NON-REVENUE WATER COMPONENT SPECIFIC DRIVERS IN LILONGWE CITY

MSc THESIS

(WATER RESOURCES MANAGEMENT AND DEVELOPMENT)

AIMLOSS GREYFIELD BANDA

(MscWRM0120)

(BSc Natural Resources Management)

MZUZU UNIVERSITY, MALAWI

JULY 2024

NON-REVENUE WATER COMPONENT SPECIFIC DRIVERS IN LILONGWE CITY

AIMLOSS GREYFIELD BANDA

(MscWRM0120)

(BSc Natural Resources Management)

A THESIS SUBMITTED TO THE FACULTY OF ENVIRONMENTAL SCIENCES, DEPARTMENT OF WATER AND SANITATION IN FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF A MASTER OF SCIENCE DEGREE IN WATER RESOURCES MANAGEMENT AND DEVELOPMENT

MZUZU UNIVERSITY

JULY 2024

DECLARATION

I hereby declare that this thesis titled "*Analysis of Non-Revenue Water Component Specific Drivers in Lilongwe City*" has been written by me and is a record of my research work. All citations, references, and borrowed ideas have been duly acknowledged. This thesis is being submitted in fulfilment of the requirements for the award of the Degree of Master of Science (MSc) in Water Resources Management and Development at Mzuzu University. None of the present work has been submitted previously for any degree or examination at any other University

Student Name: AIMLOSS GREYFIELD BANDA

Signature: ___________________________ **Date:__________________________**

CERTIFICATE OF COMPLETION

The undersigned certify that this thesis is a result of the author's work and that to the best of our knowledge, it has not been submitted for any other academic qualification within the Mzuzu University or elsewhere. The thesis is acceptable in form and content, and that satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate through an oral examination held on: ______________________________

Signature: _________________________ Date: _________________________

Main Supervisor: Dr Brighton A. Chunga (PhD)

Signature: _________________________ Date: _________________________

Co-supervisor: Eng. Eddie Mtonga

Signature: _________________________ Date: _________________________

Postgraduate Coordinator: Dr Russel C. G. Chidya (PhD)

Signature: _________________________ Date:_________________________

Head of Department (Water and Sanitation): Dr Brighton A. Chunga (PhD)

DEDICATION

This work is dedicated to my mother Miss Mellia Nellie Black and my aunt Miss Matilda Matondo whose support, encouragement and faith in my destiny made me complete this work. I further dedicate this work to Miss Rosemary Nyamwera who inspired me to work even harder and provided me with unwavering emotional support throughout the journey.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to all those who have supported and contributed to the completion of this thesis. First and foremost, I am immensely grateful to my thesis supervisors, Dr Brighton A. Chunga (PhD) and Eng. Eddie Mtonga, for their invaluable guidance, patience, and unwavering support throughout the entire research process. Their expertise, insightful feedback, and encouragement have been instrumental in shaping this work and pushing me to achieve my best.

I extend my sincere appreciation to the Lilongwe Water Board's Non-Revenue Water Manager, Mr. E. Ngaivale, and the entire Lilongwe Water Board's members of staff who generously dedicated their time and participated in this study. Their contribution and cooperation were vital in collecting the data needed for analysis and interpretation.

I am also deeply indebted to my friends and family members for their unwavering support, understanding, and encouragement throughout this challenging endeavor. Their belief in me and their continuous motivation have been a constant source of strength and inspiration.

Lastly, I would like to express my profound gratitude to all the researchers, scholars, and individuals whose work and contributions have been referenced in this thesis. Their groundbreaking research and insights have paved the way for the advancement of knowledge in this field.

ABSTRACT

The Lilongwe Water Board (LWB), one of the water utility companies in Malawi, whose water losses as of 2022 were as high as 40%, employed network rehabilitation techniques in their 2015 to 2020 strategic plan to reduce the then water losses of 38.9% to 28%. However, as Non-Revenue Water (NRW) is caused by myriad drivers, network rehabilitation activities have been proving futile. This study analysed the NRW component-specific drivers for LWB in Lilongwe City. A quantitative research design was adopted in which purposive sampling of the specific sites, Area 22 and part of Area 2 was used. R studio version 4.1.1 was used to perform descriptive statistics, pairwise Pearson Correlation tests, Minimum Night Flow Model, Fixed Effects Regression Modeling, and a feedforward backpropagation Artificial Neural Network based Improved Garson Algorithm's sensitivity analysis. The water balance framework for the two District Metered Areas (DMAs) (SZA1 and SZD3) was established, which confirmed that although post-rehabilitation NRW (38.95%) is below the 2022 national average (54.61%), it is still above the LWB's target and global average NRW range of between 30 to 35%. The water loss components analysis showed that apparent losses (AL) contributed more to water losses than real losses (RL) and unbilled authorized consumptions (UAC). Different component-specific drivers were further identified for these NRW components. Maintenance works were the main driver for UAC, while accounting errors, illegal connections, and customer non-payment drove AL. Background leakages and bursts, connection density, type of pipe materials, and population density were the main RL drivers. In conclusion, this study has revealed the component specific NRW drivers for the LWB in Lilongwe City, Malawi. It recommends the need for targeted interventions to address the identified drivers and NRW to meet the LWB's goals and global standards.

Keywords: Non-revenue water, potable water, district metered area, apparent losses, real losses, unbilled authorized consumptions

ACRONYMS AND ABBREVIATIONS

AL	Apparent Losses
ANN	Artificial Neural Network
BC	Billed Consumption
DMA	District Metered Area
ENF	Excess Night Flow
FAVAD	Fixed and Variable Area Discharge path
GIS	Geographic Information System
GoM	Government of Malawi
HSM	Hydraulic Simulation Modelling
IBNET	International Benchmarking Network
IWA	International Water Association
JICA	Japan Internation Cooperation Agency
LNF	Legitimate Night Flow
LWB	Lilongwe Water Board
MLP	Multi-Layer Perceptron
MNF	Minimum Night Flow
MZUNIREC	Mzuzu University Research Ethics Committee

NDF	Night Day Factor
NHM	Network Hydraulic Modelling
NNF	Net Night Flow
NRW	Non-Revenue Water
OECD	Organisation for Economic Co-operation and Development
PVC	Polyvinyl Chloride
RL	Real Losses
SCADA	Supervisory Control and Data Acquisition
SDG	Sustainable Development Goals
SIV	System Input Volume
UAC	Unbilled Authorized Consumption
UNICEF	United Nations Children's Fund
USA	United States of America
USAID	United State Agency for International Development
USD	United States Dollar
WDS	Water Distribution System
WRC	Water Research Commission

WSP Water and Sanitation Program

WWAP World Water Assessment Programme

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

1.1 Background

The world today suffers from an increased lack of access to an improved water supply. It was reported that, by 2020, 6% of the world's population had already been affected by the problem .It is estimated that by 2025, more than half of the world's population will be living in water scarce or water stress regions of the world (Ritchie & Roser 2021; UNICEF 2021). The United Nations Sustainable Development Goal Number 6, estimates the current global lack of safely managed drinking water is at 26% and aims at reducing the gap and achieve universal and equitable access to safe and affordable drinking water for all by 2030 (United Nations 2015). In urban settings, safe and reliable drinking water is provided for by water utilities companies. Historically, water utilities have been in operation for more than 100 years worldwide (Jang & Choi 2017). While many countries are striving to achieve the United Nations 2030's Goal Number 6 through water utilities, many of them are faced with common challenges all over the world. The main challenges affecting water utilities include increased water demand as a result of rapid population growth, water scarcity and dilapidated infrastructure due to old age (Fernandez *et al*. 2011; van den Beg 2015; Hoko & Chipwaila 2017; Farouk *et al*, 2021). As utilities are dealing with the effects of water scarcity and increased demand as a result of rapid population growth, their efforts are often crippled by challenges emanating from the ageing system infrastructure and the general utilities' management. The most prominent of these challenges is Non-revenue water.

Non-revenue Water, as defined by Elkharbotly *et al*. (2022), refers to the difference between water put into the distribution system and the amount of water billed to consumers. However, NRW may also include the billed but not paid for water in a water utility company at a particular water audit period. Globally, water productivity is less than optimal as non-revenue water is on the increase (Lambert *et al*. 2014). Van den Berg (2015) noted that an estimated cost of utility water losses through leaks (real losses), drinking water not being invoiced to customers (apparent losses) and unbilled authorized consumption is USD141 billion per year. Further to the estimated costs, Makaya (2015) asserted that 55% of the utility NRW water occur in developing countries mainly due to poor water loss management techniques such as limited ability to trace background leakages, water theft, metering errors and many others . However, based on the survey conducted in 48 cities, water losses in the Organisation for Economic Co-operation and Development (OECD) countries can be as high as 37% and in emerging economies as high as 65% (OECD, 2016), which suggests that NRW management is a global crisis. Nonetheless, differences exist between water utilities in developed and developed countries due to the operating environments and financial capacity to invest in infrastructure development.

A utility's water loss response strategy and sensitivity depends, to a greater extent, on the country's economic and governance stability and as a result, many water distribution systems in developing countries are operated in dilapidated state (Makaya & Hensel 2014 and WWAP 2014). The World Bank estimated that roughly water utilities lose 45 million cubic meters of water daily in developing countries. Considering the water supply problems experienced in these countries, the World Bank further approximated that saving half of such losses would provide enough water to serve at least 90 million people annually (World Bank Group 2016). Despite the knowledge of water loss challenges, water utilities in developing countries still have a huge gap in addressing the problem due to several factors. For instance, in their study, Gonzalez-Gomez *et al*. (2011)

argued that the slow progress in NRW reduction programs in developing countries like Malawi is mainly a result of political interferences and institutional resistance to change such that the utilities have not been ploughing back proceeds for network rehabilitation. Revenue collected by water operators is diverted to other uses than the maintenance and upgrading of water distribution systems (van den Berg & Danilenko,2017). These observations echo the findings of Meijer (2015), who concluded that water utilities in Malawi are robbed of their strength to fight non revenue water due to revenue mismanagement.

Owing to these revenue losses, high water loss levels are normally a surrogate for a poorly run water utility company that lacks 'good' governance, autonomy, accountability, and the technical and managerial skills of the management team necessary for providing reliable service to their population (Gonzalez-Gomez *et al*. 2011). As such, strengthening programs that aim at water loss reduction, does not only enhance utilities' economic stability but also general utility management and service delivery. In their study, Fernandez *et al*. (2011), also stressed that many utilities find water loss reduction as an activity of low incentive when the opportunity costs are weighed in. When the costs of reducing water loss are considered to be very high in view of poor tariff structures, water loss reduction programmes are disincentivised (Kanakoudis *et al.* 2013; Kanakoudis *et al.* 2015). Nonetheless, as also acknowledged by McKinsey (2013), Non-Revenue Water management remains an important instrument for reducing the gap between the supply of and demand for potable water.

In water distribution systems, water is lost through unauthorised consumption, metering inaccuracies, leakage on transmission and or distribution mains, utility storage tanks leakage, storage tank water overflow and service connections leakage up to customer metering (Kenias *et al*. 2019). Generally, in many water utilities, water is also lost in a distribution system through unbilled authorised consumptions such as unbilled metered consumptions and unbilled unmetered consumptions (Makaya 2015). Various researchers have categorized different factors causing Non-Revenue Water in a distribution system as physical, technical, managerial, and environmental (Dermirci *et al*. 2018; van den Berg 2015). These factors are, in most cases, interchangeably referred to as determinants, or drivers (Elkharbotly *et al.* 2022; Gonzalez-Gomez *et al*. 2012; van den Berg 2015). The NRW drivers such as age of the system, soil conditions, system servicing employee to customer ratio, topography, pressure in the system, population density and connection density are well known to influence the water lost in a distribution system (Gonzalez-Gomez *et al.* 2011). Therefore, if water utilities in developing countries like Malawi can labour to determine these drivers, the development of NRW control strategies which prioritize the most crucial drivers, may prove to be effective (Malek *et al.* 2021).

In an effort to curb NRW, water utilities employ various techniques ranging from managerial, technical, to integragrated methods (AL-Washali *et al.*, 2020). From the technical aspect, different methods and technologies have also been developed to solve NRW related problems. For instance, some of the widely adopted approaches to addressing NRW include the use of traditional statistical and network hydraulic smulation models (Elkharbotly *et al.* 2022b). While Hydraulic Simulation Modeling (HSM) with tools such as EPANET continues finding wide application worldwide (Berardi *et al.*, 2016), Serafeim *et al.* (2022) argued that this method presents other challenges due to its overreliance on the computational strength of computers for predicting and investigating the operational functionalities of Water Distribution Systems in solving NRW problems. Statistical models on the other hand, which include linear regression and time-series analysis, are also widely employed to estimate and analyze water losses based on historical data and relevant variables (Jung *et al.* 2015). However, as also suggested by Malek *et al.* (2021), while effective to some extent, statistical models often struggle to capture the intricate relationships and non-linearities inherent in water distribution systems, leading to limitations in accuracy and predictive capabilities. Consequently, reseachers such as Jembre *et al.* (2021) and, Serafeim *et al.* (2022) recommended that such methods should be used when investigating water distribution network parameter which have simple and not complex relationships with NRW. In cases of the probability of having non linear or complex relationships, many researchers and utilities have increasingly turned to Deep Learning algorithms, such as artificial neural networks (ANNs), due to their ability to handle complex and large-scale datasets, identify patterns, and provide accurate predictions (Abdulla *et al.* 2013; Makaya 2015; Makaya and Hensel 2015; Elkharbotly *et al.* 2022). Elkharbotly *et al.* (2022) also added that Deep Learning models such as the ANN are also much preferred to Hydraulic Simulation Models due to their ability to learn from existing data without needing extensive field measurements for every node, as required in hydraulic models.

Deep learning algorithms, particularly ANNs, offer significant advantages in tackling the challenges associated with NRW (Elkharbotly *et al.* 2022a). ANNs are designed to mimic the structure and functioning of the human brain, consisting of interconnected artificial neurons that process and learn from input data (Maind 2014). This enables ANNs to capture complex relationships and patterns in large datasets, even when the underlying mechanisms are not explicitly known. Makaya (2015), added that the suitability of ANNs for addressing NRW in various researches stems from their ability to analyze and integrate diverse data sources, including historical consumption patterns, network parameters, weather conditions, and socioeconomic factors. By incorporating these variables, ANNs are able to model and predict water losses more accurately, allowing utilities to identify the areas and factors contributing to NRW and implement targeted interventions. Moreover, ANN models adapt and improve their performance over time through a process known as training, in which thery are run to recognize patterns, relationships, and representations within a given dataset so that they can make accurate predictions or perform specific tasks (Maozhun and Ji 2017). By feeding the network with labeled examples, these models iteratively adjust their weights and biases to minimize errors and enhance predictions. This feature is particularly valuable in the context of NRW, where the dynamics of water systems and the contributing factors may change over time. ANNs' ability to continuously learn and update their models therefore makes them well-suited for addressing the evolving challenges associated with NRW.

In 2005, the Malawi Government enacted the National Water Policy to address all aspects of water including resource management, development and service delivery (GoM 2005). The policy articulates a water sector vision of ‘Water and Sanitation for All, Always’ and provides policy objectives to achieve sustainable delivery of water supply and sanitation services. In line with the policy, the country’s water utilities are obliged to establish periodic strategic plans for development. Despite the establishment of these strategic plans, efforts by parastatal water boards to supply water to all cities and urban centres remain challenged. The high population growth against constant ageing supply infrastructure is almost rendering the parastatals’ water supply efforts futile. Of special interest to this study is the Lilongwe Water Board (LWB), which was initially established in 1947 and later reconstituted by the Water Works Act of 1995 (JICA

2022). The utility currently supplies water to an area of about 45,000 hectares which is divided into three water supply zones within Lilongwe city urban namely; southern, central and northern zones (LWB 2018).

LWB suffers an acute water supply challenge mainly due to limited system capacity and increased water losses (WaterAid 2016). In its efforts to respond to the water supply challenges, LWB is strengthening the existing system through physical system upgrade and investments in NRW reduction programs. In one of the very notable projects on the reduction of NRW, the LWB in coordination with Lilongwe City Council is implementing a Water and Sanitation Project which commenced on 26th March 2018, with an anticipated completion date of 30th June 2023, with funding from the World Bank (USD75 million credit and USD25 million grant) and the Government of Malawi (USD2 million) (LWB 2021). Among the three major components of the project, the board envisages to rehabilitate water distribution networks, reduce non-revenue water and expand its institutional and physical capacity. Although the anticipated results of these projects promise to be a giant leap towards achieving water security in Lilongwe City, if other causes of water losses remain unchecked, water demand will continue to stifle the system's capacity to achieve water security. The LWB points to the dilapidated distribution infrastructure as the main cause of increased NRW (JICA 2020). However, as also noted by Elkharbotly *et al*. (2022), dilapidated infrastructure alone can not explain all the NRW recorded by the utility. Many drivers contribute to water losses whose impact should properly be ascertained in order to correct them in their order of effect. However, while many researchers examine NRW drivers on the total NRW value, analysing these drivers at NRW component level would eliminate the error of overlooking the impact of seemingly minor but very sensitive component specific drivers on NRW (Makaya (2015).

Generally, many studies have attempted to describe key drivers of NRW in different countries. One key study is that of van den Berg (2015) which established that for most water utilities in different countries, the major NRW drivers are population density per kilometer of network and the type of distribution network. Although her work revealed positive direction in tackling NRW systematically, like many other researchers (Murrar 2017; Murrar *et al*. 2017; Güngör-Demirci *et al*. 2018; Njerita, 2019; Malek *et al.* 2021; Elkharbotly *et al.* 2022), the examination of these drivers was on holistic level of NRW, not component specific. As stipulated, the analysis of NRW drivers on total NRW can overlook the impact of seemingly minor but very sensitive component specific drivers on NRW.

In Malawi, although previous research (Kafodya 2010; Meijer 2015; Hoko & Chipwaila 2017) has attempted to bring to light ways of reducing NRW in Malawi, the focus has been on prescribing technical methods for non-revenue water reduction such as the use of Goegraphical Information System (GIS) technology to manage leakages. No known study had been conducted to establish the main component specific drivers of Non-Revenue Water using deep learning methodologies, especially for LWB. This study, therefore, aimed to analyse the utility's Non-Revenue Water component specific drivers using a combination of statistical and deep learning algorithms.

1.2 Problem statement

LWB continues to experience high NRW (JICA 2022). As of 2022, the water board accounted for 40% of NRW of the total production volume which further constrained the quest to meet supply needs (LWB, 2022). This was despite having set a goal to reduce NRW to a rate of 28% from 38.9% by 2020 in their 2015-2020 strategic plan. According to the utility's management, the high NRW cases are a result of the dilapidated infrastructure, which consequently forced the utility to embark on system rehabilitation programs (LWB 2021). While the ongoing water distribution infrastructure rehabilitation programs could result in a considerable change in NRW, without putting in check other factors that influence such losses would only result in a temporal relief, as NRW are a function of many factors such as utility management, system operating standards and many others (Elkharbotly *et al*. 2022). Therefore, this study sought to establish the main drivers of LWB's NRW, to help the utility to devise appropriate programmes that would reduce water losses for optimal operation of the water supply system.

To avoid the error of overlooking the seemingly minor but sensitive drivers, the analysis of factors that influence water losses at NRW component level was ideal. This method resonated well with Makaya (2015), who suggested that models used to investigate factors that influence NRW should, for instance, find a way to separate losses due to Unbilled Authorized Consumption (UAC), which are in most cases accounted for as RLs. Therefore, an in-depth analysis of the main specific drivers of the three NRW components; Apparent Losses (AL), UAC and RL), for LWB was done using the Improved Garson (IGarson) algorithm based on Back Propagation Neural Network Model and the Fixed Effects Regression Model to perform a water distribution systems' sensitivity analysis. The sensitivity analysis determined and ranked the component of NRW which constituted a larger portion of the total NRW and also the main drivers for each component of LWB's NRW.

1.3 Study objectives.

The main objective of this study was to analyze the Non-Revenue Water component specific drivers for the Lilongwe Water Board

1.3.1 Specific objectives.

To achieve the main objective, the study was guided by the following specific objectives:

1. To analyse the water balance for selected District Metered Areas of Lilongwe Water Board
2. To determine the level of contribution for each Non-Revenue Water component on total Non-Revenue Water.
3. To investigate the main drivers for each Non-Revenue Water component for Lilongwe Water Board.

1.4 Research Questions.

1. How is the system's input volume of water managed in a distribution network?
2. How does the interaction of Non-Revenue Water components influence the total Non-Revenue Water value for Lilongwe Water Board?
3. What are the major factors that enhance each component of Non-Revenue Water in Lilongwe Water Board distribution system?

1.6 Study justification.

LWB, established in 1947, which is the second oldest water board in Malawi, from the Blantyre Water Board (JICA 2022), keeps recording high levels of NRW despite many interventions put in place to reduce the challenge. Coupled with its already constrained capacity to supply water to the city's high population, these high levels of NRW have the ability to cripple efforts aimed at decreasing consumer demand. Although it is not feasible to eliminate all NRW in a utility, achieving low levels is possible when reduction programs are strategic. Strategic approaches to achieving low NRW levels would not only involve the establishment of effective arrangements in the managerial and institutional environment—often requiring attention to some fundamental challenges, but also the adoption of new technical approaches. One such a technical based approach involves the establishment of the main drivers of NRW and the implementation of target-based solutions. This study therefore will help;

- LWB to target its NRW reduction solutions at the main drivers;
- NRW managers elsewhere to adopt a strategic approach to dealing with the problem;
- Utilities to save considerable amount of revenue currently being lost through NRW and various traditional approaches employed to its reduction;
- Contribute towards the achievement of SDG 6 which aims at achieving universal and equitable access to safe and affordable drinking water for all by 2030;

CHAPTER TWO: LITERATURE REVIEW

This chapter discusses relevant NRW literature and its different management techniques. Specifically, it endeavors to discuss global overview of water distribution losses, NRW in developing countries, utility water losses in Malawi, mechanisms through which water is lost in distribution systems, hydraulic and ANN modelling methodologies used to account for water losses in a water distrbution network, the water balance, the leakage theory, leakage quantification of a water supply network, traditional leak detection approaches and computational approaches of managing water losses. Finally, the chapter provides a review summary and appraisal of some of the discussed concepts and methodologies for this research

2.1 Water Distribution Losses

Since the development of the water reticulation systems, water loss management techniques have been developed by various water utilities in different countries (Elkharbotly *et al*. 2022). Despite different mechanisms put in place to reduce utility water losses, various studies have shown that utilities keep registering high levels of water losses worldwide (Güngör-Demirci *et al* 2018; Al-washali *et al.* 2019; AL-Washali *et al.* 2020). While the IWA, as qouted by Paudel (2019), classifies water losses in a distribution system as being commercial, unbilled authorised and physical, researchers such as Van den Berg (2015) and Güngör-Demirci *et al.* (2018) classify causes of these water losses into physical, technical, managerial, as well as environmental.

In their studies van den Berg (2015) and Güngör-Demirci *et al.* (2018), further categorized such physical factors that necessitate water losses as network characteristics pertaining to the age of the system, length of the distribution pipes, the type of network material, connection density and many others. While characteristics such as hours of water supply (as intermittent supply leads to increased NRW through uncontrollable water flow which creates periods of water surges when supply resumes, causing pipe bursts and other damages(Mubvaruri *et al.* 2022)), number of staff per 1000 people served, net revenue per cubic meter of water sold, planning processes and many others are classified as being managerial factors that influence levels of water losses in a distribution system (van den Berg and Danilenko 2017). The design of the distribution system, as well as the standard operating pressures and many other parameters, on the other hand, are regarded as being technical factors that influence water losses. Finally, Güngör-Demirci *et al.* (2018), stipulated that environmental factors such as the topography of the area in which a system is established, type of the soil, national levels of corruption, Gross Domestic Products per capita as well as utilities' wage rate levels are found to affect utilities' levels of water losses. Although these research findings and classfications concur with other researchers like Malek *et al*. (2021), such classifications of water loss causes have the potential to cause errors in the water balance audits. For instance, the classification of water causes into physical and technical, may result in erroneous segregation of physical or RLs from commercial losses when other water loss parameters such as metering errors are considered. This gap therefore, warranted the need for detailed analysis of NRW at component level to establish specific drivers for each componet.

2.1.1 Water losses in developed countries

Due to differences in response strategies and the sensitivity of water utilities and the governments in which they operate, NRW water management scenarios differ between utilities in developed and developing countries (Liemberger and Wyatt 2019). Utilities in developed countries have managed to reduce NRW to an acceptable and manageable level. For instance, Liemberger and Wyatt (2019) reported that by 2018 in the USA, the average NRW was estimated at 13%, while 21% was found to be lost in Belgium. Furthermore, countries like Denmark, New Zealand, Lithuania, Singapore, Australia, UK lost within the same period an average of 7%, 24%, 21%, 4%, 22% and 23 %, respectively (Liemberger and Wyatt 2019; Kizilöz and Şişman, 2021). Despite these low levels of NRW in developed countries, IBNET (2018) reports that great variations exist within different utilities of the same countries. These variations, therefore, show that, despite the advancement in terms of response strategies and sensitivity, NRW management remains a challenge even for developed countries. Kizilöz and Şişman (2021), also add that the low levels of NRW in developed countries result from continuous maintenance of water distribution systems, increased monitoring for illegal water use and physical losses. Notwithstanding these reduced NRW levels, Hui *et al.* (2020), argued that many water utilities in developed countries still find it beneficial to produce more water to increase revenue than to reduce the amount of water being lost when the cost of maintenance outweighs the revenue to be saved from such an exercise. He also adds that, the choice to produce more water to cover for NRW also comes due to lack of proper understanding of the water audit processes to pinpoint all causes of water losses. Therefore, although the water loss levels are relatively low, there still lacked systematic methods that utilities can use to target interventions on the main causes of NRW.

2.1.2 Water losses in developing countries

In their report, UN-Water (2016), stipulated that many water distribution systems in developing countries are operated under dilapidated conditions which results in supply inefficiency. In addition to the problems associated with inefficient water supply, water losses in developing countries have reached alarming rates, with non-revenue water levels in excess of 60% having been recorded in many developing countries (Makaya 2015). Makaya (2015), further argued that the slow progress in water loss reduction in developing countries is characterized by political interferences and institutional resistance to change. These observations are in tandem with the findings of Meijer (2015); van den Berg and Danilenko (2017), who found that revenue collected by water utilities is diverted to other uses than the maintenance and upgrading of water distribution systems.

In regions of developing countries like Africa, Latin America and Asia, the highest levels of NRW ranging from 40% to 60% of the System Input Volumes are recorded (IBNET 2018). For instance, developing countries such as Hungary, South Africa, Ukraine, Argentina, Bosnia and Herzegovina, Bulgaria and Manila experience water losses averaging to 32%, 34%, 36%, 42%, 49%, 61% and 65%, respectively (Kizilöz and Şişman 2021). On the other hand, NRW in African countries alone ranges from 5% in some South African towns to 70% in Liberia (van den Berg and Danilenko 2017). The high levels of NRW in developing countries indicate the seriousness of the water loss problem. Nonetheless, Makaya (2015) argued that, due to insufficient data documented by many water utilities in Africa, these levels maybe be an underestimate or overestimation of the real problem. Therefore, it is a necessity that developing countries should seriously consider monitoring and reducing their water losses in order to operate sustainably.

In an effort to curb the skyrocketing NRW levels in Africa, various initiatives have been developed. The most notable initiatives include the development of software for understanding and reducing NRW by the South African Water Research Commission (WRC), which came about when the WRC discovered that the reduction of non-revenue water was one of the key challenges facing the continent's water utilities (Mckenzie and Wegelin 2009; Farley *et al.* 2010; AL-Washali *et al.* 2020). In addition, the WRC developed a collection of models and associated documentation to assist African water utilities deal with NRW. The models currently available and in use include the SANFLOW minimum night flow analysis model, the PRESMAC pressure management model, the ECONOLEAK active leakage control assessment model and the AQUALITE water balance model (AL-Washali *et al.* 2020). Besides the development of models to aid in the reduction of NRW, 70% of the 134 utilities across Africa agreed to adopt a water loss benchmarking initiative. In their agreement, a benchmark of 0.3m^3/connection/day was set as the NRW indicator for optimum performance of water utilities in Africa (WSP, 2010). However, despite these efforts, high NRW levels keep being recorded by many utilities in Africa. Van den Berg and Danilenko (2017), attribute this continued increase to NRW levels to a lack of investment in technology and infrastructure which can foster changes in user behaviour as well as the development of integrated water loss management programs. This observation is in tandem with the findings of both WSP (2010) and (IBNET 2018) who concluded that despite having well developed water loss strategies, many utilities in developing countries do not invest in technology and infrastructure to enable the developed strategies.

2.1.3 Water Losses in Water Supply Systems in Malawi

Malawi, as a developing country, is not spared from high NRW cases. Adding to the poor governance and low tariffs which are also coupled with urbanization which strain the efficiency of the country's water utilities, NRW has been one of the most prominent challenges to water supply by the LWB, Central Region Water Board, Southern Region Water Board, Blantyre Water Board and Northern Region Water Board (Meijer 2015; Magombo and Kosamu 2016; JICA 2019, 2020). JICA (2019) stipulated that the country's water boards lose 30% to 54% of the water they produce. In light of these findings, out of the total volume produced by the Blantyre Water Board, 40 to 49% is lost, which is a bit higher than the volume of water lost by the LWB, currently at 38.9% (Magombo and Kosamu 2016; JICA 2020). On the other hand, slightly lower NRW levels are recorded by the Northern Region Water Board and the Southern Region Water Boards, at 36% and 25% respectively (Dutch Water Operators 2021; GoM 2022). Through these losses, LWB (2021) stipulates that Malawi's water boards lose about K35 billion each year which could otherwise be ploughed back into system maintenance or new infrastructural development.

Water Utilities in Malawi keep registering high NRW levels despite various interventions taking place which include network rehabilitations and adoption of monitoring systems such as GIS and SCADA in some regional water boards (Dutch Water Operators 2021; JICA 2022). This phenomenon echoes the conclusion by Elkharbotly *et al*., (2022) who postulated that rehabilitation of dilapidated infrastructure alone to deal away with NRW recorded by the water utility copmany may prove to be futile if the distribution systems are not examined to understand other contributing factors than the usual old age of the system. Moreover, Meijer (2015) and Hoko and Chipwaila(2017) in their studies found that because water utilities in Malawi heavily rely on Network Hydraulic Modells such as EPANET, in which leakage flow is assumed to be uniformly distributed along the pipeline, the calculated NRW values are in most cases flawed. Therefore, there was a need for proper system reconnaissance and analysis methodology to understand other underlaying factors that are neccestating the increase in NRW by Malawi's Water Utilities.

2.2 The Water Balance

The initial phase in strategizing NRW management techniques is to understand the magnitude of water lost by establishing the water balance or water audit (Knobloch *et al.* 2014). Through the establishment of the water balance, water utilities are able to understand not only the magnitude of the water losses but also sources and cost of such water losses (Raskar and Gawande 2014). To guide the processes of water balance establishment by different water utilities, the International Water Association (IWA) devised a standard structure for conducting water audit (Mastaller and Klingel 2017). The IWA's Water Balance methodology is recognized internationally as the best practice in identifying, measuring, and analyzing water loss from point of production to the end consumer (Canto Ríos *et al.* 2014; Charalambous *et al.* 2014; Paudel 2019; Alima 2020). Makaya (2015), explained that the IWA's water balance method in simpler terms is a methodology for water accountability and an integrated approach to water loss control. This methodology has been standardized for uniformity and understanding of NRW, location of loss, reasons for loss and the potential strategies which can be employed to achieve water utility effectiveness and efficiency (Alima 2020). The IWA's standard structure outlines that, to avoid guess estimations and achieve maximum accuracy in water auditing, the following must be put in place (Charalambous *et al.* 2014); master meters to measure the production, consumer meters to measure the billed consumption, zonal meters to determine the amount of water flowing into each zone, strategic meters for measuring authorized consumption

Before the introduction of IWA Water Balance methodology, water utilities lacked a standard way of defining NRW terminologies and monitoring indicators (Canto Ríos *et al.* 2014). The development of the IWA standards brought in the standardized set of definitions for water balance components to explain common terminologies and indicators that allow comparison and benchmarking (Radivojević. 2020; Serafeim *et al.* 2022) The IWA's water balance Table is as follows.

Table 1: International Water Association: Water Balance

<table>
<tr><td rowspan="9">System Input</td><td rowspan="4">Authorized Consumption</td><td>Billed Authorized Consumption</td><td>Billed Metered Consumption</td><td rowspan="2">Revenue Water</td></tr>
<tr><td rowspan="3">Unbilled Authorized Consumption</td><td>Billed Unmetered Consumption</td></tr>
<tr><td>Unbilled Metered Consumption</td><td rowspan="7">Non-Revenue Water</td></tr>
<tr><td>Unbilled Unmetered Consumption</td></tr>
<tr><td rowspan="5">Water Losses</td><td rowspan="2">Commercial (Apparent) Losses</td><td>Unauthorized Consumption (e.g., illegal connections)</td></tr>
<tr><td>Customer metering inaccuracies, Estimations and data handling errors</td></tr>
<tr><td rowspan="3">Physical (Real) Losses</td><td>Leakage on transmission and/or Distribution pipes</td></tr>
<tr><td>Leakage and overflows at Utility's storage tanks</td></tr>
<tr><td>Leakage on service connections up to point customer use</td></tr>
</table>

Source: IWA, Water Balance Table (2000)

Based on the IWA standards, the terminologies in the Table 1 are defined as the following;

2.2.1 System Input Volume

This refers to the total volume of water supplied by the whole or part of the utility, annually. According to Farley and Liemberger (2005), System Input Volume is measured through zonal or master meters of part or the whole distribution system. Knobloch *et al.* (2014) stipulated that the accuracy of a system's water balance depends to a greater extent on the accuracy of system input readings. If a system has unmetered sources Farley and Liemberger (2005), suggested that the annual flows have to be estimated by one or a combination of the following; temporary flow measurements using portable devices, reservoir drop tests and analysis of pump curves, pressure and average pumping hours. For LWB, a report by JICA (2019) showed that many of the utility's DMAs have multiple sources of water which in most cases also go unmetered. Only few DMAs such as SZA1 (Area 2) and SZD3 (Area 22) that have been redefined and properly metered, which makes them ideal for NRW studies.

2.2.2 Authorized Consumption

This water balance component refers to the volume of water used by the water supplier, registered customers, any other authorized water user such as fire hydrants, other government offices and many others. Authorized consumption comprises of the billed metered consumption, billed unmetered consumption, unbilled metered consumption and unbilled unmetered consumption (Farley *et al.* 2010; Knobloch, Guth and Klingel 2014; Vermersch *et al.* 2016; Mastaller and Klingel 2017).

2.2.3 Non-Revenue Water

Traditionally, many researchers such as Abd Rahman *et al.* (2019), Onyango (2021), Abbas *et al*. (2022) and Elkharbotly *et al.* (2022a) define NRW as the difference between system input volume and billed authorized consumption. Despite this definition being widely accepted however, it oversimplifies the underlying complexities of the NRW issue. As found by Ndunguru and Hoko, (2016); Hoko and Chipwaila (2017) and Liemberger and Wyatt(2019) in their researches, part of the NRW includes water billed to the customer but not paid for. The initial definition therefore holds to be true only in water distribution systems with prepaid consumer meters where no customer fails to pay for water already consumed and billed for. It can be concluded therefore that NRW refers to the difference between water put into the system and water whose revenue has been collected.

2.2.4 Commercial Losses

Also called apparent losses, commercial losses comprise unauthorized consumption and all types of inaccuracies in metering (Farouk *et al.* 2023). As explained by Mubvaruri *et al.* (2022) these losses occur when the water is consumed and used legitimately by customers, but the utility fails to bill them accurately or at all. In his study, Murrar (2017) revealed that there are several reasons behind commercial losses, such as faulty metering, billing errors, or inefficient customer management systems. However, as also presented in the previous section, in most water utilities, not all the billed consumption ends up being revenue water. Therefore, commercial losses may also include correctly billed consumption whose revenue is not collected.

2.2.5 Physical Losses

Physical losses which are also referred to as real losses or technical losses, encompass the losses in annual volumes through all types of bursts, leaks, overflows on mains, reservoirs and service connections up to the customer metering point (Mastaller and Klingel 2017; Radivojević *et al.* 2020). As stipulated by Murrar (2017), physical water losses typically occur in areas where there are pipe bursts, in faulty and damaged joints and fittings, malfunctioning or improperly closed valves, leaky or damaged hydrants, leaky DMA reservoirs and leaky pressure reducing stations. However, Jembre *et al.* (2021) also added that part of physical losses occur within the connected properties which are in most cases difficult to be accounted for in a water balance framework. To account for physical losses within connected properties therefore, one needs to estimate legitimate consumption within such properties from which physical losses are deduced by comparing the estimated legitimate consumption and the billed consumption (Knobloch *et al.* 2014; Farouk *et al.* 2023).

2.3 Establishing Water Balance

In order to fully determine the water balance, necessary steps must be considered as follows (Farley and Liemberger, 2005)

2.3.1 Determining System Input Volume

Having the system input well metered, calculation of other components of the water balance becomes easy (Paudel 2019). Firstly, by using a portable flow measuring device, the accuracy of the input meter is verified (Samwel 2020). In a circumstance where the meter readings are not in tandem with the measurements of the temporary flow metering device, a thorough investigation must be done to ascertain the problem and adjust the recorded quantities to reflect the realities on the ground (Farley and Liemberger 2005; Farley *et al.* 2010; Alima 2020). Whereas in cases of unmetered sources, the annual flows must be estimated as described in section 2.2

2.3.2 Determining Authorized Consumption

When establishing this component of the Water Balance, four sub components should be put into consideration (Mastaller and Klingel 2017).

2.3.2.1 Billed Metered Consumption

To accurately calculate this water balance (WB) component, possible billing and data handling errors, as well as information on what would constitute ALs must be detected. Consumer categories based on their consumption, such as domestic, commercial and industrial, must be mined from the billing system and analyzed while paying much attention to the group that has large number of consumers (Radivojević *et al.* 2020).

2.3.2.2 Billed Unmetered Consumption

According to Farley and Liemberger (2005) readings for this component can be obtained from the utility's billing system. However, Radivojević *et al.* (2020) argued that in order to accurately estimate the billed unmetered consumption values, unmetered domestic customers should be identified and monitored their consumption for a period of time to ascertain an accurate estimation.

2.3.2.3 Unbilled Metered Consumption

On unbilled metered consumption, Vermersch *et al.* (2016) explains that the same method used to establish billed metered consumption should be used to calculate unbilled metered consumption. However, as concluded by Thompson *et al.* (2017), unbilled metered consumption in water utilities operating in developing countries like Malawi, such as the LWB, normally comprise also of water consumed but not invoiced to the consumer through meter inaccuracies, meter reading errors, flat rate or fixed water charges which may underestimate consumption and meter tempering that leads to incorrect consumption readings. Therefore, using the same method that is used to establish billed metered consumption in the cases of the water utility company that experience unbilled metered consumption of the nature as explained by Thompson *et al.* (2017) would only result in underestimation of the unbilled metered consumption.

2.3.2.4 Unbilled Unmetered Consumption

Calculating unbilled unmetered consumption includes all volumes of water used by the utility for operational purposes such as flushing the mains and firefighting (Mastaller and Klingel 2017). Notwithstanding the real volumes that are used for such operations, AL-Washali *et al.* (2020) argued that in most cases this component of the WB is overestimated on purpose to cover up real system water losses. However, in estimating water losses as a result of unbilled unmetered consumption, each component must be identified and estimated individually by asking specific questions, for instance (Farley and Liemberger 2005); *Mains flushing*: how many times per month, for how long, how much water? *Firefighting*: Has there been fire, how much water was used? Nonetheless, Knobloch *et al.* (2014) also argued that the estimation of water losses through these activities is highly dependent on accurate measurements of the flow rate at the time the activities are taking place.

2.3.3 Determining commercial losses

While it is almost easy to determine other components of the water balance, many researchers (Farley and Liemberger 2005; Farley *et al.* 2008; Makaya 2015; Vermersch *et al.* 2016; Elkharbotly *et al*. 2022a) agree that it is a hard task to calculate commercial losses. As stipulated by Thompson *et al.* (2017), some of variables in this category include water lost through accounting errors, water theft or illegal connections, meter inaccuracies, customer non-payment, number of staff per 1000 people served within the DMA, DMA population density, corruption effects and economic effects

2.3.4 Unauthorized Consumption

Due to the nature of this component, its estimation must always be done in transparent manner and in a component-based way so that the assumptions can easily be reviewed later on (Knobloch *et al.* 2014). This form of water losses is mainly comprised of water lost through illegal connections. To estimate unauthorized consumption Vermersch *et al.* (2016) stipulated that, the type of illegal connection must be ascertained first, whether commercial or domestic. For commercial illegal connections, upon identification of the case a temporally meter should be installed to monitor consumption for specific hours. Then the monitored consumption can be used to extrapolate for the illegal consumption days. While Vermersch *et al.* (2016), suggested that for illegal domestic consumption, average area consumption can be used to estimate water lost through unauthorized consumption. Onyango (2021) argued that such generalization may lead into underestimation of water lost in cases where the illegally connected domestic property has a household size above the area household size average. As such estimation of unauthorized consumption in such cases should be based on the number of occupants per illegally connected household or property.

2.3.4.1 Customer metering inaccuracies and data handling errors

According to Knobloch *et al.* (2014) and Mastaller and Klingel (2017), the extent of customer meters inaccuracies, which may either be under- or over registration, must be determined based on tests of a representative sample of domestic meters in a distribution system. The drawn sample must also reflect the various brands and age groups of the available domestic meters (AL-Washali *et al.* 2020). When the accuracy tests have been conducted, average meter inaccuracy values (expressed as percentage of metered consumption) will be determined for different user groups. Data handling errors are sometimes a very substantial component of apparent losses (AL-Washali *et al.* 2020; Radivojević *et al.* 2020)

2.3.5 Real Losses

2.3.5.1 Calculating Real Losses

To calculate real losses, Knobloch *et al.* (2014) specified that a utility system must take into consideration the service connections leakage, leakages on distribution mains, transmission mains, overflows and reservoirs. To calculate the real losses therefore, the Apparent Losses should be subtracted from the total NRW. However, van den Berg (2015) argues that it is always possible to have errors in determining the water balance. As such, it is imperative to validate the real loss figures by using the following two methodologies (i) Component Analysis and (ii) Bottom-up real loss assessment.

2.3.5.2 Estimating real loss components

Through a component analysis, real losses can be split into small subcomponents by undertaking the following (Charalambous *et al.* 2014; Vermersch *et al.* 2016).

2.3.5.3 Leakage on transmission and/or distribution mains.

When bursts occur in transmission or distribution mains, it is always a big event which is visible and normally reported in time for repair (Alima 2020). By using data from the bursts scenarios such as the average flow rate at the time of the bursts, the duration before repair and the number of bursts within a specified period of time, leakage on transmission or distribution main component of the real losses can be estimated using the following formula (Al-washali *et al.* 2016);

Number of reported bursts **X** *Average leak flow rate* **X** *Average leak duration (say 2 days)*

Although a provision for background leakages and the undetected leaks can then be added, to make an almost accurate prediction of this form of water losses as stipulated by AL-Washali *et al.* (2020). Charalambous *et al.* (2014) argued that the estimation of water leakages is wholly depended on the time the burst is identified, as well as the response time taken to fix the leak.

2.3.5.4 Leakage and Overflows at Utility's Storage Tanks

Leakages and overflows at a utility's storage are normally known and as such should be collected from the utility (Al-washali *et al.* 2019).

2.3.5.5 Leakage on Service Connections up to Point of Customer Metering

When the leakage on transmission or distribution mains and utility's storage have been subtracted from the total volume of the real losses, the remaining is assumed to be the volume as a result of leakages on service connections (AL-Washali *et al.* 2020). Berardi *et al.* (2016) added that this volume comprises of both the reported and repaired services connection leakages as well as hidden and background losses from service connections.

In addition to this procedure, Sturm and Gasner (2017) outlined the key data that must be collected in order to perform a systems component analysis of the real losses;

- Total length of pipe network and number of service connections;
- Average service connection length between curb-stop and customer meter
- Total number of distribution mains repairs per year (reported and unreported)
- Total number of service connection repairs per year (Reported and unreported)
- Average system pressure across the entire network;
- Estimates of the time periods for Awareness, Location and Repair duration
- Estimates of utility storage tank leaks and overflows.

In their study on drivers of NRW, van den Berg (2015) and Güngör-Demirci *et al.* (2018) explained that although most of this data is readily available in well-organized water utilities, the establishment of the average pressure across the distribution network is often difficult to estimate.

2.3.6. Bottom-up Real Loss Assessment

2.3.6.1 24 Hour Zone Measurements

When a distribution system has no established DMA, areas of the utility's distribution network must be selected which can be provisionally isolated and supplied from a single or two inflow points only (Al-washali, et al. 2016). A representative sample of the system must be chosen by selecting areas situated in different parts of the distribution network. Farley *et al.* (2010); van den Berg (2015) also added that in these areas, both pressure and inflow must be measured using portable measuring data loggers for a period of 24hours at zonal the inlet points. Furthermore, data on important characteristics such as length of the mains, number of service connections, number of household properties and number and type of non-household properties must be collected to aid in the calculations and interpretation of water losses (van den Berg 2015).

2.3.6.2 Night Flow Analysis.

According to Serafeim *et al.* (2022) the minimum night flow (MNF) methodology, which seeks to analyse water leakage at the lowest customer demand period, is ranked highly in monitoring leakages within a DMA. The analysis is done when the customer demand is at its minimum and therefore the leakage component is at its largest percentage of the flow (Liemberger and Wyatt 2019). According to van den Berg (2015), the MNF in urban areas normally occurs during the early morning period, usually between around 02:00 and 04:00 hours. The estimation of the real loss component at minimum night flow is carried out by subtracting an assessed amount of legitimate night consumption for each of the customers connected to the mains in the zones being studied (Serafeim *et al.* 2022). Furthermore, Serafeim *et al.* (2022), explained that when the legitimate night consumptions are deducted from the night flow consumptions the result obtained consists principally of real losses from the distribution network. After this deduction, the daily level of real losses obtained from the minimum night flow analysis are established by applying the Fixed and Variable Area Discharge path (FAVAD) principles and simulating leakage over the full 24h period.

While there exist both top-down and bottom-up approaches to assessing real losses, many researchers (Cheung *et al.* 2010; AL-Washali *et al.* 2018; Radivojević *et al.* 2020; Serafeim *et al.* 2022) argued that combining two or more approaches produce reliable results on real losses than simply adopting one and leaving the other

2.4 Leakage Theory

2.4.1 Burst and Background Leakage

Leaks in a water distribution system are normally categorized into those that are large enough to warrant serious attention and those that are too small for such attention (AL-Washali *et al.* 2018). Although there are no rigid thresholds for leaks between countries, Charalambous *et al.* (2014) argued that the South African threshold of $0.25m^3/h$ could apply for most southern African countries like Malawi. Background leaks, which mainly occur at joints and fittings, and run continuously, but do not generate sufficient noise to be sensed by existing equipment, form one of the largest components of real losses (AL-Washali *et al.* 2020). On the other hand, burst leakages, which are of high flow rates, but short duration, contribute a moderate volume to the real losses (Serafeim *et al.* 2022). These burst leakages, result from annually occurring holes or fractures in the distribution network pipe work, which include customer service connections.

2.4.2 Leakage Control

According to Sturm and Gasner (2017), leakages are best addressed by breaking them down into: transmission mains and reservoir leakage, reticulation mains leakage, connection leakage, and service pipe leakage

McKenzie *et al.* (2003) outlined leak frequencies and average leakage rates as follows:

Table 2:Reported and unreported bursts (source: McKenzie et al, 2003)

Details	Reported Bursts		Unreported Bursts	
	Frequency	Leakage rate (m^3/h)	Frequency	Leakage rate (m^3/h)
Transmission mains	0.030/km/yr	30.0	0.00/km/yr	12.0
Reticulation mains	0.150/km/yr	12.0	0.008/km/yr	6.0
Connections	2.5/1000conn/yr	1.6	0.825/1000conn/yr	1.6
Service Pipes	2.5/1000conn/yr	1.6	0.825/1000conn/yr	1.6

Leakage control can be classified into two groups, namely passive and active leakage control.

2.4.2.1 Passive Leakage Control

This refers to reacting to reported bursts or a drop in pressure, by customers or noted by the utility's own staff while carrying out other duties than leak detection (Sturm and Gasner 2017). AL-Washali *et al.* (2020) claimed that most developing countries with low levels of technology to detect leakages rely on this methodology in leakage control. It is very likely therefore that under this leakage control measure, leakages continue to rise with little control.

2.4.2.2 Active Leakage Control

Based on the definition of Berardi *et al.* (2016), active leakage control refers to the water utility company's program that proactively pursues to discover leaks that are not reported by customers or staff members doing other activities than leak detection. Active leakage control activities usually result from a blend of flow/pressure monitoring to detect anomalies and field survey for burst identification and repair (Charalambous *et al.* 2014; Knobloch *et al.* 2014; Radivojević *et al.* 2020). The monitoring-based strategies generally involve bottom-up approaches based on water balance or through the MNF analysis per DMA. Although this method very effective in controlling leakages, Makaya (2015) explains that most water utilities in developing countries are engaged in passive leakage control, with low activities in mending only visible leaks.

2.5 Leakage Detection Computation Approaches.

2.5.1 Network Hydraulic Modelling in Leakage Assessment

As information and technology advances, many water utilities are adapting to the changes in dealing with NRW (Serafeim *et al.* 2022). Among the widely used water leakage assessment tools, Network Hydraulic Modelling (NHM) is finding wide application worldwide (Berardi *et al.* 2016). This method relies on the computational strength of computers for predicting and investigating the operational functionalities of Water Distribution Systems.

EPANET 2 is another extensively used network hydraulic modelling (NHM) program which is found in public domain(Abdelbaki *et al.* 2017). Specifically, this software is commonly used for pressure management, network zoning and decision making about pipeline replacement in leakage management. In its early years, EPANET was used to correlate water supply with network pressures empirically (Tabesh *et al.* 2014). The following mathematic formulae were used for its network analysis to establish the pressure-consumption relationship;

$$C_i = C_k^i \, a_i e^{-b_t \, p_{i/k^i}} \qquad (1)$$

where: P_i = pressure at node i.

C_i = the consumer outflow at node.

i; = the nominal consumer demand.

k, e, a_i, b_i, = constants for the particular node.

From this mathematical formulation, an extension was made in order to factor in pipe size and othe connection parameters as shown below (Tabesh *et al.*, 2014);

$$V_{ij} = C_i \left(L_{ij} D_{ij}^{d} e^{a\tau} p_{ij}^{av} \right)^{N} \qquad (2)$$

where D and τ are pipe diameter and age respectively;

d is 1 for (D < 125 mm) and is 1 for (D >125 mm);

and a is a leakage shape parameter.

$\boldsymbol{V_{ij}}$= leakage flow rate from the pipe connecting nodes i

***j*; C**$_l$= a constant depending on the network;

$\boldsymbol{L_{ij}}$= pipe length

$\boldsymbol{p_{ij}^{av}}$ = average pressure along the pipe

N1 = the pressure exponent

As noted by Makaya (2015) the only problem with this methodology is the required data. The field measurements required to determine parametric values of a_i,b_i,P_i^k for every node is the major weakness of the method. Because of these challenges many utilities in developing countries cannot afford the cost of experimental procedures (Cheung *et al.* 2010; Abdelbaki *et al.* 2017). Makaya and Hensel (2015), also argued that the method is disadvantaged in a way that leakage flow is assumed to be uniformly distributed along the pipeline. This, in reality is not the case because of the differences in types of pipeline materials and positioning of joints and fittings. The other problems with NHMs, as noted by various researchers, include the lack of ability to estimate parametric values, lack of ability to model and predict complex variations in water flow within the system network, lack of flexibility and adaptability when new data is available and many others (Makaya 2015; Radivojević *et al.* 2020; Kizilöz and Şişman 2021; Elkharbotly *et al.* 2022b). Therefore, with these challenges, looking for other nework modelling techniques such as the use of Artificial Neural Networks seems to be ideal in an effort to reduce NRW (Makaya 2015; Elkharbotly *et al.* 2022).

2.5.2 Artificial neural network algorithms

With the increased development of Artificial Intelligence, water utilities are now able to use ANN algorithms operating on quasi-static pressure and flow readings to detect leakage in pipe systems (Makaya *et al* 2015; Elkharbotly *et al.* 2022). Unlike hydraulic modelling software, ANN models are capable of learning complex relationships between input and output variables without requiring explicit equations or assumptions about the underlying system. This means that they are able to learn from existing data without needing extensive field measurements for every node, as required in hydraulic models like EPANET (Elkharbotly *et al.* 2022). Radivojević *et al.* (2020) also added that, Instead of relying solely on field measurements for determining parametric values, ANNs can be trained on available data to estimate these values. The ANN models are able to learn from historical data, sensor readings, or other available sources to infer the parametric values for nodes within the network, even in the absence of direct measurements. In addition, the assumption of uniformly distributed leakage flow along the pipeline, as noted in the EPANET model, is addressed by ANN in which the model is trained to capture variations and complexities in the system, leading to more accurate predictions and modeling of fluid flow behavior (Abdulla *et al.* 2013).

The discipline of ANN arose from the thought of mimicking the functionality of the human brain(Ciaburro and Venkateswaran 2017). Since their introduction in water loss management, ANNs have found wide application in not only leakage detection, but also water distribution network optimisation, water pipeline replacement and rehabilitation, water demand forecasting, and pressure monitoring (Abdulla *et al.* 2013).

A Neural network model takes the structure of three levels or layers; input layer, hidden layer and the output layer (Maozhun and Ji 2017). The input layer feeds the model with input variables which are processed within the hidden layer to produce model output in the output layer. The intermediate layers (hidden layers) of an ANN perform the data processing functions of the network, where weights to the neurons are adjusted by training the network based on the model's learning rule (Ciaburro and Venkateswaran 2017). When inputs are supplied to the network, the neuron transfer function plays the role of transforming the input to output for each neuron(Elkharbotly *et al.* 2022). For the data to be processed, however, the log-sigmoidal transfer function is commonly used before producing the final output; especially with the back propagation algorithm (Ciaburro and Venkateswaran 2017). Back propagation algorithms are based on multi-layered feed forward topology with supervised learning.

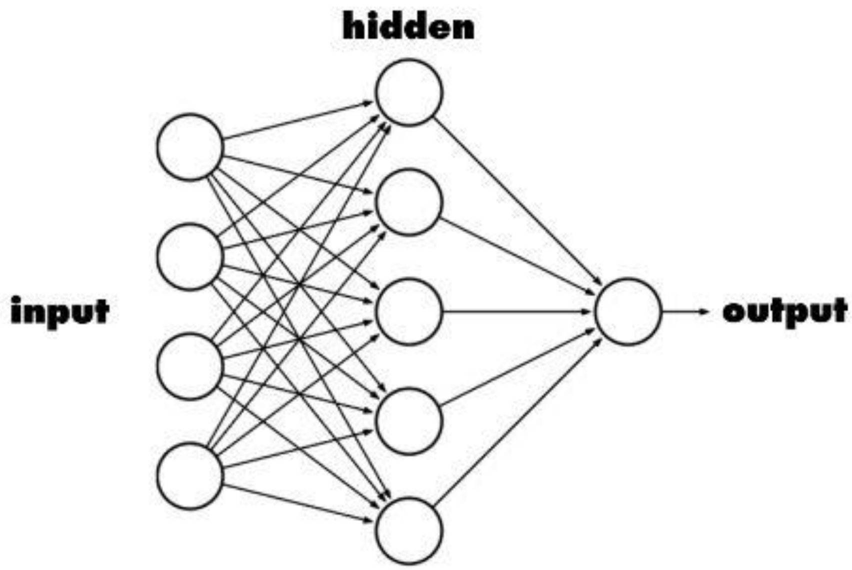


Figure 1: Atypical example of a Neural network with input, hidden and output layers (Source: Ciaburro and Venkateswaran, 2017).

While it is possible to have an ANN with single hidden layer as shown in Figure 1, the common ANNs have multiple layers. Such ANNs with multiple hidden layers are called Multi-Layer Perceptron Neural Network (MLP) (Elkharbotly *et al.* 2022). MLPs have input layers consisting of nodes that merely accept input variables such as all drivers of NRW and the outputs of neurons in a layer become the inputs to the neurons in the next layer. With the exception of the input neurons, a two steps retransformation of inputs to outputs takes place in the intermediate layers (Ciaburro and Venkateswaran 2017). The transformation is in such a way that each neuron receives a summation of weighted activations from all neurons in the preceding layer of inputs to the network. A constant term, called neuron threshold value is added to the summation to yield the net input, Y_{net}, to the neuron;

$$Y_{net}=\sum_{i=1}^{N} Y_i W_i + W_0 \qquad (3)$$

where: **N** is the total number of neurons in the preceding layer or input array.

Y_i is the neuron input received from the i^{th} neuron in the preceding layer or input array.

W_i is the connection weight or strength of the neuron to an i^{th} neuron in the preceding layer or input array.

W_o is the neuron bias/threshold value.

The second transformation involves the change of input Y_{net} into output Y_{out}

$$Y_{out} = f(Y_{net}) = f(\sum_{i=1}^{N} Y_i W_i + W_0) \quad (4)$$

In order to obtain an optimum network structure, Howard and Mark (2004) recommend the design and comparisons of different network architectures from the same data. This is because there is no rule of thumb in choosing the network structure.

In conclusion, Sonali and Maind (2014) laid out the following benefits of using ANN over the traditional models used in water loss management;

- Adaptive learning: An ability to learn how to do tasks based on the data given for training or initial experience.
- Self-Organization: An ANN can create its own organization or representation of the information it receives during learning time.
- Real Time Operation: ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.

- Pattern recognition is a powerful technique for harnessing the information in the data and generalizing about it. Neural nets learn to recognize the patterns which exist in the data set.

- The system is developed through learning rather than programming. Neural nets teach themselves the patterns in the data.
- Neural networks are flexible in a changing environment. Although neural networks may take some time to learn a sudden drastic change, they are excellent at adapting to constantly changing information.
- Neural networks can build informative models whenever conventional approaches fail. Because neural networks can handle very complex interactions, they can easily model data which is too difficult to model with traditional approaches such as inferential statistics or programming logic.
- Performance of neural networks is at least as good as classical statistical modelling, and better on most problems.

2.6 Conceptual Framework

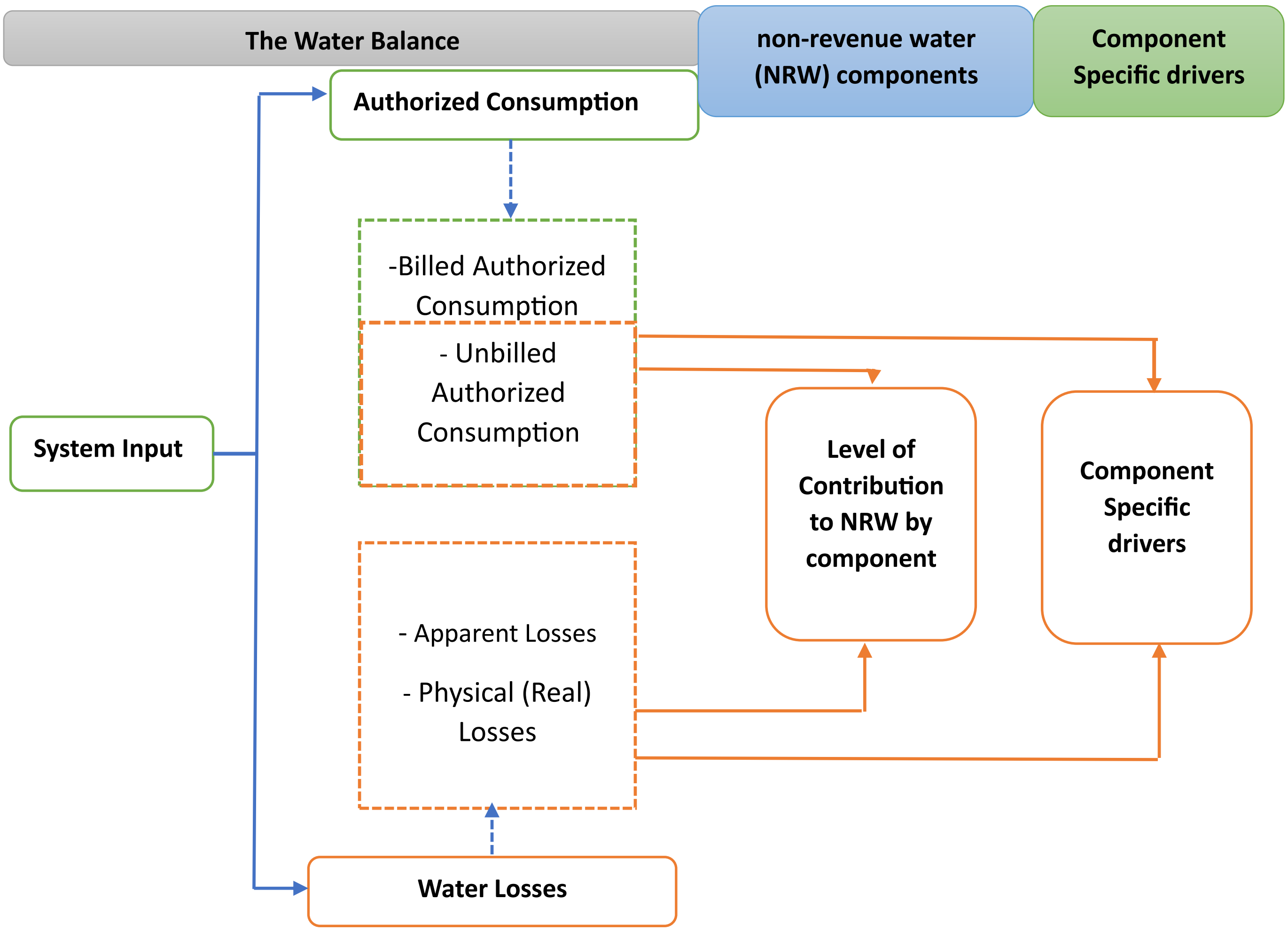


Figure 2 :Conceptual Framework

As presented in Figure 2, the conceptual framework that guided this research consisted of three main components: WB, NRW components, and NRW component specific drivers. The water balance component helped to guide the evaluation of the inflow and outflow of water in selected DMAS of LWB as discussed in the literature review. It took into consideration, the measured system input volume of water in m^3, then tracked down its usage within the water distribution system to specify the volume of water which constituted water losses. The outlined NRW components determined the level of contribution for each NRW component on total NRW, while the NRW component specific drivers investigated the main drivers for each NRW component for LWB to achieve the main objective of the study. By combining these three components, the conceptual framework provided a structured approach to comprehending the complexities of NRW within LWB's water distribution system. It served as a comprehensive guide for research methodologies and specific analyses conducted to achieve the study's objectives. Additionally, the framework drew upon the insights gathered from the literature review, ensuring a well-founded and theoretically grounded approach to addressing the research questions and enhancing the overall understanding of NRW management within the context of LWB

2.7 Conclusion

In this chapter, relevant literature has been discussed in line with water distribution losses, in both developed as well as developing countries like Malawi. To understand different aspects of the water losess in a Water Distribution System (WDS), the WB has also been discussed fully using extant literature. The discussion on WB particularly delved much into real loss component which constitutes a large portion to the overall NRW in many utilities. Different methodologies for calculating RLs have been explored to show their applicability to this study. The discussion in the chapter has also established the components that constitute leakages in a WDS. While evaluating relevant literature, it has also been revealed that most researches aimed at reducing NRW have mainly been done on the total loss volume, without necessarily investigating each NRW component to establish their specific drivers. The review also showed that machine learning algorithms like the ANN have a greater potential in NRW reduction applications. Different leakage detection computation approaches have also been reviewed in which both traditional and ANNs have been discussed to also display their applicability to this research. Finally, a conceptual framework detailing the study’s major conceptual guidelines, has been presented following the reveiewed concepts and methodologies.

CHAPTER THREE: MATERIALS AND METHODS

3.1 Introduction

In this chapter, the study area, materials, methods and the research design are discussed in detail. It discusses the main procedures and materials which were used to achieve the main objective.

3.2 Study Area

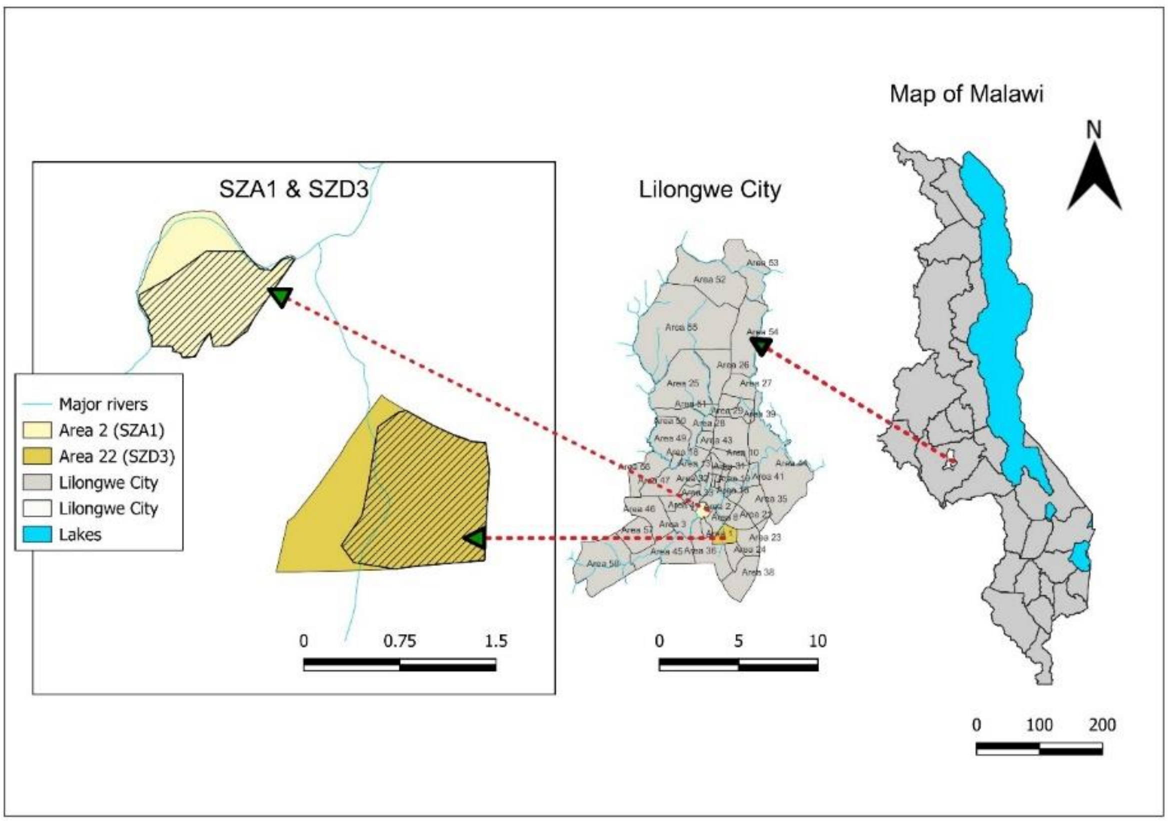


Figure 3: Map of Lilongwe City

This study used a case of the LWB's selected DMAs (SZA1 and SZD3) in Lilongwe City, Malawi. The LWB currently supplies water to an area of about 45,000 hectares which is divided into three water supply zones within Lilongwe City urban namely; southern, central and northern zones LWB 2018). Lilongwe City is located in central Malawi, at an altitude of 1,050 m (3,440 ft) above sea level (Clyde *et al*. 2022). The city's urban area houses population of more than 981,331 with an urban population growth rate of 3.8% (GoM 2018). Figure 3 shows the map of Lilongwe where the study was conducted;

3.3 Study Design

For this study, quantitative research design was followed. Specifically, the research followed diagnostic and correlational research designs. The diagnostic design was chosen due to its ability to investigate the underlying cause of a certain condition or phenomenon design (Boru 2018). It assisted in learning more about the elements that contribute to certain difficulties or challenges that were being investigated. The correlational research design, on the other hand, was incorporated in order to reveal the magnitude and/or direction of a link between variables under investigation (Robson and McCartan 2016).

While using the diagnostic and correlational research designs, both cross-sectional data and longitudinal data were used. As defined by Robson and McCartan (2016), cross-sectional data refers to the data from a representative sample collected at a specific time such as pipe material, pipe length, pipe age, consumer meter age, elevation, DMA population density and number of utility staff per 1,000 people served. While longitudinal data, which represents data collected through a repeated observation about the same variables over a short or long period of time, included the system input volumes, day and night hours pressure, minimum night flows, daily water consumption within the selected DMAs, daily pipe burst records and many others. This data was collected through key informant interviews (cross-sectional data and all the secondary data) and direct field measurements.

3.4 Sampling Methods

Two DMAs within the LWB namely SZA1 (Area 2) and SZD3 (Area 23) were purposively sampled due to their easily traceable network infrastructure. From these DMAs both primary and secondary data (before network rehabilitation) were collected. Just as with DMAs' selection, key informants were purposively sampled to provide cross-sectional as well as secondary data. Specifically, 1 NRW Manager, 1 Assistant NRW Manager and 7 meter readers from each DMA (16 in total), were selected based on their expert knowledge of the sampled DMAs. With the help from these informants, an in-depth system reconnaissance was done to collect data through reading of master meters, zonal meters and consumer meters (or its proxy; customer consumption bills data). Through their involvement, checklists and direct measurements were used to collect data on the distribution networks physical factors such as connection density, population per connection, length of distribution mains, operating pressure, flow rates, number of functional meters, age and the type of distribution pipes and many others which could as well have an impact on certain NRW components. The methodological approach is as shown in Figure 4;

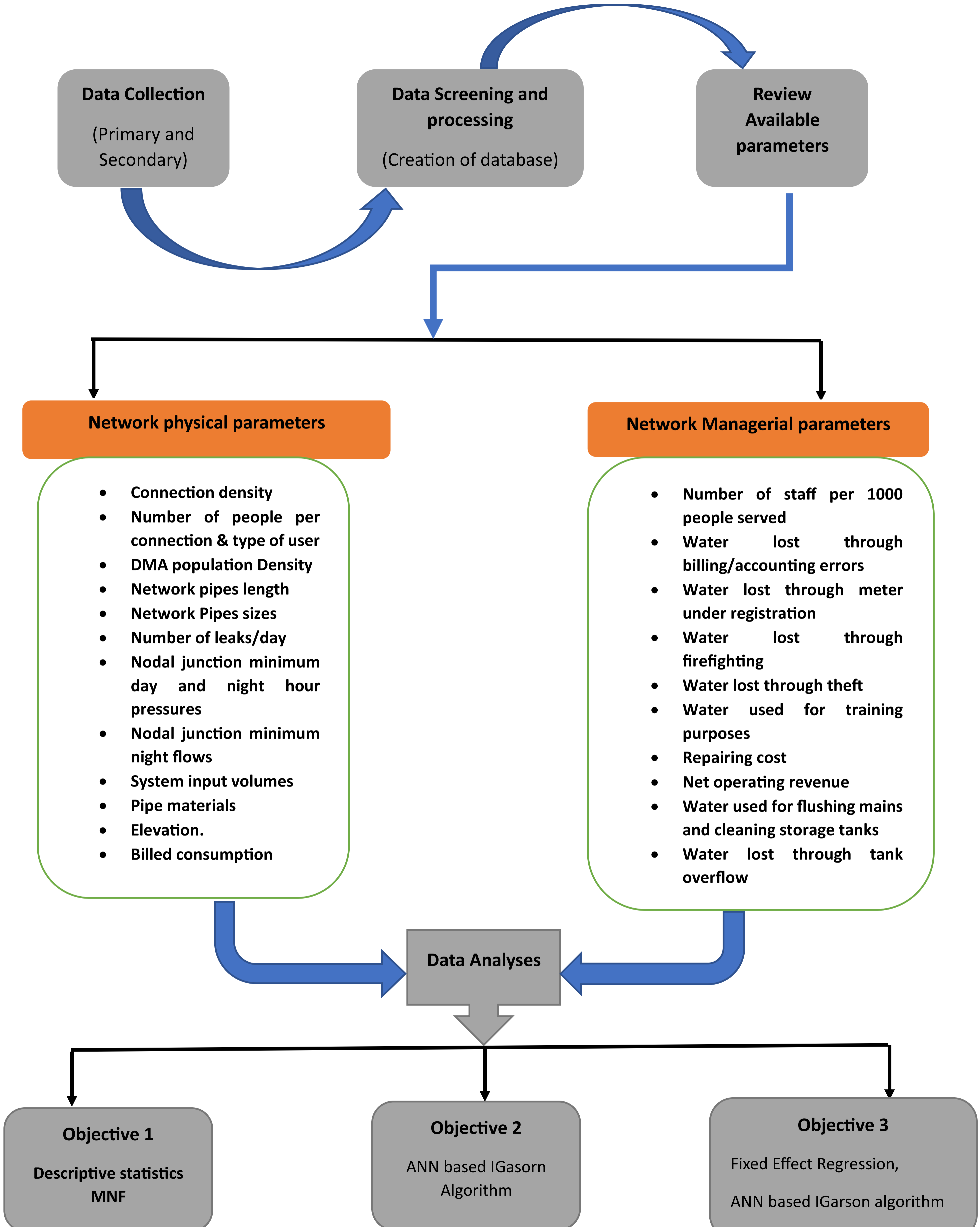


Figure 4: Diagrammatic Methodology

3.4 Data Collection

To successfully collect data for each objective of the research, various data collection methods were employed as follows;

3.4.1 Water Balance

To analyse the water balance, the researcher collected data on the total system input volume through DMA bulk meters at entry points, total billed water for the selected DMAs and longitudinal data at DMAs' nodal junctions such as net night flow, MNF, night hours' pressure, day hours' pressure and many others. The longitudinal survey was done through field measurements using flow and pressure loggers and customer consumption bills data from meter readers. Data on other managerial parameters of the distribution system such as records of water theft, estimated system underbilling or water lost through accounting errors, water lost through firefighting and flushing mains, was collected using a checklist from purposively sampled key informants. Key informants were purposively sampled to select only employees knowledgeable in NRW management. The collected data were, thereafter, used to analyse the water balance for the two selected DMAs.

3.4.2 Level of contribution for each NRW component on total NRW.

The longitudinal NRW components data, which included data on Als, Unbilled Authorized Consumption and RLs, collected sequentially from the same sampled parameters as outlined in section 3.4.1, was obtained through the distribution system's in-depth examination as well as through exploratory study of secondary data. Secondary data on the outlined target parameters for a month period before network rehabilitation was obtained from the zone managers using a checklist. Data collected for a month period on the actual daily readings of bulk meters, nodal junction meters, consumer meters, records of authorized unbilled metered consumption (such as water unpaid for by institutions and water unpaid for by individuals with disconnected service as a result of long overdue bills) and authorized unbilled unmetered water usage records (such records of water used for firefighting and flushing lines) were used to generate the three NRW components. Secondary and primary data collected were used for the comparison of the before and after system rehabilitation situation of the NRW components. This data was then used to model and test the presence of influence or contribution of each NRW component on the total NRW.

3.4.3 Investigating the main drivers for each NRW component.

To investigate the main drivers for each NRW component, both longitudinal and cross-sectional data (data collected by observing many parameters at one point or period) was collected. Data for the distribution system's physical characteristics such as connection density, number of people per connection & type of user, DMA population density, network pipes length, network pipe sizes, number of leaks/day, nodal junction minimum day and night hour pressures, nodal junction MNFs, system input volumes, pipe materials, elevation and billed consumption was collected. Similarly, the distribution system's managerial characteristics such as number of staff per 1000 people served, water lost through billing/accounting errors, water lost through meter under registration, water lost through firefighting, water lost through theft, water used for training purposes, repairing cost, net operating revenue and water used for flushing mains was collected. To investigate Real Losses drivers, network's physical parameters as outlined in Figure 4, were used. Whereas, to investigate for drivers of Als, network's managerial characteristics such as water lost through accounting errors, water theft or illegal connections, meter inaccuracies, customer non-payment and other factors such as number of staff per 1000 people served within the DMA, DMA population density, corruption effects and economic effects, were investigated. Managerial characteristics of the network such as water losses during maintenance works, water used in utility's offices, cleaning mains, institutional policy effects and system maintenance frequency, were used to investigate the main drivers of UACs. Data for both physical and managerial characteristics of the distribution system was collected as stipulated in section 3.4.1.

3.5 Data analysis

Data obtained for this study was analyzed using R studio statistical package version 4.1.1. The analysis took a form of descriptive statistics, correlational analysis as well as modeling using ANN and Fixed Effects Regression. Appropriate formulae adopted from literature on similar studies were used as follows;

3.5.1 Analysing the Water Balance

Before analysing the water balance, data on daily system input values, billed consumption, daily supply hours, minimum day and night flow hours' pressure, as well as NRW was subjected to descriptive statistical analysis in order to ascertain their measure of central tendency and dispersion. Relationships between the other mentioned variables and the total NRW were tested using Pearson Correlation test.

To analyse the water balance for LWB's selected DMAs, the total NRW for a month was calculated using the following formula;

$$NRW = SIV - BC - \text{Uncollected Billed Consumption} \quad (5)$$

Where SIV is the System Input Volume (m^3)

BC is Billed Consumption (m^3)

Uncollected Billed Consumption includes all billed water not paid for (m^3)

Billed water consumption was estimated from the total monthly water bills for the selected DMAs. To analyse the RL, the MNF analysis was used, while the AUC and AL were calculated from the balance of the differences between NRW and RL (AL-Washalia *et al*. 2020). The leaks during the MNF hour were estimated after subtracting the possible legitimate night consumption in the DMA as follows;

$$Q_{NNF}=Q_{MNF} - Q_{LNF} \tag{6}$$

Where Q_{NNF} is the net night flow (m^3/h),

Q_{MNF} is the minimum night flow (m^3/h),

Q_{LNF} is the legitimate night flow (m^3/h)

Nonetheless, the MNF represented by this equation is generally higher than the leakage rate during the other hours of the day, mainly due to the pressure–leakage relationship (Lambert 2019). To obtain the rate of the real losses during the day, the leakage rate was adjusted using a pressure correction called night–day factor (NDF), which was calculated using the following equation;

$$NDF = \sum_{i=1}^{23} \left(\frac{Pi}{Pmin}\right)^{N_1} \tag{7}$$

NDF is the night–day factor,

P_{min} is the average pressure during the minimum night hour,

P_i is the average pressure during the day hours,

N_1 is the leakage exponent that can be assumed to be 1

The daily rate of the RL in the DMA (m3/h) therefore becomes; $Q_{RL} = Q_{NNF}$ X NDF.

Having calculated the RL from the total NRW rate, the difference between NRW and RL gave the sum of AUC and AL. To separate the two (AUC and AL), the AL were calculated using data for unauthorized consumption, water lost through customer metering errors, water lost through errors in estimating unmetered consumption and errors throughout the data acquisition process. In the absence of unbilled authorized consumption therefore, the difference between NRW and RL represented the AL. Alternatively, UAC was calculated by finding the product of flow rate and the number of hours the consumption took place, and also adding water lost through other unbilled authorized consumptions (Hoko and Chipwaila 2017). Through this process therefore, the water balance for the selected LWB DMAs was analysed.

3.5.2 Establishing the level of contribution for each NRW component on total NRW.

To analyse this objective, descriptive statistical analysis for daily UAC, AL, RL and NRW was done. A pairwise Pearson Correlational test was done to establish the relationship between the three NRW components and the total NRW. While to establish the significant of differences brought by system rehabilitation on UAC, AL, RL and NRW, a paired samples t-test was performed on primary and secondary data characteristics for NRW components. The paired samples t-test was used to determine the presence of significant differences between means of three or more independent groups (Kuntner *et at.*, 1997).

The three NRW component values were then subjected to the sensitivity analysis as input variables, where total NRW value served as the dependent variable. The sensitivity analysis adopted for this study utilized the Improved Garson Algorithm (Igarson) based on feedforward Back Propagation Artificial Neural Network (BP ANN). In order to perform an effective sensitivity analysis, the quality of the built ANN model was measured by the Mean Square Error (MSE) term of the following mathematical expression;

$$MSE = \frac{\sum(y_i + \hat{y}_i)^2}{n} \tag{8}$$

Where MSE is the Mean Square Error, y_i is the observed value, $\hat{y}_i$ is the predicted value and n is the number of observations.

An ANN is a pattern recognition Machine Learning method which takes its inspiration from a human nervous system in the examination, interpretation and information exchange (Elkharbotly *et al.* 2022). The success of an ANN in modeling mainly lies in its ability to successfully define, understand and predict different issues in different disciplines (Momeni et al. 2015). In the studies of systems' efficiency, ANN algorithms analyze the influence degree of the system parameter on the output of the systems' model. One of the many algorithms widely adopted in this sort of sensitivity analysis is the Garson Algorithm which was developed by Garson in the 1990s, primarily to explain scientifically the contribution of each input factor of the neural network model according to its output (Garson 1991). However, in their study, Maozhun and Ji (2017) exposed the instability of this algorithm and its lack of precision sometimes in determining sensitivity coefficients. To correct this error, Maozhun and Ji (2017), introduced the system input attribute value to the Garson output proportional allocation algorithm as explained in the following sections. The resultant of this modification, therefore, is referred to as the Improved Garson Algorithm.

3.5.2.1 Sensitivity analysis

Sensitivity analysis is mainly used to find out the rank of input variables according to the influence on output. It mainly considers the change of a certain factor of the system models within its reasonable range, and analyzes the influence degree of the attribute on the output of the model. The influence degree of the system's input variable on the output is referred to as sensitivity coefficient (Maozhun & Ji 2017). The greater the sensitivity coefficient, the greater the influence or contribution of the input variable on the output of the system. In this study therefore, a component with much influence on the total NRW was ranked as priority for tackling issues to do with NRW reduction while also targeting the major drivers of such a component (which was established in the third objective).

The mathematical description of the sensitivity analysis is as follows; when the input factor x (daily NRW component) of the system changes slightly Δx, the corresponding y (daily total NRW) Taylor expansion follows:

$$Y(x + \Delta \mathrm{x}) = \mathrm{y(x)} + \sum_{i=1}^{n} \partial y / \partial x_1 \, \Delta x_1 + 1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} (\partial^2 \, y / \partial x_1 \partial x_j) \Delta x_1 \Delta x_j + \cdots \tag{9}$$

In this mathematical description of the sensitivity analysis, the first and second-order sensitivity coefficients are presented by the first and second-order partial differential. Practically, the first sensitivity coefficient is able to reflect the system changes and can be used as sensitivity analysis index (Ojeda et al., 2014). In this case, the sensitivity analysis expression takes the following form;

$$S^p(i) = \frac{\partial}{\partial x_i} y(x_1^p, x_2^p, \dots x_n^p) \tag{10}$$

Where $S^p(i)$ is the sensitivity coefficient of *i-th* input factor.

P is the sample of x_i

n is the number of *x*

3.5.2.2 Basic Garson Algorithm

The Garson Algorithm which was introduced by Garson and his colleagues in the 1990s, is a sensitivity analysis method which is based on connection weights of neural network (Maozhun and Ji, 2017). The structure of the neural network utilizes the input variables X_i (i=1, 2,…,n) to produce Y_k(k= 1, 2,…,n) as the network output. To convert input variables into outputs, the network utilized connection weights between the input layer and the hidden layer neurons, and also between the hidden layer and the output layers. To rank variables based on their influence importance on the output, the basic Garson Algorithm calculated the importance (significance) coefficient based on product of the connection weights among the input layer neurons, hidden layer neurons and output layer neurons of the neural network. Thus, the *i-th* input factor of *k-th* output sensitivity coefficient was calculated using the following formula;

$$S_k^p(i) = \frac{\sum_{j=1}^{n}(W_{ij}V_{jk}/\sum_{i=1}^{n} W_{ij})}{\sum_{i=1}^{n}\sum_{j=1}^{n}(W_{ij}V_{jk}/\sum_{i=1}^{n} W_{ij})} \tag{11}$$

Where, *i, j, k,* refers to input layer, hidden layer and output layer neurons respectively;

wij is the connection weights between input layer and hidden layer neuron,

vjk is the connection weights between hidden layer and output layer neurons,

n is the total number of input layer neurons.

3.5.2.3 Improved Garson Algorithm

Although the Garson method could calculate the sensitivity coefficient easily based on the connection weights of the neural network, the simplification of this algorithm reduced the accuracy and stability of sensitivity analysis. To improve this, Maozhun & Ji (2017) proposed the introduction of the system input attribute value x_i (i = 1, …, n) into Garson's output proportional allocation algorithm. This helps in calculating the sensitivity coefficient of the *k-th* output with respect to the *i-th* input sensitivity coefficient $S_k^p(i)$ in the Garson's sensitivity coefficient equation. The introduction of this system input attribute helps to correct the weakness in the Garson Algorithm which equates the gravity of the *i-th* input to the contribution of the inner function f (•) independent variable caused by the *i-th* input.

The improved Garson algorithmic equation used in this study therefore is as follows;

$$S_k^p(i)\frac{x_1 w_{i1} v_{1k}/\sum_{i=1}^{n}\ x_1 w_{ij} + \cdots + x_1 w_{ij} v_{jk}/\sum_{i=1}^{n}\ x_1 w_{ij}}{\sum_{i=1}^{n}(x_1 w_{i1} v_{jk}/\sum_{i=1}^{n}\ x_1 w_{ij} + \cdots + x_1 w_{ij} v_{jk}/\sum_{i=1}^{n}\ x_1 w_{ij})} \tag{12}$$

Where p (p =1,2, …, P) represents the sensitivity coefficient calculated with the input *p-th* sample value, and the other variables are defined as is in the basic Garson Algorithm

Through the use of this method, the research was able to establish the level of influence for each NRW component on total NRW by ranking them based on the size of their sensitivity coefficients. The bigger the sensitivity coefficient, the higher the influence the NRW component has on the total NRW.

3.5.3 Investigating the main drivers for each NRW component for LWB

As proposed in this research, the NRW drivers were investigated at component level. This made sure that the main drivers for each NRW component are established. To achieve this, the fixed effects regression model was used to examine the AUC and AL drivers, while drivers for the RL losses were investigated using the Igarson sensitivity analysis. The sensitivity analysis as well as model quality measurement for this component utilized the same formulae as outlined in the previous section (section 3.5.2).

However, before delving into specific modelling, data for all NRW components was descriptively analyzed and visualized graphically to establish their measure of central tendency. Relationships for the data used in both fixed effect regression model and the sensitivity analysis were also established using the Pearson correlation test. Similar to the model under section 3.5.2, the quality of both models for this objective was assessed by the MSE. Data for the RL factors (Network's physical parameters) was subjected to sensitivity analysis to examine the influencing factors of water losses in physical distribution system. The physical part of the distribution system formed a critical portion of the whole system. Therefore, an in-depth linear and non-linear analysis of the physical drivers which influence RL was ideal.

4.5.3.1 Fixed Effects Regression Model

To investigate the main drivers for AUC and Als, the water loss function was specified as;

$$L = f(M, O) + e + u \tag{13}$$

Where L= Water losses, M=managerial characteristics, O=Other factors that may affect AUC

The formular above, represented the relationship between AUC and/or AL as measured by the water loss (L) on one hand and a vector of management characteristics of LWB (M) and a vector of other factors (O). Added to this equation was the utility effect e and an error term (u). Each utility has its own characteristics that may influence water losses. In this study, the utility effect included; institutional, legal and regulatory frameworks; levels of corruption and many others which were discovered within the course of the study.

The estimation of the single loss reduction function ensured an insight on variables that drive water losses in form of AL or AUC. For the coefficient of this water loss function to measure the elasticity of water losses, this study specified water losses and explanatory variables in scaled forms using the Min-Max and Z-Score techniques. This helped to investigate changes in water losses relative to changes in the drivers. This model used the AUC and/or AL component(s) of NRW per DMA per day as a dependent variable, while the independent variables which were examined included; household meter accuracy (the average accuracy levels determined the amount of water that goes unmetered per day), number of staff per 1000 people served, number of illegal connections identified (which was expressed as the amount of water unauthorizedly consumed per day depending on average household size of the area), hours of system maintenance (for instance, flushing) and number of households and institutions disconnected from the system due to outstanding bills (which determined the amount of water lost due to nonpayment)

4.5.3.2 Sensitivity analysis

To investigate the RL component drivers, the distribution networks' physical parameters were considered. Although some physical parameters such as operating pressure and age of the system have a linear relationship with RL, the interaction of these parameters and other parameters such as connection density, topography of the distribution networks' area and many others, disturb their linearity. In order to effectively investigate this NRW component, therefore, an ANN based sensitivity analysis was chosen over the single loss reduction function due to its (ANN's) ability to handle complex relationships. The sensitivity analysis used the Igarson Algorithm based on BP ANN as described in the second objective's methodology, with the dependent variable being RL component of NRW. The following were the independent variables under investigation which formed the input layer of the ANN; daily average operating pressure, age of the pipes, connection density (number of connections per kilometer of a network), population per connection, population density (population served per kilometer of network), network length, topography of the distribution networks' area, system input water volume per day and water consumption per day and many others as presented in chapter 4. By ranking these factors based on the weight of their sensitivity coefficients, the main drivers of NRW's RL component for LWB were established.

3.5.4 Reliability and Validity of the Data

To ensure that the collected data was valid and reliable, knowledgeable research assistants were used. All secondary data used in this study were checked for validity and reliability. Whereas, to ensure that primary data was both valid and reliable, all instruments used in measurements were tested for their fitness and reliability for the required work. All instructions pertaining to specific measurements were also followed.

3.5.5 Ethical Consideration

Ethical clearance for this study was obtained from the Department of Water and Sanitation and the Mzuzu University Research Ethics Committee (MZUNIREC; protocol reference number MZUNIREC/DOR/23/19). The study also sought permission from the Lilongwe Water Board. All material references have been acknowledged. All utility's management data collected have been treated with the highest level of confidentiality.

3.6 Conclusion

In this chapter, materials and methods which were used in the study have been discussed in details. Data reliability and validity criteria which were followed to make sure that conclusions drawn from the research are valid have also been explained. Finally, all ethical issues related to this study were also followed as detailed in this chapter.

CHAPTER FOUR: RESULTS

4.1 Introduction

This chapter, which is divided mainly into three sections, presents the results of the study which was conducted in the two selected DMAs. The first section presents the results of the water balance. While the second and third sections, present the analysis results of the NRW components' contribution to total NRW, as well as the components' specific drivers' investigation.

4.2 Water balance analysis

System input volumes, billed consumption, NRW and other distribution network characteristics were obtained for the two selected DMAs; Area 2 (SZA1) and Area 22 (SZD3) for water balance analysis.

4.2.1 Water Distribution Network's Characteristics

Water distribution network characteristics play a major role in how water in a distribution system is managed. As such, the researcher endeavored to analyse the characteristics of the water reticulation system for the two DMA. This section, therefore, presents the critical network characteristics.

Table 3: Descriptive Statistics (Area 2: SZA1)

Descriptive Statistics						
	N	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**	**Variance**
Age of Pipe	31	28.000	65.000	51.613	12.632	159.578
Pipe Diameter	31	63.000	150.000	97.032	21.009	441.366
Pipe length	31	0.000	0.803	0.209	0.204	0.042
Functional Consumer meters	31	15.000	85.490	42.185	12.729	162.035
Non-Functional meters	31	0.000	18.000	7.198	4.452	19.820
Average Meter Age	31	0.300	3.600	1.735	0.585	0.342
Number of connections per kilometer	31	15.000	103.000	49.677	14.272	203.692
Area Population	31	25.500	175.100	84.452	24.263	588.671
Elevation	31	1033.100	1102.100	1064.548	22.884	523.661
Avg Day Hour Pressure	31	0.000	38.000	20.645	8.558	73.237
Network Repairing Costs	31	46865.000	321806.000	155208.742	44590.711	1988331496.531
Daily SIV	31	1335.000	3513.000	2181.839	785.799	617479.940
Daily BC	31	781.750	2652.100	1277.645	513.270	263446.241

Table 4:Descriptive Statistics (Area 22: SZD3)

Descriptive Statistics						
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Age of Pipe	31	4.000	46.000	34.065	15.122	228.662
Pipe Diameter	31	50.000	110.000	103.516	17.247	297.458
Pipe length (KM)	31	0.001	0.348	0.113	0.114	0.013
Functional Consumer meters	31	4.000	72.000	31.810	17.261	297.952
Non-Functional meters	31	0.000	19.000	5.404	3.732	13.931
Average Meter Age	31	0.100	7.700	1.961	1.299	1.687
Number of connections per kilometer	31	10.000	90.000	37.226	19.555	382.381
Area Population	31	145.152	1306.365	540.339	283.838	80563.780
Elevation	31	1059.600	1071.200	1064.739	4.185	17.515
Avg Day Hour Pressure	31	9.000	28.000	18.161	5.751	33.073
Network Repairing Costs	31	117630.000	1058669.000	437886.839	230020.136	52909262883.607
Daily SIV	31	301.000	1095.000	576.323	232.452	54034.092
Daily BC	31	264.920	672.950	406.298	82.876	6868.508

Based on the results of the descriptive analysis as presented in Tables 3 and 4, the research found out that, although network rehabilitation activities are taking place to reduce NRW, the water distribution system is still dominated by old pipes as the pipe mean ages in both SZA1 and SZD3 (mean = 51.61 and mean 34.065, respectively) are still relatively high. However, the standard deviations for this network characteristic show that there are high variations in the age of pipes in SZD3 (std = 15.122) as compared to SZA1 (std = 12.632), which could be as a result of various pipe replacement activities that have taken place in former, as also evidenced by a small minimum age (4.0). The pipe diameters, pipe length, functional consumer meters and meter ages in the two DMAs were found to have small mean differences, suggesting that these network characteristics are almost similar across the two DMAs, although higher variations were observed for functional consumer meters in SZD3 as compared to SZA1 (std= 17.261 and 12.729, respectively). The mean ages for consumer meters were in both DMAs found to be small (SZA1 mean= 1.735 and SZD3 mean = 1.961), which showed that a majority of consumer meters have been installed or replaced in recent years. Despite SZD3 having a higher population than SZA1 (mean= 540.339 and 84.452, respectively), the later has a higher connection density (mean=49.68 as compared to 37.226 for SZD3) mainly because of its being residential as well as business area.

Despite the two DMAs being situated at similar elevation levels (SZA1 elevation mean=1064.548, SZD3=1064.739), it was observed that SZA1 operates in slightly higher pressure levels (mean=20.645) as compared to SZD3 (mean=18.161). Adding to the higher pressure levels, SZA1 was also found to have high operating costs of repairing pipe bursts and worn out meters than SZD3, which also signaled for high network maintenance resulting from the damages caused by the high pressures. Lastly, the other critical network parameter observed was the billed consumption. As the results of the analysis show, SZA1 has higher water consumption, which also varies highly across the DMA (mean=1277.645; std= 513.270) as compared to SZD3 (mean=406.298; std.= 82.876), mainly for the same reason as presented for connection density.

4.2.2 Water losses

In order to analyse water losses for the two selected areas, water losses as a result of both managerial and physical characteristics of the distribution system were examined. Due to the absence of specific distribution tanks within the selected DMAs, two water loss assessment parameters, namely losses due to tank overflow and losses due to cleaning storage tanks, were excluded from the study. The study therefore, used the following water loss factors; losses due to accounting errors, losses due to illegal connections, losses due to meter inaccuracies, losses due to customer non-payment, Losses due to operations & other works at site, losses through office use, losses through observed bursts and losses through observed leakages. The results of the descriptive analysis on these factors are as presented below.

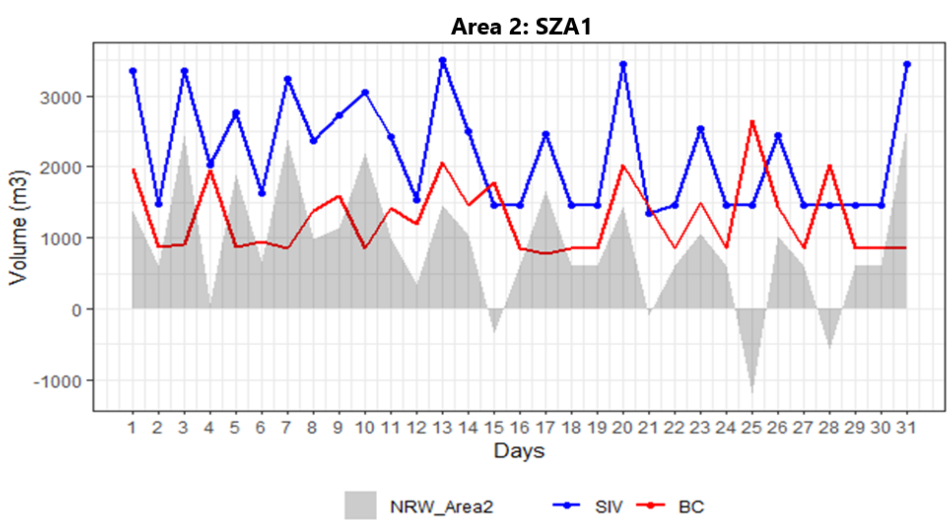


Figure 5: Water Management in SZA1

Table 5:Descriptive Statistics for water losses (Area 2: SZA1)

Descriptive Statistics						
	N	**Minimum**	**Maximum**	**Sum**	**Mean**	**Std. Deviation**
Losses Due to Accounting Errors	31	72.090	244.566	3652.398	117.819	47.332
Losses due to through Illegal connection	31	120.35	391.579	6413.871	206.899	88.794
Losses due Meter inaccuracies	31	68.085	230.979	3848.77	124.177	46.664
Losses due to Customer non-payment	31	20.025	67.935	1014.555	32.728	13.148
Losses due to operations & Other works at site	31	0.000	50.130	127.230	4.104	11.384
Losses due to For Office	31	1.469	4.982	74.401	2.400	0.964
Losses due to observed bursts	31	0.000	446.720	1440.112	46.455	116.7668
Losses due to Leakages	31	0.000	1316.890	4960.540	160.017	324.025
NRW	**31**	**-1200.100**	**3678.740**	**28030.000**	**904.194**	**953.221**

As Figure 4 and Table 5 show, the variability of water losses from the average monthly losses is quite high (NRW std.= 953.221) in SZA1, with some days indicating a negative value (Figure 4: 15th, 21st, 25th and 28th of November, 2021). Such low values were observed after days of high water input with low consumption, which meant that consumption in the following days was sustained by the huge balance of water circulating in the system. Therefore, although the following days had low supply of water, higher water consumption was still sustained, resulting into a negative NRW value for such days. This results also reveal that more water is lost through illegal connections and through observed leakages (mean=181.093 and 160.017, respectively), followed by accounting errors, meter inaccuracies, customer non-payment, pipe bursts and operations and other activities in rehabilitation sites, both having monthly water loss means of 117.819, 111.274, 32.728, 17.288 and 4.104, respectively. Water losses resulting from office usage was found to be the least among the assessed factors, showing also very low variability (mean=2.40, std.=0.964) as compared to the other water loss factors such as observed leaks, illegal connection, observed bursts and accounting errors, which have very high variability (std.= 324.025, 72.751, 67.870 and 47.332, respectively.

Similar to the observations in SZA1, the research also found that higher NRW variability exists in SZD3 despite the later DMA having low levels of total NRW. SZD3 registered a mean water loss of 170.025 m^3 with a standard deviation of 238.550 (Table 6), as compared to SZA1 which had slightly higher values. However, the low levels of water loss in SZD3 were in proportion to its cumulative system input volume (17866 m^3) which was also lower than the volume for SZA1 (67637 m^3). As recorded in Figure 5, SZD3 also registered very low NRW levels on certain days mainly due to the same reason as was in SZA1. However, as shown in Figure 5, water consumption in this DMA was slightly consistent which can be attributed to its being a residential area that may not have abrupt changes in consumption patterns as would a business area with non-domestic water users.

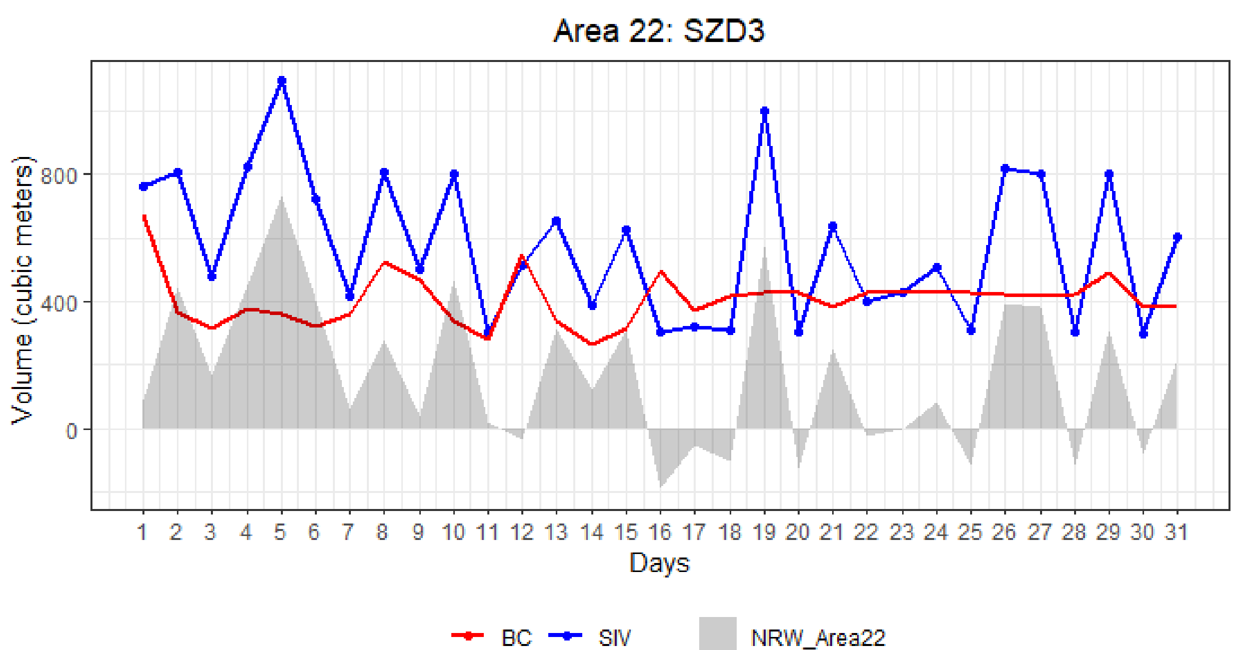


Figure 6:Water Management in SZD3

Table 6: Descriptive statistics for water losses (SZD3)

Descriptive Statistics						
	N	**Minimum**	**Maximum**	**Sum**	**Mean**	**Std. Deviation**
Losses Due to Accounting Errors	31	16.254	54.054	1072.424	34.59	18.69
Losses through Illegal connections	31	18.662	67.890	1095.292	35.332	14.708
Losses through Meter inaccuracies	31	11.739	42.705	696.774	22.477	9.066
Losses through Customer non-payment	31	5.418	19.710	317.988	10.258	4.270
Losses due to Operations & Other works at site	31	0.000	312.700	527.290	17.009	61.326
Losses through For Office	31	0.903	3.285	53.598	1.729	0.697
Losses through observed Bursts	31	0.000	261.200	261.200	8.426	46.913
Losses through Leakages	31	0.000	171.060	1214.242	39.169	46.515
Daily NRW	31	-191.340	732.640	5270.770	170.025	238.550

The water loss analysis for SZD3, as shown in Table 6, further reveals that this DMA loses more water through leakages, illegal connections and accounting errors (mean=39.169 m^3, 35.332m^3 and 28.143m^3), followed by meter inaccuracies, operations and other works at rehabilitation sites, customer non-payment, as well as pipe bursts. Similar to SZA1, water losses due to office uses contributed very small to the overall NRW for SZD3.

4.2.2.1 The relationship between network physical parameters and water losses

To further analyse the water balance for the two selected DMAs, a pairwise Pearson Correlation test was performed on cumulative characteristics for the two DMAs to establish the relationship between critical physical network features. However, before performing the correlation test, the system input values, billed consumption and the NRW water values for the two DMAs were combined as shown in Table 7 and Figure 6. The combined NRW water was found to be approximately 38.95% (33300.770) of the total input volume of water (85503.0). Owing to the higher daily NRW variability in the selected two DMAs, the combined NRW values also showed relatively higher variability of 943.906 against a mean of 1074.218 m^3.

Table 7:Descriptive results for cumulative water losses

Descriptive Statistics (Combined DMAs)						
	N	Minimum	Maximum	Sum	Mean	Std. Deviation
Total SIV	31	1753.000	4166.000	85503.000	2758.161	829.182
Total BC	31	1154.260	3078.400	52202.230	1683.943	530.803
Total NRW	31	-1317.400	2819.240	33300.770	1074.218	943.906

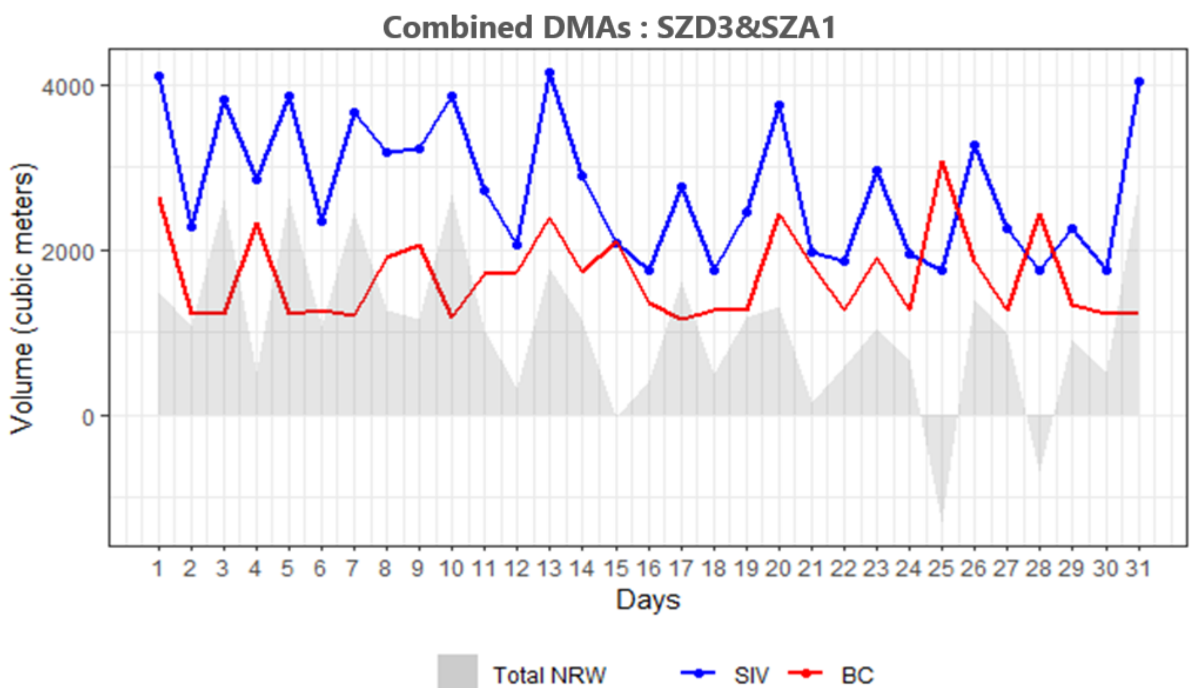


Figure 7: Cumulative water management for SZA1 &SZD3

Table 8:Pearson Correlation test

	Pearson Correlation Coefficients	t value	p value	degrees of freedom
Pipe Diameter	0.013893	0.107627	0.914651	60
Age of Pipe	0.383297	3.214511	0.002105	60
Pipe length	0.128248	1.001679	0.320521	60
Functional Consumer meters	0.266047	2.137835	0.036613	60
Non-Functional meters	0.70196	7.634424	2.06E-10	60
Average Meter Age	0.031723	0.245852	0.806636	60
Number of connections per kilometre	0.297911	2.417375	0.018688	60
Area Population	-0.38297	-3.21126	0.002125	60
Elevation	0.278313	2.24448	0.028502	60
Avg Day Hour Pressure	0.010871	0.084213	0.933168	60
Daily SIV	0.822401	11.19738	2.48E-16	60

As the results of the pairwise Pearson correlation test in Table 8 reveal, several physical parameters of the water distribution system in the two DMAs were found to be significantly correlated with NRW levels. Daily system input volumes and Non-functional meters showed a strong correlation to NRW, with high positive correlation coefficients and a very low p-values ($r=0.822$, $p<2.48E-16$ and $r=0.70196$, $p<2.06E-10$, respectively), suggesting that the amount of water put into the system per day and non-functional meters, which consequently lead to underestimated bills, contribute significantly to NRW levels. Additionally, age of pipes in the two DMAs showed a moderate positive correlation with NRW ($r = 0.38$, $p < 0.01$), while area population showed a moderate negative correlation ($r = -0.38$, $p < 0.01$). Furthermore, the number of connections per kilometer of the distribution pipes and area elevation levels above the sea level showed moderate positive correlations with NRW ($r = 0.30$, $p < 0.05$; $r = 0.28$, $p < 0.05$). Other physical parameters, such as pipe diameter, pipe length, the number of functional consumer meters, average meter age, and average day hour pressure, were found to have weak or negligible correlations with NRW levels (all p-values > 0.05). These findings suggest that some of the major contributing factors to the assessed forms of water losses in the two DMAs are billing underestimation because of non-functional meters, high population density and high daily input volume.

4.2.2.2 Water Losses through leakages.

Water in a distribution system is also lost through background leakages which may not be visible through the other means of assessments. To analyse such losses therefore, the study utilized the principles of the Minimum Night Flow model in which the following results were realized

Table 9:Estimation of Background Leakage for SZA1 Night Flow

Expected Background leakage		
Description	**Calculation**	**Value**
Main losses	0.2092253*40L/km/h.	0.008m^3/h
Connection losses	1540*3L/connection/h.	4.62m^3/h
Property losses	569 *1L/property/h.	0.569 m^3/h
Total Background leakage at 50m Pressure		**5.161 m^3/h**
Night day factor (Pressure correction factor)	(20.64516/50)$^{1.5}$	0.27
Total Background leakage at 50m Pressure		**1.40 m^3/h**
Expected Normal night use		
Domestic night use	2618*6%*10L/h	1.57086m^3/h
Small non domestic use	6*50L/h	0.3m^3/h
Total normal night use		**1.87086 m^3/h**
Description		**Value**
Expected background leakage		1.40 m^3/h
Expected normal use		1.87 m^3/h
Total Expected night use		**3.27 m^3/h**
Measured MNF		12.1 m3/h
Excess Night Flow		8.84 m3/h

Following the SANFLOW (South African Night Flow Model) guidelines as stipulated by Mckenzie (1999), water leakages in this study were estimated as expected background leakages and expected normal night usage, as presented in Tables 9 and 10. Under expected background leakages, water losses were estimated on main distribution lines (which is 40 liters per kilometer per hour), losses at connections (which is 3 liters per household connection per hour) and water leakages happening within the connected properties (which is 1 liter per property per hour). For these parameters, the study found out that the average lengths of main distribution lines in SZDA2 and SZD3 were 0.2092253 kilometers and 0.1131923 kilometers, respectively. While the number of connections and properties in SZA1 and SZD3 were found to be 1540 connections, 569 household properties and 1154 connections, 3207 household properties, respectively. Under estimated normal night usage, the study assessed water losses at household level (10 liters per hour) and small non-domestic water uses (50 liters per hour). As also stipulated in the SANFLOW guidelines, it is estimated that 6% of the population in a DMA are active during the minimum night flow hours who use approximately 10 liters of water per hour. Therefore, to find the normal domestic usage per DMA only 6% of each DMA's population were considered. Additionally, the study found out that there are 6 small non-domestic water users in SZA1, which include, schools, hospitals, police units, churches and businesses, than there are in SZD3 (3).

The total leakage flow rates for SZA1 and SZD3 at standard pressure of 50 meters were 5.161 m^3/h and 6.67 m^3/h, respectively. However, upon applying the pressure correlation factors, these flow rates became 1.40 m^3/h and 1.47 m^3/hr. Whereas for the total normal night usage, SZD3 was found to have higher flow rate (9 m^3) as compared to SZA1 (3.87 m^3/h), mainly as a result of population density differences as well. Consequent to the higher Expected Night Flow rate, the total Excess Night Flow (ENF) rate for SZD3 (1.64 $m^3/hr.$) was found to be lower than the average DMA's minimum night flows for the two DMAs (12.11 $m^3/hr.$). Similarly, lower Expected Night

Flow rate for SZA1 resulted into a higher ENF (8.84 m^3/hr.) as shown in Figure 4, indicating higher levels of background leakages than in SZD3.

Table 10:Estimation of Background Leakage for SZD3 Night Flow

Expected Background leakage		
Description	**Calculation**	**Value**
Main losses	0.1131923 *40L/km/h.	0.0045 m^3/h
Connection losses	1154*3L/connection/h.	3.462 m^3/h
Property losses	3207*1L/property/h.	3.207m^3/h
Total Background leakage at 50m Pressure		**6.67 m^3/h**
Night day factor (Pressure correction factor)	$(18.16129/50)^{1.5}$	0.22
Total Background leakage at 50m Pressure		**1.47 m^3/h**
Expected Normal night use		
Domestic night use	14750*6%*10L/h	8.85 m^3/h
Small non domestic use	3*50L/h	0.15 m^3/h
Total normal night use		**9 m^3/h**
Description		**Value**
Expected background leakage		1.47 m^3/h
Expected normal use		9 m^3/h
Total Expected night use		**10.47 m^3/h**
Measured MNF		12.1 m^3/h
Excess Night Flow		1.64 m^3/h

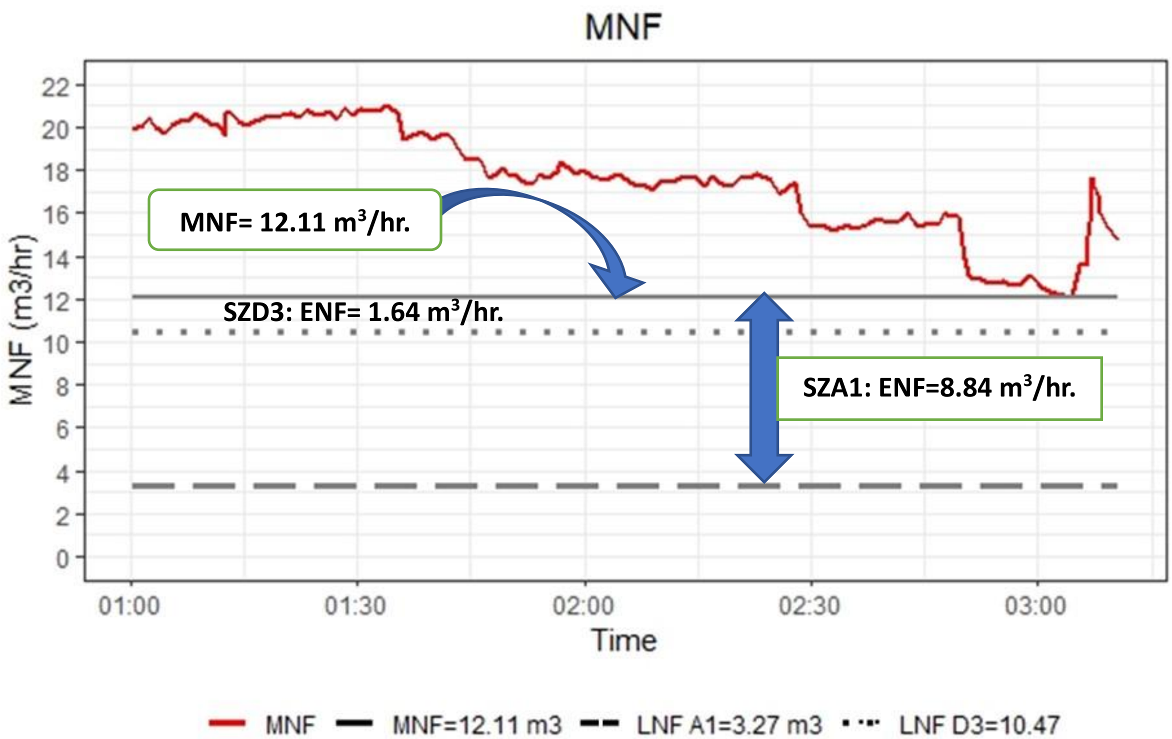


Figure 8: Minimum Night Flows for SZA1 & SZD3

Table 11: leakage water losses

Water Supply Area	ENF (m^3/hr.)	Monthly Water Leakages	Avg Daily Area Supply	Estimated background Water Leakages		
				% loss on supply	M^3/conn/month	L/conn/day
SZA1	8.84	5508.204 m^3	2181.839	**6.44%**	3.57	115.38
SZD3	1.64	1021.884 m^3	576.323	**1.20%**	0.89	28.57
Total		6530.088	2758.162	**7.64%**	4.46	143.95

Based on the estimated Excess Night Flows therefore, the two DMAs were found to lose 7.464% of the water put into the system through invisible background leakages, claiming 19.61% from the total water. Although the invisible leakage losses estimates were found to be very low, which can also be attributed to the lower average Minimum Night Flow as shown in Figure 7, the cumulative losses resulting from leakages in general (observed and invisible leakages), were found to be higher comprising of 38.15% of the total NRW.

4.2.3 Combined Water Balance

From the assessments as presented in the prior section, the following water balance Table for the two DMAs was established.

Table 12:Water Balance Table for SZA1 & SZD3

<table>
<tr><th colspan="4">SZA1 & SZD3</th><th>Basis</th><th>Amount (m^3)</th><th>%</th><th>%</th><th>RW/NRW</th></tr>
<tr><td rowspan="11">System Input Volume (85503 m^3)</td><td rowspan="4">Authorized consumption</td><td rowspan="2">Billed Authorized Consumption</td><td>Billed Metered Consumption</td><td>Billing system</td><td>52202.230</td><td>61.05</td><td rowspan="2">61.05</td><td rowspan="2">61.05</td></tr>
<tr><td>Billed Unmetered Consumption</td><td>none</td><td>0</td><td>0</td></tr>
<tr><td rowspan="2">Unbilled Authorized Consumption</td><td>Unbilled Metered Consumption</td><td>Water use of Office. etc.</td><td>127.9987</td><td>0.15</td><td rowspan="2">0.92</td><td rowspan="9">38.95</td></tr>
<tr><td>Unbilled Unmetered Consumption</td><td>O&M works at site. etc.</td><td>654.52</td><td>0.77</td></tr>
<tr><td rowspan="7">Water losses</td><td rowspan="4">Apparent Losses</td><td>Unauthorized Consumption</td><td>Illegal connection</td><td>7509.163</td><td>8.78</td><td rowspan="4">21.18</td></tr>
<tr><td>Customer Meter Inaccuracies</td><td>Meter inaccuracies</td><td>4545.544</td><td>5.31</td></tr>
<tr><td>Data Handling Errors</td><td>Accounting errors</td><td>4724.822</td><td>5.53</td></tr>
<tr><td>Customer nonpayment</td><td>Customer nonpayment</td><td>1332.543</td><td>1.56</td></tr>
<tr><td rowspan="3">Real losses</td><td>Reported bursts</td><td>Reported bursts</td><td>1701.312</td><td>1.99</td><td rowspan="3">16.85</td></tr>
<tr><td>Visible Leakage</td><td>Reported leakages</td><td>6174.782</td><td>7.22</td></tr>
<tr><td>Invisible bursts and Leakage</td><td>MNF</td><td>6530.088</td><td>7.64</td></tr>
<tr><td colspan="5">Total</td><td>85503</td><td>100</td><td>100</td><td>100</td></tr>
</table>

Table 12, presents a summary of water management activities within the two selected DMAs. Upon accounting for the water put into the system, the primary audit failed to account for water amounting 1303.46 m^3 in order to balance usage and input volumes. However, further analysis showed that due to consumer meter inaccuracies in some sections of Area 2 and 22, meter readers had underestimated the volume of water lost through illegal connections. The re-examination therefore helped to allocate the balance to the right segments of water losses (illegal connections). The final water balance therefore, as presented in Table 10, indicated that 38.95 % of water put into the system is lost. Although this NRW level is slightly below the overall Lilongwe Water Board's current NRW (40%), it is still above the intended threshold of 28%. The research has also found out that LWB losses more water through illegal connections and invisible leakages (8.78 and 7.64, respectively). Although water lost through illegal connections in SZA1 (mean = 206.899 m^3) was found to be more than water lost in SZD3 (mean= 35.332 m^3) (as reported earlier in Tables 3 and 4), the illegal connection cases were found to be more prominent in the later DMA. Through the established water balance, it is also evident that the two DMAs are losing more water through visible leakages, accounting errors and meter inaccuracies (7.22%, 5.53% and 5.32%, respectively). Although many of the visible leakages are reported, it was evident during this research that most water leakage cases are left unreported, especially in densely populated areas of 22. Some of the leakage points were also used to steal water from the system. Finally, the research has also revealed that the utility in these two DMAs lose water through pipe bursts (1.99%), customers who fail to pay for their water bills (1.56%), other management and operations at rehabilitation sites (0.77%) and at the utility's offices (0.15%).

Although the volumes lost through these factors may seem to be low, cumulatively as a water utility company, such volumes are able to cause huge losses if not controlled.

4.3 Establishing the level of contribution for each NRW component on total NRW

In order to achieve the main objective, the research also sought to establish the level of contribution that each NRW component has on total NRW for LWB. The first stage of the analysis sought to reveal the effects that various ongoing network rehabilitation programmes have on the overall NRW components. To do this, descriptive statistical analyses were performed on the System's Input Volume, Billed Consumption, daily total NRW, daily UAC, AL and RL for the two DMAs before and after network rehabilitation and the results are as presented in Tables 13 and 14.

Table 13:Descriptive statistics of SZA1 &SZD3 's water balance components before network rehabilitation

Descriptive Statistics: SZA1 &SZD3						
	N	**Minimum**	**Maximum**	**Sum**	**Mean**	**Std. Deviation**
SIV Before	62	123.00	4523.77	78148.00	1260.45	1031.62
Billed Consumption before	62	78.23	2649.03	46215.68	745.41	597.94
DailyTotal NRW before	62	-627.14	1874.74	31932.32	515.04	496.58
Daily AUC before	62	0.98	95.00	1503.54	24.25	23.63
Daily AL before	62	20.91	778.09	13401.28	216.15	178.10
Daily RL before	62	22.88	1001.56	16907.51	272.70	233.95

Table 14:Descriptive statistics of SZA1 &SZD3 's water balance components After network rehabilitation

Descriptive Statistics: SZA1 & SZD3						
	N	**Minimum**	**Maximum**	**Sum**	**Mean**	**Std. Deviation**
SIV After	62	301.00	3513.00	85503.00	1379.08	992.59
Consumption After	62	264.92	2652.10	52202.23	841.97	570.85
Daily NRW After	62	-1200.10	2600.91	33300.77	537.11	732.73
Daily AUC After	62	0.90	314.22	782.52	12.62	44.12
Daily AL After	62	52.25	819.39	18112.07	292.13	223.00
Daily RL After	62	26.44	422.10	14406.18	232.36	110.46

As the results show, the mean SIV for the two DMAs (combined) increased from 1260.45 m^3, before the rehabilitation of mains, to 1379.08 m^3 after network rehabilitation. However, the variations in daily supply decreased as indicated by the change in their standard deviations (1031.62 to 992.59). The analysis also showed that after the rehabilitation of the mains in the two DMAs, the daily water consumption increased in the two DMAs, whereas a slight decrease was observed in daily variations as well (mean =46215.68 m^3, sd=597.94 to 841.97 m^3, sd=570.85). Despite the changes to the distribution network however, the analysis showed that the total NRW increased slightly from 515.04 m^3 daily NRW average to 537.11 m^3. These results therefore, confirmed the primary concern that gave rise to this study, which stipulated that, despite the network rehabilitation, NRW for LWB keeps rising. Contrary to the increasing total NRW in the two DMAs, the descriptive statistics show that there is a sharp decrease in the highly varying average daily UAC from 24.25 m^3 to 12.62 m^3. This decrease in average daily UAC however did not come as surprise, as the previous analysis in objective 1 showed that UAC for these two DMAs is dominated by water lost through operations and management works at rehabilitation sites. Therefore, having rehabilitated the mains water losses as a result of this factor decreased, which consequently also decreased UAC volumes in general. Lastly, the descriptives statistics also show that although there is an increase in average daily AL (from 216.15m^3 to 292.13m^3), the average daily RL have slight decreased after the rehabilitation from 272.70 m^3 to 232.36 m^3.

4.3.1 Effects of network rehabilitation on NRW components

While the descriptive statistics showed that changes emerged among the NRW components and the total NRW after network rehabilitation, the statistical significance of such changes was not revealed. Therefore, to ascertain the significance of such changes, the NRW components and other water balance parameters were subjected to a paired samples t-test and the results are as shown in Table 15.

Table 15:Paired Samples t-test for before and after rehabilitation

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference: Lower	95% Confidence Interval of the Difference: Upper	t	df.	Sig. (2-tailed)
Pair 1	**SIV After - SIV Before**	118.63	573.81	72.87	-27.09	264.35	1.63	61	0.11
Pair 2	**BC After - BC before**	96.56	402.63	51.13	-5.69	198.81	1.89	61	0.06
Pair 3	**NRW After - NRW before**	22.07	655.21	83.21	-144.32	188.47	0.27	61	0.79
Pair 4	**AUC After - AUC before**	-11.63	52.74	6.70	-25.02	1.76	-1.74	61	0.09
Pair 5	**AL After - AL before**	75.98	121.28	15.40	45.18	106.78	4.93	61	0.00
Pair 6	**RL After - RL before**	-40.34	545.18	69.24	-178.79	98.11	-0.58	61	0.56

The results of the Paired Samples t-test as shown above, reveal that although the initial descriptive statistics showed that there were changes in the SIV, BC, NRW, AUC and RL after network rehabilitation, their mean differences were not statistically significant at 95% confidence interval (p>0.05). However, the t-test confirms that a statistically significant difference for Apparent Losses (p<0.05) occurred after the distribution network rehabilitation. The analysis also shows that although a decrease in AUC was observed after network rehabilitation (mean = -11.63), there was 8% probability that the observed difference was due to a random chance. Therefore, it can be concluded that, although the utility's distribution network rehabilitation works intended to reduce NRW, such an anticipated change never happened, as an insignificant increase resulted instead.

4.3.2 The relationship between NRW components and the total NRW

To further understand the relationship between the three NRW components and the total NRW after the network rehabilitation, a Pearson Correlation test was performed as follows;

Table 16:Pearson Correlation Test for NRW components and the total NRW after network rehabilitation

	Pearson Correlation Coefficients	t value	p value	degrees of freedom
Daily AUC After	-0.09179	-0.71404	0.477971	60
Daily AL After	0.741456	8.55928	5.48E-12	60
Daily RL After	0.96618	29.02247	5.12E-37	60

The results of the Pearson Correlation test in Table 16 show that after the rehabilitation of the network, the daily UAC have a weak negative relationship (r = -0.09179) with the total NRW in the two DMAs. Although the relationship is statistically insignificant (p=0.477971), which could mean the results were by chance, the descriptive analysis in the previous section has shown that despite the increasing total NRW, the daily UAC has decreased because of the reasons as previously provided (section 4.3). This was, therefore, an early indicator for the negative relationship between the two parameters. Contrary to the negative relationship between UAC and total NRW, a very significant positive relationship exists between the daily AL and RL, and the total NRW (r= 0.741456; p=5.48E-12 and r=0.96618; p=5.12E-37, respectively). Therefore, it can be concluded that, despite the network rehabilitation activities in the two DMAs, there exist statistically significant positive relationships between the total utility water losses and the AL and

RL. However, the Pearson Correlation being a test that does not signify a cause-and-effect relationship, the research further examined the level of contribution or influence of each NRW component on the total NRW through a sensitivity analysis using an improved version of the Garson Algorithm based on a feed forward backpropagation Artificial Neural Network as follows.

4.3.2.1 ANN Architecture for NRW components

Having performed a series of model parameter adjustments, an optimal network architecture comprising of 2 hidden layer and 7 neurons, as shown in Figure 8, was developed using a "neuralnet's" back propagation algorithm (backprop). Although nodes in the hidden layer were fitted with sigmoidal (logistic) activation functions, the model was tuned to linear output as the response variable (NRW) was continuous. In order to achieve a well performing network, the water loss components data after network rehabilitation were used as held-out test set, which is a data sample that is not used during the training of a machine learning model. The purpose of the held-out test was to evaluate the performance of the designed model on unseen data which would eventually give accurate sensitivity coefficients for the improved Garson algorithm. The model was trained on NRW components data collected before network rehabilitation at a ratio of 70% for training. Because of the high variations in both training and testing data, Min-Max normalization technique was used to scale the input data into a range of between 0 and 1 by subtracting the minimum value from each value in the input data, then dividing the result by the range. After 15 epochs, the model reduced the value of the error function to 0.049123 in 84456 iterations as shown in Figure 5. This error function value gave the best optimization of the model performance with a least MSE of 0.00061 and an R-Squared value of 0.9845962 as shown in Table 15. This meant that using the established network architecture, the model was able to explain

98.45% of the variations in the response variable (NRW after network rehabilitation). After applying the model to the test set, the predicted and the actual (test) response data (NRW) were scaled back to original values using a reverse Min-Max normalization technique and were plotted as shown in Figure 9. As shown in the graph, the model predicted the response variable (NRW after network rehabilitation) correctly, which set a good platform for further investigations using the Improved Garson Algorithm to establish the level of contribution for each NRW component on the total NRW.

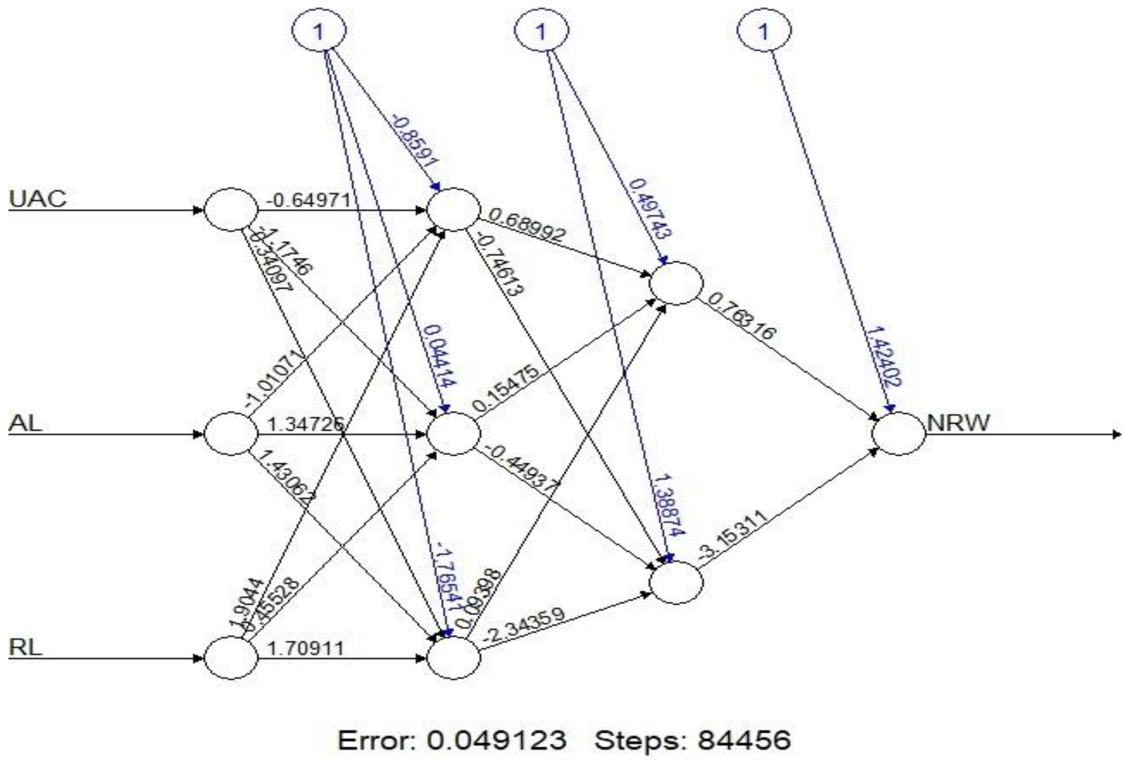


Figure 9: Feed Forward Back Propagation Neural Network Architecture

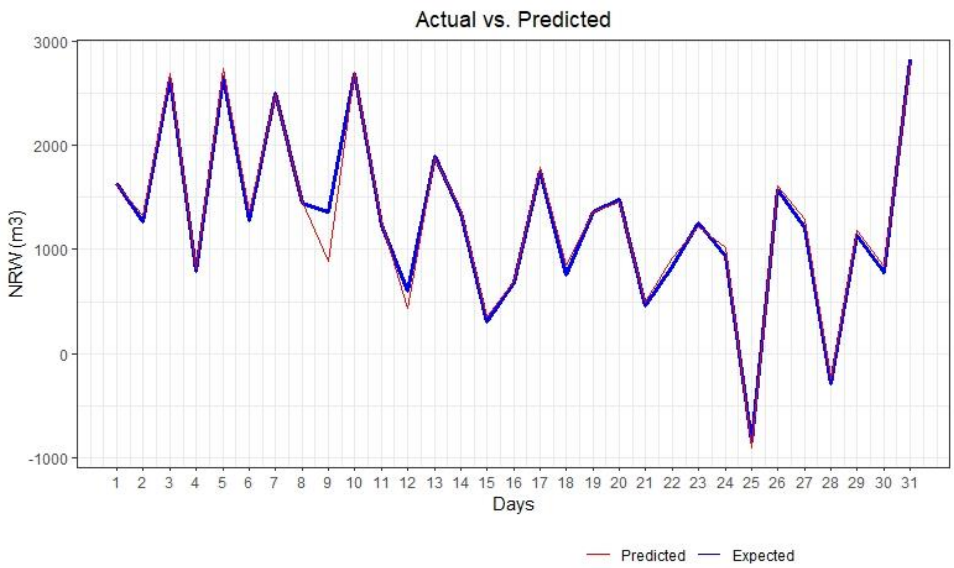


Figure 10: Predicted versus Actual NRW

Table 17:ANN Model Fitness

Actul mean (m^3)	Predicted mean (m^3)	Mean devation %	MSE	R-squared
1074.22	1281.59	4.75	0.00061	0.98

4.2.2.2 Components level of contribution on NRW

Upon the successful building of the feed forward back propagation Artificial Neural Network, an analysis on the level of contribution by every NRW component on the total NRW was done. This analysis was performed to ascertain variable importance in determining the non decreasing levels of NRW after the rehabilitation of the destribution netwok's mains. Table 18 and Figure 10 show the results of the sensistivity analysis based on the Improved Garson Algorithm.

Table 18: IGarson Coefficients

Level of importance	variable	IGarson Coefficient ($S_k^p(i)$)	IGarson %
1	AL	1.32	43.94
2	RL	0.95	31.68
3	UAC	0.73	24.39

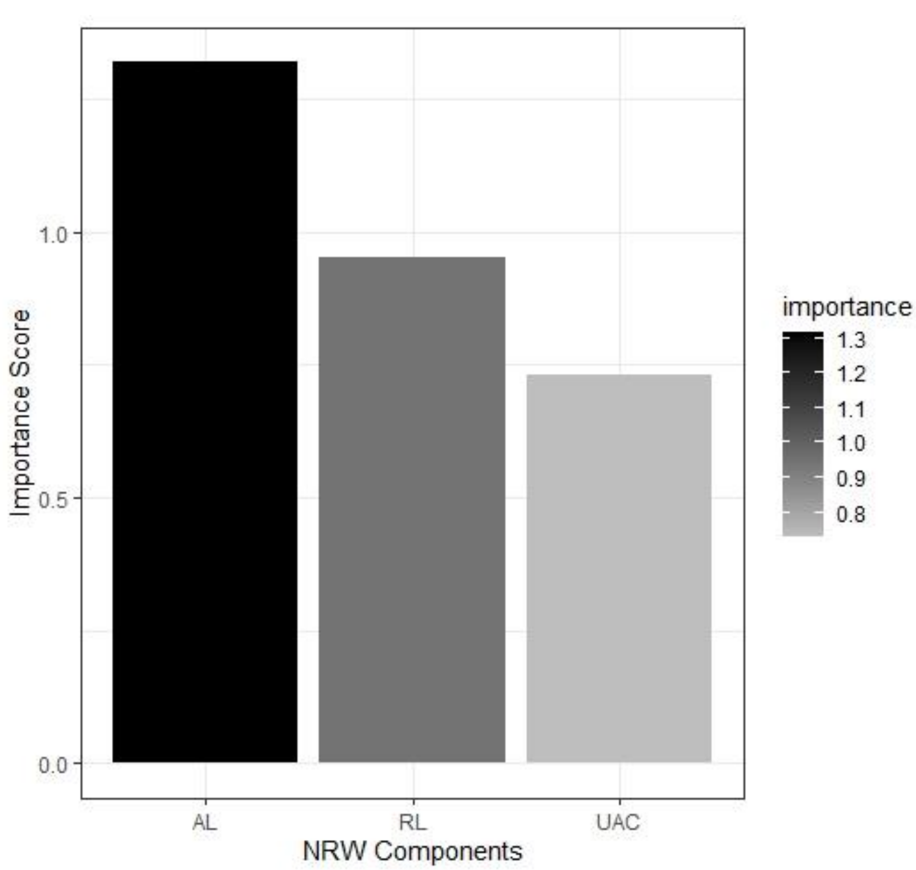


Figure 31: IGarson components contribution graph

To determine components contribution to the total NRW, connection weights and biases for the inputs, hidden and output layers were exported and prepared to be used in the IGarson fomular. The processing of the connection weights and biases included the calculation of absolute values and normalization, which was followed by the introduction of the training predictor variables (AL, UAC and RL) to the IGarson algorithm to calculate sensitivity coefficients for each NRW component. The results of the simulation, as shown in Table 18 and Figure 10, revealed a higher IGarson coeffient for AL (1.318049), than the two other NRW water components (0.950335 and 0.731616, for RL and UAC, respectively).

The IGarson sensitivity analysis showed that although the preliminary Pearson correlation test showed a very strong linear correlation between RL and the total NRW than the AL, more water is lost through AL (43.94%) than RL (31.68%). These results were found to be consistent with the water balance findings in objective one which showed that more water is lost through AL than RL, followed by UAC. However, despite the water balance analysis in objective 1 showing a very meagre percentage of NRW contribution by UAC, the IGarson sensitivity analysis, which took into account the non-linearity and interactivity of input attributes in influencing the total NRW, showed that UAC still indirectly contributes more (24.39%) than what was revealed in objective 1. The IGarson sensitivity analysis was able to unravel this level of contribution through the examination of the interaction between the three NRW components. It was discovered that, although the direct contribution of the UAC to the total NRW is quite minimal, there exists an important influence that this component has on the other components which leads to an increased level of NRW. For instance, as it was previously explained, a drastic decrease in UAC which would indicate reduced levels of maintenance works, increases the total NRW through an increase in RL. Therefore, although the final account of NRW resulting from this sort of interaction is netted on RL in water balance audits, AL share the contribution indirectly. In conclusion therefore, the analysis of component contribution on the total NRW using the IGarson algorithm based on the feedforward backpropagation ANN, found that while the initial analysis in objective 1 hinted on the level of contribution by the three components, the indirect contribution or influence of UAC on the total NRW through its interaction with other NRW components was underestimated. The analysis has also shown that NRW in the two DMAs is mostly influenced by apparent losses, followed by real losses and unbilled authorized consumptions, respectively.

4.4 Investigating the main drivers for each NRW component for LWB

To investigate the main drivers for each NRW component, the descriptive and correlational analyses were firstly conducted on different NRW components and various water distribution network characteristics before the ANN modelling. From the preliminary descriptive analysis, the mean differences for the three components were observed to be in tandem with the second objective's key findings that there are more AL (mean= 292.13 m^3) in the distribution system than RL and UAC (mean= 232.36m^3 and 12.62m^3, respectively) as shown in Table 19. On the other hand, the Pearson Correlation test showed that a significantly strong positive relationship exists between the total UAC values and the cleaning exercises of the distribution network's main pipes (r =0.97745, p=3.15E-42), as well as a very strong positive relation between UAC and water losses emanating from the distribution network's maintenance works (r =0.97367, p=3.12E-40). Moderate positive relationships between UAC and System Maintenance Frequency and water used by the utility were also observed (r =0.629385, p=4.27E-08 and r =0.461872, p=0.000158, respectively). The moderate positive relationship between UAC and the system maintenance frequency however, gave an early indication that although there might be few maintenance works, there are huge losses during such few exercises as evidenced by the strong relationship between UAC and water lost through maintenance works.

Table 19:Descriptive statistics for NRW components and the Pearson Correlation test

Descriptive Statistics: SZA1 & SZD3						
	N	**Minimum**	**Maximum**	**Sum**	**Mean**	**Std. Deviation**
UAC After	62	0.90	314.22	782.52	12.62	44.12
AL After	62	52.25	819.39	18112.07	292.13	223.00
RL After	62	26.44	422.10	14406.18	232.36	110.46

	Correlation Coefficients	t value	p value	degrees of freedom
UNBILLED AUTHORIZED CONSUMPTION				
Maintenance losses	0.97367	33.08463	3.12E-40	60
Office Use	0.461872	4.033661	0.000158	60
Cleaning Mains	0.97745	35.85485	3.15E-42	60
System Maintenance Frequency	0.629385	6.273635	4.27E-08	60
APPARENT LOSSES				
Accounting Errors	0.504275	4.523333	2.93E-05	60
Illegal connections	0.889547	15.08266	4.43E-22	60
Meter inaccuracies	0.363693	3.024258	0.003665	60
Customer non-payment	0.773447	9.451756	1.73E-13	60
No. of Staff 100/ppl. served	-0.02121	-0.1643	0.870045	60
Population density	-0.27436	-2.21002	0.03093	60

The correlation analysis for AL parameters also showed that there is a significantly strong relationship between the cumulative AL and the presence of illegal connections within the two assessed DMAs, as well as the presence of customers who fail to pay for their water consumption bills (r=0.889547, p=4.43E-22; r=0.773447, p=1.73E-13, respectively). While the analysis revealed moderate and week positive relationships between accounting errors (r=0.504275, p=2.93E-05) and meter inaccuracies (r=0.363693, p=0.003665), and the cumulative AL, a very weak significant negative relationship was observed between the AL and the DMAs population density. Although this result was unexpected, it was found to be consistent with the previous analyses which revealed that although SZA1 has a low population density as compared to SZD3 (for the same reasons as provided earlier), there are more apparent water losses in the former than there are in SZD3. Therefore, combined losses for the two DMAs affected the direction of the relationship between population density and AL. Finally, although not significant, another weak relationship between number of ground staff in each section of the DMA and the total AL was observed, which could signal the impact of the presence of utility staff on the ground to monitor AL indicators such as illegal connections.

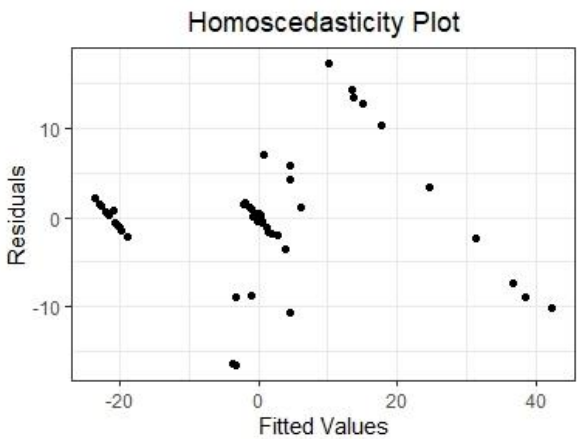
4.4.1 UAC and AL Components specific Drivers

Having examined the relationship between UAC and AL components and their different parameters, the research endeavored to further establish the contribution effect relationships among them using the fixed effects regression analysis as follows.

4.4.1.1 Fixed Effects Regression Analyses

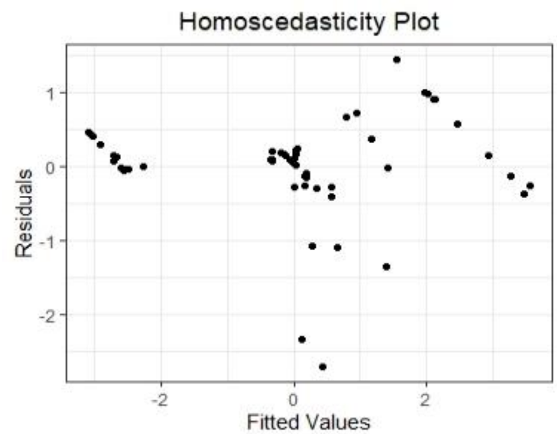
In order to perform robust fixed effects regression analyses, homoscedasticity tests were done using the Breusch-Pagan test for heteroscedasticity. This was to make sure that variations of model errors were constant across all levels of independent variables. The Breusch-Pagan test for heteroscedasticity posits a null hypothesis that the variance of the errors is constant across all levels of the independent variables. While its alternative hypothesis, stipulates that the variance of the errors is not constant across observations. The first model for UAC, in which UAC was the response variable and water used by the utility (Office Use), losses as a result of maintenance works (Maintenance Losses), water lost through cleaning mains (Cleaning Mains), water loss effects as a result of maintenance frequency (System Maintenance Frequency) and other AL losses due to utility management policies (Institutional Policy Effects), were predictor variables, showed that there were variations of residual error across levels of independent variables (Figure 11, BP p-value = 0.0008257). This BP result therefore indicated a violation of the fixed effects regression model's assumption of homoscedasticity. To correct this violation and optimize the model consequently, a square root transformation on the data was performed. This helped to achieve homoscedasticity at last, as evidenced by the BP p-value of 0.06211 (Figure

12). The resultant model for the investigation of UACs was finally able to explain significantly ($p < 2.22e-16$) more than 85.18% of the variations in the response variable and had a small Mean Square Error of 0.4699512. This meant that the model was now ready to be used for predictions and the results of the fixed effects regression analysis for UAC components are as shown in Table 21.



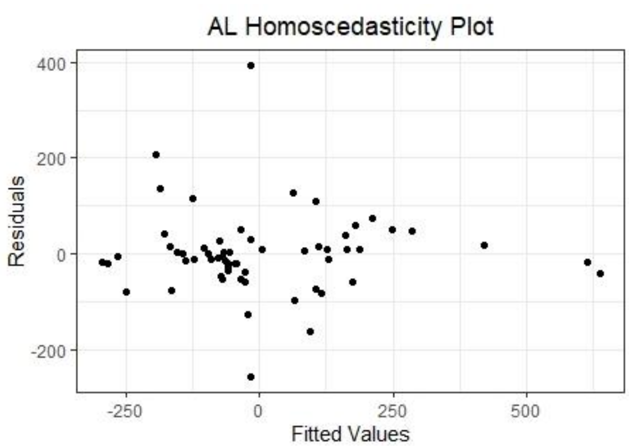
BP = 20.956, df = 5, p-value = 0.0008257

Figure 12: First model residuals



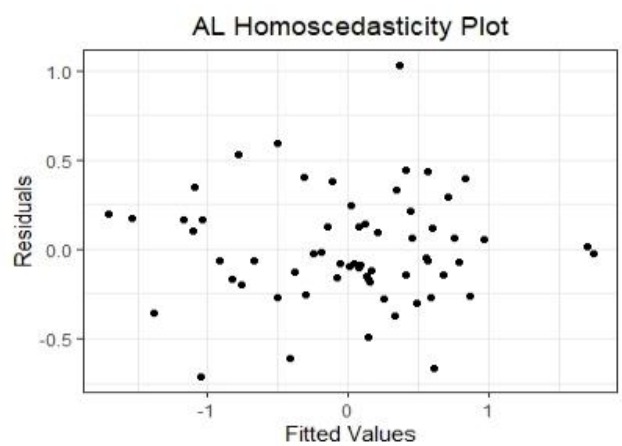
BP = 10.963, df = 5, p-value = 0.06211

Figure 13:Residuals after data square root transformation



BP = 8.0797, df = 8, p-value = 0.4257

Figure 44: Residuals of the first model



BP = 8.8432, df = 8, p-value = 0.3557

Figure 55: Residuals for log transformed model

Table 20:Residuals descriptive statistics for UAC and AL's Fixed effects regression analyses

Model Residuals	Min.	1st Qu.	Median	3rd Qu.	Max.
UAC	-2.6992	-0.12663	0.068581	0.193696	1.431369
AL	-0.712795	-0.164574	-0.058602	0.164001	1.027006

Table 21:UAC's Fixed Effects Regression Analysis results

Coefficients:					
	Estimate	Std. Error	t-value	Pr(>|t|)	
Maintenance losses	1.086566	0.085463	12.7139	< 2e-16	***
Office Use	1.168865	0.725951	1.6101	0.1131	
Cleaning Mains	0.004031	0.043869	0.0919	0.92713	
Institutional Policy effects	0.35688	0.196894	1.8125	0.07536	.
System Maintenance Frequency	0.159076	0.218519	0.728	0.46972	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Total Sum of Squares: 193.46
Residual Sum of Squares: 25.847
R-Squared: 0.8664
Adj. R-Squared: 0.85182
F-statistic: 71.3328 on 5 and 55 DF, p-value: < 2.22e-16
MSE: 0.4699512

Based on the results of the fixed effects regression model for UAC (Table 21), the research found that the major contributing factor or driver for UAC in SZA1 and SZD3 is water lost during maintenance works as shown by its significant coefficient of 1.086566 (p=< 2e-16). As already observed through the Pearson Correlation test however, the effects of system maintenance frequency on the total UAC were found to be insignificantly small (β=0.159076, p=0.46972), which meant that it is not necessarily the number of times maintenance works take place that increases NRW through UAC, rather the magnitude of water losses that are incurred during such exercises have a greater impact. Therefore, while there maybe few maintenance works taking place within the distribution system, the way water losses are controlled during such exercises results in increased UAC. Although not significant (p=0.1131), water losses as a result of utility office consumption were found to increase UAC with 1.169 units by every 1 unit increase in water usage. The results of the analysis also showed that both losses resulting from cleaning main pipes in the distribution system and losses made due to the impacts of the established institutional policies on the management of UAC, are not significant drivers of UAC (β =0.004031, p=0.92713; β =0.35688, p=0.46972, respectively). In conclusion therefore, the research found out that UAC are mostly driven by losses that are incurred during system rehabilitation and maintenance works, than the consumption at the utility's office and cleaning mains. It has also established that although the maintenance works frequency and the utility's policies were identified as some of the drivers of UAC, their influence on the total UAC is minimal.

Just as in the first fixed effects regression model for UAC, to improve the performance of the first AL's fixed effects regression model, a square root normalization was applied to the model's inputs. This was done despite the initial Breusch-Pagan test proving a condition of homoscedasticity for model errors (p= 0.4257), in order to scale down outliers which compromised the model's MSE. The resultant model, whose response variable was AL and the predictors were water losses resulting from account errors (Accounting.Errors), water losses resulting from illegal connections (Illegal.connections), water lost through meter inaccuracies (Meter.innaccuracies), water lost as result of customers not paying for their consumption (Customer.nonpayment), water lost due to the impacts of the number of ground staff in proportion 1000 people served in a DMA (No. Staff.1000ppl.served), water lost in a DMA due to the effects of population DMA's population density (Population), water lost in the DMA due to effects of increased national corruption levels and water lost in the assessed DMAs which could be as a result of effects of an ailing national economy (Economy Effect), showed a reduced MSE of 0.1108351 at BP's p value of 0.3557 (Figure 14). The built model, whose predictive ability was found to be significant (p<2.22e-16) at 95% confidence interval, was finally able to explain more than 84.54% of the variations in AL as shown in Table 22.

Table 22:AL's Fixed Effects Regression analysis results

Coefficients:					
	Estimate	**Std. Error**	**t-value**	**Pr(>|t|)**	
Accounting.Errors	0.500722	0.101266	4.9446	8.36E-06	***
Illegal.connections	0.284214	0.091137	3.1185	0.002961	**
Meter.innaccuracies	0.082057	0.055076	1.4899	0.142298	
Customer.nonpayment	0.206602	0.086836	2.3792	0.021055	*
No. Staff.1000ppl.served	0.0675	0.064685	1.0435	0.301537	
Population	0.133437	0.120787	1.1047	0.274363	
Corruption effects	0.42946	0.216149	1.9869	0.052216	.
Economy Effect	0.045123	0.248877	0.1813	0.856833	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Total Sum of Squares: 37.288
Residual Sum of Squares: 5.7634
R-Squared: 0.84543
Adj. R-Squared: 0.81868
F-statistic: 35.5534 on 8 and 52 DF, p-value: < 2.22e-16
MSE: 0.1108351

As the results of the analysis show in Table 22, three significant AL drivers were revealed. Despite the water balance audit results as presented in objective one which indicated that illegal connections comprise a huge chank of the AL, the fixed effects regression analysis showed that the most significant driver for the apparent losses is data handling or accounting errors made during monthly consumption bills consolidation and processing (p=8.36E-06). Nonetheless, illegal connections were also found to be one of the three significant AL drivers for the Lilongwe Water Board in the two DMA's, followed by water lost through customer non-payment (p=0.002961 and 0.021055, respectively). The analysis also showed that the increase in national corruption levels have an almost significant impact on the AL component (*p*= 0.052216). However, meter inaccuracies were found not to be a significant driver for AL (*p*=0.052216), which could mainly be attributed to the recent rehabilitation works that replaced old and damaged consumer meters especially in SZD3. Similarly, the effects of the number of ground staff per 1000 customers served, DMA's population density and the ailing economic conditions were found to be insignificant AL drivers for the two assessed DMAs at 95% confidence interval (*p*=0.301537, 0.274363 and 0.856833, respectively). In conclusion therefore, the research has found that, there are three main AL component specific drivers for LWB in the two DMAs namely losses resulting from account errors, water losses resulting from illegal connections and water lost as result of customers not paying for their consumption.

4.4.2 RL Component Specific Drivers

Before establishing the main RL component specific drivers, a linear relationship between some of the factors that influence RL in water distribution systems, was examined using a pairwise Pearson Correlation test. The test which was done on data collected from the two DMAs showed that a significantly strong and positive relationship exists between pipe age and RL (*r*=0.788075, *p*=2.93E-14) as shown in Table 23. It was also revealed that moderately significant positive relationships exist between RL and the presence of unobserved background leakages and bursts, observed leakages, number of connections per kilometer of a distribution network, and daily billed consumptions (*r*=0.454325, *p*=0.000208; *r*=0.434109, *p*=0.000423; *r*=0.433425, 0.000433; *r*=0.41491, *p*=0.000799, respectively). While weak significant relationships were also observed between RL and observed bursts, average day hours operating pressure, daily system input volume and the specific area elevation, a negative significant relationship was also observed between RL and pipe diameter (*r*= -0.39061, *p*=0.001697) and population density of an area (*r*=-0.38234, *p*=0.002164). The rest of the other factors did not show any significant relationship with RL at 95% confident interval as to warrant for any conclusive existing linear correlation

Table 23:Pearson Correlation test between RL and its influential factors

Variable	Pearson Correlation Coefficients	t value	p value	degrees of freedom
Age.of.Pipe	0.788075	9.916526	2.93E-14	60
Pipe.Diameter	-0.39061	-3.28672	0.001697	60
Pipe.length	0.105221	0.819588	0.415695	60
Functional.Consumer.meters	0.155481	1.219181	0.227547	60
Non.Functional.meters	0.017092	0.132412	0.895102	60
Average.Meter.Age	-0.13429	-1.04969	0.298072	60
Conne.per.kilometer	0.433425	3.725403	0.000433	60
Area.Population	-0.38234	-3.20511	0.002164	60
Elevation	0.281201	2.269762	0.026829	60
Daily.SIV	0.375541	3.138658	0.002632	60
Daily.BC	0.41491	3.532265	0.000799	60
Avg.Day.Hour.Pressure	0.283559	2.290451	0.025525	60
Observed.Leakages	0.434109	3.73265	0.000423	60
Observed.Bursts	0.395821	3.338693	0.00145	60
Background.Leak.Busts	0.454325	3.950434	0.000208	60

4.4.2.1 ANN architecture for RL drivers

After assessing the direction and magnitude of the relationship between RL and some of its common influencing factors, the research further assessed the magnitude of influence of each factor on RL using the IGarson algorithm based on a feedforward backpropagation ANN. Just as in the development of the optimal network architecture in section 4.2.2.1, a “neuralnet” package in R was used to develop a feedforward backpropagation ANN with two hidden layers and 11 neurons as shown in Figure 15. The hidden layer and the output nodes were also fitted in the same way as done in objective 2. Having performed several network parameter adjustments to achieve optimal coefficient of determination and MSE, the research utilized a data augmentation technique to add noise to the model. This was also done to necessitate the use of part of the data collected after network rehabilitation as a held-out test set. The model was finally trained at the ration of 70% of data for training and 30% for testing. To minimize variations in both training and testing data, a Z-score normalization technique was used to scale the data points to a mean of zero and standard deviation of 1. This was done to reduce the influence of outliers on the response variable, RL. In 21 epochs, the model reduced the error function to 0.00391 in 37502 iterations with a minimal MSE of 0.041 and the optimal coefficient of determination (R^2) of 89.92%. Through this network therefore, the model was able to explain more than 89% of variations in the response variable. Before applying the IGarson algorithm to the model results, the predicted and actual values were scaled back to their original state by reversing the Z score normalization and were ploted as shown in Figure 16. As the graph shows, the model was able to predict most of the data points without overfitting. This stage therefore, warranted the application of the IGarson algorithm to establish the main drivers of RL in the two DMAs

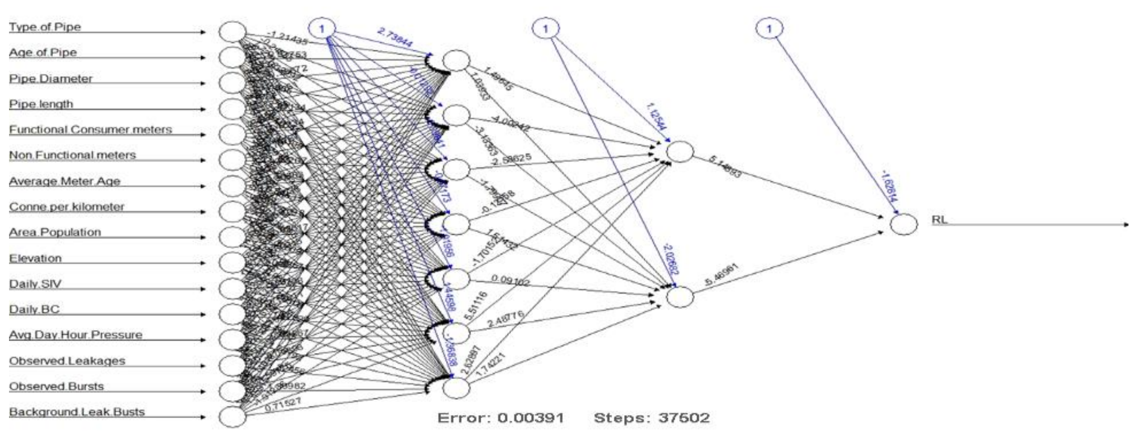


Figure 16: ANN architecture for RL drivers' analysis

Actul mean (m^3)	Predicted mean (m^3)	Mean devation %	MSE	R-squared
232.36	211.34	9.047338808	0.041	0.899208

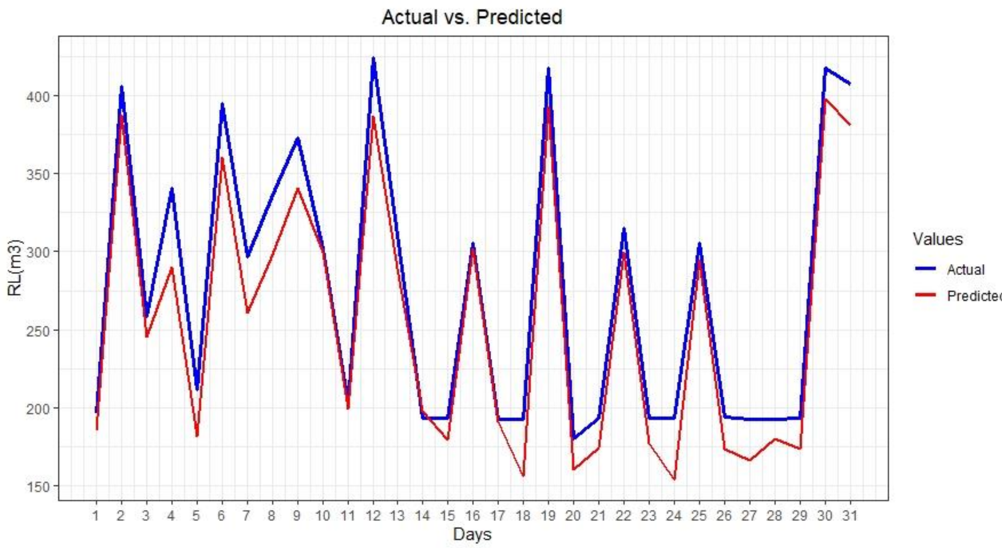


Figure 17: Predicted vs Actual RL

4.4.2.2 Drivers' level of influence on RL

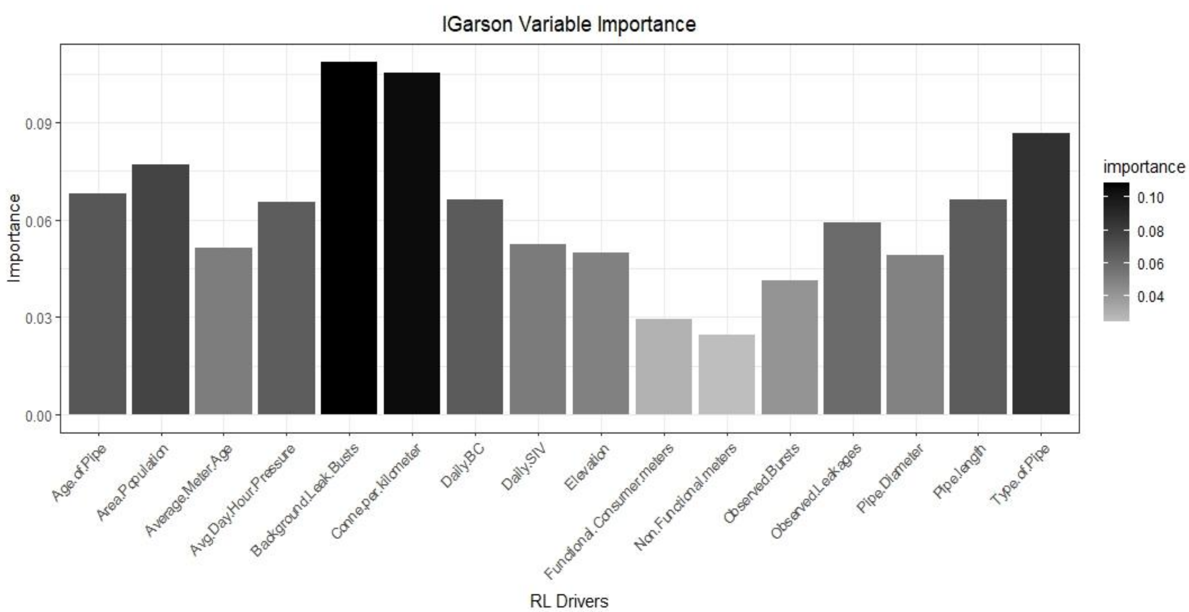


Figure 18: RL Drivers

Table 24:IGarson variable importance for LR drivers

Level of importance	**Variable**	**IGarson importance ($S_k^p(i)$)**	**IGarson %**
1	Background.Leak.Busts	0.109	10.866
2	Conn.per.kilometer	0.105	10.511
3	Type.of.Pipe	0.086	8.647
4	Area.Population	0.077	7.697
5	Age.of.Pipe	0.068	6.814
6	Daily.BC	0.066	6.620
7	Pipe.length	0.066	6.602
8	Avg.Day.Hour.Pressure	0.066	6.557
9	Observed.Leakages	0.059	5.900
10	Daily.SIV	0.052	5.243
11	Average.Meter.Age	0.051	5.140
12	Elevation	0.050	4.980
13	Pipe.Diameter	0.049	4.904
14	Observed.Bursts	0.041	4.144
15	Functional.Consumer.meters	0.029	2.926
16	Non.Functional.meters	0.024	2.449

As shown in both Table 24 and Figure 17, 16 common influencing factors of RL were assessed. Just as found in the NRW audit as presented in objective one, the analysis for the RL component specific drivers showed that background leakages and bursts which are mostly not observed, have a great influence on RL in the two DMAs ($S_k^p(i)$= = 0.109). The analysis also showed that the number of connections per a kilometer of a distribution network is one of influential drivers for RL, followed by the type pipe material for water distribution ($S_k^p(i)$= 0.109 and 0.086, respectively). Of equally great impact was the DMA's population density on RL. While it was observed in the initial Pearson Correlation test that a negative relationship exists between area population and RL, IGarson sensitivity analysis showed that population density is one of the major RL drivers ($S_k^p(i)$= 0.077). The observed negative relationship was found to be a resultant of

SZA1's having more connections per a kilometer of a network but with low population density, while losing more water as compared to SZD3 which has a high population density. The other moderately strong RL drivers as revealed by this research included the age of the distribution pipes, volume of water consumed in a DMA per day (Daily BC), distribution pipe length, average day pressure, observed leakages, daily system input volume of water, average meter age and area topography (elevation) ($S_k^p(i)$=. 0.068, 0.066, 0.066, 0.066, 0.059, 0.052, 0.051 and 0.050, respectively). The distribution pipe diameter, which was found to have a negative correlation with the RL in the previous analysis, was found to be another relatively important driver of RLs followed by the number of leakages observed and reported. The condition of meters being functional or not, was found to be the weakest of all drivers which were assessed in this research. Therefore, it can be concluded that, the LWB in the DMAs has unobserved leaks and bursts, connection density, type of pipe material, population density, age of pipes, daily BC, pipe length and the average day operating pressures, as major drivers of RL.

4.4.3 Conclusion

Based on the data which was collected and the methods as proposed in chapter 3, this chapter has successfully presented and offered interpretation of results of the study. By examining in details the data collected, each specific objective has been answered, which has consequently helped to address the study's research questions and the main objective.

CHAPTER FIVE: DISCUSSION

5.1 Introduction

Based on the results presented in the previous chapter, this chapter discusses the key findings of the research. The chapter discusses the results based on the respective objectives, starting with water balance analysis, NRW components' contribution to total NRW, to NRW components' specific drivers.

5.2 The water balance.

The study has revealed that, although SZD3 has a higher population than SZA1, its water consumption is lower than the later, mainly due to the presence of small non-domestic water users within the SZA1's business areas. As also stipulated by Mubvaruri *et al.* (2022), many urban areas in developing countries who comprise of both business and residential properties tend to experience high water consumption patterns in comparison to purely residential areas like SZD3. The water balance analysis has also shown that of the total 85503 m^3 of water put into the two DMAs, only 61.05% is accounted for as revenue water. The utility loses 38.95% of its system input volume in the two DMAs, which, although being below the 54.61% national NRW average (IBNET 2021), is above the 30% to 35% average global NRW as reported by Liemberger and Wyatt (2019). It does not come as a surprise therefore that the two DMAs have also a higher NRW per capita per day (61.85 liters) similar to the sub-Saharan average NRW level per capita per day (64 liters). However, when compared to the average NRW per capita per day levels of developed regions like USA and Canada (119 liters), this research concurs with Makaya (2015) who concluded that despite the advancements in response strategies and sensitivity, the NRW management remains a challenge even for developed countries.

5.2.1 Water losses

The analysis has shown that, although the Lilongwe Water Board is undertaking network rehabilitation activities of the main distribution lines in the two DMAs, the distribution system is still broadly dominated by old pipes as indicated by the relatively high mean pipe ages in both SZA1 and SZD3. As a result of the aging infrastructure, water losses through leakages and pipe bursts, as observed in the Minimum Flow Analysis, is prominent in the two DMAs. These findings were found to be consistent with other researchers such as Liemberger and Wyatt (2019) and Mubvaruri *et al.* (2022) who concluded that aging water distribution infrastructure, especially pipes, which leads to frequent bursts and leakages, is one of the major causes of serious water losses in developing countries. Based on the Pearson Correlation test's results, it is quite evident that water is not just lost as a result of aging pipes but also as a result of nonfunctional meters. The research has found that monthly consumptions for most customers whose meters are nonfunctional are underestimated, which consequently results into revenue water losses, a condition which is quite synonymous with many utilities in developing countries as reported by Al-washali *et al.* (2019). However, despite other areas having functional meters, the analysis showed a weaker relationship between NRW and functional meters which indicated the prevalence of data handling errors and meter inaccuracies within the two DMAs which result into increased NRW levels.

Adding to the aging pipes and meter reading inaccuracies, this research has also found that the utility loses more water through leakages and bursts in areas with high operating pressures. The research has found that the LWB has higher operating pressures in areas with high connection density such as area 2, a case which Jung *et al.* (2015) argues that is done deliberately in an effort to curb low flow rates due to the high demand. Despite having a low population, SZA1 was found to have a higher connection density with high water consumption, a condition that can better be explained by its being a mixed area comprising of residential houses and business properties. Previous research (Morote *et al.* 2017; Onyango 2021; Abbas *et al.* 2022) have shown that urban areas that tend to have mixed occupancy, experience high water consumption due to either legal or illegal small non domestic water users such as clinics, events gardens, public swimming pools, schools, car wash businesses and many others, which in turn require high operating pressures to maintain standard flow rates. As noted by Ekwule & Utsev (2019) however, utilities that increase operating pressures in an already aged system in order to improve flow rates to densely connected urban areas, further constrain the old pipes into bursts and leakages, a case that explains why there is more water lost through leakages in Area 2 than Area 22.

Finally, the research has also found that the other large proportion of the utility's NRW water is incurred through theft and billed customer nonpayment. It has been observed that the prominent water theft cases in SZD3 were a result of a higher population in the DMA which has fewer connections, while a large amount of water lost through few illegal connections in SZA1 indicated the presence of illegal small non-domestic water users. These findings agree with the findings of Banda and Mwale (2018) and Ziemendorff & Kersting (2020) who also concluded that areas with high population density but fewer number of connections tend to experience high levels of water theft due to unmet water demand.

5.3 Contribution of NRW components on the total NRW

In order to satisfy the main objective, the NRW components' contribution on the total NRW was established as presented in the results section. Based on the findings, it is quite evident that, although there are various network rehabilitation projects such as the rehabilitation of mains, lowering of main pipes and replacement of consumer meters under the ongoing Lilongwe Water and Sanitation project as report by (LWB 2022) and JICA (2022), the positive impacts of such projects are mainly seen in the slight reduction of the physical or real losses. As evidenced by both t-test and the results from the IGarson Algorithm, after the rehabilitation of the main parts of the distribution system, the order of contribution by the NRW components on the total NRW is topped by AL, followed by RL and the UAC.

The findings show that while there is a slight significant decrease in RL, not much is being done to reduce apparent losses, as this component continues to increase, making it a number one influential NRW component for LWB in the two DMA's amidst the ongoing projects. As stipulated by Malek *et al.* (2021), while it is of great benefit to rehabilitate the old distribution network infrastructure in an effect to reduce NRW, when other components of water losses such as AL and UAC are not well taken into consideration, NRW reduction efforts prove to be futile. This key finding, therefore, was found to be in consistent with the extant literature which shows that NRW is a function of many parameters and as such, projects aimed at its reduction must be holistic in nature to all its other drivers (Vermersch *et al.* 2016; Güngör-Demirci *et al.* 2018; Elkharbotly *et al.* 2022). The observed trend in NRW through its components was also found to be in liaison with the findings of Shushu *et al.* (2021), who found that most utilities in developing countries like Malawi, tend to focus on the rehabilitation of the physical system while neglecting other forms of water losses, which in turn continue to stifle the efforts made.

Although the research has also found that there is a decrease in unbilled authorized consumption, Kleiner (1997) noted that decreased UAC in other instances may not signal a decrease in frequency of maintenance works, as there might be well handled maintenance cases that would result in few water losses. In conclusion therefore, the research has revealed that despite the different ongoing water distribution infrastructure rehabilitation works taking place in the two DMAs, no substantial NRW reduction is being achieved mainly due to the increases in AL. It has also revealed that despite the ongoing works, RL are still a major contributing factor to NRW as compared to UACs which have been found to have a small contribution.

5.4 Investigating the main drivers for each NRW components.

While the analysis of the second objective showed that there is no significant change in the reduction of NRW after various network rehabilitation activities in the two DMAs, a further investigation carried out in objective three revealed the major drivers that are necessitating such a phenomenon. This section, therefore, discusses the key findings as presented in the previous chapter as follows.

5.4.1 Unbilled Authorized Consumption

Although UAC may not be regarded as major contributing factor to the overall NRW values in the two DMAs, its aggregated influence at utility level has the ability to cause great losses. As revealed through the fixed effects regression analysis, the most influential driver for this component is water being lost during maintenance works. However, as also found through the correlation test, the strength of this driver is not in the number of times maintenance works take place, but rather the magnitude of losses that take place through a single instance. The relationship between this driver and the UAC can therefore be attributed to poor workmanship at maintenance sites which results in high volumes of water being lost. This research result was found to be in agreement with Zyl (2014), who also concluded that most utilities in developing countries do not just lose water due to dilapidated infrastructure, but also as a result of poor workmanship at maintenance sites such as the use of substandard pipe fitting materials and prolonged maintenance time.

As also found by the analysis of UAC drivers, although some of the key informants opined that water used by the utility in its offices contributes a lot to the high levels of unbilled authorized consumptions, it is quite evident that this driver, among other drivers like the institutional policies of reducing water losses through UAC and losses made due to the activities of cleaning main distribution pipes, has less influence on the total UAC values. This may be a result of the utility's strong policies that negate wasteful uses of water within the utility offices. As observed by Frauendorfer and Liemberger (2010), water utilities that have strong policies against wasteful usage of water within their premises are found to be more efficient in reducing NRW in the form of UAC. Although the insignificance of water losses as a result of cleaning mains may be viewed as a positive achievement by the LWB in the two DMAs, a critical investigation showed that such exercises do not take place as many times as they should; a condition which also postulates a great risk to the quality of water delivered to utility's customers. As pointed out by Heibati *et al.* (2017), many water utilities, in an effort to curb UACs, tend to forego the necessity for main pipe clean up exercises which eventually lead to compromised quality of water delivered to the final consumers. Therefore, the reduction of water losses resulting from the reduced number of mains cleanups, which has positive results on the fight against NRW, on the other hand, poses a great potential for public health concerns in the two DMAs as well In conclusion, this research has found that, while there are mainly four identified drivers of UAC for LWB in the DMAs, the most influential and significant driver which should be targeted in reducing water loses in the form of UACs, is water lost through maintenance works.

5.4.2 Apparent Losses

As revealed by the previous analyses in this study, AL, are the major contributing factor to NRW in the two DMAs. Further to such analysis results, this research has shown that there are three mainly significant drivers of this NRW component among many others, namely accounting error, illegal connections and customer nonpayment for water consumption bills. Although illegal connections were found to be highly correlated with AL in the initial Pearson correlation test of this component, further investigations using the fixed effects regression analysis show that the most significant driver for AL in SZA1 and SZD3 is water being lost due to accounting errors. The high correlation between illegal connections and AL was found to be grounded in the fact that there were more illegal connections in SZA1, an area that also experiences more commercial losses, as compared to SZD3. However, on aggregate, the impact of illegal connections, despite also being very significant, is a bit lesser than the impact of water consumption data accounting errors. As also explained by the UN-Habitat (2012) in its Water Audit Manual, as water utilities that rely on manual data collection and processing are focused on reducing NRW through other drivers, most of them tend to incur serious losses due to post data collection handling mistakes such as accounting errors. Farouk *et al.* (2023), also argues that when a water distribution system has many non-functional consumer meters and that a water utility company relies on consumption estimates, serious data handling error occur as the utilities try to balance up their water audits.

Despite the fixed effects regression results showing that economic changes did not really have a significant impact on AL, this research reveals that customer non-payment for their water consumption bills is the third significant commercial losses' driver in the two DMAs. The disassociation of these drivers on AL was found to be contradictory with other researchers such as Sualihu *et al.* (2017) and Güngör-Demirci *et al.* (2018) who both found out that significant economic effects tend to increase the likelihood of customers not paying for their bills. However, further investigation of such contradicting findings showed that, economic effects on AL could not be captured significantly by the model due to the small sampling timeframe. Therefore, although the fixed effects model results show an insignificant economic effect on AL, it is safe to deduce that based on the drastic changes in the economic outlook and the consequential rising of cost of living in Malawi, as reported by Sherillyn Raga (2023), most customers fail to pay for their water bills due to economic hardships. Additionally, as also argued by Sualihu *et al.* (2017), sometimes water losses resulting from customer non-payment, especially in developing countries signal massive corruption activities being practiced by data collectors who connive with customers to still leave them connected to uninterrupted supply upon receiving bribery in potential cases of disconnection. But as the results of the Pearson correlation test showed, there exists a negative relationship between the presence of ground staff like meter leaders and AL, which may be taken as an indicator of good dealings between meter readers and utility customers. Therefore, the condition of corrupt practices influencing AL, can indeed as revealed by the fixed effects model, be treated as being insignificant. In conclusion, therefore, AL in SZD3 and SZA1 are mainly driven by accounting errors, illegal connections and customer non-payment, in their respective order.

5.4.3 Real Losses

Although objective 2 has shown that there is a slight decrease in RL after various network rehabilitation works in SZA1 and SZD3, RL continues to be one of the major contributing factors to the total NRW. As such, in order to understand the main drivers of real losses in the two LWB's DMAs, many factors were assessed as presented in section 4.3.2.2. However, the four main RL drivers identified were background bursts and leakages, connection density, the type of main distribution pipes and population density. While the initial Pearson correlation test revealed a very strong and positive significant relationship between RL and the unobserved background bursts and leakages, the non-linear sensitivity analysis confirmed the correlation analysis results. Despite the network rehabilitation works such as main pipes replacement, lowering of main pipes, which were supposed to reduce water losses that result from unobserved background leakages, this RL component driver continues to dominate its influence on RL. However, further investigations showed that although the main pipes have been rehabilitated, many water pipes that convey water from the main distribution pipes are not rehabilitated. Therefore, as results in objective one have shown that more background bursts and leakages occur at and within a property connection, the presence of old pipes connecting from the mains are the ideal culprits of the high RL contribution to the total NRW. The results of this analysis thus complement the observations in other research (USAID 2015; Abd Rahman *et al.* 2019; Shushu *et al.* 2021) which also concluded that water distribution systems that have well rehabilitated main distribution pipes tend to equally experience high RL when their subbranches that connect to respective properties are still in a dilapidated state.

As also revealed by this study, one of the major RL drivers in the assessed DMAs is the type of main distribution pipe material. The two DMAs are dominated by PVC pipes as shown in objective one, which according to (WPS 2010) and Jones *et al.* (2021) has a very high susceptibility to breakages and leakages as compared to other pipe materials. Therefore, as the sensitivity analysis also found that the age of pipes, especially pipes that branch from the mains, is one of the moderate drivers of RL, the combined effect of susceptible main pipe materials and pipe age explains why RL are the second NRW component in SZA1 and SZD3.

The connection density which has also been found to be one of the major RL drivers, was also observed by other researchers such as González-Gómez *et al.* (2012) to be one of the main factors that influence RL in water distribution systems. As González-Gómez *et al.* (2012) puts it, when there are many connections within a distribution system, more water is lost through leakages at connection points. He further points out that, the increased number of connections within a certain area makes it even more difficult to track areas of major leakage concerns due to fluctuations in pressure levels. This therefore explains the observed high RLs in SZA1 which has a higher connection density as compared to SZD3. Despite the high connection density however, as observed in the previous analyses, SZA1 has a low population density which was quite reflective in the negative Pearson Correlation coefficient between population density and RL. Therefore, while population density has also been revealed to be a major RL driver in the two DMAs, the observed negative relationship in the Pearson Correlation does not negate the findings of other researchers such as van den Berg (2015) who concluded that high population density leads to increased RL. As also Murrar (2017) suggests, high real water losses are synonymous with high population density due to the fact that as water utilities try to compensate for the reduced water pressure resulting from high water demand, with high operating pressures, there are frequent pipe bursts and leakages.

Apart from the four main RL drivers, there are other moderate drivers unveiled by the research such as pipe length, SIV, average operating pressure, observed bursts and leakages, average consumer meter age, topography and average pipe diameter, which were all in tandem with the findings of many researchers who also found them to be among the main factors that influence RLs (Paper *et al.* 2014; Berg 2015; Güngör-Demirci *et al.* 2018; Mubvaruri *et al.* 2022). In conclusion, despite the ongoing network rehabilitation activities, the two selected LWB's DMAs are mostly affected by background leakages and bursts, connection density, type of pipe materials and population density, among the many identified moderated drivers, which influence the RL component of NRW.

5.5 Conclusion

In order to make scientifically reasonable conclusions, this chapter has discussed the results as presented in chapter 4 with respect to the existing literature. Arguments have been made which have shown the agreement as well as disagreements of the findings of this study with other studies worldwide. Being cognizant of the uniqueness of this study, new knowledge established has also been discussed.

CHAPTER 6: CONCLUSION AND RECOMMENDATIONS

6.1 Introduction

Through the designed objectives of this study, key findings on the NRW components specific drivers have been unveiled. This chapter therefore concludes the research findings and provides recommendations.

6.2 Conclusion

In this study, the specific drivers of NRW components have successfully been identified using the Fixed Effects Regression Models and the ANN based Improved Garson Algorithms. The most influential component on the total NRW has also been identified as the Apparent Losses, followed by the Real Losses and the AUC, respectively.

6.2.1 The water balance analysis

Despite the ongoing distribution network rehabilitation activities taking place at LWB, the water balance audit conducted in this research shows that NRW keeps crippling such efforts as it is above the targeted levels. Such NRW has been found to emanate not only from pipe bursts and leakages, but also mainly through apparent losses such as illegal connections, accounting error and many others. The research's water balance analysis activities have also shown that despite the popular expectation that more water losses would be present in densely populated domestic areas like Area 22 (SZD3), mixed-use areas (having both domestic and small non-domestic water uses) like part of Area 2 (SZA1) which was examined in this research, tend to experience high water losses due to the presence of non-domestic consumers. Therefore, despite having low population densities, such areas are prone to have high NRW levels. In conclusion, as the research endeavored to analyse the water balance for the two selected LWB DMAs through this objective, the total NRW has been found to mainly be comprised of losses resulting from illegal connections, invisible bursts and leakages, visible or reported leakages, accounting errors, meter inaccuracies, reported bursts, customer nonpayment and many other water loss ways, in their respective order.

6.2.2 NRW Components Contribution on the total NRW

As the main purpose of this objective was to examine both linear and non-linear contribution of each NRW component on the total NRW, the research has found that the network rehabilitation works have only managed to reduce slightly the RL and the UACs. The AL continue to increase at the expense of the decreasing RL and UACs which could signal the neglect to control factors that contribute to the commercial losses. Therefore, the order of contribution by NRW components on the total NRW has been found to be from AL, RL to UAC.

6.2.3 NRW Components Specific Drivers

Having established the water balance and investigating the contribution of each NRW component on the total NRW, the research has also successfully unveiled NRW components' specific drivers for the two selected DMAs. Although there were many drivers identified within the course of study, different analysis techniques employed in this research were able to specify the most crucial drivers for NRW components, starting with UAC, AL through RL. For UAC, the research has concluded that while four drivers were identified as being influential, water losses resulting from maintenance works were identified as the main driver for unbilled authorized consumption losses. The Apparent Losses which were preliminarily found to be the major contributing factor to the total NRW, were also found to have many drivers. However, the analytical methods used in the investigation of this component were able to pick the most influential drivers as accounting errors, illegal connections and customer non-payment. Whereas for Real Losses, background leakages and bursts, connection density, type of pipe materials and population density, were found to be the major drivers. In conclusion, the identified components specific drivers were found to be in agreement with findings of other researchers as presented in the previous sections.

6.3 Recommendations

Based on the findings and limitations of this research, this study recommends the following to both water utilities and further researchers;

6.3.1 Water Analysis.

Based on the water balance audit conducted in this research, the following are recommended;

a) Rehabilitation activities should run simultaneously with activities aimed at the reduction of other forms of water losses than just RLs.

b) Deliberate efforts should be directed towards comprehensive strategies that identify and mitigate all contributing factors to NRW, including meter inaccuracies, reported bursts, customer nonpayment, and other forms of water loss.

c) That water utilities in urban areas should pay very close attention to mixed-use areas as these are areas prone to many leakages as well as illegal activities which escalate water losses. By adopting a holistic approach to NRW reduction, the LWB and other utilities can enhance the effectiveness of their rehabilitation activities and achieve significant improvements in their water balance.

6.3.2 NRW Components Contribution on the total NRW.

Given the significant contribution of apparent losses (AL) to the total non-revenue water (NRW), the researcher recommends the following;

a) To prioritize efforts aimed at addressing and reducing AL by implementing robust monitoring systems and control mechanisms to identify and address factors contributing to commercial losses such as regular inspections, leak detection technologies, and metering accuracy checks, water losses can be mitigated promptly.

b) Collaboration and partnerships as methods of reducing AL. Collaboration with stakeholders such as local authorities, community organizations, and non-governmental organizations to raise awareness about the importance of reducing NRW and partnerships to implement water management practices and initiatives that promote responsible water usage and discourage illegal connections, have the ability to curb commercial losses.

c) Regular assessments of the effectiveness of NRW reduction measures and adjust strategies.

6.3.3 NRW components Specific Drivers

Based on the different identified components specific drivers in this research, the study recommends;

a) addressing maintenance-related water losses.

b) strengthening measures to reduce illegal connections, accounting errors and customer non-payment

c) improve on leakage detection and repair, and consider factors that affect population density and connection density.
d) prioritize proactive maintenance practices.
e) the LWB and other water utilities should also enhance billing systems, improve customer engagement and education, and establish strict enforcement against illegal connections to reduce AL.

6.3.4 Further research

While this research has successfully fulfilled its main objective, the researcher recommends the following areas for further study:

a) further assessments of the adequacy of existing infrastructure to accommodate population growth and ensure appropriate connection density to optimize water supply efficiency.
b) Study at a larger spatial and temporal scale to fully understand all other drivers which might have been missed due to spatial and temporal limitations.
c) The use and application of Deep Learning and Artificial Intelligence models like the Artificial Neural Networks to pinpoint and predict background bursts and leakages in areas where MNF estimations may not be effective, such as in areas where non-domestic consumptions are taking place illegally.

References

Abbas, M., Kazama, S. and Takizawa, S., 2022, 'Water Demand Estimation in Service Areas with Limited Numbers of Customer Meters—Case Study in Water and Sanitation Agency (WASA) Lahore, Pakistan', *Water (Switzerland)*, 14(14), doi:10.3390/w14142197.

Abd Rahman, N.; Muhammad, N.S.; Abdullah, J.; Wan Mohtar, W.H.M., 2019 'Model Performance Indicator of Aging Pipes in a Domestic Water Supply Distribution Network', *Water*, 11(11), doi:10.3390/w11112378.

Abdelbaki, C., Benchaib, M.M., Benziada, S., Mahmoudi, H., Goosen, M., 2017, 'Management of a water distribution network by coupling GIS and hydraulic modeling: a case study of Chetouane in Algeria', *Applied Water Science*, 7(3), 1561–1567. doi:10.1007/s13201-016-0416-1.

AL-Washali, T.; Sharma, S.; AL-Nozaily, F.; Haidera, M.; Kennedy, M.., 2019, 'Monitoring Nonrevenue Water Performance in Intermittent Supply', 1–15.

AL-Washali, T.; Sharma, S.; AL-Nozaily, F.; Haidera, M.; Kennedy, M., 2018, 'Modelling the leakage rate and reduction using minimum night flow analysis in an intermittent supply system', *Water (Switzerland)*, 11(1), doi:10.3390/w11010048.

AL-Washali, T.; Sharma, S.; AL-Nozaily, F.; Haidera, M.; Kennedy, M., 2020, 'Assessment of water losses in distribution networks: Methods, applications, uncertainties, and implications in intermittent supply', *Resources, Conservation and Recycling*, 152(September 2019), 104-515, doi:10.1016/j.resconrec.2019.104515.

Al-washali, T., Sharma, S. and Kennedy, M., 2016, 'Methods of Assessment of Water Losses in Water Supply Systems : a Review', *Water Resources Management* [Preprint]. doi:10.1007/s11269-016-1503-7.

AL-Washali, T.; Sharma, S.; AL-Nozaily, F.; Haidera, M.; Kennedy, M., 2020, 'A review of nonrevenue water assessment software tools', *WIREs Water*, 7(2), doi:10.1002/wat2.1413.

Banda, E.M. and Mwale, F., 2018, 'Utility performance in supplying water to informal settlements: A case study from Malawi', *Utilities Policy*, 55, 151–157.

Berardi, L., Daniele B., Laucelli., Simone, A., Mazzolani, G., Giustolisi, O., 2016, 'Active Leakage Control with WDNetXL', *Procedia Engineering*, 154, 62–70. doi:10.1016/j.proeng.2016.07.420.

Berg, C. Van Den, 2015, 'Drivers of non-revenue water : A cross-national analysis *', *Utilities Policy*, 36, 71–78. doi:10.1016/j.jup.2015.07.005.

van den Berg, C. and Danilenko, A., 2017, 'Performance of Water Utilities in Africa', *Performance of Water Utilities in Africa* [Preprint]. doi:10.1596/26186.

Board, L.W., 2021, *Non-Revenue Water Costs Water Boards*, *Lilongwe Water Board*, viewed 6 June 2022, from https://www.lwb.mw/2021/06/09/non-revenue-water-costs-water-boards/

Boru, T., 2018, 'Research design and methodology; research methodology ; University of South

Africa , PHD Thesis', *Reuters.com*, 41(December 2018), 1–41. doi:10.13140/RG.2.2.21467.62242.

Canto Ríos, C J., Santos-Tellez, R.U., Hansen Rodríguez, P., Antúnez Leyva, E., Nava Martínez, V., 2014, 'Methodology for the identification of apparent losses in water distribution networks', *Procedia Engineering*, 70, 238–247. doi:10.1016/j.proeng.2014.02.027.

Charalambous, B., Foufeas, D. and Petroulias, N., 2014, 'Leak detection and water loss management', *Water Utility Journal*, 8, 25–30.

Cheung, P., &Girol, G.V., & Abe, N.,Propato, M., 2010, 'Night flow analysis and modeling for leakage estimation in a water distribution system', *Integrating Water Systems - Proceedings of the 10th International on Computing and Control for the Water Industry, CCWI 2009*, (January), 509–513.

Chimote, K., 2012, 'Water Audit', *Water*. 4th edn. Edited by Lake Victoria Water and Sanitation Initiative team, (11103), 17–20, vied 16 December 2022, from https://unhabitat.org/sites/default/files/download-manager-files/Water Audit Manual Volume 4.pdf.

Ciaburro, G. and Venkateswaran, B., 2017, *Neural network with R: Smart models using CNN, RNN, deep learning, and artificial intelligence principles*, *Packt*. Packt Publishing. Available at: https://medium.com/@nadhifasofia/1-convolutional-neural-network-convolutional-neural-network-merupakan-salah-satu-metode-machine-28189e17335b.

Dutch Water Operators, 2021, *Malawi NRWB Water Efficiency Project*, *Vie*., views 6 June 2022, from https://www.vei.nl/projects/malawi-nrwb-water-efficiency-project

Ekwule, O.R., & Utsev, J.T., 2019, *Evaluation of a Municipal Water Distribution Network Using waterCAD and waterGEMS.*

Elkharbotly, M.R., Seddik, M. and Khalifa, A., 2022a, 'Toward Sustainable Water : Prediction of non-revenue water via Artificial Neural Network and Multiple Linear Regression modelling approach in Egypt', *Ain Shams Engineering Journal*, 13(5), 101673. doi:10.1016/j.asej.2021.101673.

Elkharbotly, M.R., Seddik, M. and Khalifa, A., 2022b, 'Toward Sustainable Water: Prediction of non-revenue water via Artificial Neural Network and Multiple Linear Regression modelling approach in Egypt', *Ain Shams Engineering Journal*, 13(5), 101673. doi:10.1016/j.asej.2021.101673.

Farley, M., Wyeth, G., Bin Md. Ghazali, Z., Istandar, A., Singh, A., 2008, 'The Manager's Non-Revenue Water Handbook: A Guide to Understanding Water Losses. United States Agency for International Development (USAID).', 1–110.

Farley, M., Wyeth, G., Bin Md. Ghazali, Z., Istandar, A., Singh, A., 2010, 'The Manager's Non-Revenue Water Handbook For Africa: A Guide To Understanding Water Losses', *USAID Ranhill Utilities Berhad*, (March), 1–109.

Farley, M. and Liemberger, R., 2005, ‘Developing a non-revenue water reduction strategy: Planning and implementing the strategy’, *Water Science and Technology: Water Supply*, 5(1), 41–50. doi:10.2166/ws.2005.0006.

Farouk, A.M., Rahman, R.A. and Romali, N.S., 2023, ‘Non-revenue water reduction strategies: a systematic review’, *Smart and Sustainable Built Environment*, 12(1), 181–199. doi:10.1108/SASBE-04-2021-0071.

Frauendorfer, R. and Liemberger, R., 2010, *The Issues and Challenges of Reducing Non-Revenue Water*, *Asian Development Bank*.

G., G.D., 1991, ‘Interpreting neural network connection weights[J]’, *AI Expert*, 6(4), 47–53.

Gamze Güngör-Demirci, Juneseok Lee, Jonathan Keck, R.G., 2018, ‘Determinants of non-revenue water for a water utility in California, 270–278. doi:10.2166/aqua.2018.152.

GoM, G. of M., 2022, *Sourthern Region Water Board*, *Public Service Reforms*. viewd 6 June 2022, from http://www.reforms.gov.mw/psrmu/southern-region-water-board

González-Gómez,F., Martínez-Espiñeira, R., García-Valiñas, M.A., García-Rubio, M.A., 2012, ‘Explanatory factors of urban water leakage rates in Southern Spain’, *Utilities Policy*, 22, 22–30. doi:10.1016/j.jup.2012.02.002.

Government of Malawi, 2005, ‘National Water Policy’. Lilongwe, Malawi: Ministry of Irrigation and Water Development.

Heibati, M., Stedmon, C.A., Stenroth, K., Rauch, S., Toljander, J., Säve-Söderbergh, M., Murphy, K.R., 2017, ‘Assessment of drinking water quality at the tap using fluorescence spectroscopy’, *Water Research*, 125, 1–10. doi:https://doi.org/10.1016/j.watres.2017.08.020.

Hoko, Z. and Chipwaila, J.A., 2017, ‘Investigating unaccounted for water and its components in Zomba city water supply system, Malawi’, *Journal of Water Sanitation and Hygiene for Development*, 7(3), 495–506. doi:10.2166/washdev.2017.003.

Howard, D. and Mark, B., 2004,‘Neural Network Toolbox Documentation’, *Neural Network Tool*, 846.

Lai, C.H., Tan, D., Roy, R., Chan, N.W., Zakaria, N., 2020, ‘Systems thinking approach for analysing non-revenue water management reform in Malaysia’, 22, 237–251. doi:10.2166/wp.2020.165.

IBNET, 2018, *The international benchmarking networks*. viewed 6 June 2022, from https://www.ib-net.org/.

IBNET, 2021, *Non Revenue Water Percentage*. viewed 19 March 2023, from https://database.ib-net.org/Reports/Indicators/HeatMap?itemId=27

Jembre, Y.Z., Nugroho, Y.W., Khan, M.T., Attique, M., Paul, R.K., Shah, S.H., Kim, B., 2021, ‘Evaluation of Reinforcement and Deep Learning Algorithms in Controlling Unmanned Aerial Vehicles’, *Applied Sciences*, 11(16), doi:10.3390/app11167240.

JICA, 2019, *LWB And JICA To Hold National Workshop On Non-Revenue Water Reduction*, *Lilongwe Water Board*. viewed 6 June 2022, from https://www.lwb.mw/2021/05/31/lwb-and-jica-to-hold-national-workshop-on-non-revenue-water-reduction/

JICA, 2020, 'Sector position paper: Water Sector', (May 2020), 6.

JICA, 2022, 'Project for Strengthening the Capacity of Non-Revenue Water Reduction for Lilongwe Water Board (LiSCaP)', *Development, Water Change, Climate*, 1–10.

Jones, L.J.N. *et al.*, 2021, 'Non-Revenue Water in Malaysia: Influence of Water Distribution Pipe Types', *Sustainability*, 13(4), doi:10.3390/su13042310.

Jung, J.D., Kang,D., Liu, J., Lansey,K., 2015, 'Improving the rapidity of responses to pipe burst in water distribution systems: A comparison of statistical process control methods', *Journal of Hydroinformatics*, 17(2), 307–328. doi:10.2166/hydro.2014.101.

Kizilöz, B. and Şişman, E., 2021, 'Exceedance probabilities of non-revenue water and performance analysis', *International Journal of Environmental Science and Technology*, 18(9), 2559–2570. doi:10.1007/s13762-020-03018-y.

Kleiner, Y., 1997, 'WATER DISTRIBUTION NETWORK REHABILITATION: Selection and Scheduling of Pipe Rehabilitation Alternatives', 852–864.

Knobloch, A., Guth, N. and Klingel, P., 2014, 'Automated water balance calculation for water distribution systems', *Procedia Engineering*, 89, 428–436. doi:10.1016/j.proeng.2014.11.208.

Liemberger, R. and Wyatt, A., 2019, 'Quantifying the global non-revenue water problem', *Water Science and Technology: Water Supply*, 19(3), 831–837. doi:10.2166/ws.2018.129.

LWB, 2022, *2021/2022 NRW Annual Report*. Lilongwe.

M. B. Abdulla, R. O. Herzallah and M. A. HammadM. B. Abdulla, R.O.H., 2013, 'Pipeline leak detection using artificial neural network: Experimental study', in *5th International Conference on Modelling, Identification and Control (ICMIC)*, 328-332.

Magombo, P.U. and Kosamu, I.B.M., 2016, 'Challenges of water accessibility in the urban centres of Malawi: a case study of Blantyre City', *African Journal of EnvironmentalScience and Technology*, 10(10), 380–385. doi:10.5897/AJEST2015.2126.

Makaya, E., 2015, 'Water loss management strategies for developing countries: Understanding the dynamics of water leakages'.

Makaya, E. and Hensel, O., 2015, 'Modelling flow dynamics in water distribution networks using artificial neural networks - A leakage detection technique', *International Journal of Engineering, Science and Technology*, 7(1), 33–43. doi:10.4314/ijest.v7i1.4.

Malek, H.A., Zakaria, M.H. and Zulkifli, M.L. ,2021, 'Determinants of Non-Revenue Water', 6(1), 642–649.

Manoj Arun Raskar, S.G., 2014, 'Water Loss Identification in Water Distribution System', (5), 156–159.

Maozhun, S. and Ji, L., 2017, 'Improved Garson Algorithm based on Neural Network Model', 4307–4312.

Mastaller, M. and Klingel, P., 2017, 'Adapting the IWA water balance to intermittent water supply and flat-rate tariffs without customer metering', *Journal of Water Sanitation and Hygiene for Development*, 7(3), 396–406. doi:10.2166/washdev.2017.116.

McKenzie, RS,Wegelin, WA & Meyer, N., 2003, *Water Demand Management and Cookbook. Water Resource Planning and Conservation Report. 236p. UN- habitat/Rand Water/WRP/MIYA.Soth Africa*, viewed 7 June 2022, from http://www.unhabitant.org.

Mckenzie, R., 1999, 'Development of a standardised approach to evaluate burst and background losses in water distribution systems in South Africa', (June).

Mckenzie, R.S. and Wegelin, W., 2009, 'Implementation of pressure management in municipal water supply systems', *Iwa*, 03(February), pp. 1–18.

Meijer, S.W., 2015, 'GIS support to reduce Non-Revenue Water in Mzuzu, Malawi.', *Utrecht University.* [Preprint].

Michael H. Kuntner, Christopher J. Nachtsheim, John Neter, W.L., 1997, *Applied Linear Statistical Models*, *Journal of Quality Technology*. doi:10.1080/00224065.1997.11979760.

Momeni E, Nazir R, Jahed Armeghani D, M.H., 2015, 'Application of Artificicial Neural Network for Predicting Shaft and Tip Resistance of Concrete Piles.', *Earth Sci*, 19(1), 85–93.

Morote, Á.F., Saurí, D. and Hernández, M., 2017, 'Residential Tourism, Swimming Pools, and Water Demand in the Western Mediterranean', *Professional Geographer*, 69(1), 1–11. doi:10.1080/00330124.2015.1135403.

Mubvaruri, F., Hoko, Z., Mhizha, A., Gumindoga, W*.,* 2022, 'Investigating trends and components of non-revenue water for Glendale, Zimbabwe', *Physics and Chemistry of the Earth, Parts A/B/C*, 126, 103145. doi:https://doi.org/10.1016/j.pce.2022.103145.

Murrar, A., 2017, 'The Determinants of Non-Revenue Water in Balkan Countries', *American Journal of Water Science and Engineering*, 3(2), 18. doi:10.11648/j.ajwse.20170302.11.

Ndunguru, M.G. and Hoko, Z., 2016, 'Assessment of water loss in Harare, Zimbabwe', *Journal of Water Sanitation and Hygiene for Development*, 6(4), 519–533. doi:10.2166/washdev.2016.157.

Onyango, T.A., 2021, 'Examining the Level of Illegal Water Consumption in High- and Low-Density Estates of Kisumu City, Kenya', *International Journal of Innovative Research and Development*, 10(2), 204–213. doi:10.24940/ijird/2021/v10/i2/158144-389335-1-sm.

Paudel, S.P., 2019, 'Pilot Application of IWA Water Balance', *ADB*, 1–10. viewed 29 March 2022, from https://www.adb.org/results/pilot-application-iwa-water-balance-methodology.

Radivojević, D., Blagojević, B. and Ilić, A., 2020, 'Water supply system performance improvement in the town of pirot using water balance IWA methodology and numerical simulations', *Tehnicki Vjesnik*, 27(3), 970–977. doi:10.17559/TV-20180514195054.

Robson, C. and McCartan, K., 2016, *Real World Research: A Resource for Users oof Social Research Methods in Applied Settings*, *Wiley*.

Samwel, A., 2020, *Assessment of factors which contribute to non-revenue water in kenya and their mitigation: case of meru water supply*. University of Nairobi.

Serafeim, A.V., Kokosalakis, G., Deidda,R., Karathanasi, I., Langousis, A., 2022, 'Probabilistic estimation of minimum night flow in water distribution networks: large-scale application to the city of Patras in western Greece', *Stochastic Environmental Research and Risk Assessment*, 36(2), 643–660. doi:10.1007/s00477-021-02042-9.

Sherillyn Raga, 2023, *Malawi: Macroeconomic and Trade Profile*. Lilongwe. viewed 19 April 2023, from https://cdn.odi.org/media/documents/Malawi_macroeconomic_and_trade_profile_2023_final.pdf

Shushu, U.P., Komakech, H.C., Dodoo-Arhin, D.,Ferras, D., Lal Kansal, M., 2021, 'Managing non-revenue water in Mwanza, Tanzania: A fast-growing sub-Saharan African city', *Scientific African*, 12, e00830. doi:10.1016/j.sciaf.2021.e00830.

Sonali. B. Maind, P.W., 2014, 'Research Paper on Basic of Artificial Neural Network', *International Journal on Recent and Innovation Trennds in Computing and Communication*, 2(1), 96–100.

Sturm, R., Gasner, K., 2017, 'Real Loss Component Analysis: Whats your leakage profile', *AWWA Sustainable Water Management*, 1–18.

Sualihu, M.A., Rahman, M.A. and Tofik-Abu, Z., 2017,'The Payment Behavior of Water Utility Customers in the Greater Accra Region of Ghana: An Empirical Analysis', *SAGE Open*, 7(3), p. 2158244017731494. doi:10.1177/2158244017731494.

Tabesh, M., Shirzad, A., Arefkhani, V., & Mani, A., 2014, 'A comparative study between the modified and available demand driven based models for head driven analysis of water distribution networks', *Urban Water Journal*, 11(3), 221–230. doi:10.1080/1573062X.2013.783084.

Thompson, K., Skeens, B. and Liggett, J., 2017, 'Nonrevenue water', *Internet of Things and Data Analytics Handbook*, 467–480. doi:10.1002/9781119173601.ch28.

UN-Water, 2016, *The United Nations World Water Development Report 2016 on 'Water and Jobs''*', *Leaflet: UN Development Report*.

USAID, 2015, *Africa Non-Revenue Water Program Synthesis Report*. viewed 22 May 2022, from https://www.globalwaters.org/sites/default/files/africa-nrw-program-synthesis-report.pdf.

van den Berg, C., 2014, The Drivers of Non-Revenue Water: How Effective are Non-Revenue Water Reduction Programs?. World Bank Policy Research Working Paper No. 6997, Available at SSRN: https://ssrn.com/abstract=2476706

Vermersch, M,. Carteado, F., Rizzo, A., Johnson, E., Arregui, F., Lambert., A., 2016, 'Guidance Notes on Apparent Losses and Water Loss Reduction Planning’, *The LEAKSSuite Library*, (September), 1–77.

WPS, 2009, *Water Operators Partnerships: African Utility Performance Assessment.*, *Water and Sanitation Program (WSP)*. Africa, Nairobi, Kenya.

WPS, 2010, *The State of African Utilities: Performance Assessment and Benchmarking Report*. Nairobi, Kenya, viewed 20 June 2022, from https://www.ib-net.org/docs/Africa Utility Assessment II.PDF.

Ziemendorff, S. and Kersting, M., 2020, ‘Comparison of technologies for location of illegal drinking water connections’, *ResearchGate*, 1–20.

Zyl, J. van, 2014, *Introduction to Operation and Maintenance of Water Distribution Systems*. 1st edn. Pretoria: The Water Research Commission.

Annex 1

Mzuzu University Research Ethics Committee (MZUNIREC)

Informed Consent Form for Research on

Analysis of Non-Revenue Water Component Specific Drivers in Lilongwe City

Introduction

I am **Aimloss Greyfield Banda** from the **Department of Water and Sanitation, Mzuzu University**. We are doing research on **Analysis of Non-Revenue Water Component Specific Drivers in Lilongwe City**

This consent form may contain words that you do not understand. Please ask me to stop as we go through the information and I will take time to explain. If you have questions later, you can ask them of me or of another researcher.

Purpose of the research

The main purpose of this research is to analyse Non-Revenue Water component specific drivers for Lilongwe Water Board in Lilongwe City.

Type of Research Intervention

This research will involve your participation in an individual interview.

Participant Selection

You are being invited to take part in this research because **the information that you will provide for this research would be valuable as a person who works in the area of interest and have knowledge of the LWB's Non-Revenue Water issues and/or the distribution system's characteristics.**

Voluntary Participation

Your participation in this research is entirely voluntary. It is your choice whether to participate or not. You may skip any question and move on to the next question.

Duration

The research takes place for a period of **1 month.**

Risks

You do not have to answer any question or take part in the discussion/interview/survey if you feel the question(s) are too personal or if talking about them makes you uncomfortable.

Reimbursements

You will not be provided any incentive to take part in the research.

Sharing the Results

The knowledge that we get from this research will be shared with you and the water board before it is made widely available to the public. Following which, we will publish the results so other interested people may learn from the research.

Who to Contact

If you have any questions, you can ask them now or later. If you wish to ask questions later, you may contact: Mr Aimloss Greyfield Banda. (Mzuzu University, Department of Water and Sanitation; +265888224773 /997036033, bandaaim@gmail.com).

This proposal has been reviewed and approved by Mzuzu University Research Ethics Committee (MZUNIREC) which is a committee whose task is to make sure that research participants are protected from harm. If you wish to find about more about the Committee, contact Mr. Gift Mbwele, Mzuzu University Research Ethics (MZUNIREC) Administrator, Mzuzu University, P/Bag 201, Luwinga, Mzuzu 2, Phone: 0999404008/0888641486

Do you have any questions?

Part II: Certificate of Consent

I have been invited to participate in research about the Analysis of Non-Revenue Water Component Specific Drivers in Lilongwe City. **I have read the foregoing information, or it has been read to me. I have had the opportunity to ask questions about it and any questions I have been asked have been answered to my satisfaction. I consent voluntarily to be a participant in this study**

Print Name of Participant__________________

Signature of Participant ___________________

Date ___________________________

Day/month/year

If illiterate [1]

I have witnessed the accurate reading of the consent form to the potential participant, and the individual has had the opportunity to ask questions. I confirm that the individual has given consent freely.

Print name of witness____________ **Thumb print of participant**

Signature of witness _____________

Date ________________________

Day/month/year

Statement by the researcher/person taking consent

I have accurately read out the information sheet to the potential participant, and to the best of my ability made sure that the participant understands the research project. I confirm the participant

[1] A literate witness must sign (if possible, this person should be selected by the participant and should have no connection to the research team). Participants who are illiterate should include their thumb print as well.

was given an opportunity to ask questions about the study, and all the questions asked by the participant have been answered correctly and to the best of my ability. I confirm that the individual has not been coerced into giving consent, and the consent has been given freely and voluntarily.

Signature of Researcher /person taking the consent__________________________

Date ___________________________

Day/month/year

Anex 2. Checklist

Name of participant ……………………………………

Participant's Department ……………………………………

Participant's Department ……………………………………

Date ……………………………………

Part A: [For NRW managers]

I. : Managerial Characteristics of the Distribution System.

1. Number of staff per 1000 people served									
Area 2 DMA					**Area 22 DMA**				
Section 1	**Section 2**	**Section 3**	**Section 4**	**Nth Section**	**Section 1**	**Section 2**	**Section 3**	**Section 4**	**Nth Section**
2. Estimated water lost through billing/accounting errors									
3. Estimated water lost through meter under registration									
4. Daily estimated Water lost through firefighting									
5. Daily estimated water lost through theft (driven from reported cases of water theft and duration)									
6. Daily estimated water used for training purposes									
7. Estimated network repairing cost									
8. Net operating revenue (An estimate from DMA's total operating revenue)									
9. Water used for flushing mains									
10. Water used for cleaning storage tanks									
11. Water lost through tank overflow									

II. : Physical Characteristics of the Distribution System

12. Elevation									
Area 2 DMA					**Area 22 DMA**				
Section 1	**Section 2**	**Section 3**	**Section 4**	**Nth Section**	**Section 1**	**Section 2**	**Section 3**	**Section 4**	**Nth Section**
13. Pipe materials									
14. Daily system input volumes									
15. Billed consumption									
16. Network Pipes sizes (Pipe diameter)									
17. Network pipes length									
18. Other factors that affect NRW management									
i.									
Magnitude		High	Medium	Low	**Magnitude**		High	Medium	Low
ii.									
Magnitude		High	Medium	Low	**Magnitude**		High	Medium	Low
iii.									
Magnitude		High	Medium	Low	**Magnitude**		High	Medium	Low
iv.									
Magnitude		High	Medium	Low	**Magnitude**		High	Medium	Low

Part B: [For Meter Readers]

19. Daily nodal junction minimum night flow rates									
20. Daily nodal junction minimum day hour pressure									
21. Daily nodal junction minimum night hour pressure									
22. Number of reported or observed leaks/day									
23. Number of people per connection									
24. Connection density									

25. DMA population Density									
26. Other factors that affect NRW management									
i.									
Magnitude		High	Medium	Low	**Magnitude**		High	Medium	Low
ii.									
Magnitude		High	Medium	Low	**Magnitude**		High	Medium	Low
iii.									
Magnitude		High	Medium	Low	**Magnitude**		High	Medium	Low
iv.									
Magnitude		High	Medium	Low	**Magnitude**		High	Medium	Low

Thank you for your participation.

Anex 3: Research Approval Letter from the Mzuzu University Research Committee

MZUZU UNIVERSITY

DIRECTORATE OF RESEARCH

Mzuzu University
Private Bag 201
Luwinga
Mzuzu 2
MALAWI
TEL: 01 320 722
FAX: 01 320 648

MZUZU UNIVERSITY RESEARCH ETHICS COMMITTEE (MZUNIREC)

Ref No: MZUNIREC/DOR/23/19 **11/03/2023.**

Aimloss Banda,
Mzuzu University,
P/Bag 201, Luwinga,
Mzuzu 2.

bandaaim@gmail.com

Dear Aimloss,

RESEARCH ETHICS AND REGULATORY APPROVAL AND PERMIT FOR PROTOCOL REF NO: MZUNIREC/DOR/23/19: ANALYSIS OF NON-REVENUE WATER COMPONENT SPECIFIC DRIVERS IN LILONGWE CITY

Having satisfied all the relevant ethical and regulatory requirements, I am pleased to inform you that the above referred research protocol has officially been approved. You are now permitted to proceed with its implementation. Should there be any amendments to the approved protocol in the course of implementing it, you shall be required to seek approval of such amendments before implementation of the same.

This approval is valid for one year from the date of issuance of this approval. If the study goes beyond one year, an annual approval for continuation shall be required to be sought from the Mzuzu University Research Ethics Committee (MZUNIREC) in a format that is available at the Secretariat. Once the study is finalised, you are required to furnish the Committee with a final report of the study. The Committee reserves the right to carry out compliance inspection of this approved protocol at any time as may be deemed by it. As such, you are expected to properly maintain all study documents including consent forms.

Wishing you a successful implementation of your study.

Committee Address:
Secretariat, Mzuzu University Research Ethics Committee, P/Bag 201, Luwinga, Mzuzu 2; Email address: mzunirec@mzuni.ac.mw

Yours Sincerely,

Gift Mbwele

SENIOR RESEARCH ETHICS ADMINISTRATOR

For: CHAIRMAN OF MZUNIREC