was implemented using Visual Basic for Applications (VBA) programming language and on an Intel Core i7 PC with 2.70GHz and 16GB RAM. An extract of codes written in developing the ANN-DE model is presented in Appendix 2. A full version of the codes would be deposited in the institutional repository and made accessible to potential users upon request from the appropriate authority. The SVM, ANN-CG and MLR models were implemented using the DTREG platform (Sherrod, 2003) while ES_m was implemented using the Data Analysis Tool pack in Microsoft Excel. Hence, a total of five modeling approaches namely MLR, ES_m, ANN-CG, ANN-DE and SVM were implemented in this study. Four of the modelling approaches (MLR, ANN-CG, ANN-DE and SVM) were implemented for the two scenarios described in Equations (6) and (7) and their performance tested against ES.

Summary of key information that governs the development and run of each of the modeling techniques are presented in Table 4-2.

Table 4-2: Summary of key information used for model development

MLR	ES_m	ANN-CG	ANN-DE	SVM
Confidence interval: 95%	Optimization of damping factor: [0.1, 0.9] Incremental function: 0.1	Model type: Multilayer Perceptron Number of network layers: 3 (1 hidden) Optimization of hidden layer neurons: [1, 10] Stepping function: 1 Overfitting prevention: Hold out 20% of training rows Hidden layer activation function: Logistic Output layer activation function: Linear CG Parameters: Convergence tries: 4 Maximum iterations: 10000 Iterations without improvement: 100 Convergence tolerance: 1.000e-005 Minimum improvement delta: 1.000e-005 Minimum gradient: 1.000e-005 Training method: Scaled-conjugate gradient	Model type: Multilayer Perceptron Number of network layers: 3 (1 hidden) Optimization of hidden layer neurons: [1, 10] Stepping function: 1 Overfitting prevention: Yes, Early stopping Hidden layer activation function: Logistic-sigmoidal [0, 1] & re-scaling of inputs: [0.1, 0.9] Output layer activation function: Linear DE Parameters: Pop. Size, $\boldsymbol{NP} = \boldsymbol{D} * \mathbf{10}$ where $\boldsymbol{D}$ = number of weights and biases Sensitivity analysis: Yes Crossover rate, $\boldsymbol{CR}$: [0.5, 0.9] interval Mutation rate, $\boldsymbol{F}$: [0.5, 0.9] interval Stepping value for $\boldsymbol{CR}$ and $\boldsymbol{F}$: 0.1 Number of generations: 1000	Model type: Epsilon SVR Kernel function: RBF Stopping criteria: 0.001 Parameter optimization: Grid search: [10, 1] Pattern search: Intervals: 10 Tolerance: 1e-008 % rows to use for search: 100 Cross-validate: 4 folds Model Parameters: C: [0.1, 5000] Gamma: [0.1, 50] P: [0.0001, 100]

4.6 MODEL EVALUATION

To assess the predictive capabilities of the models developed using the baseline scenario and feature selection techniques, three statistical measures were applied namely root mean-square error (RMSE), mean absolute percent error (MAPE) and coefficient of determination (R^2). The RMSE is an assessment of the error variance in the model prediction, while the MAPE scores the absolute differences between observed and predicted output values (Amaranto *et al.* 2018). R^2 is a function of the extent of collinearity between observed and modeled values, thereby defining the amount of variance in observed values as explained by the models. Both RMSE and MAPE indicate an improved model as their values tend towards 0 while R^2 indicate an improved model as its value tend towards 1. The mathematical expression for the three statistical measures is expressed in Equations (8) to (10):

$$\boldsymbol{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(\boldsymbol{O_i} - \boldsymbol{P_i})^2}{\boldsymbol{N}}} \tag{8}$$

$$\boldsymbol{MAPE} = \mathbf{100}\frac{\mathbf{1}}{\boldsymbol{N}}\sum_{I=1}^{N}\frac{|\boldsymbol{O_i} - \boldsymbol{P_i}|}{\boldsymbol{O_i}} \tag{9}$$

$$\boldsymbol{R^2} = \left[\frac{\sum_{i=1}^{N}(\boldsymbol{O_i} - \overline{\boldsymbol{O}})(\boldsymbol{P_i} - \widetilde{\boldsymbol{P}})}{\sqrt{\sum_{i=1}^{N}(\boldsymbol{O_i} - \overline{\boldsymbol{O}})^2 \cdot \sum_{i=1}^{N}(\boldsymbol{P_i} - \widetilde{\boldsymbol{P}})^2}}\right]^2 \tag{10}$$

where $\boldsymbol{N}$ is the number of instances in the set, and $\boldsymbol{P_i}$, $\boldsymbol{O_i}$, $\widetilde{\boldsymbol{P}}$ and $\overline{\boldsymbol{O}}$ are the predicted and observed values, and their respective average values.

4.7 RESULT AND DISCUSSIONS

The performance of the models developed both in the baseline scenario and scenario 2 were evaluated based on learning accuracy and presented in Table 4-3 and Table 4-4 respectively. Results for the baseline scenario show satisfactory performance during training as a high degree of convergence can be observed in all the models. However, it can be noticed that all the models suffered from overfitting during testing. Comparing the performance of the modeling techniques irrespective of the overfits, it can be seen from Table 4-3 that the ANN-DE model produced the lowest RMSE and MAPE values of 5 160 Mℓ and 1.8% respectively, as well as the highest R^2 values (0.8576), while the SVM, MLR and ANN-CG models followed in that order. The ANN-CG model produced the highest error estimates of 8 655 Mℓ and 2.8% for RMSE and MAPE respectively, as well as the lowest R^2 value of 0.6614. Interestingly, the ANN-DE and conventional MLR models produced the lowest performance differences (percentage overfits) between the training and testing phases, with the MLR models having a slight edge over the ANN-DE models in terms of R^2 and MAPE. The percentage overfit can be mathematically expressed as:

$$\frac{\boldsymbol{P_t - P_v}}{\boldsymbol{P_t}} * \mathbf{100} \tag{11}$$

where $\boldsymbol{P_t}$ is the value of the performance metric observed during training while $\boldsymbol{P_v}$ is the value of the performance metric observed during validation.

The percentage overfit obtained by the ANN-DE model was calculated to be 5.4%, 8.3% and 2.8% for R^2, RMSE and MAPE respectively, while the MLR produced 3.4%, 8.8% and 1.8% for R^2, RMSE and MAPE respectively. The SVM model had the highest percentage overfit, estimated to be 16.8%, 51.2% and 128.4% for R^2, RMSE and MAPE respectively. Generally, the overfitting problems encountered in the baseline scenario suggest that some of the explanatory variables considered may be irrelevant and/or redundant in determining the water consumption profile for the City of Ekurhuleni. The scatter plots presented in Figure 4-2 through Figure 4-5 further illustrate performance of the baseline scenario models wherein some degree of under- and over-estimations can be observed.

Table 4-3: Performance of models developed for baseline scenario

Baseline	Training	Testing	Training	Testing	Training	Testing
Techniques	R^2	R^2	RMSE	RMSE	MAPE	MAPE
MLR	0.7268	0.7030	7 449	8 107	2.6699	2.7181
ANN-CG	0.7236	0.6614	7 492	8 655	2.5906	2.7959
ANN-DE	0.9038	0.8576	4766	5 160	1.7892	1.8398
SVM	0.8842	0.7568	4 850	7 336	0.9789	2.2359

Table 4-4: Performance of models developed for scenario 2

Optimal Dataset	Training	Testing	Training	Testing	Training	Testing
Techniques	R^2	R^2	RMSE	RMSE	MAPE	MAPE
MLR	0.6765	0.7201	8 106	7 430	2.6823	2.4282
ANN-CG	0.6835	0.7122	8 017	7 507	2.5472	2.4490
ANN-DE	0.8812	0.9233	5 092	4 172	1.6650	1.5090
SVM	0.8609	0.8678	5 315	5 296	1.8775	1.6655

Table 4-5: Performance of models developed using exponential smoothing

Optimal Dataset	Training	Testing	Training	Testing	Training	Testing
Techniques	R^2	R^2	RMSE	RMSE	MAPE	MAPE
ES_m	0.9038	0.6103	4 348	8 682	0.9458	1.3188

In scenario 2, it can be observed that the learning accuracies of all the models are superior to those obtained in the baseline scenario. Each of the models recorded improved R^2, RMSE and MAPE estimates. The percentage improvements across the four modeling techniques employed range from 7%-15%, 9%-39% and 12%-34% for R^2, RMSE and MAPE respectively. The results show that all models produced in scenario 2 were more generalized and not plagued by overfitting. The scatter plots show that the data points were much closer to the line of equality than in the baseline scenario (Figure 4-2 to Figure 4-5: scenario 2). These improvements suggest that the adoption of correlation analysis was successful in finding the optimal subset of input variables required to model the water consumption data. Remarkably, the optimal subset comprising four input variables ($\boldsymbol{HH, P, HDI, WS}$) were good enough to adequately represent the water consumption profile of the City as opposed to eight variables used in the baseline scenario. This implies that overparameterization effects were totally avoided in the scenario 2 models by incorporating a screening technique, thereby identifying and removing irrelevant and redundant variables, and consequently, reducing the dimensionality of the input vector space. The performance metrics of the scenario 2 models are presented in Table 4-4. In a similar but more accurate manner to the baseline scenario, the ANN-DE model, during testing, outperformed other models, producing the lowest error estimates of 4 172 Mℓ and 1.5% for RMSE and MAPE respectively as well as highest R^2 value of 0.9233. The SVM model produced the second-best performance with R^2, RMSE and MAPE estimates of 0.8678, 5 296 Mℓ and 1.7% and was followed by the conventional MLR with corresponding estimates of 0.7201, 7 430 Mℓ and 2.4%. The ANN-CG models had the least performance with its R^2, RMSE and MAPE values estimated to be 0.7122, 7 507 and 2.4% respectively.

Table 4-5 and Figure 4-6 present results obtained from the ES_m technique which was implemented by varying the damping factor between 0.1 and 0.9 using an incremental function of 0.1. Upon initial analysis, the damping factor of 0.1 (i.e. smoothing factor of 0.9) yielded the least error estimates, and thus, was regarded as the optimal damping factor. The performance indices presented in Table 4-5 show a remarkable model performance during the training phase of the simulation, however, the performance depreciated during the testing phase resulting in poor generalization, and consequently, a model characterized by overfitting. This is also evident in **Figure** 4**-**6 wherein two significant underestimated points can be observed. The percentage overfit in the ES model is estimated to be 32.5%, 99.7% and 39.4% for R^2, RMSE and MAPE respectively. These results are inferior to those presented in Table 4-4. Notably, the performance metrics of the ANN-DE model (developed in scenario 2) show a better generalized model, not plagued by overfitting. The ANN-DE model can thus be regarded as a better model for water demand forecasting than the standard time series-based ES_m technique in this case study. These results further demonstrate the efficacy of evolutionary-based soft computing techniques over conventional methods such as time series analysis and linear regression-based methods.

A comparative evaluation of the architecture of the ANN models show that the ANN-DE model exhibited lesser complexity compared to the ANN-CG model. Upon varying the number of hidden layer neurons in each of the ANN models between 1 and 10, the optimal architecture of the ANN-CG models comprised nine hidden layer neurons, while that of the ANN-DE model comprised only four, A higher number of hidden layer neurons in the ANN-CG model consequently implies a higher computational demand and time than the ANN-DE model. It is interesting to note that the ANN-DE model, with lesser complex architecture and lower computational demand and time, achieved a higher degree of accuracy than the ANN-CG model. This observation agrees with the submission of Adeyemo et al. (2018) in a river flow forecasting study that continuous training (i.e. higher computational time) is not a guarantee to achieving better generalization. Findings from this study thus suggests that the DE algorithm is more robust and reliable in training and optimizing ANN network parameters than the CG algorithm as it offers a good compromise between accuracy and complexity. Ultimately, this study demonstrates the potential of DE-trained ANN in evolving more accurate water demand forecasting models than the conventionally-used MLR and ES_m, SVM and CG-trained ANN.

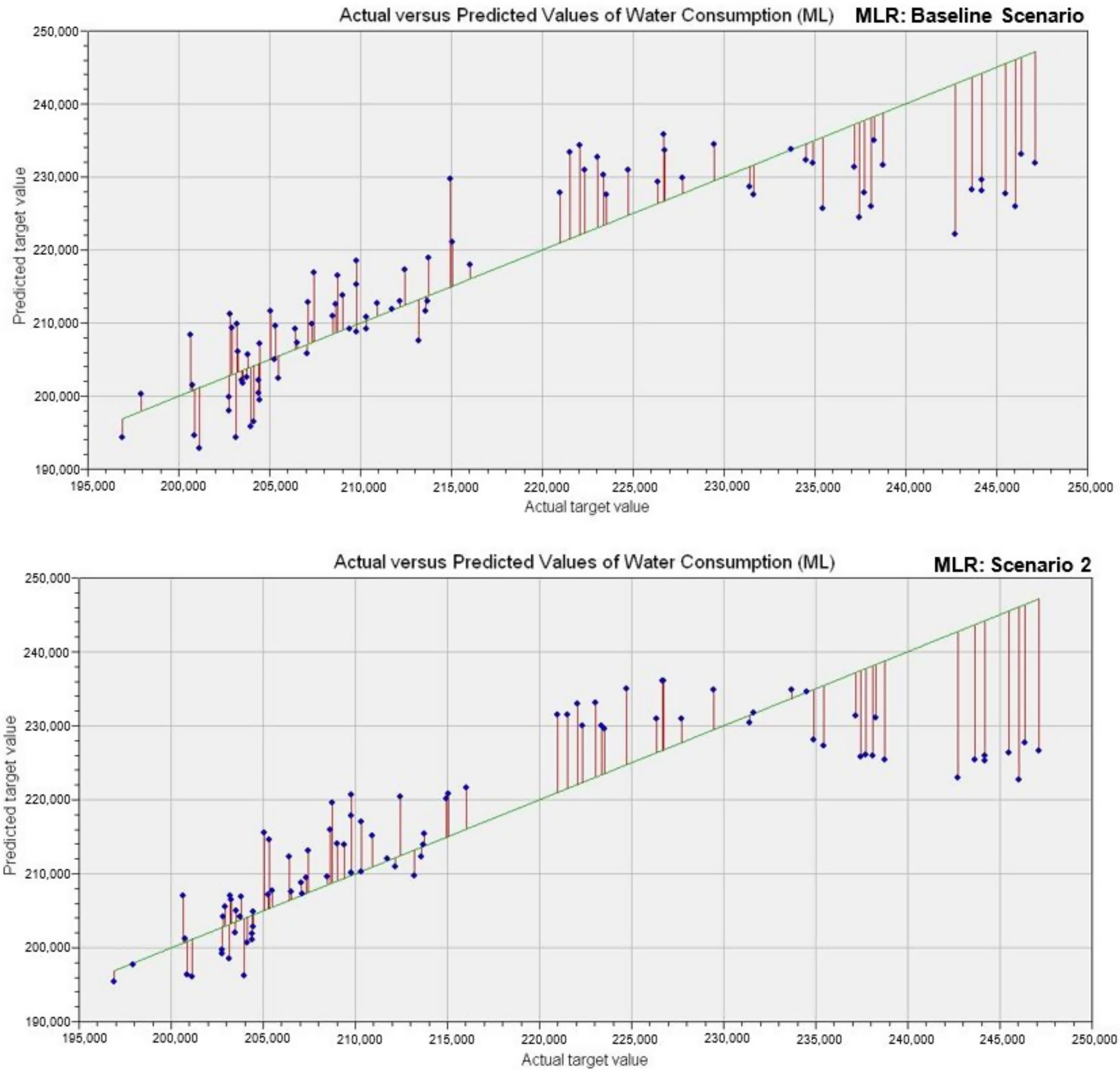


Figure 4-2: Scatter plots of observed and MLR-predicted water demand for both scenarios

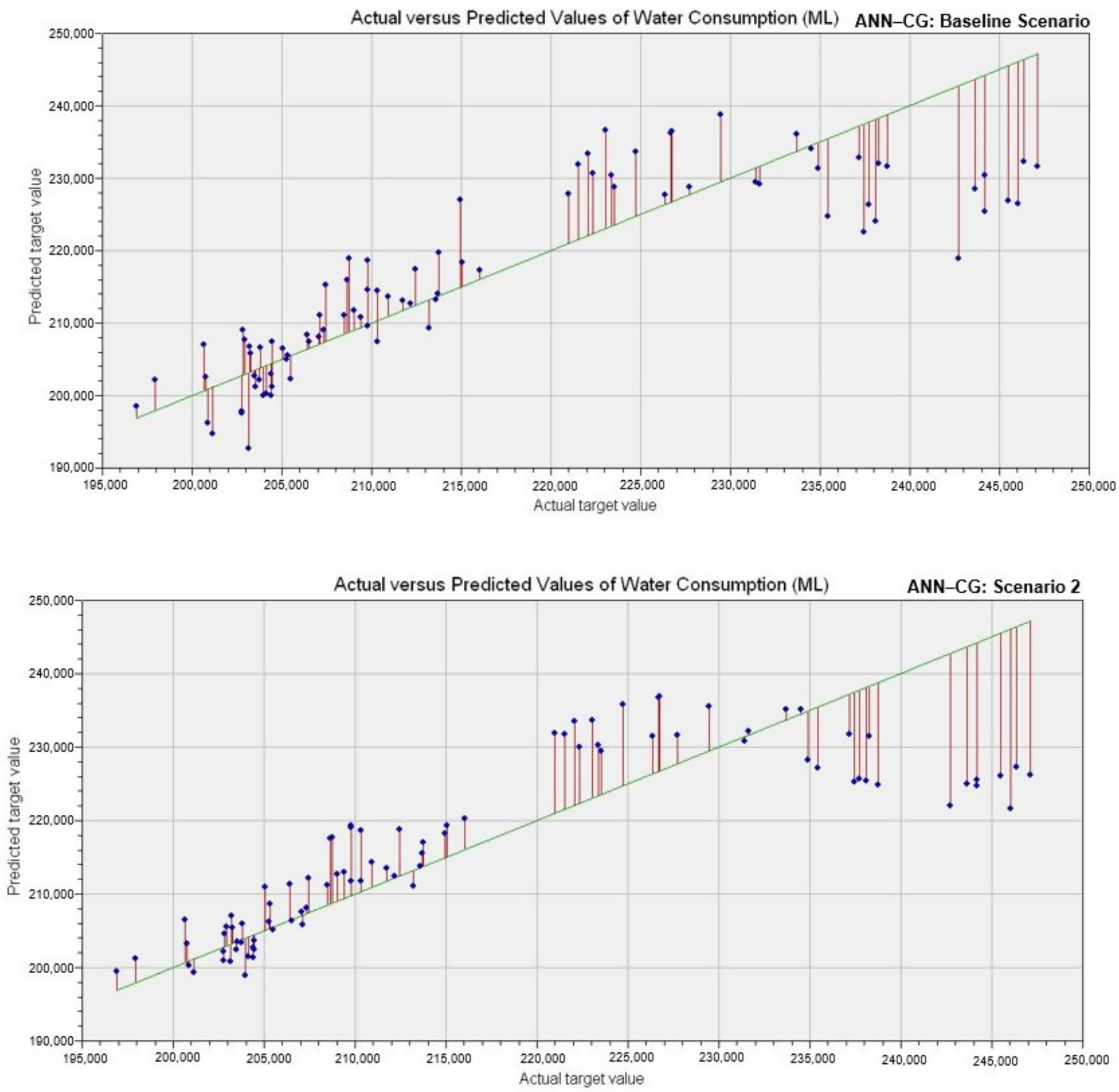
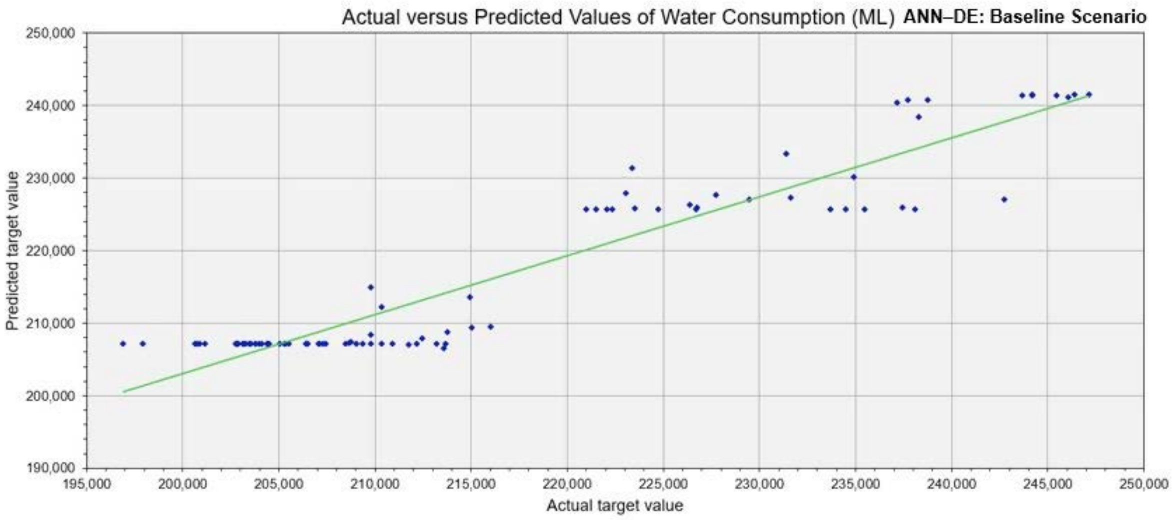


Figure 4-3: Scatter plots of observed and ANN-CG-predicted water demand for both scenarios



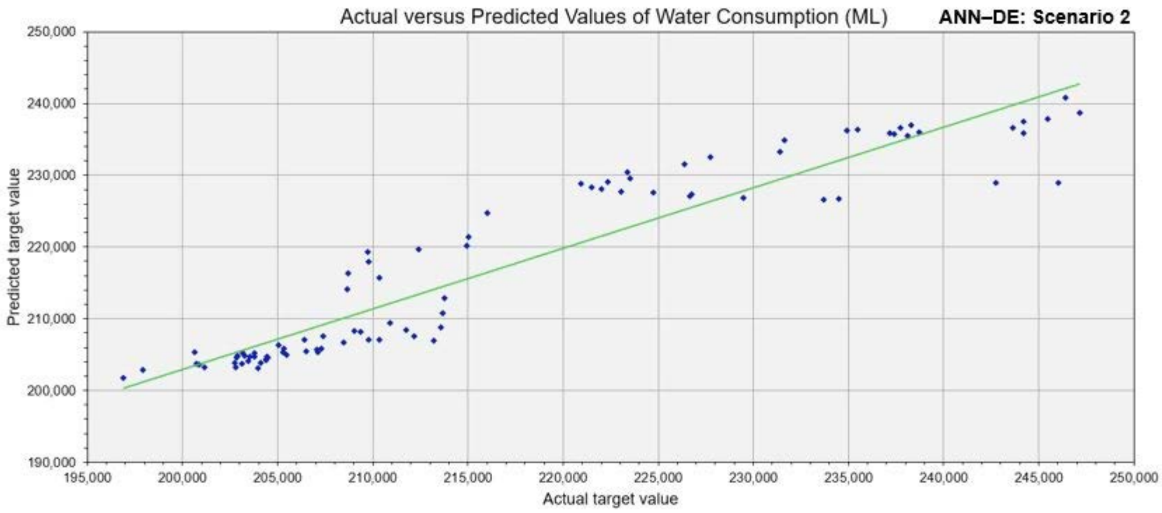


Figure 4-4: Scatter plots of observed and ANN-DE-predicted water demand for both scenarios

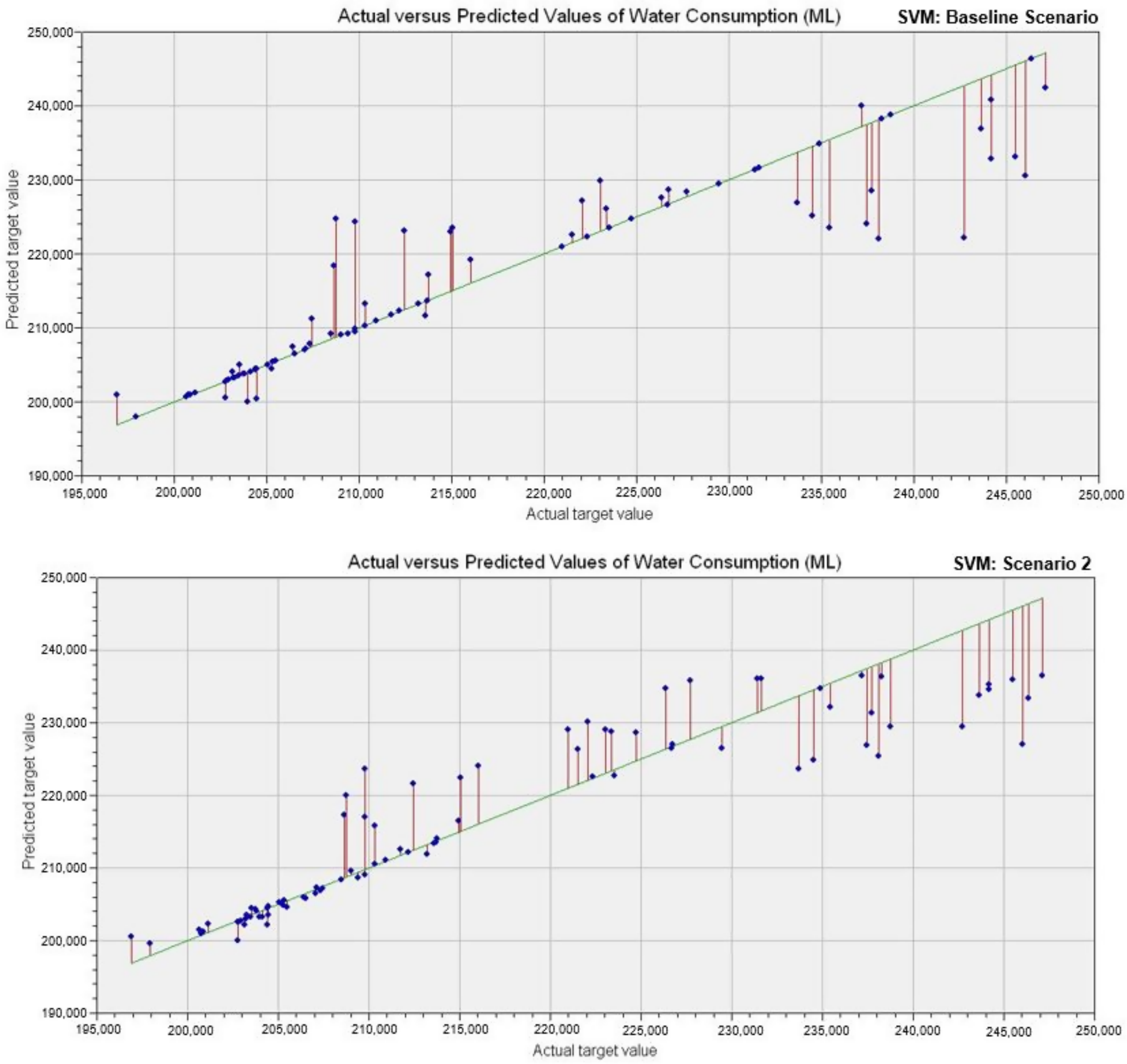


Figure 4-5: Scatter plots of observed and SVM-predicted water demand for both scenarios

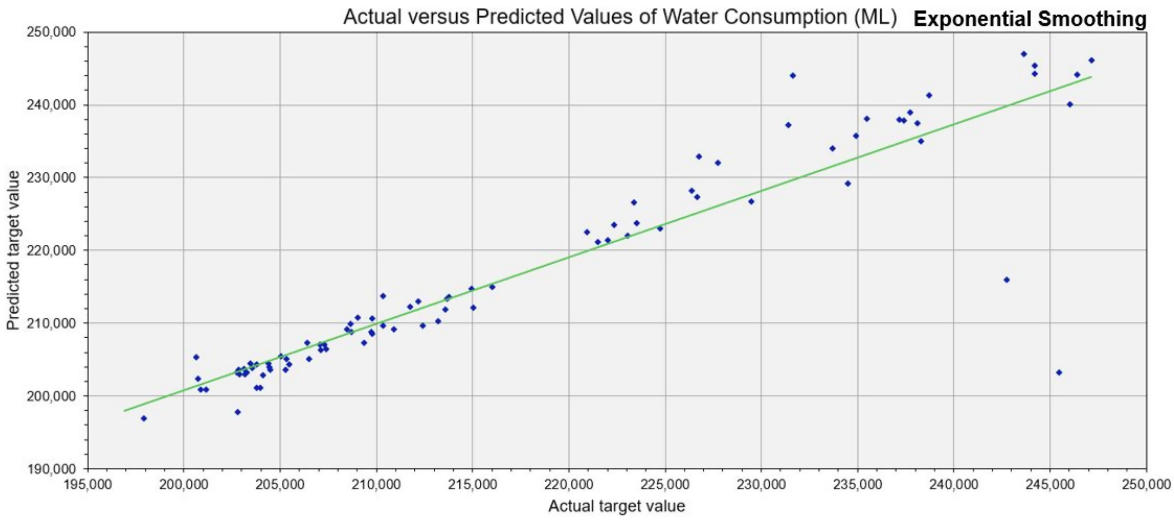


Figure 4-6: Scatter plots of observed and ES_m-predicted water demand

4.8 CONCLUSION

This study has investigated the potential of three soft computing techniques, namely, ANN-CG, ANN-DE and SVM against the conventional ES_m and MLR in estimating the water consumption of the City of Ekurhuleni. The aim was to determine the most superior technique in terms of accuracy for water demand forecasting in the City of Ekurhuleni using three performance evaluation metrics (R^2, RMSE and MAPE). Two scenarios were implemented in the study. The first scenario referred to as the baseline scenario involved the use of all potential explanatory variables that could influence water consumption in the City, while the second employed correlation analysis for model input selection. Eight models were developed in total; four for each scenario. Results showed that all the baseline scenario models suffered from overfitting problems suggesting parameter redundancy or irrelevancy in the input vector space of the models. The ANN-DE model however produced the best performance across the three performance evaluation metrics adopted in predicting billed water consumption. Results also showed that all the scenario 2 models, generally, outperformed the baseline scenario models. This suggests that the adoption of correlation analysis in scenario 2 was successful in reducing the dimensionality of the models, thereby preventing overparameterization. The ANN-DE model was the most superior of the four models in terms of accuracy, while the SVM, conventional MLR and ANN-CG came second, third and last respectively. Results also show that the performance of the ANN-DE model was superior to that obtained from

ES_m – a standard time series model. Furthermore, the results obtained show that the DE algorithm exhibited superiority and robustness over the CG algorithm in terms of ANN learning, producing a model with higher accuracy and less complex network architecture. This study therefore proves that the integration of evolutionary computation techniques like DE can be beneficial to the water demand modeling community as it may assist in providing sustainable solutions to complex water demand and supply problems. Future studies could investigate the potential of other specialized computations such as Bayesian Optimization for model architecture determination and hyperparameter configuration.

4.9 RESEARCH OUTPUT

1. **Oyebode, O.** & Ighravwe, D. E. (2019). Urban water demand forecasting: A comparative evaluation of conventional and soft computing techniques. Resources 8 (3): 156. https://doi.org/10.3390/resources8030156

CHAPTER 5

GENERAL CONCLUSIONS AND RECOMMENDATIONS

5.1 GENERAL CONCLUSIONS

The primary focus of this research was to develop an artificial intelligent model for municipal water demand forecasting to improve on current practices in predictive analysis of water demand.

As mentioned in chapter 1, this study has five specific objectives which were:

1. To conduct an extensive review of the extent to which evolutionary-inspired artificial intelligent have been employed in water demand modelling.
2. To identify and analyze the factors that affect municipal water demand.
3. To develop an intelligent system model for municipal water demand prediction.
4. To evaluate the performance of the intelligent model.
5. To propose a framework for sustainable allocation of water resources.

Specific objective 1 was addressed in chapter 2 wherein a comprehensive state-of-art-review of EC techniques in water demand forecasting applications was presented. The review unearthed challenges and current gaps in knowledge, within the predictive analytics and water demand forecasting domains, that needed to be solved and bridged respectively. The challenges and gaps in knowledge identified in the reviews can be summarized as: (i) understanding the internal mechanism of ANNs, with specific emphasis of model accuracy, complexity and architecture, when deployed to solve real-world problems; (ii) inability of conventional and traditional water demand forecasting models to account for inherent nonlinearities in explanatory variables and accurately predict water demand; (iii) need to incorporate weather and economic factors as model inputs to enable water demand models cater for the impacts of weather and socioeconomic factors on water demand; (iv) need to shift focus to medium- to long-term water demand forecasting as existing studies have majorly focused on short-term forecasting; (iv) a lack of studies focusing on the application of EC techniques in water demand modelling in developing countries, especially in Africa; (v) need for EC techniques to extend their significance beyond advisory roles and be positioned as an effective tool for developing proper standard operating procedures in the water sector; and (vi) need for a sustainable framework that syncs conventional EC techniques with social aspects of the society for sustainable water resource allocation.

To address the gap relating to the integration of social aspects of society with EC techniques, a novel sustainable framework (IWDMMF) was proposed in the latter section of chapter 2. IWDMMF offers a platform for extending the reach of EC techniques beyond just ensuring optimality in water management, to assessing their wider impacts on the socioeconomic aspect of society in line with equity and justice requirements. Chapter 2 therefore satisfies specific objectives 1 and 5 of this study.

In chapter 3, a DE-inspired ANN model (ANN-DE) was developed and applied to forecast water demand in the City of Ekurhuleni, Johannesburg, South Africa. The development of the ANN-DE model comprises a combination of genetic operations relating to a classic DE algorithm with feature selection and early stopping techniques, thereby introducing a new scheme and systematic approach for improving the accuracy and optimizing the complexity of ANN models. The model development also entails the use of weather and socioeconomic variables as input variables. Results obtained show that the models developed produced a good representation of the water demand pattern in the City of Ekurhuleni; reproducing the peaks and troughs including the sharp spikes. The DE algorithm enabled a smooth and adaptive learning process in the ANN model. A sensitivity analysis of the DE algorithm ensured selection of optimal control parameters (CR and F) that, in turn, produced a model with high accuracy and minimal complexity. The integration of an early stopping criterion also ensured that the model did not experience overfitting. A computation of the contribution of the impacts of each input variable show that weather and socioeconomic variables could have significant impacts on water demand. The results further show that the application of feature selection techniques enabled a reduction in dimensionality, thereby eliminating parameter irrelevancy or redundancy, and consequently preventing the model from suffering from overparameterization. The work undertaken in this chapter and the results obtained therein addresses the specific objectives 2 and 3 of this study.

In chapter 4, the performance of the optimal ANN-DE model developed in chapter 3 was compared against four techniques comprising the conventionally-used MLR and ES_m, and two other prominent soft computing techniques – a SVM and a conjugate gradient-trained multilayer perceptron (ANN-CG). Results show that the DE-inspired ANN model was superior to the other four techniques across two input combination scenarios in terms of model accuracy and complexity. DE showcased robustness in fine-tuning algorithm parameter values thereby producing good performance in terms of the solution efficiency and quality. It was found that the ANN-DE model, with lesser complex architecture and lower computational demand and time, achieved a higher accuracy than the ANN-CG model. This suggests that DE could be a better training algorithm for ANNs than CG as it offers a good compromise between accuracy and complexity. This chapter thus addresses specific objective 4 of this study.

Generally, it is concluded that the synergetic use of feature selection techniques, DE algorithm and early stopping criterion could be used in overcoming the limitations of ANN and developing an improved and more reliable water demand forecasting model. The ANN modelling approach suggested in this study serves as an alternative to the simplistic and thumb-rule concept which forms the basis of conventional water demand forecasting techniques, thereby making them to be not accurate or reliable enough for every situation. The methodologies, principles and techniques behind this study fosters sustainable development and thus could be adopted in planning and management of water resources towards the overall prosperity of the society.

5.2 NOVELTIES AND CONTRIBUTIONS TO THE BODY OF KNOWLEDGE

The following novelties and contributions to the general body of knowledge are accomplished and published as enumerated in chapter one:

1. A state-of-art review of the extent of use of EC techniques in water demand modelling, identifying important research challenges and future directions while recommending strategies for their use by policy-makers in meeting sustainable development goals (SDGs).

2. Pioneering the use of DE algorithm, in water demand forecasting, to optimize the learning process and model architecture of multilayer feedforward neural networks, and comparing its performance with conventional and emerging soft computing techniques.

3. Incorporation of climatic and socioeconomic variables for long-term water demand forecasting to account for the impact of climate and socioeconomic variations. The study introduced a new variable (HDI) that accesses a municipality or city's overall achievement in terms of socioeconomic dimensions including life expectancy, education and income levels.

4. Introduction of a new methodology that comprises a combination of evolutionary computation, early stopping criterion and feature selection in developing water demand forecasting models.

5. Development of a novel integrated water demand and management modelling framework (IWDMMF) that enables water policy-makers to assess the wider impact of water demand management decisions through the principles of equity and justice. The novel framework provides a platform for integrating conventional EC techniques with social aspects of the society, and fosters the realization of the UN SDGs.

5.3 RECOMMENDATIONS AND FUTURE RESEARCH

The following recommendations suggest new research ideas that could be developed based on the outcome of this work.

1. Additional input variables like land use, water price could be investigated in future research for long-term water demand forecasting.

2. Future research could extend the application EC techniques to forecast water demand at other spatial and temporal scales including in-house end-use profiles.

3. Future research could also focus on extending the role of EC techniques to assessing the impacts of non-revenue water and nature-based solutions as well as water conservation and reuse strategies on water demand. These may include simulating the impacts of the use of water efficient

appliances, consumer behaviour, alternative water sources on water demand.

4. This study was limited by availability of data. Additional simulations of water demand using lengthier data samples may be undertaken when relevant information becomes readily available.
5. Application of IWDMMF for resolving real-world multi-objective water demand problems and conflicts is left for further studies.

REFERENCES

Abdul-Kader, H. 2009. Neural networks training based on differential evolution algorithm compared with other architectures for weather forecasting34. International Journal of Computer Science and Network Security 9 (3): 92-99.

Adamowski, J. and Karapataki, C. 2010. Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: evaluation of different ANN learning algorithms. Journal of Hydrologic Engineering 15 (10): 729-743.

Adeyemo, J., Oyebode, O. and Stretch, D. 2018. River Flow Forecasting Using an Improved Artificial Neural Network, EVOLVE-A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation VI. Springer, pp. 179-193.

Ali, A.M., Shafiee, M.E. and Berglund, E.Z. 2017. Agent-based modeling to simulate the dynamics of urban water supply: Climate, population growth, and water shortages. Sustainable cities and society 28: 420-434.

Almutaz, I., Ali, E., Khalid, Y. and Ajbar, A.H. 2013. A long-term forecast of water demand for a desalinated dependent city: case of Riyadh City in Saudi Arabia. Desalination and Water Treatment 51 (31-33): 5934-5941.

Amaranto, A., Munoz-Arriola, F., Corzo, G., Solomatine, D.P. and Meyer, G. 2018. Semi-seasonal groundwater forecast using multiple data-driven models in an irrigated cropland. Journal of Hydroinformatics: (in press).

Arbués, F., Garcıa-Valiñas, M.Á. and Martınez-Espiñeira, R. 2003. Estimation of residential water demand: a state-of-the-art review. The Journal of Socio-Economics 32 (1): 81-102.

Ardabili, S.F., Najafi, B., Shamshirband, S., Minaei Bidgoli, B., Deo, R.C. and Chau, K.-w. 2018. Computational intelligence approach for modeling hydrogen production: a review. Engineering Applications of Computational Fluid Mechanics 12 (1): 438-458.

Atkinson, R., Van Zyl, J., Walters, G. and Savic, D. 2000. Genetic algorithm optimisation of level-controlled pumping station operation. Water network modelling for optimal design and management: 79-90.

Azhagusundari, B. and Thanamani, A.S. 2013. Feature selection based on information gain. International Journal of Innovative Technology and Exploring Engineering (IJITEE) 2 (2): 18-21.

Babel, M., Gupta, A.D. and Pradhan, P. 2007. A multivariate econometric approach for domestic water demand modeling: an application to Kathmandu, Nepal. Water Resources Management 21 (3): 573-589.

Babel, M.S. and Shinde, V.R. 2011. Identifying prominent explanatory variables for water demand prediction using artificial neural networks: a case study of Bangkok. Water resources management 25 (6): 1653-1676.

Bäck, T., Rudolph, G. and Schwefel, H.-P. 1993. Evolutionary programming and evolution strategies: Similarities and differences, In Proceedings of the Second Annual Conference on Evolutionary Programming. Citeseer.

Bakker, K. 2007. The “commons” versus the “commodity”: Alter-globalization, anti-privatization and the human right to water in the global south. Antipode 39 (3): 430-455.

Bakker, M., Van Duist, H., Van Schagen, K., Vreeburg, J. and Rietveld, L. 2014. Improving the performance of water demand forecasting models by using weather input. Procedia Engineering 70: 93-102.

Bárdossy, G., Halász, G. and Winter, J. 2009. Prognosis of urban water consumption using hybrid fuzzy algorithms. Journal of Water Supply: Research and Technology-AQUA 58 (3): 203-211.

Behera, S., Sahoo, S. and Pati, B. 2015. A review on optimization algorithms and application to wind energy integration to grid. Renewable and Sustainable Energy Reviews 48: 214-227.

Bennett, C., Stewart, R.A. and Beal, C.D. 2013. ANN-based residential water end-use demand forecasting model. Expert systems with applications 40 (4): 1014-1023.

Bhagwat, P.P. and Maity, R. 2013. Hydroclimatic streamflow prediction using least square-support vector regression. ISH Journal of Hydraulic Engineering 19 (3): 320-328.

Bi, W., Maier, H.R. and Dandy, G.C. 2016. Impact of starting position and searching mechanism on the evolutionary algorithm convergence rate. Journal of Water Resources Planning and Management 142 (9): 04016026.

Billings, R.B. and Jones, C.V. 2008. Forecasting urban water demand, 2nd ed. American Water Works Association.

Blokker, E., Vreeburg, J. and Van Dijk, J. 2010. Simulating residential water demand with a stochastic end-use model. Journal of Water Resources Planning and Management 136 (1): 19-26.

Blokker, M., Vloerbergh, I. and Buchberger, S. 2012. Estimating peak water demands in hydraulic systems II - future trends, WDSA 2012: 14th Water Distribution Systems Analysis Conference, 24-27 September 2012. Engineers Australia, Adelaide, South Australia, pp. 1138-1147.

Bowden, G.J., Dandy, G.C. and Maier, H.R. 2005. Input determination for neural network models in water resources applications. Part 1—background and methodology. Journal of Hydrology 301 (1-4): 75-92.

Buchberger, S.G. and Wu, L. 1995. Model for instantaneous residential water demands. Journal of hydraulic engineering 121 (3): 232-246.

Cabral, M. 2014. Water Demand Projection in Distribution Systems using a Novel Scenario Planning Approach, Department of Civil Engineering and Architecture. Technical University of Lisbon, Lisbon, p. 131.

Caiado, J. 2009. Performance of combined double seasonal univariate time series models for forecasting water demand. Journal of Hydrologic Engineering 15 (3): 215-222.

Capeetc 2018. Water Crisis: Cape Town water map goes live. Available: http://www.capetownetc.com/news/green-light-for-cape-town-water-map-despite-objections/ [Accessed: 04 March 2018].

Casini, M., Bastianoni, S., Gagliardi, F., Gigliotti, M., Riccaboni, A. and Betti, G. 2019. Sustainable Development Goals indicators: A methodological proposal for a Multidimensional Fuzzy Index in the Mediterranean area. Sustainability 11 (4): 1198.

Ch, S., Anand, N., Panigrahi, B.K. and Mathur, S. 2013. Streamflow forecasting by SVM with quantum behaved particle swarm optimization. Neurocomputing 101: 18-23.

Chadalawada, J., Havlicek, V. and Babovic, V. 2017. A Genetic Programming Approach to System Identification of Rainfall-Runoff Models. Water Resources Management 31 (12): 3975-3992.

Chang, F.-J., Tsai, Y.-H., Chen, P.-A., Coynel, A. and Vachaud, G. 2015. Modeling water quality in an urban river using hydrological factors–Data driven approaches. Journal of environmental management 151: 87-96.

Chang, L.-C., Chang, F.-J., Wang, K.-W. and Dai, S.-Y. 2010. Constrained genetic algorithms for optimizing multi-use reservoir operation. Journal of Hydrology 390 (1-2): 66-74.

Chomba, M., Hill, T., Nkhata, B. and Förster, J. 2017. Paradigms for water allocation in river basins: a society-science-practice perspective from Southern Africa. Water Policy: wp2017130.

CoCT 2018. City of Cape Town: Water and Sanitation Tariffs. Available: http://resource.capetown.gov.za/documentcentre/Documents/Forms%2c%20notices%2c%20tariffs%20and%20lists/Water%20Sanitation_restriction_Tariffs_L6.pdf [Accessed: 04 March 2018].

Coomes, P., Rockaway, T., Rivard, J. and Kornstein, B. 2010. North America residential water usage trends since 1992. Water Research Foundation and USEPA, Colorado.

CSIR 2019. The Neighbourhood Planning and Design Guide (Red Book) – Creating Sustainable Human Settlements, Version 1.1 ed. The South African Government, South Africa, p. 56.

Dhungel, R. and Fiedler, F. 2014. Price elasticity of water demand in a small college town: an inclusion of system dynamics approach for water demand forecast. Air, Soil and Water Research 7: ASWR. S15395.

Di Nardo, A., Di Natale, M., Gisonni, C. and Iervolino, M. 2015. A genetic algorithm for demand pattern and leakage estimation in a water distribution network. Journal of Water Supply: Research and Technology-Aqua 64 (1): 35-46.

Ding, S., Li, H., Su, C., Yu, J. and Jin, F. 2013. Evolutionary artificial neural networks: a review. Artificial Intelligence Review 39 (3): 251-260.

Donkor, E.A., Mazzuchi, T.A., Soyer, R. and Alan Roberson, J. 2012. Urban water demand forecasting: review of methods and models. Journal of Water Resources Planning and Management 140 (2): 146-159.

Dos Santos, C.C. and Pereira Filho, A.J. 2014. Water demand forecasting model for the metropolitan area of São Paulo, Brazil. Water Resources Management 28 (13): 4401-4414.

Dragoi, E.-N., Curteanu, S., Leon, F., Galaction, A.-I. and Cascaval, D. 2011. Modeling of oxygen mass transfer in the presence of oxygen-vectors using neural networks developed by differential evolution algorithm. Engineering Applications of Artificial Intelligence 24 (7): 1214-1226.

Ehteram, M., Karami, H., Mousavi, S.-F., Farzin, S. and Kisi, O. 2017. Evaluation of contemporary evolutionary algorithms for optimization in reservoir operation and water supply. Journal of Water Supply: Research and Technology-Aqua: jws2017109.

Elshorbagy, A., Corzo, G., Srinivasulu, S. and Solomatine, D. 2010. Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology-Part 1: Concepts and methodology. Hydrology and Earth System Sciences 14 (10): 1931-1941.

Enitan, A.M., Adeyemo, J., Olofintoye, O.O., Bux, F. and Swalaha, F.M. 2014. Multi-objective Optimization of Methane Producing UASB Reactor Using a Combined Pareto Multi–objective Differential Evolution Algorithm (CPMDE), EVOLVE-A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation V. Springer, pp. 321-334.

Escrig-Olmedo, E., Muñoz-Torres, M.J., Fernández-Izquierdo, M.Á. and Rivera-Lirio, J.M. 2017. Measuring corporate environmental performance: a methodology for sustainable development. Business Strategy and the Environment 26 (2): 142-162.

Fagiani, M., Squartini, S., Gabrielli, L., Spinsante, S. and Piazza, F. 2015. Domestic water and natural gas demand forecasting by using heterogeneous data: A preliminary study, Advances in Neural Networks: Computational and Theoretical Issues. Springer, pp. 185-194.

Ferreira, C. 2001. A New Adaptive Algorithm for Solving Problems. Complex Systems. Gene Expression Programming 13 (2): 87-129.

Ferreira, C. 2006. Gene expression programming: mathematical modeling by an artificial intelligence. Springer, Berlin-Heidelberg.

Firat, M., Yurdusev, M.A. and Turan, M.E. 2009. Evaluation of artificial neural network techniques for municipal water consumption modeling. Water resources management 23 (4): 617-632.

Fjeldstad, O.H. 2004. What's trust got to do with it? Non-payment of service charges in local authorities in South Africa. The Journal of Modern African Studies 42 (4): 539-562.

Fox, C., McIntosh, B. and Jeffrey, P. 2009. Classifying households for water demand forecasting using physical property characteristics. Land Use Policy 26 (3): 558-568.

Gaudin, S. 2006. Effect of price information on residential water demand. Applied economics 38 (4): 383-393.

Gaudin, S., Griffin, R.C. and Sickles, R.C. 2001. Demand specification for municipal water management: evaluation of the Stone-Geary form. Land Economics 77 (3): 399-422.

Ghalehkhondabi, I., Ardjmand, E., Young, W.A. and Weckman, G.R. 2017. Water demand forecasting: review of soft computing methods. Environmental monitoring and assessment 189 (7): 313.

Giustolisi, O. and Savic, D.A. 2006. A symbolic data-driven technique based on evolutionary polynomial regression. Journal of Hydroinformatics 8 (3): 207-222.

Govindaraju, R.S. 2000. Artificial neural networks in hydrology. I: Preliminary concepts. Journal of hydrologic Engineering 5 (2): 115-123.

Gubuza, D. 2017. Water Conservation and Water Demand Management in the City of Ekurhuleni: On-Site Leak Repair, in: Ekurhuleni (Ed.), Presentation at Rand Water Services Forum, Johannesburg.

Guven, A. and Aytek, A. 2009. New approach for stage–discharge relationship: gene-expression programming. Journal of Hydrologic Engineering 14 (8): 812-820.

Guven, A. and Kişi, Ö. 2011. Estimation of suspended sediment yield in natural rivers using machine-coded linear genetic programming. Water resources management 25 (2): 691-704.

Hall, M.A. 1999. Feature selection for discrete and numeric class machine learning. 1-16.

Hanke, M. 2017. Conjugate gradient type methods for ill-posed problems, 1st Edition ed. Routledge, New York.

Hassan-Esfahani, L., Torres-Rua, A. and McKee, M. 2015. Assessment of optimal irrigation water allocation for pressurized irrigation system using water balance approach, learning machines, and remotely sensed data. Agricultural Water Management 153: 42-50.

House-Peters, L.A. and Chang, H. 2011. Urban water demand modeling: Review of concepts, methods, and organizing principles. Water Resources Research 47 (5).

Hu, T., Wu, F. and Zhang, X. 2007. Rainfall–runoff modeling using principal component analysis and neural network. Hydrology Research 38 (3): 235-248.

Huang, W.C., Yuan, L.C. and Lee, C.M. 2002. Linking genetic algorithms with stochastic dynamic programming to the long-term operation of a multireservoir system. Water Resources Research 38 (12): 1-9.

Huntra, P. and Keener, T. 2017. Evaluating the Impact of Meteorological Factors on Water Demand in the Las Vegas Valley Using Time-Series Analysis: 1990–2014. ISPRS International Journal of Geo-Information 6 (8): 249.

Husselmann, M. and Van Zyl, J. 2006. Effect of stand size and income on residential water demand. Journal of the South African Institution of Civil Engineering 48 (3): 12-16.

Hyndman, R., Koehler, A.B., Ord, J.K. and Snyder, R.D. 2008. Forecasting with exponential smoothing: the state space approach. Springer Science & Business Media.

IDP 2018. Integrated Development Plan of City of Ekurhuleni 2017/2018 to 2020/2021, Annexure A ed. City of Ekurhuleni, City of Ekurhuleni.

Ilonen, J., Kamarainen, J.-K. and Lampinen, J. 2003. Differential evolution training algorithm for feed-forward neural networks. Neural Processing Letters 17 (1): 93-105.

IPART 2004. Residential water use in Sydney, the Blue Mountains and Illawarra: Results from the 2003 household survey. Independent Pricing and Regulatory Tribunal of New South Wales, New South Wales.

Jacobs, H., Geustyn, L., Loubser, B. and Van Der Merwe, B. 2004. Estimating residential water demand in Southern Africa. Journal of the South African Institution of Civil Engineering= Joernaal Van Die Suid-Afrikaanse Instituut Van Siviele Ingenieurswese 46 (4): 2-13.

Jacobs, H. and Haarhoff, J. 2007. Prioritisation of parameters influencing residential water use and wastewater flow. Journal of Water Supply: Research and Technology—AQUA 56 (8): 495-514.

Ji, G., Wang, J., Ge, Y. and Liu, H. 2014. Urban water demand forecasting by LS-SVM with tuning based on elitist teaching-learning-based optimization, The 26th Chinese Control and Decision Conference (2014 CCDC). IEEE, Changsha, pp. 3997-4002.

Johnson, H., South, N. and Walters, R. 2016. The commodification and exploitation of fresh water: Property, human rights and green criminology. International Journal of Law, Crime and Justice 44: 146-162.

Kalteh, A.M., Hjorth, P. and Berndtsson, R. 2008. Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application. Environmental Modelling & Software 23 (7): 835-845.

Karimi, S., Shiri, J., Kisi, O. and Shiri, A.A. 2016. Short-term and long-term streamflow prediction by using'wavelet–gene expression'programming approach. ISH Journal of Hydraulic Engineering 22 (2): 148-162.

Karimi, Z., Kashani, M.M.R. and Harounabadi, A. 2013. Feature Ranking in Intrusion Detection Dataset using Combination of Filtering Methods. International Journal of Computer Applications 78 (4): 21-27.

Kim, J.H., Hwang, S.H. and Shin, H.S. 1999. An optimal neural network model for daily water demand forecasting, 29th Annual Water Resources Planning and Management Conference (WRPMD '99). ASCE, Arizona, pp. 1-10.

Kim, J.H., Hwang, S.H. and Shin, H.S. 2001. A neuro-genetic approach for daily water demand forecasting. KSCE Journal of Civil Engineering 5 (3): 281-288.

Kişi, Ö. 2010. River suspended sediment concentration modeling using a neural differential evolution approach. Journal of Hydrology 389 (1-2): 227-235.

Kisi, O. and Cigizoglu, H.K. 2007. Comparison of different ANN techniques in river flow prediction. Civil Engineering and Environmental Systems 24 (3): 211-231.

Kofinas, D., Mellios, N., Papageorgiou, E. and Laspidou, C. 2014. Urban water demand forecasting for the island of Skiathos. Procedia Engineering 89: 1023-1030.

Kornfeld, I.E. 2012. Water: A Public Good or a Commodity?, Proceedings of the ASIL Annual Meeting. Cambridge University Press, pp. 49-52.

Koza, J.R. 1994. Genetic programming as a means for programming computers by natural selection. Statistics and computing 4 (2): 87-112.

Kozłowski, E., Kowalska, B., Kowalski, D. and Mazurkiewicz, D. 2018. Water demand forecasting by trend and harmonic analysis. Archives of Civil and Mechanical Engineering 18 (1): 140-148.

Lehner, A., Erlacher, C., Schlögl, M., Wegerer, J., Blaschke, T. and Steinnocher, K. 2018. Can iso-defined urban sustainability indicators be derived from remote sensing: An expert weighting approach. Sustainability 10 (4): 1268.

Lewis, A. and Randall, M. 2017. Solving multi-objective water management problems using evolutionary computation. Journal of Environmental Management 204: 179-188.

Londhe, S. and Charhate, S. 2010. Comparison of data-driven modelling techniques for river flow forecasting. Hydrological Sciences Journal–Journal des Sciences Hydrologiques 55 (7): 1163-1174.

Londhe, S.N. and Narkhede, S. 2017. Forecasting stream flow using hybrid neuro-wavelet technique. ISH Journal of Hydraulic Engineering: 1-10.

Mafi, S. and Amirinia, G. 2017. Forecasting hurricane wave height in Gulf of Mexico using soft computing methods. Ocean Engineering 146: 352-362.

Maier, H., Razavi, S., Kapelan, Z., Matott, L., Kasprzyk, J. and Tolson, B. 2019. Introductory overview: Optimization using evolutionary algorithms and other metaheuristics. Environmental Modelling & Software.

Maier, H.R. and Dandy, G.C. 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environmental modelling & software 15 (1): 101-124.

Maier, H.R., Jain, A., Dandy, G.C. and Sudheer, K.P. 2010. Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions. Environmental Modelling & Software 25 (8): 891-909.

Maier, H.R., Kapelan, Z., Kasprzyk, J., Kollat, J., Matott, L.S., Cunha, M., Dandy, G.C., Gibbs, M.S., Keedwell, E. and Marchi, A. 2014. Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions. Environmental Modelling & Software 62: 271-299.

Makropoulos, C., Natsis, K., Liu, S., Mittas, K. and Butler, D. 2008. Decision support for sustainable option selection in integrated urban water management. Environmental modelling & software 23 (12): 1448-1460.

Mangal, A. and Holm, E.A. 2018. A comparative study of feature selection methods for stress hotspot classification in materials. Integrating Materials and Manufacturing Innovation: 1-9.

Marques, J., Cunha, M. and Savić, D. 2018. Many-objective optimization model for the flexible design of water distribution networks. Journal of environmental management 226: 308-319.

Marques, J., Cunha, M. and Savić, D.A. 2015. Multi-objective optimization of water distribution systems based on a real options approach. Environmental Modelling & Software 63: 1-13.

Martí, P., Shiri, J., Duran-Ros, M., Arbat, G., De Cartagena, F.R. and Puig-Bargués, J. 2013. Artificial neural networks vs. gene expression programming for estimating outlet dissolved oxygen in micro-irrigation sand filters fed with effluents. Computers and electronics in agriculture 99: 176-185.

Menapace, A., Avesani, D., Righetti, M., Bellin, A. and Pisaturo, G. 2018. Uniformly Distributed Demand EPANET Extension. Water Resources Management 32 (6): 2165-2180.

Mirghani, B.Y., Mahinthakumar, K.G., Tryby, M.E., Ranjithan, R.S. and Zechman, E.M. 2009. A parallel evolutionary strategy based simulation–optimization approach for solving groundwater source identification problems. Advances in Water Resources 32 (9): 1373-1385.

Modaresi, F., Araghinejad, S. and Ebrahimi, K. 2018. A Comparative Assessment of Artificial Neural Network, Generalized Regression Neural Network, Least-Square Support Vector Regression, and K-Nearest Neighbor Regression for Monthly Streamflow Forecasting in Linear and Nonlinear Conditions. Water Resources Management 32 (1): 243-258.

Mohamed, M.M. and Al-Mualla, A.A. 2010. Water demand forecasting in Umm Al-Quwain (UAE) using the IWR-MAIN specify forecasting model. Water resources management 24 (14): 4093-4120.

Mohammadi, R., Ghomi, S.F. and Zeinali, F. 2014. A new hybrid evolutionary based RBF networks method for forecasting time series: a case study of forecasting emergency supply demand time series. Engineering Applications of Artificial Intelligence 36: 204-214.

Montgomery, J. and Chen, S. 2010. An analysis of the operation of differential evolution at high and low crossover rates, Evolutionary Computation (CEC), 2010 IEEE Congress on. IEEE, pp. 1-8.

Monyei, C. and Adewumi, A. 2017. Demand Side Management potentials for mitigating energy poverty in South Africa. Energy Policy 111: 298-311.

Monyei, C.G., Adewumi, A.O., Obolo, M.O. and Sajou, B. 2017. Nigeria's energy poverty: Insights and implications for smart policies and framework towards a smart Nigeria electricity network. Renewable and Sustainable Energy Reviews 81 (1): 1582-1601.

Mtembu, N. 2018. Cape Town water woes: Dialysis patients at risk. IOL - Weekend Argus. Available: https://www.iol.co.za/weekend-argus/cape-town-water-woes-dialysis-patients-at-risk-13466781 [Accessed: 01 March 2018].

Muleta, M.K. and Nicklow, J.W. 2004. Joint application of artificial neural networks and evolutionary algorithms to watershed management. Water resources management 18 (5): 459-482.

Mun, J. 2006. Modeling risk: Applying Monte Carlo simulation, real options analysis, forecasting, and optimization techniques. John Wiley & Sons.

Nasseri, M., Moeini, A. and Tabesh, M. 2011. Forecasting monthly urban water demand using Extended Kalman Filter and Genetic Programming. Expert Systems with Applications 38 (6): 7387-7395.

Neal, M.J., Lukasiewicz, A. and Syme, G.J. 2014. Why justice matters in water governance: some ideas for a 'water justice framework'. Water Policy 16 (S2): 1-18.

Nel, N., Jacobs, H.E., Loubser, C. and Du Plessis, K.J. 2017. Supplementary household water sources to augment potable municipal supply in South Africa. Water SA 43 (4): 553-562.

Nicklow, J., Reed, P., Savic, D., Dessalegne, T., Harrell, L., Chan-Hilton, A., Karamouz, M., Minsker, B., Ostfeld, A. and Singh, A. 2009. State of the art for genetic algorithms and beyond in water resources planning and management. Journal of Water Resources Planning and Management 136 (4): 412-432.

Nouiri, I. 2014. Multi-objective tool to optimize the water resources management using genetic algorithm and the Pareto optimality concept. Water resources management 28 (10): 2885-2901.

Olofintoye, O., Otieno, F. and Adeyemo, J. 2016. Real-time optimal water allocation for daily hydropower generation from the Vanderkloof dam, South Africa. Applied Soft Computing 47: 119-129.

Olofintoye, O.O. 2015. Real time optimal water allocation in the Orange River catchment in South Africa, Department of Civil Engineering and Surveying. Durban University of Technology, Durban, p. 160.

Oshima, N. and Kosuda, T. 1998. Distribution reservoir control with demand prediction using deterministic-chaos method. Water science and technology 37 (12): 389-395.

Oyebode, O. and Adeyemo, J. 2014. Reservoir inflow forecasting using differential evolution trained neural networks, EVOLVE-A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation V. Springer, pp. 307-319.

Oyebode, O., Adeyemo, J. and Otieno, F. 2014a. Monthly streamflow prediction with limited hydro-climatic variables in the upper Mkomazi River, South Africa using genetic programming. Fresenius Environmental Bulletin 23 (3): 708-719.

Oyebode, O., Babatunde, D.E., Monyei, C.G. and Babatunde, O.M. 2019. Water demand modelling using evolutionary computation techniques: Integrating water equity and justice for realization of the sustainable development goals. Heliyon 5 (11): e02796.

Oyebode, O. and Stretch, D. 2019. Neural network modeling of hydrological systems: A review of implementation techniques. Natural Resource Modeling 32 (1): e12189.

Oyebode, O.K. 2014. Modelling streamflow response to hydro-climatic variables in the Upper Mkomazi River, South Africa, Masters Thesis: Department of Civil Engineering and Surveying. Durban University of Technology, Durban.

Oyebode, O.K., Otieno, F.A.O. and Adeyemo, J. 2014b. Review of three data-driven modelling techniques for hydrological modelling and forecasting. Fresenius Environmental Bulletin. 23 (7): 1443-1454.

Papageorgiou, E.I., Poczęta, K. and Laspidou, C. 2016. Hybrid model for water demand prediction based on fuzzy cognitive maps and artificial neural networks, 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). IEEE, pp. 1523-1530.

Perea, R.G., Poyato, E.C., Montesinos, P. and Díaz, J.R. 2015. Irrigation demand forecasting using artificial neuro-genetic networks. Water Resources Management 29 (15): 5551-5567.

Phyu, T.Z. and Oo, N.N. 2016. Performance Comparison of Feature Selection Methods, MATEC Web of Conferences. EDP Sciences, p. 06002.

Piotrowski, A.P. 2014. Differential evolution algorithms applied to neural network training suffer from stagnation. Applied Soft Computing 21: 382-406.

Piotrowski, A.P. and Napiorkowski, J.J. 2011. Optimizing neural networks for river flow forecasting–Evolutionary Computation methods versus the Levenberg–Marquardt approach. Journal of Hydrology 407 (1-4): 12-27.

Polebitski, A.S. and Palmer, R.N. 2009. Seasonal residential water demand forecasting for census tracts. Journal of Water Resources Planning and Management 136 (1): 27-36.

Pulido-Calvo, I. and Gutierrez-Estrada, J.C. 2009. Improved irrigation water demand forecasting using a soft-computing hybrid model. Biosystems Engineering 102 (2): 202-218.

Pulido-Calvo, I., Montesinos, P., Roldán, J. and Ruiz-Navarro, F. 2007. Linear regressions and neural approaches to water demand forecasting in irrigation districts with telemetry systems. Biosystems Engineering 97 (2): 283-293.

Qi, C. and Chang, N.-B. 2011. System dynamics modeling for municipal water demand estimation in an urban region under uncertain economic impacts. Journal of environmental management 92 (6): 1628-1641.

Qian, B., Wang, L., Hu, R., Huang, D. and Wang, X. 2009. A DE-based approach to no-wait flow-shop scheduling. Computers & Industrial Engineering 57 (3): 787-805.

Qu, J., Cao, L. and Zhou, J. 2010. Differential evolution-optimized general regression neural network and application to forecasting water demand in Yellow River Basin, 2010 2nd International Conference on Information Science and Engineering (ICISE) IEEE, pp. 1129-1132.

Rangel, H.R., Puig, V., Farias, R.L. and Flores, J.J. 2017. Short-term demand forecast using a bank of neural network models trained using genetic algorithms for the optimal management of drinking water networks. Journal of Hydroinformatics 19 (1): 1-16.

Raskutti, G., Wainwright, M.J. and Yu, B. 2014. Early stopping and non-parametric regression: an optimal data-dependent stopping rule. The Journal of Machine Learning Research 15 (1): 335-366.

Reed, P.M., Hadka, D., Herman, J.D., Kasprzyk, J.R. and Kollat, J.B. 2013. Evolutionary multiobjective optimization in water resources: The past, present, and future. Advances in water resources 51: 438-456.

Roa Garcia, M.C. 2014. Equity, efficiency and sustainability in water allocation in the Andes: Trade-offs in a full world. Water alternatives (WaA) 7 (2): 298-319.

Robnik-Šikonja, M. and Kononenko, I. 2003. Theoretical and empirical analysis of ReliefF and RReliefF. Machine learning 53 (1-2): 23-69.

Romano, M. and Kapelan, Z. 2014. Adaptive water demand forecasting for near real-time management of smart water distribution systems. Environmental Modelling & Software 60: 265-276.

Rosemeyer, M.C. 2017. The Growing Importance of SDG 6 for Minimizing Conflict and Promoting Peace, Health, and Equitable Development, in: Decoster, K., Van Damme, W., Scholtalbers, R., Affun-Adegbulu, C., Arora, R. and Saluja, D. (Eds.), International Helath Policies. Institute of Tropical Medicine, Antwerp.

Sadoff, C.W., Hall, J.W., Grey, D., Aerts, J.C.J.H., Ait-Kadi, M., Brown, C., Cox, A., Dadson, S., Garrick, D., Kelman, J., McCornick, P., Ringler, C., Rosegrant, M., Whittington, D. and Wiberg, D. 2015. Securing Water, Sustaining Growth: Report of the GWP/OECD Task Force on Water Security and Sustainable Growth, UK, p. 180.

Saladini, F., Betti, G., Ferragina, E., Bouraoui, F., Cupertino, S., Canitano, G., Gigliotti, M., Autino, A., Pulselli, F. and Riccaboni, A. 2018. Linking the water-energy-food nexus and sustainable development indicators for the Mediterranean region. Ecological Indicators 91: 689-697.

Savic, D.A. and Walters, G.A. 1997. Genetic algorithms for least-cost design of water distribution networks. Journal of Water Resources Planning and Management 123 (2): 67-77.

Sedki, A. and Ouazar, D. 2012. Hybrid particle swarm optimization and differential evolution for optimal design of water distribution systems. Advanced Engineering Informatics 26 (3): 582-591.

Sette, S. and Boullart, L. 2001. Genetic programming: principles and applications. Engineering applications of artificial intelligence 14 (6): 727-736.

Shabani, S., Candelieri, A., Archetti, F. and Naser, G. 2018. Gene Expression Programming Coupled with Unsupervised Learning: A Two-Stage Learning Process in Multi-Scale, Short-Term Water Demand Forecasts. Water 10 (2): 142.

Shabani, S., Yousefi, P., Adamowski, J. and Naser, G. 2016. Intelligent soft computing models in water demand forecasting, Water Stress in Plants. InTech.

Shabani, S., Yousefi, P. and Naser, G. 2017. Support vector machines in urban water demand forecasting using phase space reconstruction. Procedia Engineering 186: 537-543.

Shahin, M.A., Jaksa, M.B. and Maier, H.R. 2008. State of the art of artificial neural networks in geotechnical engineering. Electronic Journal of Geotechnical Engineering 8: 1-26.

Shang, Y., Lu, S., Gong, J., Shang, L., Li, X., Wei, Y. and Shi, H. 2017. Hierarchical prediction of industrial water demand based on refined Laspeyres decomposition analysis. Water Science and Technology 76 (11): 2876-2887.

Shaw, W.D. 2007. Water resource economics and policy: an introduction. Edward Elgar Publishing.

Sherrod, P.H. 2003. DTREG predictive modeling software. Software available at http://www. dtreg. com.

Siksnelyte, I., Zavadskas, E., Streimikiene, D. and Sharma, D. 2018. An overview of multi-criteria decision-making methods in dealing with sustainable energy development issues. Energies 11 (10): 2754.

Smith, S.L. 2017. Developing an ecumenical framework for water justice, in: Groenfeldt, D. and Ziegler, R. (Eds.), Global Water Ethics: Towards a Global Ethics Charter, London, pp. 243-253.

Soltani, F., Kerachian, R. and Shirangi, E. 2010. Developing operating rules for reservoirs considering the water quality issues: Application of ANFIS-based surrogate models. Expert Systems with Applications 37 (9): 6639-6645.

Stats-SA 2018. Statistical Release P0302: Mid-year population estimates, 2018, in: Africa, S.S. (Ed.), Pretoria.

Steyerberg, E.W. 2008. Clinical prediction models: a practical approach to development, validation, and updating. Springer Science & Business Media.

Storn, R. and Price, K. 1997. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization 11 (4): 341-359.

Su, Y., Gao, W., Guan, D. and Su, W. 2018. Dynamic assessment and forecast of urban water ecological footprint based on exponential smoothing analysis. Journal of cleaner production 195: 354-364.

Suribabu, C. 2010. Differential evolution algorithm for optimal design of water distribution networks. Journal of Hydroinformatics 12 (1): 66-82.

Tabesh, M. and Dini, M. 2009. Fuzzy and neuro-fuzzy models for short-term water demand forecasting in Tehran. Iranian Journal of Science and Technology Transcation B- Engineering 33 (B1): 61-77.

Tabesh, M. and Hoomehr, S. 2009. Consumption management in water distribution systems by optimizing pressure reducing valves' settings using genetic algorithm. Desalination and Water Treatment 2 (1-3): 96-102.

Tanyimboh, T.T. and Seyoum, A.G. 2016. Multiobjective evolutionary optimization of water distribution systems: Exploiting diversity with infeasible solutions. Journal of Environmental Management 183: 133-141.

Tomić, A.Š., Antanasijević, D., Ristić, M., Perić-Grujić, A. and Pocajt, V. 2018. Application of experimental design for the optimization of artificial neural network-based water quality model: a case study of dissolved oxygen prediction. Environmental Science and Pollution Research: 1-11.

Toth, E., Bragalli, C. and Neri, M. 2018. Assessing the significance of tourism and climate on residential water demand: Panel-data analysis and non-linear modelling of monthly water consumptions. Environmental Modelling & Software 103: 52-61.

Uca, Toriman, E., Jaafar, O., Maru, R., Arfan, A. and Ahmar, A.S. 2018. Daily Suspended Sediment Discharge Prediction Using Multiple Linear Regression and Artificial Neural Network, Journal of Physics: Conference Series. IOP Publishing, p. 012030.

UN 2010. The human right to water and sanitation, 64/292, 64th Session ed, United Nations, pp. 1-3.

UNDESA 2013. World Population Prospects: The 2012 Revision. United Nations for Department of Economic and Social Affairs, Population Division, New York.

UNDP 2018. Sustainable Development Goals. United Nations Development Programme. Available: http://www.undp.org/content/undp/en/home/sustainable-development-goals.html [Accessed: 01 March 2018].

UNEP 2012. A Glass Half Empty: Regions at Risk Due to Groundwater Depletion, in: Pengra, B. (Ed.), Ecosystem Management and Resource Efficiency. United Nations Environment Programme, Sioux Falls, South Dakota.

UNESCO 2015. The United Nations World Water Development Report 2015: Water for a Sustainable World. United Nations World Water Assessment Programme - WWAP. UNESCO, Paris.

UNESCO 2016. The United Nations World Water Development Report 2016: Water and Jobs – Facts and Figures. United Nations World Water Assessment Programme - WWAP, Paris.

USCB 2012. International Programs. World Population, United States Census Bureau.

Van Zyl, J.E., Savic, D.A. and Walters, G.A. 2004. Operational optimization of water distribution systems using a hybrid genetic algorithm. Journal of water resources planning and management 130 (2): 160-170.

Vapnik, V. 2013. The nature of statistical learning theory. Springer science & business media.

Vapnik, V.N. 1999. An overview of statistical learning theory. IEEE transactions on neural networks 10 (5): 988-999.

Varahrami, V. 2010. Application of genetic algorithm to neural network forecasting of short-term water demand, International Conference on Applied Economics—ICOAE. Citeseer, pp. 783-787.

Vasan, A. and Simonovic, S.P. 2010. Optimization of water distribution network design using differential evolution. Journal of Water Resources Planning and Management 136 (2): 279-287.

Vásquez, W.F. 2015. Nonpayment of water bills in G uatemala: Dissatisfaction or inability to pay? Water Resources Research 51 (11): 8806-8816.

Vijayalaksmi, D. and Babu, K.J. 2015. Water supply system demand forecasting using adaptive neuro-fuzzy inference system. Aquatic Procedia 4: 950-956.

Vorster, J., Geustyn, L., Loubser, E., Tanner, A. and Wall, K. 1995. A strategy and master plan for water supply, storage and distribution in the East Rand region. Journal of the South African Institution of Civil Engineers 37 (2): 1-5.

Walker, D., Creaco, E., Vamvakeridou-Lyroudia, L., Farmani, R., Kapelan, Z. and Savić, D. 2015. Forecasting domestic water consumption from smart meter readings using statistical methods and artificial neural networks. Procedia Engineering 119: 1419-1428.

Wang, W., Xu, D., Qiu, L. and Ma, J. 2009. Genetic programming for modelling long-term hydrological time series, Fifth International Conference on Natural Computation, 2009. ICNC'09. . IEEE, Tianjin, pp. 265-269.

White, S., Robinson, J., Cordell, D., Jha, M. and Milne, G. 2003. Urban water demand forecasting and demand management: Research needs review and recommendations. Water Services Association of Australia, Sydney.

WHO 2016. Water Safety in Distribution. World Health Organization, Geneva.

Witten, I.H., Frank, E., Hall, M.A. and Pal, C.J. 2016. Data Mining: Practical machine learning tools and techniques, Fourth ed. Morgan Kaufmann Massachusetts.

Wu, Z.Y. and Yan, X. 2010. Applying genetic programming approaches to short-term water demand forecast for district water system, Water Distribution Systems Analysis 2010, pp. 1498-1506.

WWAP 2018. The United Nations World Water Development Report 2018: Nature-based Solutions for Water. UNESCO, Paris, p. 154.

Yang, X.-S. 2009. Harmony search as a metaheuristic algorithm, Music-inspired harmony search algorithm. Springer, pp. 1-14.

Yao, X. 1993. A review of evolutionary artificial neural networks. International Journal of Intelligent Systems 8 (4): 539-567.

Yousefi, P., Shabani, S., Mohammadi, H. and Naser, G. 2017. Gene Expression Programing in Long Term Water Demand Forecasts Using Wavelet Decomposition. Procedia Engineering 186: 544-550.

Zeng, Y., Cai, Y., Jia, P. and Jee, H. 2012. Development of a web-based decision support system for supporting integrated water resources management in Daegu city, South Korea. Expert Systems with Applications 39 (11): 10091-10102.

Zhai, C., Zhang, H. and Zhang, X. 2009. Application of system dynamics in the forecasting water resources demand in Tianjin polytechnic university, International Conference on Artificial Intelligence and Computational Intelligence (AICI '09). IEEE, Shanghai, pp. 273-276.

Zheng, F., Zecchin, A.C. and Simpson, A.R. 2012. Self-adaptive differential evolution algorithm applied to water distribution system optimization. Journal of Computing in Civil Engineering 27 (2): 148-158.

APPENDIX 1

EXTRACT OF CODES WRITTEN IN DEVELOPING ANN-DE MODULES

ANN_Run

```
Public Function InputRangesOK(ByVal NumSamples, ByRef Array2DToCheck) As
Boolean
    'Used to check input ranges to avoid extrapolation
    Dim SampleMin As Double
    Dim SampleMax As Double
    Dim echNode As Integer

    InputRangesOK = True    'Assume all the samples are ok
    For echNode = 1 To numInputNodes
        SampleMin = getMin(echNode, NumSamples, Array2DToCheck)
        SampleMax = getMax(echNode, NumSamples, Array2DToCheck)
        If (SampleMin < MinInputVal(echNode)) Then InputRangesOK = False 'sample
exceeds limit
        If (SampleMax > MaxInputVal(echNode)) Then InputRangesOK = False
    Next echNode
End Function

Public Sub RunANNonSingleSample(ByVal SamplePoint As Integer, ByRef
InputSample2DArray)
    'Runs the ANN with a single input sample point and specified input array
    Dim echNode As Integer
    'Load the sample into the 1D input array
    For echNode = 1 To numInputNodes
        InputVal(echNode) = InputSample2DArray(SamplePoint, echNode)
    Next echNode
    ''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
    'Now run the ANN from the input nodes to the output
    ComputeInputNodesActivations InputVal
    ComputeHiddenNodesActivations
    ComputeOutputNodesActivations
    ComputeOutputNodesValues
End Sub

Public Sub RunANNonEntireSampleSet(ByVal NumOfSamples As Integer _
                                , ByRef InputArray2D _
                                , ByRef OutputArray2D)
'Runs ANN on entire sample set in a 2D array and stores the output in a 2D array
    Dim echSample As Integer
    Dim eachNode As Integer

    For echSample = 1 To NumOfSamples
        RunANNonSingleSample echSample, InputArray2D
        'The ANN output is stored in the outputvals array
        'now copy to the proper output 2D array
        For eachNode = 1 To numOutputNodes
```

```
            OutputArray2D(echSample, eachNode) = OutputVal(eachNode)
        Next eachNode
    Next echSample
End Sub
```

ANN_Setup

```
Option Explicit

Public numInputNodes As Integer '= Number of inputs
Public NumHiddenNodes As Integer
Public numOutputNodes As Integer    '= Number of outputs
Public NumWtsAndBiases As Integer
Public NumRunSamples As Integer

Public WtPlusBiasArray(1 To 1600) As Double
Public InputNodesActivations(1 To 15) As Double
Public HiddenNodesActivations(1 To 50) As Double
Public OutputNodesActivations(1 To 15) As Double

Public InputVal(1 To 15) As Double
Public MinInputVal(1 To 15) As Double
Public MaxInputVal(1 To 15) As Double

Public OutputVal(1 To 15) As Double
Public MinOutputVal(1 To 15) As Double
Public MaxOutputVal(1 To 15) As Double


Public HiddenNodesStartIndexs(1 To 50) As Integer
Public OutputNodesStartIndexs(1 To 15) As Integer

Dim ConnectionsPerHiddenNode As Integer
Dim ConnectionsPerOutputNode As Integer

Public RunSamples2DArray(1 To 1000, 1 To 15) As Double '1000 points by 15
nodes
Public RunSamplesOutput2DArray(1 To 1000, 1 To 15) As Double '1000 points
by 15 nodes
Dim RunnerSheet As Worksheet
Dim TrainerSheet As Worksheet

Public Sub Setup_ANNRun()
    Dim echSamplePoint As Integer
    Dim echNode As Integer
    Set TrainerSheet = ThisWorkbook.Worksheets("ANNTrainer")
    Set RunnerSheet = ThisWorkbook.Worksheets("RunSheet")
    'get from worksheet
    numInputNodes = TrainerSheet.Range("E9").Value
    NumHiddenNodes = TrainerSheet.Range("E10").Value
```

```
    numOutputNodes = TrainerSheet.Range("E11").Value
    NumRunSamples = RunnerSheet.Range("W2").Value
    NumWtsAndBiases = ComputeNumWtAndBiases
    ComputeHiddenNodesStartIndexes
    ComputeOutputNodesStartIndexes
    LoadWtsAndBiases
    '''''''''''''''''''''''''''''
    'Load the Run samples
    For echSamplePoint = 1 To NumRunSamples
       For echNode = 1 To numInputNodes     'Load the input space
          '1st sample on row 10, column 3
          RunSamples2DArray(echSamplePoint, echNode) = _
                RunnerSheet.Cells(echSamplePoint + 9, echNode + 2).Value
       Next echNode
    Next echSamplePoint
    '''''''''''''''''''''''''''''''''''
    'Setup the run ranges
    getInputRanges '
    getoutputRanges
End Sub

Public Sub LoadWtsAndBiases()
    Dim mRow As Integer
    Set TrainerSheet = ThisWorkbook.Worksheets("ANNTrainer")

    For mRow = 1 To NumWtsAndBiases
       WtPlusBiasArray(mRow) = _
                TrainerSheet.Range("O" & Trim(Str(mRow + 1))).Value
    Next mRow
End Sub

Public Function ComputeNumWtAndBiases() As Integer
    ComputeNumWtAndBiases = ((numInputNodes + 1) * NumHiddenNodes) + _
                    ((NumHiddenNodes + 1) * numOutputNodes)
End Function

Public Sub ComputeHiddenNodesStartIndexes()
    Dim EachHiddenNode As Integer

    ConnectionsPerHiddenNode = numInputNodes + 1 'Add the bias
    For EachHiddenNode = 1 To NumHiddenNodes
       HiddenNodesStartIndexs(EachHiddenNode) = _
                (ConnectionsPerHiddenNode * (EachHiddenNode - 1)) + 1
    Next EachHiddenNode
End Sub

Public Sub ComputeOutputNodesStartIndexes()
    Dim EachOutputNode As Integer
    Dim OutputNodesStartOffset As Integer
```

```
   OutputNodesStartOffset = (ConnectionsPerHiddenNode * NumHiddenNodes)
+ 1
   ConnectionsPerOutputNode = NumHiddenNodes + 1 'Add the bias
   For EachOutputNode = 1 To numOutputNodes
      OutputNodesStartIndexs(EachOutputNode) = _
               (ConnectionsPerOutputNode * (EachOutputNode - 1)) + _
               OutputNodesStartOffset
   Next EachOutputNode
End Sub

Public Sub ComputeInputNodesActivations(ByRef InputValsArray)
   'Input Nodes activations are the inputs normalized btwn 0.1 - 0.9
   Dim EachInputNode As Integer

   For EachInputNode = 1 To numInputNodes
      InputNodesActivations(EachInputNode) = _
         (0.8 * (InputValsArray(EachInputNode) - MinInputVal(EachInputNode)) /
_
            (MaxInputVal(EachInputNode) - MinInputVal(EachInputNode))) + 0.1
   Next EachInputNode
End Sub
Public Sub ComputeOutputNodesValues()
   'Compute output values from Nodes activations normalized btwn 0.1 - 0.9
   Dim EachOutputNode As Integer

   For EachOutputNode = 1 To numOutputNodes
      OutputVal(EachOutputNode) = _
         (((OutputNodesActivations(EachOutputNode) - 0.1) * _
            (MaxOutputVal(EachOutputNode) - MinOutputVal(EachOutputNode)))
/ 0.8) + _
            MinOutputVal(EachOutputNode)
   Next EachOutputNode
End Sub

Public Sub ComputeHiddenNodesActivations()
   Dim EachInputNode As Integer
   Dim EachHiddenNode As Integer
   Dim NodeOffset As Integer
   Dim wtPos As Integer
   Dim yj As Double
   Dim wjk As Double
   Dim Sk As Double    'Propagation rule
   Dim SumWtdInputs As Double

   For EachHiddenNode = 1 To NumHiddenNodes
      NodeOffset = HiddenNodesStartIndexs(EachHiddenNode)
      wtPos = NodeOffset
      SumWtdInputs = 0

      For EachInputNode = 1 To numInputNodes
```

```
      yj = InputNodesActivations(EachInputNode)
      wjk = WtPlusBiasArray(wtPos)
      SumWtdInputs = SumWtdInputs + (wjk * yj)
      wtPos = wtPos + 1 'Next weight or the bias
    Next EachInputNode

    Sk = SumWtdInputs + WtPlusBiasArray(wtPos)  'Add the bias
    HiddenNodesActivations(EachHiddenNode) = Sigmoid(Sk)
  Next EachHiddenNode
End Sub

Public Sub ComputeOutputNodesActivations()
  Dim EachHiddenNode As Integer
  Dim EachOutputNode As Integer
  Dim NodeOffset As Integer
  Dim wtPos As Integer
  Dim yj As Double
  Dim wjk As Double
  Dim Sk As Double    'Propagation rule
  Dim SumWtdInputs As Double

  For EachOutputNode = 1 To numOutputNodes
    NodeOffset = OutputNodesStartIndexs(EachOutputNode)
    wtPos = NodeOffset
    SumWtdInputs = 0

    For EachHiddenNode = 1 To NumHiddenNodes
      yj = HiddenNodesActivations(EachHiddenNode)
      wjk = WtPlusBiasArray(wtPos)
      SumWtdInputs = SumWtdInputs + (wjk * yj)
      wtPos = wtPos + 1 'Next weight or the bias
    Next EachHiddenNode

    Sk = SumWtdInputs + WtPlusBiasArray(wtPos)  'Add the bias
    OutputNodesActivations(EachOutputNode) = PureLin(Sk)
  Next EachOutputNode
End Sub

Public Function Sigmoid(ByVal Sk As Double) As Double
  If Sk < -709 Then   'To avoid overflow error
    Sigmoid = 0
    Exit Function
  End If
  Sigmoid = 1 / (1 + Exp(-Sk))
End Function

Private Function PureLin(ByVal Sk As Double) As Double
  PureLin = Sk
End Function
```

```
Public Sub getInputRanges()
   'Gets the ranges of input and output from the worksheet
   Dim echNode As Integer
   Dim RangeMin As Double
   Dim RangeMax As Double
   Dim mRow, mCol As Integer

   Set TrainerSheet = ThisWorkbook.Worksheets("ANNTrainer")
   Set RunnerSheet = ThisWorkbook.Worksheets("RunSheet")
   'Scan the ranges
   For echNode = 1 To numInputNodes    'get input ranges
      RangeMin = RunnerSheet.Cells(6, echNode + 2).Value
      RangeMax = RunnerSheet.Cells(7, echNode + 2).Value
      MinInputVal(echNode) = RangeMin
      MaxInputVal(echNode) = RangeMax
   Next echNode
End Sub

Public Sub getoutputRanges()
   'Gets the ranges of input and output from the worksheet
   Dim echNode As Integer
   Dim RangeMin As Double
   Dim RangeMax As Double
   Dim mRow, mCol As Integer

   Set RunnerSheet = ThisWorkbook.Worksheets("RunSheet")
   'Scan the ranges
   For echNode = 1 To numOutputNodes    'get output ranges
      RangeMin = RunnerSheet.Cells(6, (echNode + 2) + numInputNodes).Value
      RangeMax = RunnerSheet.Cells(7, (echNode + 2) + numInputNodes).Value
      MinOutputVal(echNode) = RangeMin
      MaxOutputVal(echNode) = RangeMax
   Next echNode
End Sub
Public Sub OutputRunResultsToGUI()
   Dim echSamplePoint As Integer
   Dim echNode As Integer
   'Output the run samples results
   For echSamplePoint = 1 To NumRunSamples
      For echNode = 1 To numOutputNodes
         '1st result on row 10, column after last input node
         RunnerSheet.Cells(echSamplePoint + 9, (echNode + 2) +
numInputNodes).Value = _
                      RunSamplesOutput2DArray(echSamplePoint, echNode)
      Next echNode
   Next echSamplePoint
End Sub
```

ANN_Trainer

```
Option Explicit
Public NumLearnSamples As Integer
Public NumValidationSamples As Integer
Public HiddenMin As Integer 'For varying number of hidden nodes
Public HiddenMax As Integer

   Public bestETest As Double
   Public bestELearning As Double
   Public bestNumHidNodes As Integer
   Public bestWtPlusBiasArray(1 To 3500) As Double 'for 15 inpt, 50 hidden, 15
out

Public LearnSamplesInput2DArray(1 To 1000, 1 To 15) As Double '1000 samples
by 15 nodes
Public LearnSamplesExpOutput2DArray(1 To 1000, 1 To 15) As Double '1000
samples by 15 nodes
Public LearnSamplesActualOutput2DArray(1 To 1000, 1 To 15) As Double '1000
points by 15 nodes
Public ValidationSamplesInput2DArray(1 To 1000, 1 To 15) As Double '1000
points by 15 nodes
Public ValidationSamplesExpOutput2DArray(1 To 1000, 1 To 15) As Double
'1000 points by 15 nodes
Public ValidationSamplesActualOutput2DArray(1 To 1000, 1 To 15) As Double
'1000 points by 15 nodes
Dim TrainerSheet As Worksheet
Dim RunnerSheet As Worksheet
Dim LearningSamplesSheet As Worksheet
Dim ValidationSamplesSheet As Worksheet

Public Sub LoadTrainerSetup()
   Dim echSamplePoint As Integer
   Dim echNode As Integer
   Set TrainerSheet = ThisWorkbook.Worksheets("ANNTrainer")
   'Set RunnerSheet = ThisWorkbook.Worksheets("RunSheet")
   'Set TrainedSheet = ThisWorkbook.Worksheets("TrainedANN")
   Set LearningSamplesSheet = ThisWorkbook.Worksheets("LearningSamples")
   Set ValidationSamplesSheet =
ThisWorkbook.Worksheets("ValidationSamples")

   numInputNodes = TrainerSheet.Range("E3").Value
   TrainerSheet.Range("E9").Value = numInputNodes
   numOutputNodes = TrainerSheet.Range("E4").Value
   TrainerSheet.Range("E11").Value = numOutputNodes
   HiddenMin = TrainerSheet.Range("G5").Value
   HiddenMax = TrainerSheet.Range("I5").Value

   NumLearnSamples = TrainerSheet.Range("I3").Value
   NumValidationSamples = TrainerSheet.Range("I4").Value

   'Load the learning samples
```

```
    For echSamplePoint = 1 To NumLearnSamples
        'Load the input nodes
        For echNode = 1 To numInputNodes
            '1st sample on row 10, column 3
            LearnSamplesInput2DArray(echSamplePoint, echNode) = _
                    LearningSamplesSheet.Cells(echSamplePoint + 9, echNode +
2).Value
        Next echNode

        'Load the expected output nodes
        For echNode = 1 To numOutputNodes
         '1st sample after the last input node
            LearnSamplesExpOutput2DArray(echSamplePoint, echNode) = _
                    LearningSamplesSheet.Cells(echSamplePoint + 9,
((numInputNodes + 2) + echNode)).Value
        Next echNode
    Next echSamplePoint

    'Load the validation samples
    For echSamplePoint = 1 To NumValidationSamples
        'Load the input nodes
       For echNode = 1 To numInputNodes
            '1st sample on row 10, column 3
            ValidationSamplesInput2DArray(echSamplePoint, echNode) = _
                    ValidationSamplesSheet.Cells(echSamplePoint + 9, echNode +
2).Value
       Next echNode

       'Load the expected output nodes
        For echNode = 1 To numOutputNodes
         '1st sample after the last input node
            ValidationSamplesExpOutput2DArray(echSamplePoint, echNode) = _
                    ValidationSamplesSheet.Cells(echSamplePoint + 9,
((numInputNodes + 2) + echNode)).Value
        Next echNode
    Next echSamplePoint
End Sub

Public Function getMin(ByVal NodeNumbr As Integer, ByVal NumSamples As
Integer, _
                        ByRef Array2D) As Double
'Used in getting the minimum sample for each node
    Dim echSamplPt As Integer
    Dim minMan As Double
    minMan = Array2D(1, NodeNumbr) 'Pick on the first sample

    For echSamplPt = 1 To NumSamples    'run through the samples
        If (Array2D(echSamplPt, NodeNumbr) < minMan) Then
            minMan = Array2D(echSamplPt, NodeNumbr)
        End If
```

```
    Next echSamplPt
    getMin = minMan
End Function

Public Function getMax(ByVal NodeNumbr As Integer, ByVal NumSamples As
Integer, _
                  ByRef Array2D) As Double
'Used in getting the maximum sample for each node
    Dim echSamplPt As Integer
    Dim maxMan As Double
    maxMan = Array2D(1, NodeNumbr) 'Pick on the first sample

    For echSamplPt = 1 To NumSamples    'run through the samples
        If (Array2D(echSamplPt, NodeNumbr) > maxMan) Then
            maxMan = Array2D(echSamplPt, NodeNumbr)
        End If
    Next echSamplPt
    getMax = maxMan
End Function

Public Sub SetInputRanges()
    'Gets the ranges of input and output from the learning data and set on the Run
sheet
    Dim echNode As Integer
    Dim RangeMin As Double
    Dim RangeMax As Double

    Set RunnerSheet = ThisWorkbook.Worksheets("RunSheet")
    'Clear the Data ranges on the Run Sheet.
    RunnerSheet.Range("C6:AF7").ClearContents


    For echNode = 1 To numInputNodes    'Set input ranges
        RangeMin = getMin(echNode, NumLearnSamples,
LearnSamplesInput2DArray)
        RangeMax = getMax(echNode, NumLearnSamples,
LearnSamplesInput2DArray)
        MinInputVal(echNode) = RangeMin
        MaxInputVal(echNode) = RangeMax
        RunnerSheet.Cells(6, echNode + 2).Value = RangeMin
        RunnerSheet.Cells(7, echNode + 2).Value = RangeMax
    Next echNode
End Sub

Public Sub SetoutputRanges()
    'Gets the ranges of input and output from the learning data and set on the
trained sheet
    Dim echNode As Integer
    Dim RangeMin As Double
    Dim RangeMax As Double
```

```
   Set RunnerSheet = ThisWorkbook.Worksheets("RunSheet")
   For echNode = 1 To numOutputNodes    'Set output ranges
      RangeMin = getMin(echNode, NumLearnSamples,
LearnSamplesExpOutput2DArray)
      RangeMax = getMax(echNode, NumLearnSamples,
LearnSamplesExpOutput2DArray)
      MinOutputVal(echNode) = RangeMin
      MaxOutputVal(echNode) = RangeMax
      RunnerSheet.Cells(6, (echNode + 2) + numInputNodes).Value = RangeMin
      RunnerSheet.Cells(7, (echNode + 2) + numInputNodes).Value = RangeMax
   Next echNode
End Sub

Public Function ComputeEp(ByVal SamplePoint As Integer, _
                      ByRef ExpectedOutputsArray2D, _
                      ByRef ActualOutputsArray2D) As Double
'Function to compute Ep = Error per sample
   Dim dp As Double 'expected node output
   Dim yp As Double 'actual node output
   Dim echOutput As Integer
   Dim sumError As Double

   sumError = 0
   For echOutput = 1 To numOutputNodes
      dp = ExpectedOutputsArray2D(SamplePoint, echOutput)
      yp = ActualOutputsArray2D(SamplePoint, echOutput)
      sumError = sumError + ((dp - yp) ^ 2)
   Next echOutput
   ComputeEp = sumError / 2
End Function

Public Function ComputeAveEp(ByVal NumSamples As Integer, _
                      ByRef ExpOutArray2D, _
                      ByRef ActOutArray2D) As Double
'Function to compute average error per sample
   Dim echSample As Integer
   Dim sumEp As Double
   Dim Ep As Double

   sumEp = 0
   For echSample = 1 To NumSamples
      Ep = ComputeEp(echSample, ExpOutArray2D, ActOutArray2D)
      sumEp = sumEp + Ep
   Next echSample
   ComputeAveEp = sumEp / NumSamples
End Function

 Public Sub OutputTrainResultToGUI()
   Dim mRow As Integer
```

```
    Dim epslon As Double
    Set TrainerSheet = ThisWorkbook.Worksheets("ANNTrainer")
    Set RunnerSheet = ThisWorkbook.Worksheets("RunSheet")
    'ClearOutput
    TrainerSheet.Range("O2:O2000").ClearContents 'wts and biases space
    'Output the parameters
    For mRow = 1 To NumberOfParameters
       TrainerSheet.Range("O" & Trim(Str(mRow + 1))).Value = _
       bestWtPlusBiasArray(mRow)
    Next mRow
    'Also Ouput the Fittest Values
    TrainerSheet.Range("C12").Value = bestELearning
    TrainerSheet.Range("C13").Value = bestETest
    epslon = Abs(bestELearning - bestETest)
    TrainerSheet.Range("C14").Value = epslon
    TrainerSheet.Range("E10").Value = bestNumHidNodes

    'Also Output for optimization monitoring
    ThisWorkbook.Worksheets("Optimizer").Range("K15").Value = bestELearning
    ThisWorkbook.Worksheets("Optimizer").Range("K16").Value = bestETest
    ThisWorkbook.Worksheets("Optimizer").Range("K18").Value =
bestNumHidNodes

    TrainerSheet.Range("I8").Value = NumberOfVectors
    TrainerSheet.Range("I9").Value = NumberOfParameters
    TrainerSheet.Range("I10").Value = MaxGen
    TrainerSheet.Range("H11").Value = bestCr
    TrainerSheet.Range("J11").Value = bestF
 End Sub
```

DE

```
Option Explicit
'Variables declaration
Public MaxGen As Integer          'Maximum number of evolutionary generations
Public NumberOfVectors As Integer   'Number of DE population members

Public CurPopParameters(1 To 2000, 1 To 2000) As Double    '20000 current
vectors, 2000 parameters
Public TrialPopParameters(1 To 2000, 1 To 2000) As Double  '2000 trial
vectors, 2000 parameters

Public CurPopObjective(1 To 2000) As Double    'Objective functions values
Public TrialPopObjective(1 To 2000) As Double

Public Cr As Single     'DE's crossover rate
Public CrMin As Single
Public CrStep As Single
Public CrMax As Single
```

```
Public F As Single     'DE's mutation scale parameter
Public FMin As Single
Public FStep As Single
Public FMax As Single

Public Sub GenerateRandomizedPopulation()
'Generate a randomized initial poupulation
   Dim EachVector As Integer
   For EachVector = 1 To NumberOfVectors
      GenerateRandomizedVector EachVector
   Next EachVector
End Sub

Public Sub GenerateRandomizedVector(VertorIndx As Integer)
   Dim EachParameter As Integer
   'Generates a randomized current vector at the given index
   'Generate random parameters for the vector
      For EachParameter = 1 To NumberOfParameters
         CurPopParameters(VertorIndx, EachParameter) = _
                     RandomizeBetween(LowerParameterBound(EachParameter),
_
                                 UpperParameterBound(EachParameter))
      Next EachParameter
      'Compute objective
      CurPopObjective(VertorIndx) = _
               ComputeObjective(VertorIndx, CurPopParameters)
End Sub

Public Sub Optimize()
   Dim gKounter As Integer 'To count generations
   Dim w As Integer
   LoadSetup

   'The for-next system of VBA has a bug, it fails when variables are in
   'the fractional area, especially the step. Therefore we employ a series
   'of do...loops here to vary the DE parameters Cr and F.
   'We also use the single data type instead of double for Cr and F.
   If FirstValidationRun = True Then
      getInitialStartupBestVector 'Avoid overwriting best solution so far
   End If
   Cr = CrMin
   F = FMin
   ''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
   Cr = CrMin - CrStep
   F = FMin - FStep
   Do
      Cr = Cr + CrStep
      If Cr > CrMax Then Cr = CrMax
      Do
         F = F + FStep
```

```
        If F > FMax Then F = FMax
        'Output the parameter progress to the GUI'''''''''''''
        ThisWorkbook.Worksheets("Optimizer").Range("C15").Value = Cr
        ThisWorkbook.Worksheets("Optimizer").Range("C16").Value = F
        ''''''''''''''''''''''''''''''''''''''''''''
        'Main optimizer section'''''''''''''''''''''''''''
        GenerateRandomizedPopulation        'Initialize population for every
parameter change

        ''''''''''''''''''''''''''''''

        gKounter = 0    'reset the generation counter
        Do
           GenerateTrialPopulation 'generate a trial population
           SelectNextGeneration    'SELECTION
           SaveBestOfGeneration
     UpdateValidation (gKounter)
     'Dim eValidation As Double 'Best Validation error ever found in the whole
search
           gKounter = gKounter + 1
           'Display Progress
           ThisWorkbook.Worksheets("Optimizer").Range("D15").Value = _
           Str(gKounter) & " of " & MaxGen
           DoEvents
        Loop Until gKounter >= MaxGen
        '''''''''''''''''''''''''''''''''''''''''''''

        'Check for a better result in the current population
        'For w = 1 To NumberOfVectors
        '    If (CurPopObjective(w) <= OutputerObjective) Then
        '        CopySolutionForOutput w, CurPopParameters, CurPopObjective
        '    End If
        'Next w
        'OutputVectorToGUI
     Loop Until F = FMax
     If Cr < CrMax Then F = FMin - FStep
  Loop Until Cr = CrMax
End Sub

Public Sub SaveBestOfGeneration()
  Dim w As Integer
  'Dim crspdngEtest As Double
  Dim BetterObjectiveValue As Double
  Dim bestIndx As Integer 'Index of best vector in current population
  'During call to select next generation, all better vectors have been copied to
the
  'Current population.
  'Now check for a better result in the current population and save to the outputer
  bestIndx = 1    'Pick on the first solution
  BetterObjectiveValue = OutputerObjective
```

```
    For w = 1 To NumberOfVectors
        If (CurPopObjective(w) <= BetterObjectiveValue) Then   'minimization
            BetterObjectiveValue = CurPopObjective(w)
            bestIndx = w
        End If
    Next w
    If (BetterObjectiveValue <= OutputerObjective) Then 'Improvement
        CopySolutionForOutput bestIndx, CurPopParameters, CurPopObjective
        OutputVectorToGUI  'Output better solution
    End If

        'Output the corresponding eTest on the sheet
        'crspdngEtest        =        RunANNonEntireSampleSet(NumTestSamples,
TestSamples2DArray, _
                                        TestSamplesActualOutput2DArray _
                                        , ValDayIndx)
        'ThisWorkbook.Worksheets("Optimizer").Range("Q6").Value                    =
crspdngEtest
End Sub

Public Sub GenerateTrialPopulation()
    Dim i As Integer     'Vector ID
    Dim j As Integer     'Parameter ID
    Dim jrand As Integer 'Random crossover point
    Dim r0 As Integer    'Base vector index
    Dim r1 As Integer    'Difference vector index 1
    Dim r2 As Integer    'Difference vector index 2
    Dim EachConstraint As Integer

    'Generate a trial vector population (Parameters)
    For i = 1 To NumberOfVectors   'For each vector
        'Make sure indices are unique
        Do: r0 = RandomInteger(1, NumberOfVectors): Loop While (r0 = i)
        Do: r1 = RandomInteger(1, NumberOfVectors): Loop While ((r1 = r0) Or (r1
= i))
        Do: r2 = RandomInteger(1, NumberOfVectors): Loop While ((r2 = r1) Or (r2
= r0) Or (r2 = i))
        jrand = RandomInteger(1, NumberOfParameters)
        ''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
        For j = 1 To NumberOfParameters 'generate a trial vector
            Randomize
            If ((Rnd <= Cr) Or (j = jrand)) Then
                TrialPopParameters(i, j) = _
                    CurPopParameters(r0, j) + _
                        F * (CurPopParameters(r1, j) - CurPopParameters(r2, j))
                'Check Trial vector for out of bound constraint
                EnforceTrialVectorParameterBounds i, r0, j
            Else
                TrialPopParameters(i, j) = CurPopParameters(i, j)
            End If
```

```
      Next j
      'Compute objective
      TrialPopObjective(i) = ComputeObjective(i, TrialPopParameters)
   Next i
End Sub

Private Sub SelectNextGeneration()
   Dim i As Integer    'Vector ID
   For i = 1 To NumberOfVectors   'For each vector index
      If (TrialPopObjective(i) <= CurPopObjective(i)) Then
         'Copy the trial vector to the current vector for the next generation
         CopyTrialVectorToCurrent i
         'Check for better result in the population
         'CopySolutionForOutput i, CurPopParameters, CurPopObjective
         'OutputVectorToGUI
      End If
   Next i
End Sub

Private Sub EnforceTrialVectorParameterBounds(ByVal TrialVectorIndex As
Integer, _
                                    ByVal TrialBaseVectorIndex As Integer, _
                                       ByVal ParameterIndex As Integer)
   'Enforces parameter bounds on parameters requiring compulsory boundary
constarint
   'This procedure employs the bounce-back stategy, (Storn and Price, 2004)
   If EnforceParameterBound(ParameterIndex) = True Then
      If (TrialPopParameters(TrialVectorIndex, ParameterIndex) <
LowerParameterBound(ParameterIndex)) Then
      'Lower bound exceeded
         TrialPopParameters(TrialVectorIndex, ParameterIndex) = _
            RandomizeBetween(LowerParameterBound(ParameterIndex), _
               CurPopParameters(TrialBaseVectorIndex, ParameterIndex))
      End If
      If (TrialPopParameters(TrialVectorIndex, ParameterIndex) >
UpperParameterBound(ParameterIndex)) Then
      'Upper bound exceeded
         TrialPopParameters(TrialVectorIndex, ParameterIndex) = _
            RandomizeBetween(CurPopParameters(TrialBaseVectorIndex,
ParameterIndex), _
               UpperParameterBound(ParameterIndex))
      End If
   End If
End Sub

Private Sub CopyTrialVectorToCurrent(ByVal VectrIndx As Integer)
   'Copies the trial vector to the target at the given index
   Dim j As Integer
   For j = 1 To NumberOfParameters 'Each parameter
      CurPopParameters(VectrIndx, j) = TrialPopParameters(VectrIndx, j)
```

```
    Next j
    'Copy the objective
        CurPopObjective(VectrIndx) = TrialPopObjective(VectrIndx)
End Sub
```

DE_Outputer

```
Dim OutSheet As Worksheet
Public OutputerObjective As Double
Public OutputerParametersArray(1 To 2000) As Double
Public bestCr As Single
Public bestF As Single

Public Sub CopySolutionForOutput(ByVal SolnIndx As Integer, _
                                ByRef PopParameterArray, _
                                ByRef PopObjectiveArray)
    'Copies the specified solution from the given array to the output array
    Dim j As Integer

    'Copy the parameters
    For j = 1 To NumberOfParameters
        OutputerParametersArray(j) = PopParameterArray(SolnIndx, j)
    Next j
    'Copy the objectives
    OutputerObjective = PopObjectiveArray(SolnIndx)
    ''''''''''''''''''''''''''''''''''''''''''
    bestCr = Cr 'Parameters at the instance of copy
    bestF = F
End Sub

Public Sub OutputVectorToGUI()
    Dim mRow As Integer

    'ClearOutput
    Set OutSheet = ThisWorkbook.Worksheets("Optimizer")
    'Output the parameters
    For mRow = 1 To NumberOfParameters
        OutSheet.Range("N" & Trim(Str(mRow + 5))).Value = _
        OutputerParametersArray(mRow)
    Next mRow
    'Also Ouput the Fittest Values
    OutSheet.Range("Q8").Value = bestCr
    OutSheet.Range("Q9").Value = bestF
    'Best objective function
    OutSheet.Range("P5").Value = OutputerObjective
 End Sub
```

DE_Setup

```
Option Explicit
Public NumberOfParameters As Integer
Public LowerParameterBound(1 To 2000) As Double
Public UpperParameterBound(1 To 2000) As Double
Public EnforceParameterBound(1 To 2000) As Boolean

Public Sub LoadSetup()
   MaxGen = ThisWorkbook.Worksheets("Optimizer").Range("D6").Value

   ''''''''''''''''''''''''''''''''''''''''
   NumberOfParameters = NumWtsAndBiases
   'NumberOfVectors                                            =
ThisWorkbook.Worksheets("Optimizer").Range("D4").Value
   NumberOfVectors = NumberOfParameters * 10    'Advise of Storn and Price
(19xx)
   'NumberOfVectors = 100   'Overule
   ''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''

   CrMin = ThisWorkbook.Worksheets("Optimizer").Range("C10").Value
   CrMax = ThisWorkbook.Worksheets("Optimizer").Range("C11").Value
   CrStep = ThisWorkbook.Worksheets("Optimizer").Range("C12").Value

   FMin = ThisWorkbook.Worksheets("Optimizer").Range("D10").Value
   FMax = ThisWorkbook.Worksheets("Optimizer").Range("D11").Value
   FStep = ThisWorkbook.Worksheets("Optimizer").Range("D12").Value


   ClearOutput
   LoadParameterBounds
End Sub

Public Sub LoadParameterBounds()
   'Loads the Lower and Upper bounds of each vector parameter from the
worksheet
   Dim EachParameter As Integer ' For looping through the parameters

   For EachParameter = 1 To NumberOfParameters
      LowerParameterBound(EachParameter) = _

ThisWorkbook.Worksheets("Optimizer").Range("I6").Value
      UpperParameterBound(EachParameter) = _

ThisWorkbook.Worksheets("Optimizer").Range("J6").Value
      EnforceParameterBound(EachParameter) = _

ThisWorkbook.Worksheets("Optimizer").Range("k6").Value
   Next EachParameter
End Sub
```

```
Public Sub ClearOutput()
    'Clears the output area
    ThisWorkbook.Worksheets("Optimizer").Range("P5").Value = ""
    ThisWorkbook.Worksheets("Optimizer").Range("M6:N2000").ClearContents
 End Sub

 Public Sub getInitialStartupBestVector()
 'This function generates initial vector and outputs as the best vector
    GenerateRandomizedVector (1)
    CopySolutionForOutput 1, CurPopParameters, CurPopObjective
    OutputVectorToGUI
 End Sub
```

General Functions

```
Public Function RandomizeBetween(ByVal MinValue As Double, ByVal
MaxValue As Double) As Double
    'Here we get a random values between the lower and upper boundaries
    Randomize
    RandomizeBetween = ((MaxValue - MinValue) * Rnd) + MinValue
End Function

Public Function RandomInteger(ByVal MinValue As Integer, ByVal MaxValue As
Integer) As Integer
    'Generate a random integer between the lower and upper boundaries
    Randomize
    RandomInteger = Int((MaxValue - MinValue + 1) * Rnd + MinValue)
End Function
```

Objective Function

```
Option Explicit
'This module contains functions that computes the objective functions
'Users should make necessary modifications here

Public Function ComputeObjective(ByVal VectorNumber As Integer, _
                                ByRef VectorPopArray) As Double
    'First get the 1D parameter array
    Dim EachParameter As Integer
    Dim Elearn As Double 'Learning error rate to be minimized

    For EachParameter = 1 To NumberOfParameters
        'Copy the parameters for ANN computations
        WtPlusBiasArray(EachParameter) = VectorPopArray(VectorNumber,
EachParameter)
    Next EachParameter
    ''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
    'Now use the prameter array for the necessary computations
    'The wt n bias array is used for ANN computaion, the output is stored in
    'the specified 2D array.
```

```
    RunANNonEntireSampleSet                                     NumLearnSamples,
LearnSamplesInput2DArray, _
                                                LearnSamplesActualOutput2DArray

    Elearn                    =                        ComputeAveEp(NumLearnSamples,
LearnSamplesExpOutput2DArray, _
                                               LearnSamplesActualOutput2DArray)
    'return Elearning
    ComputeObjective = Elearn
End Function
```

Validation

```
'Validation Module
Option Explicit
Public FirstValidationRun As Boolean

Public Sub UpdateValidation(ByVal genNumber As Integer)
    Dim eLearning As Double
    Dim eTest As Double
    Dim echParam As Integer

    'At each generation in optimize, this sub is called
    'During call to optimize, the best Elearning and weights and biases
    'for the given number of hidden nodes are stored on the DE_Outputer
    'Copy for further processing

        'First, we get the testing error
        For echParam = 1 To NumberOfParameters
            'Copy the best parameters for ANN computations
            WtPlusBiasArray(echParam) = OutputerParametersArray(echParam)
        Next echParam

        RunANNonEntireSampleSet                                NumValidationSamples,
ValidationSamplesInput2DArray, _
                                                ValidationSamplesActualOutput2DArray
        eTest                    =                   ComputeAveEp(NumValidationSamples,
ValidationSamplesExpOutput2DArray, _
                                               ValidationSamplesActualOutput2DArray)
        eLearning = OutputerObjective
        'Check for first run or better eTest''''''''''''''''''''''''''''''''''''''''''''
        If (FirstValidationRun = True) Or (eTest < bestETest) Then
            bestETest = eTest
            bestELearning = eLearning
            bestNumHidNodes = NumHiddenNodes
            For echParam = 1 To NumberOfParameters
                'Copy the best parameters for ANN computations
                bestWtPlusBiasArray(echParam) = WtPlusBiasArray(echParam)
            Next echParam
            OutputTrainResultToGUI
```

```
        'Output the generation best solution was found
        ThisWorkbook.Worksheets("Optimizer").Range("K17").Value =
genNumber
        FirstValidationRun = False
    End If
End Sub
```