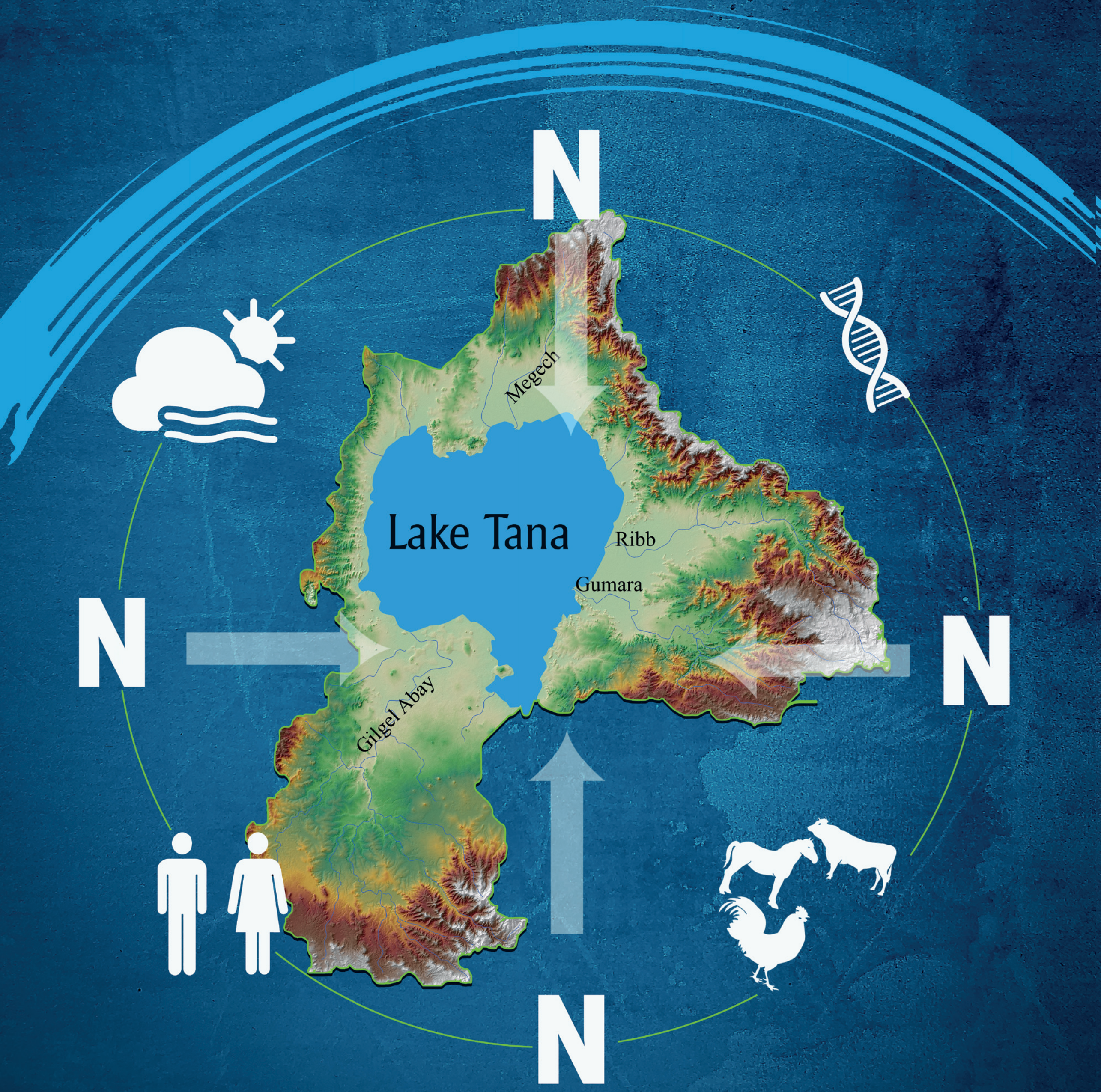
# ANALYSING WATER QUALITY AND ITS IMPACT TOWARDS SUSTAINABLE BASIN MANAGEMENT

## Modelling and experimental approaches in Lake Tana Basin, Ethiopia




Goraw Goshu

**Propositions**

1. Managing eutrophication in large lakes requires understanding the spatial heterogeneity of the lake.
   (this thesis)
2. Model validation in data-poor regions is only possible when experiments are part of the study design.
   (this thesis)
3. Effective science is boosted through science-policy debates.
4. Most Scientific publications have an insignificant impact on the livelihoods of the rural community in developing countries.
5. Large transboundary infrastructure projects in Africa require accounting for cultural differences between the involved countries.
6. The expansion of social media enlarges the difference between rural and urban communities.

Propositions belonging to the thesis, entitled

Analysing water quality and its impact towards sustainable basin management:

Modelling and experimental approaches in Lake Tana Basin, Ethiopia

These propositions are considered opposable and defendable and have been approved as such by the promotors prof. dr. AA Koelmans (bart.koelmans@wur.nl) and Prof. dr. C. Kroeze (carolien.kroeze@wur.nl) and co-promotors J.J.M de Klein and M. Strokal.

Goraw Goshu Yemer

Wageningen,

18 October,2022

# Analysing water quality and its impact towards sustainable basin management: Modelling and experimental approaches in Lake Tana Basin, Ethiopia

**Goraw Goshu Yemer**

**Thesis Committee**

**Promotors**
Prof. Dr AA Koelmans
Professor of Aquatic Ecology and Water Quality Management
Wageningen Universiteit & Research

Prof. Dr C Kroeze
Professor of Water systems and Global Change
Wageningen Universiteit & Research

**Copromotors**
Dr JJM de Klein
Assistant Professor, Aquatic Ecology and Water Quality Management
Wageningen Universiteit & Research

Dr M Strokal MSc
Assistant Professor, Water Systems and Global Change
Wageningen Universiteit & Research

**Other members**
Prof. Dr AJF Hoitink, Wageningen Universiteit & Research
Dr G. Gettel, UNESCO-IHE, Delft, the Netherlands
Prof. Dr H Smidt Wageningen Universiteit & Research
Dr T. Tang, IIASA, Vienna, Austria

This research was conducted under the auspices of the Graduate School Wageningen Institute for Environment and Climate Research.

# Analysing water quality and its impact towards sustainable basin management: Modelling and experimental approaches in Lake Tana Basin, Ethiopia

**Goraw Goshu Yemer**

**Thesis**
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University,
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Tuesday 18 October 2022
at 11 a.m. in the Omnia Auditorium.

Goraw Goshu Yemer
Analysing water quality and its impact towards sustainable basin management: Modelling and experimental approaches in Lake Tana Basin, Ethiopia
219 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands
(2022)
With references, with summary in English

ISBN: 978-94-6447-365-0
DOI: 10.18174/575340

Dedicated to the memory of my late father and mother

# Table of contents

Chapters 2, 3 and 5 have been published as peer reviewed scientific articles. Chapter 4 is accepted to the journal of Hydrology. The text, figures and tables of the published and the submitted articles have been adjusted to the PhD thesis format. References should be made to the original articles.

# 1

# General Introduction and Thesis Outline

## 1.1 Background

### 1.1.1 Global issues of water quality

Water resources across the world are under pressure due to human activities that have resulted in the pollution of surface waters. Globally, eutrophication and faecal pollution remain major problems and pose continuing risks to human health, limit food production, and hinder economic growth. Additionally, newly emerging pollutants like pharmaceuticals, pesticides, and industrial and household chemicals represent a new water quality challenge, with unknown long-term impacts on human health and ecosystems (Zimmerman et al., 2008). Moreover, changing climate patterns have a potential impact on water quality. The problems are undeniably greatest in the developing world, where traditional sources of water pollution, such as industrial emissions, poor sanitation, inadequate waste management, and contaminated water supplies, affect large numbers of people (Briggs, 2003).

One such source of water pollution is eutrophication, which causes damage to many of the world's freshwater and coastal ecosystems (Chislock et al., 2013). It has undesired effects on an ecosystem's state and services, public health, and socioeconomic activities. Some of these environmental effects in lakes include the risk of harmful algal blooms, which lead to a decrease in water clarity, an increase in hypoxia and fish kills, the alteration of biogeochemical processes, and a disruption in the aquatic food web (Fetahi, 2019). Eutrophication is caused by an excess supply of nutrients, particularly nitrogen (N) and phosphorous (P), to surface waters (Goshu et al., 2020; Lürling and van Oosterhout, 2013). The extent of eutrophication could increase as a result of increased nutrient inputs from the intensification of agriculture and the increase of the human population. Eutrophication is broadly caused by non-point and point sources of pollution. Non-point source pollution is defined as pollution that can be caused by a variety of activities that have no specific point of discharge, such as agriculture and urban areas. Point source pollution refers to the pollution that occurs from a single identifiable source, often sewage systems or industry sites.

The faecal pollution of water bodies is another water quality challenge. Faecal pollution is considered one of the worst forms of pollution in water bodies due to the potential spread of waterborne diseases that can cause public and ecosystem health problems (Bianco et al., 2020; Paruch et al., 2019). It is estimated that each year more than 842,000 people die from diarrhoea globally (UN-Water, 2017). There are millions of annual deaths from water-related illnesses and a growing risk of regional and international conflicts over scarce, shared water supplies (Carlton et al., 2012). Faecal pollution is caused by the contamination of water with faecal matter originating either from animals or humans (i.e., due to agriculture and urban land use, respectively). The contamination of a water body with faecal matter may occur at any time; however, the survival of microbial contaminants largely depends on the physical and chemical conditions of the water.

### 1.1.2 Water quality issues in Ethiopia

Ethiopia has huge surface and groundwater resources: about 124.4 billion $m^3$ of river water, 70–88 billion $m^3$ lake water, and 30 billion $m^3$ groundwater resources, with a potential to develop 3.8 million ha of irrigation and 45,000 MW of hydropower production (Berhanu et al., 2014). However, at the same time, water resource development in Ethiopia faces one or a combination of natural, topographic, and social challenges. The natural challenge consists of the spatial and temporal variability of the water resources' availability. In regard to topography, there is land degradation on the upper stream of sub-basins and improper human settlement on the highlands. Finally, the social challenge is the pollution of water resources from diffuse and point sources that reduces the use of water for socio-economic developments and poses public and ecosystem health risks.

The major water quality issues in Ethiopia are eutrophication (Fetahi, 2019; Goshu et al., 2020), faecal pollution (Abera et al., 2017; Abera et al., 2014; Goshu et al., 2021; Mengesha et al., 2004), high sediment load (Belete, 2013; Degife et al., 2021; Goshu and Aynalem, 2017; Lemma et al., 2019; Yitaferu, 2007), and contamination with agrochemicals, pesticides and trace metals (Mengistie et al., 2017; Merga et al., 2021; Teklu et al., 2021).The sources of eutrophication in Ethiopia's water bodies are agriculture, urban use, and industrial activities (Ayele and Atlabachew, 2021; Fetahi, 2019). The use of fertilizers in Ethiopia has grown significantly, from 3,500 tons in the early 1970s to approximately 34,000 tons in 1985 and then from 140,000 tons in the early 1990s to approximately 650,000 tons in 2012 (Diao et al., 2013). The amount of fertilizer used in 2021 has reached approximately 1.92 million tonnes (CSA, 2021). Moreover, the livestock system in Ethiopia is open grazing, and Ethiopia has the fifth largest livestock (i.e. cattle, sheep, and goats) population in the world; animal manure is thus a significant source of nutrients and faecal organisms for lakes and reservoirs (Goshu et al., 2020). From 1995/1996 to 2012/2013, Ethiopian livestock numbers grew from 54.5 million to over 103.5 million heads with an average annual increase of 3.4 million heads. Furthermore, the number of livestock is projected to exceed 157.4 million heads by 2024/25 (Leta and Mesele, 2014). Therefore, more animals likely signify more eutrophication and faecal pollution.

The effects of high nutrient loads are visible in lakes and reservoirs across Ethiopia. These include decreased water transparency, the production of harmful algae (toxins), a bad odour and taste, and disrupted oxygen balances (Fetahi, 2019; Wondie et al., 2007). The effects of faecal pollution in Ethiopia primarily consist of public and ecosystem health risks (Goshu et al., 2010b; Mushi et al., 2021). The major sources of faecal pollution in Ethiopia are agriculture, urbanization, and domestic and industrial wastes (Goshu et al., 2010a; Goshu et al., 2021; Mushi et al., 2021). Sediment transport is also a serious threat to many water bodies in Ethiopia (Lemma et al., 2019). Like faecal pollution, the presence of agrochemicals, pesticides, and trace metals in Ethiopian water bodies has also affected public and ecosystem health (Merga and Van den Brink, 2021; Nigatu et al., 2016).

### 1.1.3 Lake Tana basin characteristics and water quality problems

**Lake Tana basin characteristics**

Lake Tana is located in northern Ethiopia (see Fig. 1-1) and forms the source of the Blue Nile. With six major rivers, 40 seasonal streams, and extensive wetlands and flood plains, it is the major freshwater resource in the country. The basin provides economic, social, ecological, and environmental benefits to the local community. The fresh water and wetlands support fisheries, recession agriculture, communal grazing, water supply (for domestic and industrial use, hydropower generation, and irrigation), reed harvesting, navigation and boating, sand mining, urbanisation, settlement, and investment in agriculture, tourism, and industry. The non-consumptive uses include environmental education, scientific research, ecological conservation and monitoring, use as a cultural museum and spiritual place, and for recreation and tourism (Stave et al., 2017). Finally, hydropower development depends largely on the water resources of Lake Tana.

The Lake Tana basin has an agricultural land area of 234,948 ha that can potentially be irrigated, though the total developed area under irrigation is less than 4% of the potential (Molla and Minilik, 2004). Moreover, Lake Tana is home to 26 fish species, of which 15 are unique fish species flocks (Mengistu et al., 2017). About half the endemic fish species migrate upstream of the Lake Tana tributaries during the rainy season for reproduction (Goshu et al., 2010c). Lake Tana is rich in plant biodiversity and has also some old, indigenous trees in the monasteries and along the rivers. The lake is an important area for birds; a total of 213 species have been recorded, though total numbers are thought to exceed 20,000 seasonally (Francisco and Shimles, 2007). There are also approximately 16 large mammals and one large reptile, namely the Nile crocodile, living within the basin.

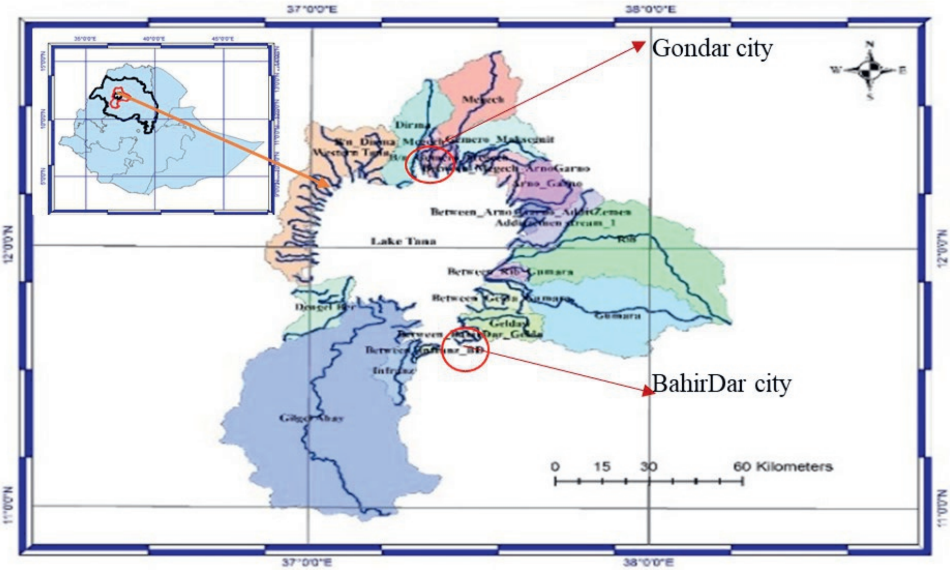



Fig. 1-1–The drainage area of the Lake Tana basin and the 20 sub-basins draining into the lake.

Source: modified from Goraw et al. (2020).

**Lake Tana's current situation and problems**

Economic reform, environmental degradation, and population growth are the major drivers of high nutrient and bacteria concentrations in Lake Tana (Stave et al., 2017), although the concentrations demonstrate spatial and temporal variability (Dersseh et al., 2019; Goshu et al., 2017; Stave et al., 2017). Agriculture is the major economy of Ethiopia; it directly supports 80% of the Ethiopian population in terms of employment and livelihood, and it contributes about 50% of Ethiopia's gross domestic product (MoFED, 2008). Nonetheless, Ethiopian agriculture has remained subsistent, leaving little if any surplus for sale or trade, for a long time (Dorosh and Rashid, 2013). Food insecurity has forced the government to look for strategies to transform Ethiopian agriculture from its current subsistent orientation into a market-orientated production system. Some of the interventions that the government have undertaken to increase agricultural production include more use of agricultural fertilizers that ultimately find their way into the water system.

Moreover, the floriculture industry is booming in the Lake Tana basin. Some of the farms take water directly from Lake Tana, and some have constructed deep wells as their source of water. However, in return, the farm waste also pollutes the water with nutrients and other chemicals.

In addition to the above drivers of current Lake Tana's situation and problem, there is watershed degradation in the Lake Tana basin that is manifested in different forms, such as extensive soil erosion, sediment load, soil fertility losses, declining land productivity, the disappearance of wetlands, the drying of streams in the upper slopes, and the seasonal flooding of the plains (Stave et al., 2017; Yitaferu, 2007). Currently, Lake Tana faces multiple problems in regard to its water quality. One such problem is the occurrence of toxic algal blooms on the lake's shore areas and river mouths, primarily during the pre-and post-rainy seasons. The algal bloom is caused by excess loads of nutrients. Lake Tana is an oligotrophic lake (Wondie et al., 2007; Teshale et al., 2002; Wudneh, 1998; Nagelkerke, 1997), and based on Cunha et al.'s (2013) trophic status classification, the shore areas and river mouths can be considered to be eutrophic.

Another manifestation of Lake Tana's water quality deterioration is the occurrence and expansion of water hyacinth in the north and north-eastern parts of the lake. The coverage of the weed has increased yearly since it was first observed in 2010 (Wondie et al., 2012). The east of the lake once had very good coverage of native macrophytes, which are now almost non-existent. At its most invasive, the weed has intruded approximately 200 meters and has expanded from the north-eastern to the eastern shore of the lake, increasing from 278 ha in 2015 to 2505 ha in 2019 (Dersseh et al., 2020).Moreover, faecal pollution presents another water quality problem in the lake. There are frequent outbreaks of diarrheal diseases on the Lake Tana islands and the lake adjacent areas where the community largely depends on raw water.

A further water quality problem in the Lake Tana basin is the high sediment load. The tributary rivers, river mouths, and shore areas of the lake appear very turbid during the rainy and post-rainy seasons. This is due to the high sediment load exported from the poorly protected watershed to the receiving surface water bodies found in the lake basin. Sedimentation is a threat to the lake and its associated ecosystems.

The aforementioned problems are exacerbated by the population pressure in the Lake Tana basin (Berisso, 1995; Yitaferu, 2007; Stave et al., 2017), which showed a significant increase between 1994 and 2007 (Anteneh, 2017). According to Ethiopia's 1994 and 2007 Central Statistical Agency (CSA) data, the population density in the Lake Tana basin was 166 people per $km^2$ in 1994 and 228 people per $km^2$ in 2007. It appears that population pressure in the Lake Tana basin has increased more than three times that of the national average (Anteneh, 2017). The consequences are more food production (i.e. higher nutrient loads) and more domestic waste (i.e. nutrients and faecal pollution).

Notably, there have already been ecological (e.g. water hyacinth infestation) and public health (e.g. frequent outbreaks of diarrhoea) problems and socio-economic impacts (e.g. restricted access to fish landing sites, fish stock decline, and reduced tourism) in the Lake Tana basin as a result of eutrophication and faecal pollution (Anwar Nuru and Yimer, 2012; Asmare, 2017; Gezie et al., 2018). To sustainably manage the water resources in general and the water quality in particular, understanding the existing situation, analysing future scenarios, and developing a basin plan is crucial. However, water management and the design of nutrient and faecal pollution management strategies have been challenged by a lack of spatially and temporally representative data alongside the tools necessary to predict and analyse future pollution loads. Data on nutrient export, retention, and loads to the receiving water bodies and their effect on the ecology of the lake are lacking. The performance of faecal pollution indicators of water quality is not adequately known in tropical waters. It is hence difficult to design sufficient measures to reduce eutrophication effects and the level of pollution in the surface waters of the

Lake Tana basin. It can conclude that there was no extensive dataset of water quality in the Lake Tana basin available until now.

### 1.1.4 Current state of research and modelling

**Nitrogen and phosphorous biogeochemical pathways**

Nitrogen and phosphorus in soils are present in a variety of forms, and the distribution of N and P between these forms changes with time and soil development. The eventual erosion of soil material and transport by rivers delivers N and P to the water bodies. Riverine N and P occur in two main forms: particulate and dissolved (inorganic and organic).

Nitrogen and phosphorus are essential for the maintenance of human activities due to their role in the production of food, feed, and synthetic chemicals. Artificial N and P are the main components of the earth's biogeochemical cycles (Bouwman et al., 2013). In recent decades, these N and P cycles have been strongly influenced by human activities (Galloway et al., 2008). Between 1961 and 2009, the global N input to agricultural land increased by a factor of 4.4 (Lassaletta et al., 2016). The global spread of anthropogenic phosphorous loads to freshwater from agriculture, industry, and domestic sectors for the period 2002–2010 is indicated in Fig. 1-2.

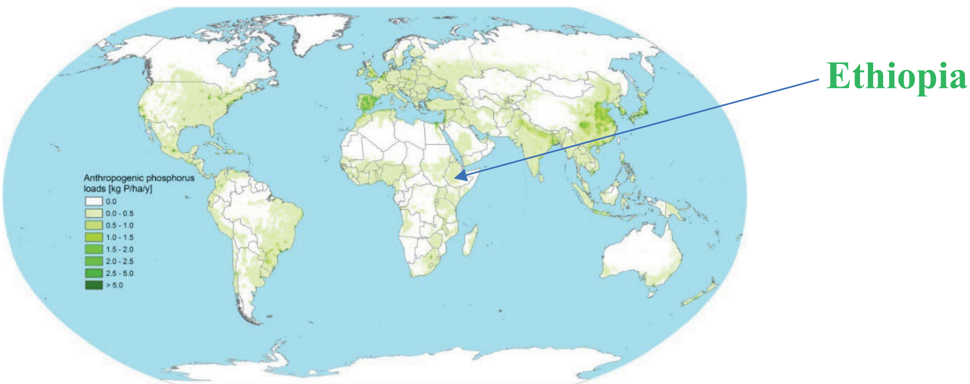


Fig. 1-2 – Global spread of anthropogenic phosphorus loads to freshwater from agriculture, industrial, and domestic sectors at a 5 × 5 arcmin grid. Period: 2002–2010. Source: modified from Mekonnen and Hoekstra (2017).

It is estimated that half the N and over half the P fertilizers applied to agricultural land are lost to the atmosphere, the soil, and the receiving surface and groundwater system (Nan et al., 2019). The amount of N and P that enter the aquatic ecosystem is determined by the geology and land use within the watershed. Nutrient transport is driven by several complex processes, including rainfall-runoff patterns, manure and fertilizer application, soil-water interactions, and crop and forage growth. It is a challenge to understand all these processes and quantify the eventual loading to water bodies. Nutrient export models (such as Global NEWS-2, MARINA, and NEWS2-DIN-S) can be useful, but it is still unclear if and how these can be used efficiently.

**Alternative stable states and critical loads of nitrogen and phosphorous**

One consequence of high nutrient loads is that lakes switch from a clear state with submerged macrophytes to a turbid state dominated by phytoplankton (alternative stable states; Scheffer, 1998). Each stable state has several buffer systems that keep it stable through top-down and bottom-up processes (Scheffer, 1998). The switch from the clear to the turbid state develops if P and N loads exceed their critical loads (Janse et al., 2008).To estimate the critical nutrient load of a specific lake, it is important to know which type of load-response curve to employ. The literature distinguishes three types of load-response curves that differ in linearity and the presence of hysteresis (Fig. 1-3; e.g. Scheffer et al., 2001), the third of which is most relevant for shallow lakes (Fig. 1-3c).

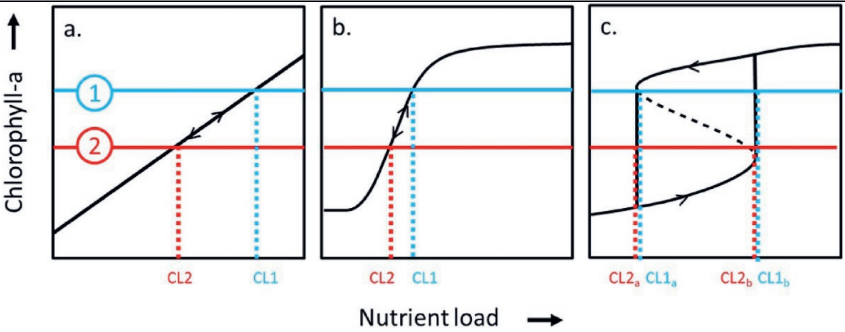



Fig. 1-3 – Three types of load-response curves: a) linear; b) nonlinear without hysteresis; c) nonlinear with hysteresis (i.e. alternative stable states). The horizontal, solid lines depict two human-defined restrictions for the maximum allowable chlorophyll-a that are strict (blue, number 1) and less strict (red, number 2). The blue and red dotted lines show the corresponding critical nutrient load (CL). Source: Scheffer et al. (2001).

As Fig. 1-3c indicates, the third load-response curve is nonlinear but includes hysteresis that emerges from strong positive feedback leading to two alternative stable states. Lakes exhibiting a nonlinear load-response curve with hysteresis have two critical nutrient loads, one for eutrophication and one for oligotrophication (denoted by CLa and CLb in Fig. 1-3c). In this case, the critical nutrient load is nearly independent of human-defined chlorophyll-a limits due to the abrupt character of the shift (denoted by the blue and red lines in Fig. 1-3c, with no differences between the critical nutrient loads). A good application of lake ecosystem models to quantify critical loads is very much needed in Lake Tana as this will help to set targets for future loads.

**Models for nutrient exports and lake eutrophication**

The export of nutrients and critical loads and their effect on the receiving water bodies can be studied using land- and water-based models. Different nutrient-related models exist and can be used to better understand water quality issues and their causes and solutions. Moreover, models can be used to project the future, allowing for the assessment of the effects of policy measures, and can also help to support decision-making, especially when observation data is scarce. The models in question vary in regard to their purpose, level of complexity, data requirements, and spatial and temporal aggregations. Some models that are commonly used to assess the nutrient export from watersheds are Global NEWS-2 (Global Nutrient Export from WaterSheds; Mayorga et al., 2010b), MARINA (Model to Assess River Inputs of Nutrients to seAs [or lAkes]; Strokal et al., 2016; Li et al., 2019; Wang et al., 2020; Strokal et al., 2021), NEWS2-DIN-S (Nutrient Export from Watersheds-Dissolved Inorganic Nitrogen-Seasonal model; McCrackin et al., 2014), SWAT (Soil and Water Analysis Tool; Douglas-Mankin et al., 2010), and SPARROW (SPAtially Referenced Regressions On Watershed attributes; Schwarz et al., 2006).

The aforementioned models can be used to explore solutions based on scenarios and to serve as decision support tools. Most of the models above quantify river exports of nutrients while accounting for human activities on land and nutrient retentions during the movement from land to rivers and seas. The temporal resolution of Global NEWS-2, MARINA, and SPARROW models is annual, whereas NEWS2-DIN-S and SWAT allow a more seasonal analysis of nutrient fluxes. The spatial resolution is basin-scale for models such as Global NEWS-2, SPARROW, and NEWS2-DIN-S, whereas MARINA can take a sub-basins scale. The spatial scale of SWAT is formed by hydrological response units that are generally smaller than sub-basins. Moreover, Global NEWS-2, MARINA, NEWS2-DIN-S, and SPARROW are steady-state models. In contrast, SWAT is an example of a dynamic model that shows the changes in nutrient transport over time.

Regarding source attribution, SPARROW and SWAT focus on the diffuse-source pollution from agriculture. Uncollected human waste is not accounted for in Global NEWS-2, and SPARROW provides the source attribution for nutrient river exports. The recent version of MARINA is applied to sub-basins globally and focuses on point-source pollution from sewage systems and open defecation (Strokal et al., 2021).

Furthermore, as previously stated, the different models have different complexities and data requirements. Global NEWS-2 has some level of complexity, whereas MARINA and NEWS2-DIN-S take lumped approaches to represent processes of nutrient retention and nutrient export from land to water. The data requirement of Global NEWS-2, SWAT, and SPARROW is extensive, and that of MARINA and NEWS2-DIN-S is less extensive.

Nutrient models also take different approaches to quantify nutrient flows from land to water. An example is a steady state, namely a state or condition of a system or process that does not change over time or a condition that changes only negligibly over a specified time. Global NEWS-2 is one of the models that is based on the steady-state approach (Mayorga et al., 2010). Global NEWS-2 considers denitrification and reservoirs for nutrient losses from and retentions in rivers, and it also considers water consumption. Similarly, MARINA, NEWS2-DIN-S, and SPARROW are steady-state models, whereas SWAT is an example of a dynamic model.

In addition, the nutrient forms and applications vary among different models. For example, the Global NEWS-2 model quantifies river exports of multi-elements in different forms for past and future years. The model considers N, P, silicon, and carbon in dissolved organic and inorganic and particulate forms for the period 1970–2050. The model also quantifies the potential for coastal eutrophication. Future trends are analysed based on four Millennium Ecosystem Assessment (MEA) scenarios (Alcamo et al., 2007; Seitzinger et al., 2010). MARINA quantifies total nitrogen (TN), total phosphorous (TP), dissolved inorganic phosphorous (DIP), and dissolved inorganic nitrogen (DIN; Janse et al., 2010; Strokal et al., 2016). NEWS2-DIN-S quantifies dissolved inorganic nitrogen (DIN; McCrackin et al., 2014).

In previous literature, Global NEWS-2 has been widely applied for both global and regional analyses of river nutrient export. The regional studies of riverine nutrient export to coastal waters include those of Douglas-Mankin et al. (2010), Schwarz et al. (2006), and Mayorga et al. (2010). However, the basin scale limits spatially explicit analyses of nutrient sources within a basin (Mayorga et al., 2010; Steiniger, 2010). The spatial resolution of the Global NEWS-2 model has been improved by MARINA, which is a downscaled version of Global NEWS. The NEWS 2-DIN-S model is a seasonal version of Global NEWS-2 for DIN.

For Lake Tana, many large-scale existing models for nutrient export by rivers are coarse, and most of them are annual. Other models that are applied regionally or locally (e.g. SWAT) are more detailed and require considerable amounts of input data, making them difficult to apply in data-poor regions, such as Lake Tana. There are two exceptions. First, MARINA-Lakes, a sub-basin model for lakes, was successfully applied to a number of lakes in China (Yang et al., 2019; Li et al., 2019; Wang et al., 2019; Strokal et al., 2021; Ma et al., 2020). The MARINA-Lakes model considers the spatial variability in human activities and hydrology for river export of nutrients to lakes from sub-basins. There are also seasonal applications of MARINA-Lakes for two lakes (Li et al., 2019; Wang et al., 2019) and one river (Chen et al., 2019) in China. However, the annual and seasonal versions of the MARINA model have not been developed for Lake Tana. Second, a seasonal modelling approach exists for global rivers: NEWS 2-DIN-S (McCrackin et al., 2014). This modelling approach is for large basins and yet it does not consider the spatial variability among basins. Therefore, the development of an integrated approach aiming to assess the seasonality and sources of DIN from sub-basins to Lake Tana is crucial.

The details (i.e. purpose, modelling approach, spatial and temporal scales, complexity, data requirement, nutrient form, and nutrient source) and strengths and weaknesses of the models used in this thesis are described in the general discussion and conclusion (Chapter 6). We model nutrients – but not pathogens – in the Lake Tana basin, and we collected bacteria data through a monitoring program. Notably, the Lake Tana basin is a data-poor region in general. For modelling nutrients, measurements are available for the year we set up the model; however, fewer applicable models are available for pathogens, and, additionally, data for the tributary rivers is scarce.

To explore ecological state transitions and set critical loads in lakes, lake quality and ecosystem models are required. There are different lake models (such as PCLake+, NiRReLa, SiRReLa, VEMALA v3, and Delft3D-WAQ/ECO) that have been reviewed by Janssen et al. (2015; 2019). The authors concluded that the models have different complexities, data requirements, spatial and temporal scales, among other factors. The models are grouped into two categories: statistical models and dynamic models. The statistical models are generally simple and may point to causal relationships. Statistical techniques do not necessarily reveal an understanding of the true underlying biological processes. Consequently, using statistical models for projections of algal blooms is generally not recommended (Janssen et al., 2019).However, unlike statistical models, the process-based models are complex but have the advantage of capturing the biological response. Moreover, Janssen et al. (2019) asserted that dynamic models better capture nonlinear responses and that spatial heterogeneity is worth considering. This is relevant to our study because the food web model is needed to address spatial heterogeneity and state transitions and to identify critical loads. From the dynamic models, we selected PCLake+ to study the critical loads and spatial heterogeneity of eutrophication in Lake Tana. We prefer PClake+ to other aquatic models because PClake+ is a complete food web model with feedback, and it gives us the ability to set critical thresholds. In addition, our supervisory team has expertise and experience in PClake+.

In extant research, most lake models are 0D and do not consider heterogeneity in the lake. Analysing the effects of N and P nutrients on shallow, non-stratified lakes and showing the importance of the spatial heterogeneity of eutrophication in a large tropical lake have not yet been undertaken for Lake Tana. This information gap requires the novel approach of coupling the ecosystem model PCLake+ with a more spatially explicit water flow model, in this case Duflow (Clemens et al., 1993). Duflow is a modelling platform for 1D flow and quality and is used in this research to define Lake Tana's impact zones (i.e. the areas of the lake that are under the specific influence of the feeding tributary river). To define the impact zones, we constructed a 2D application of Lake Tana using Duflow. The models we selected to study Lake Tana's state transition and critical loads need seasonal inputs of nutrients from gauged and ungauged catchments. These seasonal inputs of nutrients were compiled by introducing seasonality in the nutrient transport models.

**Faecal Indicator Bacteria, Pollution Risk Mapping, and Microbial Source Tracking**

The concentration of faecal indicator bacteria is useful for pollution risk mapping. The use of faecal coliforms *E. coli*, *C. perfringens*, and Enterococci, among others, as indicator bacteria for the assessment of faecal pollution and possible water quality deterioration in various freshwater sources is a widely used and accepted concept in temperate regions (APHA, 1995; Toranzos and McFeters, 1997). However, recent studies (Hazen, 1988; Rivera et al., 1988; Byappanahalli and Fujioka, 1998; Ahmed et al., 2008) of tropical freshwater have shown that high proportions of faecal Coliform-positive isolates may be of non-faecal origin, and presumptive *E. coli* can become a normal inhabitant of tropical waters, as reported for pristine environments in some tropical waters. This apparent unreliability of traditional faecal pollution indicators in tropical conditions should prompt the performance evaluation of the standard indicators that are not well known in tropical countries like Ethiopia.

Pollution risk mapping as applied by Kavka and Poetsch (2006) was often based on one parameter, most frequently on presumptive *E.coli* quantification. However, taking into account the uncertainty in the indication value of *E.coli* and the nature of the water sources, the development of a dual faecal mapping system is very much necessary.

The persistent problem of faecal pollution and its consequences for public health are not only due to the lack of reliable indicators but also partly to the inability of the indicators to identify the source of faecal pollution (Bernhard, 2000). To this end, a plethora of genetic faecal markers have been developed for temperate waters; these include BacR (ruminant-associated faecal pollution), PigIIBac (pig-associated faecal pollution), and BachHUM or HF 183 Taqman (human-associated faecal pollution). The performance of these qPCR assays has not yet been evaluated in a high land tropical country, such as Ethiopia, and the following are the knoweldge gaps to be addressed in this PhD thesis.

## 1.2 Knowledge gaps

The current ecological condition of Lake Tana, as described above in regard to water quality, clearly indicates that there is nutrient pressure from the local land uses. The nutrient pressure is manifested in the form of eutrophication effects on the current ecological condition of the lake, especially in the shore and river mouth areas. It is unclear how the lake will evolve in the face of high population growth, increased socioeconomic development, and climate change, which will have a synergistic effect in terms of creating nutrient pressure for Lake Tana. Therefore, it is necessary to develop nutrient modelling tools for the quantification of nutrients sources for the current and future situations of Lake Tana, preferably on a seasonal and sub-basin scale. There are some nutrient export models available; however, as asserted above, these lack either temporal or spatial resolution. This study must therefore fill this gap by developing a seasonal and sub-basin scale model for Lake Tana Basin.

A lake with the ability to maintain its state in the face of both internal change and external shocks and disturbances is said to be resilient. A key research question of this thesis is to what extent Lake Tana can resilient before it collapses. The analysis of Lake Tana's resilience requires an ecological model, such as PCLake+. Thus far, PCLake+ has been used as a 0D model and has only very limited applications that take into account heterogenous lakes like Lake Tana. There is therefore a need to apply PCLake+ in a more dimensional setting to account for spatial heterogeneity. In doing do, proper critical nutrient loadings can be set, and the lake's resilience in future conditions can be analysed.

The main knowledge gaps that this study aims to address in regard to water quality variations (diffuse and point sources of pollution) and options for sustainable management of the Lake Tana basin are as follows;

1. There is no systematic overview of the available water quality data in the Lake Tana basin;
2. There is a limited understanding of seasonal effects on river export of N and P to Lake Tana from sub-basins and their sources. Such information (i.e. monitoring data on N and P loads) is limited for Lake Tana, especially for ungauged sub-basins;
3. Data on river export of N and P to Lake Tana are lacking. The setting of critical thresholds of lakes and reservoirs needs N and P load data from gauged and ungauged sub-basins. This requires a modelling approach, so it is therefore complicated to set critical thresholds. Moreover, data on the spatial variability of nutrient loading and its effect on Lake Tana is lacking;
4. There is a lack of information on the performance (i.e. discrimination ability) of faecal pollution indicators in tropical waters. Furthermore, the level of faecal pollution and its risk to surface and groundwater systems in the Lake Tana basin are not well known.

## 1.3 Research objectives

This Ph.D. research aims to address the above-mentioned knowledge gaps. The overall research goal is, therefore, to understand better river export of nutrients to Lake Tana by sub-basin, season, and source and to analyse the importance of the spatial heterogeneity of eutrophication, the performance evaluation of faecal pollution indicators, and the risk mapping of faecal pollution.

Four research sub-objectives were formulated to achieve the overall aim. These are as follows:

1. Perform a systematic overview of the water quality status of the Lake Tana basin through a literature review;

2. Develop and apply a basin model for quantifying river export of N and P to Lake Tana by sub-basin, season, and source;

3. Assess the spatial variability of eutrophication in tropical shallow lakes, taking Lake Tana as a case study and determining critical loads of N and P for different zones of the lake;

4. Evaluate the performance of rapid and practical techniques to monitor faecal pollution in Lake Tana by testing them on a pollution gradient in low, mid-, and high altitudinal environments and water types.

## 1.4 Research approach

To address the above objectives, the following research approaches have been followed. In Chapter 2, I address sub-objective 1. I present a systematic overview of water quality variations in the Lake Tana basin based on available data and previous works in this area.

In Chapter 3, I discuss sub-objective 2. The main novelty here is the development of a new model that addresses the seasonal export of DIN from sub-basins to a lake. I developed a new model for river export of nutrients by sub-basin, source, and season by combining two existing modelling approaches with modifications for our study area. The two existing modelling approaches were MARINA's sub-basin scale modelling approach (Strokal et al., 2016; Li et al., 2018; Yang et al., 2019; Wang et al., 2019) and the seasonal modelling approach of the NEWS.2-DIN-S model (McCrackin et al., 2014). In this study, I added seasonality into the sub-basin scale MARINA based on the approach of McCrackin et al. (2014); this resulted in a seasonal version of MARINA for sub-basins and rivers discharging to Lake Tana. I used our newly developed model to generate sub-basin, season, and source-specific DIN data for 20 sub-basins of the Lake Tana basin, which may assist in designing effective nutrient management in the Lake Tana basin. I also added a new method of accounting for open defecation as a non-point source of pollution.

In Chapter 4, I explore sub-objective 3. I determine the spatial variability of eutrophication for a large, heterogeneous lake, namely Lake Tana. I identify critical thresholds of N and P for the impact zones in Lake Tana using a novel modelling approach that results in a 2D flow and quality model of Duflow and PCLake+. I also couple Duflow, PCLake+, and seasonal MARINA models.

Finally, in Chapter 5, I address sub-objective 4. I present new insights on the performance of faecal pollution indicators and pollution risks to the water systems. I aim to evaluate the performance of faecal pollution indicators, including total coliforms (TC), presumptive *E. coli* (EC), intestinal enterococci (IEC), presumptive Clostridia perfringens spores (CP), and qPCR assays (all Bac R and BacR). I analyse these in different surface and groundwater systems that are located at altitudes ranging from 1,100 m.a.s.l. (i.e. lowlands) to 3,835 m.a.s.l. (i.e. highlands) in a highland tropical country, namely Ethiopia, and I map the pollution risk based on consensus faecal indicator parameters.

## 1.5 Ph.D. thesis outline

The thesis has six main topics, each of which is assigned one thesis chapter (Fig. 1-4). This first chapter has provided a general introduction that contains background information and the thesis's objectives. Chapter 2 provides an overview of water quality variations and options for sustainable management. The third chapter assesses the seasonal export of DIN inputs to Lake Tana as influenced by human activities, the climate, and hydrology in the basin. Chapter 4 quantifies the potential impact of N and P inputs on the ecological status of the lake, which was achieved by applying the PCLake+ model to assess the critical N and P loads. Chapter 5 evaluates the performance of faecal pollution indicators to produce a "consensus picture" of faecal pollution and to monitor faecal pollution. Finally, the sixth chapter is a general discussion and conclusion of the thesis; it includes the main findings and novelties of the thesis, reflects on the modelling and experimental methods used in the thesis, and reaches a didactic conclusion to present a future outlook for the Lake Tana basin.

It is my hope that the results obtained from this thesis will contribute to the sustainable management of nutrients, faecal pollution, and drinking water in the Lake Tana basin. In addition, the new models devised for this thesis can be applied in other regions that experience similar environmental problems.

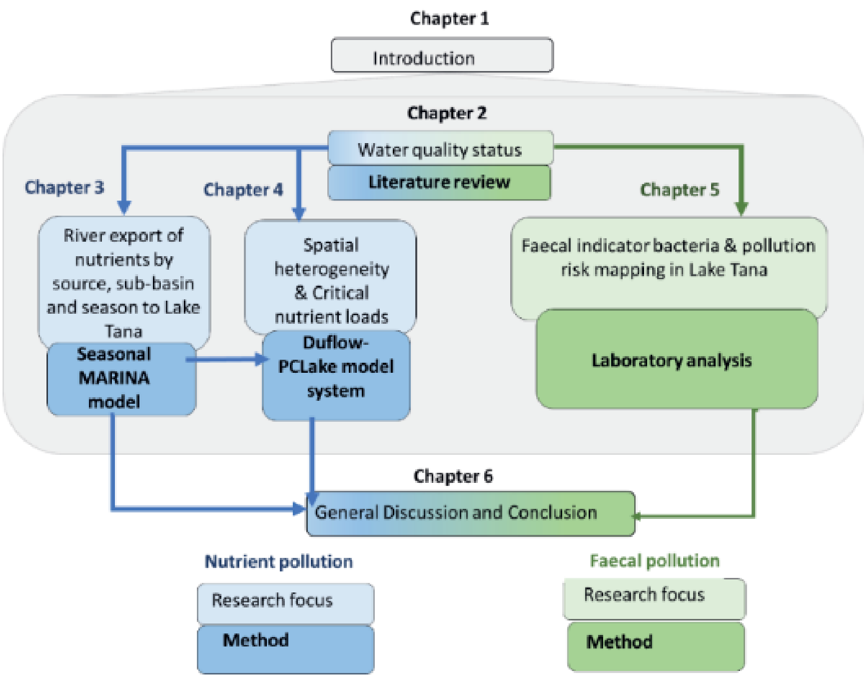



Fig. 1-4 – An overview of the Ph.D. thesis structure that provides the focus of chapters and the methods. Sub-objectives are given in Section 1.3.

# 2

# Water quality of Lake Tana Basin, Upper Blue Nile, Ethiopia. A review of available data

This chapter is published in revised form as:

Goshu, G., Koelmans, A.A., de Klein, J.J.M. (2017). Water Quality of Lake Tana Basin, Upper Blue Nile, Ethiopia. A Review of Available Data. In: Stave, K., Goshu, G., Aynalem, S. (eds) Social and Ecological System Dynamics. AESS Interdisciplinary Environmental Studies and Sciences Series. Springer, Cham. ISBN:978-3-319-45753-6, https://doi.org/10.1007/978-3-319-45755-0_10

**Abstract**

Water is at the forefront of the economic agenda of the Ethiopian government and the Tana basin has been identified as a major economic corridor because of the basin's immense water resource potential for socio-economic development. For effective and sustainable utilization of water resources in the basin, it is essential to assess the water quality in spatial and temporal dimensions. Nonetheless, scientific information on Ethiopian water bodies is rare and the available ones are of expeditious origin. In the Tana basin, there is no detailed and systematic characterization of water quality based on the long term and spatially representative data, due to the absence of a sustainable monitoring program. Despite the fact that data and studies are fragmentary, the available information on Lake Tana indicates that the trophic status of the lake has gradually changed to mesotrophic and eutrophic in some places due to nutrient loads. Sedimentation is a threat to Lake Tana. Faecal pollution and toxigenic cyanobacteria are detected in the lake especially in the shores and river mouths. Although the current dataset on water and Lake bottom sediment characteristics and quality impacts is spatially and temporally limited, the available information indicates the occurrence of potential anthropogenic pollution mainly on river mouths and shore areas. Impairment of water quality has been going on for years, which has significantly affected the basin's potential for agriculture, industry, hydroelectric, ecosystem, water supply, and recreation sectors. Sedimentation, eutrophication, faecal pollution, wetland encroachment, and hydrological alterations have been identified as main issues of water quality management in the Lake Tana basin. Efficient monitoring programs based on a practical selection of robust water quality indicators are recommended for the basin.

## 2.1 Introduction

The surface water potential of Ethiopia is estimated to encompass about 7400 $km^2$ of the lake area and a total river length of about 7000 km (Wood and Talling, 1988). The storage capacity of the lakes is in excess of 88 billion cubic meters with Lake Tana basin comprising more than 52 % of the total surface area or 33% of the total volume. Tana basin is the second-largest sub-basin of the Blue Nile.

Lake Tana has many unique characteristics. It is the source of the Blue Nile, home of unique fish species flock, feeding, and breeding ground for birds and it is the home of century-old Island monasteries dating back to the $13^{th}$ century with a unique cultural heritage.

Given the importance of Lake Tana, its tributary rivers, and groundwater resources to the region's and nation's economy, ecosystems, human health, and livelihoods, understanding its water quality across spatiotemporal scales for selected indicators is very critical for sustainable development.

Water quality is a dynamic balance of physical, chemical, biological, and hydrological characteristics and processes occurring in an aquatic system that can be defined in terms of certain user functions (Koelmans et al. 2001). Water quality indicators are useful tools for monitoring ecosystem integrity and as an indicator of the appropriateness of the water for the intended uses.

Scientific information about the water quality of Ethiopian lakes including Lake Tana is generally rare and the available ones are of expedition origin. Chemical and related biological features of Ethiopian water bodies have been reported in the literature over more than fifty years, but the studies have overall not been systematic nor sustained. Relatively, a suite of water quality studies has been conducted in South Ethiopia where the physico-chemical characteristics of the Rift valley lakes have been studied systematically and, in much detail, (e.g., Kebede etal.1992; Zinabu et al.2002). The physical, chemical, and biological characteristics of most water bodies in the Lake Tana basin, however, are not characterized systematically.

Most of our current understanding of water characteristics in the Lake Tana basin comes from irregular monitoring and short-term biological and hydrological expeditions conducted in the recent decade. Therefore, it is not possible to unambiguously assess whether the magnitude and direction of changes in the water bodies are due to natural or anthropogenic sources. Consequently, it is difficult to predict the possible consequences.

These water resources of Lake Tana are not utilized sustainably for the socio-economic development of the region and the nation at large. The development intervention outweighs environmental protection and management, which leads to resource degradation (Teshale etal.2002; Tadesse 2010). The sustainable management of Lake Tana basin water quality is challenged by easily perceivable extensive basin degradation, sediment load, competing for development needs, soil fertility losses and declining land productivity, the disappearance of wetlands, drying of streams in the upper slopes, and seasonal flooding of the plains. These problems are exacerbated by population pressure in the basin.

Lake Tana used to be an oligotrophic lake (Wondie etal.2007; Teshale etal.2002; Wubneh 1998; Nagelkerke 1997) but its trophic status has changed gradually. Especially river mouths have experienced seasonal eutrophication (Goraw 2012). This is caused by increased concentrations of phosphorous and nitrogen from hotels, recreation centres, and nonpoint sources with negative consequences on public and ecosystem health as well as on various sectors of the economy. Microbial contamination is also a factor for public and ecosystem health risks and the recurrent water-related disease outbreaks in the basin are most likely to be caused by such problems.

This review is the first to consolidate different data and information related to water quality and bring forward relevant development, research, and policy intervention ideas for sound management of water quality and water resources in the Lake Tana basin.

## 2.2 Review of Available Data

### 2.2.1 Morphometric characteristics of Lake Tana basin

Lake Tana is the largest water body (ca. 3,094 km$^2$) in Ethiopia (*Table 2-1*). It is located at 12°N, 37°15′E, and 1,830 m altitude in the north-western highlands of Ethiopia. The Lake Tana basin area covers 16,111 km$^2$ (*Table 2-1*). The distance from north to south is approximately 84 Km and east to west is 66 Km. It is a shallow lake with a mean depth of 8 m and a maximum depth of 14 m (*Table 2-1*). More than seven large, permanent rivers and approximately 40 small seasonal rivers feed the lake. The main tributaries to the lake are the Gilgel Abay (Little Nile River), Megech River, Gumara River, and the Rib River. Together they contribute more than 95% of the total annual inflow (Lamb et al. 2007). The Blue Nile is the only outflowing river.

Lake Tana's bottom substrate is volcanic basalt mostly, covered with a muddy substratum with a low organic matter content of 1% in 1994 (Howell & Allan 1994) and reached to 14 % in 2011(Goraw 2011a). The lake has high silt concentrations with a loading rate of 8.96 - 14.84 M tons of soil per year (Yitaferu 2007) and the trophic status is oligotrophic to mesotrophic (Ilona etal. 2011; Wondie etal.2007; Teshale etal.2002; Wudneh 1998; Nagelkerke 1997). The Lake Tana area has warm temperatures, and the mean annual rainfall is about 1564 mm, of which 59 percent falls in July and August when the rainfall can be 444-483 mm per month. The seasonal rains cause the lake level to fluctuate regularly with an average difference between the minimum, in May-June, and maximum in September-October of about 1.5 m before the Tana Beles hydroelectric power plant starts operation.

The land cover of the Lake Tana basin in 2013 includes farmland (32.23%), waterbody (19.57%), built-up area (18.6%), grassland (11.3%), forestland (7.3%), wetland (4.26%), shrub land (3.66%), woodland (2.1%) and plantation forest with natural trees (0.87%) (Wubneh and Goraw 2013 Land use/ cover change in Lake Tana basin, unpublished).

The lake is believed to be created by basalt outflow in the Pleistocene cutting of the basin at the southern extreme of the lake at Chara-Chara in Bahir Dar. In other words, it is thought to have originated by volcanic blocking of the Blue Nile River two million years ago (Mohr 1962 in Vijverberg etal.2009). Geologic evidence shows that these volcanic activities resulted from tectonic movements. Lake Tana is formed by damming of the Abay River by 50km long quaternary volcanic (Aden Volcanoes) basalt flow (Chorowicz etal. 1998).

The soils of islands, peninsulas, and surrounding wetlands and dry uplands of the lake are thought to be dominated by Nitosols, Luvisols, and Vertisols, but require further investigation (Mekonnen 2011).

The permanent and seasonal rivers and streams flow down in different geologic landscapes contributing to different habitat structures. All these structures support diverse riverine organisms, especially serving as the endemic fish labeobarbs spawning grounds. The rivers are also essential for the development of agriculture, transportation, and fishing, especially from the perspective of tourism, as they are rich in different tourist attractions resources, and they are essential for various tourism activities.

Table 2-1. Morphometric characteristics of Lake Tana and or Tana Basin (after various sources)

| Morphometric characteristics | Values | Reference |
|---|---|---|
| Maximum depth (m) | 14 | Vijverberg etal.2009 |
| Mean depth (m) | 9 | " |
| Latitude Longitude | 12°N, 37°15′E | " |
| Altitude (m) | 1800 | " |
| Lake area (AL) ($km^2$) | 3111 | Wubneh and Goraw 2013 |
| Basin area (Ac) ($km^2$) | 15114 | " |
| Lake volume (VL) ($km^3$) | 28.4 | Seifu etal.2005 |
| Water residence time (years) | 3 | " |
| Runoff coefficient (k) | 0.22 | |

### 2.2.2 Water Characteristics

**Physico-chemical characteristics of major Tributary rivers**

The physico-chemical characteristics of the major tributaries of Lake Tana are summarized in *Table 2-2*.There is limited data on spatially and temporally variability of the physicochemical water characteristics, hence not possible to describe in detail. Nonetheless, the available pooled data showed that there is significant pH, conductivity, total dissolved solids, dissolved oxygen, and temperature differences among the Lake Tana basin rivers (($P < 0.05$, n=3 -119, Kruskal-Wallis h test). The above differences are more attributed to the differences in the background geology. The pH, electrical conductivity, total dissolved solids, dissolved oxygen and temperature ranges from 6.43-8.93;10 to 1000 $\mu Scm^{-1}$ ;5 to 490 mg $L^{-1}$;1.1 to 8.6 mg $L^{-1}$ and 11.7 to 28.6 ºC respectively (*Table 2-2*). The pH of most natural waters can range from 6.0 to 8.5, although lower values can occur in dilute waters high in organic content, and higher values in eutrophic waters, groundwater brines, and salt lakes (APHA 1995). The measured pH values in the rivers are in the normal range but it does not mean that the rivers are free from contamination. The nutrient concentrations especially nitrogen and phosphorous are high and the turbidity is also very high. The inherent self-purification capacity of rivers should be also taken into account.

Table 2-2.*Physico-chemical characteristics of major tributary rivers of Lake Tana, analysis from different water samples collected during various times in the year 2010-2015.Abbreviations: Min-minimum, Max-maximum values, SD-standard deviation, and N-refers to the number of samples.*

| | G/Abay | | | | Gumara | | | | Rib | | | | Dirma | | | | Megech | | | | Arno | | | | Garno | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Min | Max | SD | N | Min | Max | SD | N | Min | Max | SD | N | Min | Max | SD | N | Min | Max | SD | N | Min | Max | SD | N | Min | Max | SD | N |
| $NO_3$ – N ($mgL^{-1}$) | 0.027 | 9.0 | 2.32 | 14 | 0.0 | 1.45 | 0.55 | 10 | 0.08 | 1.55 | 0.56 | 8 | 0.15 | 2.420 | 0.72 | 8 | 0.01 | 1.4 | 0.68 | 5 | 0.25 | 2.8 | 1.80 | 3 | 0.09 | 2.1 | 1.42 | 3 |
| $NO_2$ – N ($mgL^{-1}$) | 0.0 | 0.9 | 0.26 | 18 | 0.0 | 11 | 3.01 | 13 | 0.00 | 11 | 3.42 | 10 | 0.01 | 1.01 | 0.42 | 10 | 0.02 | 1.0 | 0.45 | 6 | 0.0 | 0.01 | 0.0 | 3 | 0.00 | 0.02 | 0.01 | 3 |
| $NH_3$ – N ($mgL^{-1}$) | 0.04 | 6.60 | 2.34 | 13 | 0.02 | 3.3 | 1.38 | 6 | 0.02 | 1.40 | 0.77 | 3 | 0.03 | 5.5 | 3.13 | 3 | 0.10 | 2.7 | 1.50 | 3 | 0.09 | 1.3 | 0.7 | 3 | 0.02 | 0.5 | 0.25 | 3 |
| $PO_4$-P($mgL^{-1}$) | 0.080 | 20.6 | 7.26 | 19 | 0.03 | 12 | 3.46 | 13 | 0.06 | 9.70 | 3.44 | 10 | 0.04 | 13.8 | 4.55 | 10 | 0.07 | 9.4 | 3.89 | 8 | 0.36 | 3.2 | 1.47 | 3 | 0.36 | 2.40 | 1.06 | 3 |
| Sulphate ($mgL^{-1}$) | 5 | 320.0 | 83.06 | 18 | 5.00 | 170 | 48.59 | 13 | 5.00 | 300 | 93.85 | 10 | 3 | 380 | 119.05 | 10 | 5.00 | 560 | 193.65 | 8 | 0.0 | 5 | 2.89 | 3 | 0.0 | 7 | 4.04 | 3 |
| Sulphide ($mgL^{-1}$) | 0.0 | 22.0 | 6.27 | 12 | 0.02 | 12 | 4.18 | 8 | 0.02 | 7 | 2.76 | 6 | 0.04 | 5 | 1.97 | 6 | 0.02 | 8.0 | 3.33 | 5 | 0.0 | 0.7 | 0.49 | 3 | 0.0 | 0.7 | 0.49 | 3 |
| $Fe^{2+}$( $mgL^{-1}$) | 0.0 | 11.0 | 4.35 | 17 | 0.00 | 6.5 | 2.00 | 10 | 0.00 | 5.80 | 2.08 | 7 | 0 | 18 | 7.95 | 7 | 0.00 | 35.2 | 13.98 | 6 | 0.0 | 1.52 | 0.82 | 3 | 0.0 | 1.05 | 0.54 | 3 |
| Total hardness ($mgL^{-1}$) | 60 | 1080.0 | 286.27 | 19 | 22.00 | 500 | 182.41 | 13 | 47.00 | 1350 | 391.67 | 10 | 80 | 1080 | 377.84 | 10 | 110 | 600 | 177.23 | 8 | 102 | 1080 | 562.92 | 3 | 108 | 380 | 153.16 | 3 |
| Alkalinity ($mgL^{-1}$Ca $CO_3$) | 35 | 1080.0 | 250.89 | 18 | 47.00 | 520 | 162.12 | 12 | 73.00 | 675 | 209.22 | 9 | 90 | 1730 | 620.42 | 9 | 132 | 2050 | 649.1 | 8 | 112 | 1300 | 679.35 | 3 | 153 | 800 | 363.86 | 3 |
| Temperature($^0$C) | 13.0 | 25.8 | 3.23 | 35 | 11.70 | 27.6 | 4.87 | 18 | 12.10 | 28.60 | 5.61 | 13 | 12.40 | 24.8 | 3.39 | 14 | 17.70 | 25.4 | 2.92 | 11 | 19.4 | 27.4 | 2.86 | 7 | 19.7 | 25.7 | 2.18 | 7 |
| $p^H$ | 6.63 | 8.8 | 0.59 | 35 | 7.40 | 8.83 | 0.39 | 18 | 7.12 | 8.87 | 0.49 | 13 | 6.40 | 8.68 | 0.57 | 14 | 7.20 | 8.93 | 0.57 | 11 | 7.88 | 8.86 | 0.44 | 7 | 7.81 | 8.51 | 0.28 | 7 |
| Dissolved Oxygen ($mgL^{-1}$) | 1.1 | 5.1 | 1.31 | 21 | 3.11 | 7.90 | 1.58 | 11 | 3.26 | 8.60 | 2.05 | 7 | 2.17 | 7.2 | 1.84 | 7 | 3.73 | 5.9 | 1.16 | 4 | 3.01 | 4.0 | 0.57 | 4 | 2.96 | 4.1 | 0.55 | 4 |
| Electrical Conductivity ($\mu Scm^{-1}$) | 12 | 284.0 | 72.15 | 34 | 10.0 | 280 | 75.97 | 20 | 11 | 490 | 153.92 | 14 | 100 | 1000 | 240.74 | 14 | 110.00 | 380 | 100.92 | 10 | 76.0 | 271 | 61.42 | 7 | 180 | 633 | 163.35 | 7 |
| Total Dissolved Solids ($mgL^{-1}$) | 0 | 180.0 | 46.50 | 39 | 5.0 | 180 | 46.83 | 20 | 6 | 240 | 75.83 | 14 | 40 | 490 | 118.56 | 14 | 50.00 | 190 | 51.19 | 11 | 49.0 | 198 | 45.42 | 7 | 100 | 411 | 110.17 | 7 |
| Salinity(ppt) | 0.0 | 0.13 | 0.04 | 15 | 0.01 | 0.1 | 0.03 | 6 | 0.01 | 0.09 | 0.04 | 3 | 0.05 | 0.18 | 0.08 | 3 | 0.05 | 0.14 | 0.06 | 3 | 0.08 | 0.1 | 0.01 | 3 | 0.09 | 0.31 | 0.13 | 3 |
| Turbidity (NTU) | 5 | 1002.0 | 280.50 | 31 | 8.45 | 993 | 355.36 | 17 | 3.06 | 869.00 | 347.83 | 11 | 2.96 | 1002 | 411.14 | 11 | 3.89 | 962 | 286.61 | 11 | 2.45 | 1002 | 397.89 | 6 | 2.13 | 1002 | 394.85 | 6 |

**Physico-chemical characteristics of Hand-dug wells and protected pumps**

The physico-chemical water characteristics of traditional hand-dug wells and protected hand pumps in the Lake Tana basin is presented in *Table 2-3* The examined water bodies are significantly different in temperature, conductivity, $P^H$, and total hardness ($P < 0.05$, $n = 40$, Kruskal- Wallis H test). The conductivity and total hardness of two water points F3 and F4 were significantly different from other sampling sites ($P < 0.05$, $n = 40$, Kruskal- Wallis H test). The pH of F4 was significantly different from the traditional hand-dug well sites except for H2. There was also a significant difference among traditional hand-dug wells in terms of conductivity, pH, total hardness, and total dissolved solids. Within protected hand pumps, sites were significantly different from each other's in terms of conductivity, pH, total suspended solids, total hardness, and total dissolved solids ($P < 0.05$, $n=39$, Kruskal-Wallis h test).

Generally, all traditional hand-dug wells and F3, F4 water points from protected hand pumps showed significantly distinct patterns in terms of total hardness and conductivity and pH to a lower extent. This pattern was caused by different levels of anthropogenic influence on the sampling sites, which correlated excellently with the different kinds of usage and pollution of the different water sampling sites. It was possible to see a range of differing habitats, land use, and pollution pattern in the investigated water systems.

Table 2-3. Physicochemical characteristics of the examined water habitats. Abbreviations: THD — traditional hand-dug wells, PHPS — protected hand pumps, MD - median, To- temperature, DO - dissolved oxygen, Con - conductivity, TSS - total suspended solids, Cl - chloride, TH - total hardness, $NO_3$ - nitrate, $NO_2$ - nitrite, NH3 - ammonia, TDS - total dissolved solids, $SAK_{254}$ - specific absorption at 254 nm. *Significant difference (P < O.05, n = 20 - 40, Kruskal-Wallis H test). Source: Goraw 2007

**Water Resources**

| Parameters | **THDS (n = 20)** | | | **PHPs (n = 20)** | | |
|---|---|---|---|---|---|---|
| | Range | | MD | Range | | MD |
| T (°C)* | 21 | 23.6 | 22.1 | 22.5 | 23.9 | 23.5 |
| DO ($mgL^{-1}$) | 1.3 | 4.0 | 2.6 | 1.3 | 5.0 | 3.3 |
| Con (µS/*cm*) * | 50 | 347* | 166 | 50 | 1516* | 386 |
| pH (-) * | 5.8 | 6.7* | 6.3 | 6.0 | 7.6* | 6. 7 |
| TSS ($mgL^{-1}$) | 1 | 32 | 2.5 | 1 | 10* | 4 |
| CI ($mgL^{-1}$) | 0 | 50 | 19 | 0 | 40 | 10 |
| TH ($mgL^{-1}$) * | 0 | 180* | 80 | 10 | 700* | 255 |
| $NO_3$($mgL^{-1}$) | 0.18 | 7.66 | 2.02 | 0.08 | 8.89 | 1.45 |
| $NO_2$ ($mgL^{-1}$) | 0.00 | 0.36 | 0.02 | 0.00 | 0.63 | 0.01 |
| $NH_3$($mgL^{-1}$) | 0.0 | 12.00 | 0.42 | 0.0 | 12.0 | 0.54 |
| $SAK_{254}$ ($m^{-1}$) | 9.9 | 35.2 | 19.0 | 13.7 | 21.8 | 18.3 |
| TDS ($mgL^{-1}$) * | 0.02 | 0.16* | 0.09 | 0.02 | 0.35* | 0.20 |

**Physico-chemical water characteristics of Lake Tana**

The physicochemical characteristics of Lake Tana are summarized in *Table 2-4*. Lake Tana has relatively low water temperatures, varying only within small limits (range: 20.8 – 28.6°C) (*Table 2-4*). The conductivity of Lake Tana ranges from 100 -1,000 µS $Cm^{-1}$. Higher values of electrical conductivity were reported in river mouths and shore areas in adjacent towns. The conductivity of most freshwaters ranges from 10 - 1,000 µS $Cm^{-1}$(APHA 1995) but may exceed 1,000 µS $Cm^{-1}$, especially in polluted waters, or those receiving large quantities of land run-off (APHA 1995). The pH of Lake Tana ranges from 6.8 - 8.3, which is common for most natural waters. The pH of most natural waters is between 6.0 and 8.5, although lower values can occur in dilute waters high in organic content, and higher values in eutrophic waters, groundwater brines, and salt lakes (APHA 1995). The dissolved oxygen concentration in Lake Tana ranges from 3 - 7.6 mg $L^{-1}$.

The trophic status of Lake Tana is described as oligotrophic (2.6 mg $m^{-3}$ to 8.5 mg $m^{-3}$ Chl-a) having low nutrient concentrations (Vijverberg etal. 2009; Wondie etal. 2007; Dejen etal. 2003). However, in rivers, river mouths, and shore areas of the lake, increasing concentrations of chlorophyll-a (maximum of 50.46 µg$L^{-1}$ ) were measured from time to time, indicating an increase of trophic status (Ilona et al. 2011; Goraw, 2011b; Vijverberg et al. 2009). The increased algal biomass in major tributary rivers and river mouths was the result of increasing inputs of nitrate and phosphate from the basin. A similar pattern was also observed in Rift Valley lakes (Zinabu etal.2002). In Lake Tana, the nitrate concentration ranges from n.d. to 3.66 mg $L^{-1}$ and reactive phosphorous from 0.1 - 9 mg $L^{-1}$. The transparency of Lake Tana is low due to the high silt load of the inflowing rivers during the rainy seasons (May to October), and due to daily re-suspension of sediments in the inshore zone.

Although our data is spatially limited to the Gulf of Lake Tana, Biological Oxygen Demand on which incubation was done for five days at 20ºC ($BOD_5$ 20) and Specific absorbance coefficient at 254 nm wavelength ($SAK_{254nm}$) measurements are indicative for organic pollution in the Gulf. $BOD_5$ 20 ranges from 8.5 - 226.3 mg$L^{-1}$and $SAK_{254nm}$ from 8.2 - 48 ($m^{-1}$). Data on heavy metal concentrations was based on sensitive graphite furnace atomic absorption spectrometry (AAS) and total aqueous Pb concentrations ranged from 0.04 - 42.6 (ng $L^{-1}$); Cd from 2 - 19.8 (ng $L^{-1}$) and Cr from 11 - 18 (ng $L^{-1}$) (Hirut 2014).

The concentration of toxigenic bacteria (i.e., microcystins) is available only from Ilona et al. (2011) and the concentration ranged from n.d. (no detected) to - 2.65 (µg $L^{-1}$) whereas cyanobacteria biomass density was assessed as 188.18 mg $L^{-1}$.

Table 2-4.Some physico-chemical features of Lake Tana (after various sources). *The cyanobacteria biomass and microcystin concentration measurements were taken only in post rainy season in November and were only from limited samples.

| Parameter | Mean/range | Reference |
|---|---|---|
| Rainfall (mm $yr^{-1}$) | 1,326 | SMEC 2008 |
| Annual evaporation (mm) | 1,675 | " |
| Temperature (ºC) | 20.8 - 28.6 | Goraw 2011 |
| Turbidity (NTU) | 11.2 – 125 | " |
| Conductivity (µS $cm^{-1}$) | 100 – 1000 | " |
| Total dissolved solids ($mgL^{-1}$) | 148 -178 | " |
| pH | 6.8 - 8.3 | " |
| Dissolved Oxygen ($mgL^{-1}$) | 3 - 7.6 | " |
| $Ca^{2+}$( $mgL^{-1}$) | 14 – 15 | " |
| $Mg^{2+}$ ($mgL^{-1}$) | 12 – 17 | " |
| $Fe^{2+}$( $mgL^{-1}$) | 2.2 | " |
| Total hardness ($mgL^{-1}$) | 22-390 | " |
| Alkalinity ($mgL^{-1}$Ca $CO_3$) | 35-440 | " |
| Color (pt COU) | 64 – 1140 | " |
| $NO_3$ – N ($mgL^{-1}$) | n.d. - 3.66 | " |
| $NO_2$ – N ($mgL^{-1}$) | n.d. - 0.366 | " |
| $NH_3$ – N ($mgL^{-1}$) | n.d. -12 | " |
| SAK $_{254}$ ($m^{-1}$) | 8.2 – 48 | " |
| $PO_4$-P (mgL-1) | 0.1 – 9.1 | " |
| Transparency/Secchi disk ( cm) | 31-182 | Nagelkerke 1997 |
| $Na^{+1}$ ($mgL^{-1}$) | 7 – 9 | "Rzoska 1976 a |
| Sulphide ($mgL^{-1}$) | 0.3 | " |
| Sulphate ($mgL^{-1}$) | 16.3 | " |
| Silicate ($mgL^{-1}$) | 0.93 | " |
| Chlorophyll a (µg $L^{-1}$) | 2.6-50.46 | Wondie etal. 2007; Dejen etal. 2003 and Ilona etal.2011 |
| BOD $_5$ 20 ($mgL^{-1}$) | 8.5 – 226.3 | Tenagne 2009 |
| TN ($mgL^{-1}$) | 41 | " |
| TP ($mgL^{-1}$) | 1 | " |
| Pb ($ngL^{-1}$) | 0.04-42.6 | Hirut 2014 |
| Cd (ng $l^{-1}$) | 2.0 -19.8 | " |
| Cr (ng $l^{-1}$) | 11-18 | " |
| Microcystins (µg $L^{-1}$) | n.d. – 2.65 | Ilona etal.2011* |
| Cyanobacteria biomass ($mgL^{-1}$) | 188.18 | " |

**Bacteriological characteristics of Lake Tana**

The bacteriological characteristics of Lake Tana are summarized in *Table 2-7* and Fig. 2-1.The bacteriological parameters that have been researched include total coliforms, faecal coliforms, and *E. coli*. Goraw et al. (2011) and Yemenu (2005) reported faecal pollution levels that were significantly increased and discernible in the Bahir Dar Gulf locations, as compared to presumptively anthropogenic uninfluenced reference locations near the outlet of the Blue Nile River of Lake Tana (Fig. 2-1).

The minimum, maximum, median, and percentage occurrence of microbial parameters at the investigated sites for pooled data are shown in *Table 2-7*. Total coliforms (TC), faecal coliforms (FC), *Escherichia coli* (EC), and *Clostridium perfringens* (CP) were detected in 100%, 86%, 82% and 90% of all sampling sites analysed throughout the sampling period, respectively. The pooled median value of Standard heterotrophic plate count (HPC) was log 2.3 CFU per ml (Fig. 2-1).

*Table 2-5.*GPS reading of water sampling sites with a description of locations

| **Site** | **Northing** $00^{0}$ 00’ 00” | **Easting** $00^{0}$ 00’ 00” | **Relative location** |
|---|---|---|---|
| L1 | $11^{0}$37.1’66” | $37^{0}$21.8’ 34” | Near Medhalelem orphanage |
| L2 | $11^{0}$36’37.1” | $37^{0}$22’24.7” | Felege Hiwot Referral Hospital |
| L3 | $11^{0}$36’37.3” | $37^{0}$22’40.2” | Fish production & marketing enterprise |
| L4 | $11^{0}$35’56.4” | $37^{0}$23’4” | Hidar 11 recreational center |
| L5 | $11^{0}$35’47.7” | $37^{0}$2322.5” | Near Mango recreational center |
| L6 | $11^{0}$36’21.6” | $37^{0}$24’5.2” | Near regional prison |
| L7 | $11^{0}$37’28.7” | $37^{0}$24’5.2” | Near Blue Nile River outlet |
| L8 | $11^{0}$38’05” | $37^{0}$24’25.5” | Near Blue Nile River outlet |

*Table 2-6.*Levels of faecal indicators (log CFU / 100 ml), HPC (log CFU / 1 ml). n.d. refers to non-detectable (n = 22). Total coliforms (TC), faecal coliforms (FC), Escherichia coli (EC), Clostridium perfringens spores (CP) in log (CFU /100 ml) and standard heterotrophic plate count (HPC) in log (CFU /1 ml) for sampling sites

| Statistic | TC | FC | *E. coli* | CP | HPC |
|---|---|---|---|---|---|
| Maximum | 6.3 | 6.2 | 6.1 | 4 | 4 |
| Minimum | 2.4 | n.d. | n.d. | n.d. | 1.1 |
| Median | 3.1 | 1.4 | 1.3 | 1.6 | 2.3 |
| % Occurrence | 100 | 86 | 82 | 90 | 100 |

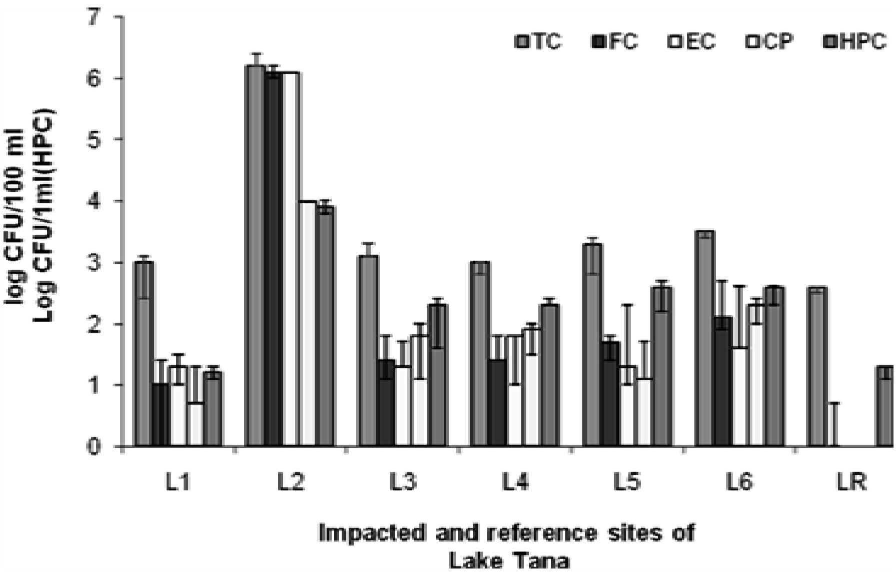



Fig. 2-1– Abundance of total coliforms (TC), faecal coliforms (FC), Escherichia coli (EC), Clostridium perfringens spores (CP) in log (CFU /100 ml) and standard heterotrophic plate count (HPC) in log (CFU /1 ml) for sampling sites. The error bars indicate the maximum, minimum and median concentration. For description of abbreviation of sampling sites please refer *Table 2-7*. Source: Goraw et al. 2010

### 2.2.3 Physico-chemical characteristics of Lake Tana bottom Sediment

The physico-chemical characteristics of Lake Tana sediment sampled in the period 2010 /2011 are summarized in *Table 2-7*. The physico-chemical characteristics of Lake Tana bottom sediment were described from 97 sediment samples collected from 33 sites all over Lake Tana at different times during 2010,2011,2013. The sites are distributed over different stations representing lakeshore (adjacent to cites/towns including anthropogenically influenced ones), pelagic, and river mouths.

The pH of Lake Tana bottom sediment ranges from slightly acidic to slightly alkaline. The minimum was noted in Bahir Dar Gulf of Lake Tana where the influence from anthropogenic activities has been very high. Especially, a sewerage line from Kebele 03 all the way down to Lake Tana is inflowing into the Bahir Dar gulf area and Tana Fish production and marketing enterprise also disposes of a lot of fish offal into the Lake. The maximum pH was noted in the Delgi area.

The electrical conductivity of Lake Tana bottom sediment ranges from 40 to 1450 ($\mu S\ cm^{-1}$). The minimum was noted in Delgi, and the maximum was noted in the Bahir Dar Gulf area.

The total nitrogen content of Lake Tana bottom sediment ranges from 0.02 % to 3.5 % and the minimum was noted in the pelagic part of Lake Tana and the maximum was noted in the Gorgora shore area.

Low values of organic matter were noted in the Korata area (0.07%), and the maximum value of organic matter was noted in Bahir Dar gulf close to the dense stands of wetland vegetation (14.05%). The minimum values of Lake bottom sediment available phosphorous were noted in Bahir Dar Gulf near the mango recreational center(1.07ppm). On the other hand, the maximum available phosphorous (80.64 ppm) was recorded in Bahir Dar gulf where fish offal has been dumbed. The minimum % organic carbon was noted in the Delgi area (0.65%), and the maximum was noted in the Bahir Dar gulf area (4.87%).

The physico-chemical characteristics of the Lake Tana bottom sediment indicate Lake Tana has been influenced significantly by anthropogenic activities of adjacent cities.

*Table 2-7.*The Physicochemical characteristics of the Lake Tana Sediment sampled from March 2010 to 2013, the number of samples, minimum, maximum, mean, and standard error of the mean. TN-refers total nitrogen, OM - organic matter, OC-organic carbon, and Av p-refers available phosphorous.

| Parameter | N | Min | Max | Mean | Std. Error |
|---|---|---|---|---|---|
| pH | 97 | 5.74 | 7.96 | 6.81 | 0.051 |
| Conductivity (µS $cm^{-1}$) | 97 | 0.04 | 1.45 | 0.22 | 0.022 |
| TN (%) | 96 | 0.05 | 3.51 | 0.38 | 0.060 |
| OM (%) | 97 | 0.07 | 14.05 | 3.87 | 0.310 |
| AVP (mg $kg^{-1}$) | 97 | 1.07 | 80.64 | 20.98 | 1.45 |
| OC (%) | 40 | 0.65 | 4.87 | 2.05 | 0.16 |

## 2.3 Conclusions

The water quality indicators are important to assess the suitability of the water for the intended water uses. Understanding the water characteristics is also vital input for water resources development and environmental protection plans and management. Nevertheless, to date, there is no clear and comprehensive water quality assessment available for Lake Tana and its environs. The data is not systematically organized. The water characteristics in the Lake basin are affected by anthropogenic and natural processes but mainly by human activities from the point and diffuse sources.

Although only a limited data set in terms of time and space has been established, anthropogenic activities are significantly influencing water quality in and around Lake Tana. This was evidently shown by a set of multiple faecal indicators, organic pollution by Saprobity determination using HPC and $SAK_{254nm}$ measurement and microcystin concentration measurement and measurement of other chemical and physical water quality and pollution indicators.

Sedimentation, eutrophication, faecal pollution, wetland encroachment, and hydrological alterations have been identified as main issues of water quality management of the Lake Tana basin. Present results urgently call for further research and continuous data generation concerning the whole lake area and the basin. As a first immediate step in establishing good knowledge on the general situation of the Lake and basin water characteristics, simple monitoring programs based on a good and practical selection of robust physical, biological, hydrological indicators and basic chemical parameters are urgently needed.

## 2.4 Recommendations

As basin water resources are of crucial importance - serving multiple purposes and being the largest freshwater body in Ethiopia - the following recommendations are given for further work and sustainable development:

- Quantify contaminants load to Lake Tana from the basin and bottom sediment
- Quantify and characterize organic contaminants from large scale floriculture and agriculture farms
- More data on physical, chemical, biological, and microbiological characteristics of the surface and groundwater
- Identifying point and diffuse pollution sources and erosion hotspots
- Development of decision support tools for lake ecosystem management
- Determination of phosphorous and nitrogen critical loads
- Evaluation of technical and socioeconomic competitiveness of different wastewater treatment and quality management technologies
- Conduct a microbial risk assessment
- Conduct pesticide risk assessment
- As a first and immediate step in establishing good knowledge on the general situation of the water quality in the basin, simple monitoring programs based on a good and practical selection of robust physical, bacteriological, biological indicators and basic chemical parameters are urgently needed. The programs should account for the spatial and temporal variations.

# 3

# Assessing seasonal nitrogen export to large tropical lakes

**Goraw Goshu,[a, b, 1] M. Strokal,[c] C. Kroeze,[c] A.A. Koelmans,[a] J.J.M de Klein[a]**

[a] *Aquatic Ecology and Water Quality Management Group, Department of Environmental Sciences, Wageningen University &Research, P.O. Box, 47,6700AA Wageningen, The Netherlands*

[b] *College of Agriculture and Environmental Sciences and Blue Nile Water Institute, Bahir Dar University, P.O. Box 1701, Bahir Dar, Ethiopia*

[c] *Water Systems and Global Change Group, Department of Environmental Sciences, Wageningen University & Research, P.O. Box, 47, 6700AA Wageningen, The Netherlands*

This chapter is published in revised form as:

Goraw Goshu, M. Strokal, C. Kroeze, A.A. Koelmans, J.J.M. de Klein, Assessing Seasonal nitrogen export to large tropical lakes, Science of The Total Environment, Volume 731, 2020, 139199, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2020.139199.

**Abstract**

Rivers are exporting increasing amounts of nitrogen (N) to lakes, which is leading to eutrophication. However, the seasonality apparent in nutrient loading, especially in tropical areas, is thus far only partially understood. This study aims to better understand the seasonality and the sources of dissolved inorganic nitrogen (DIN) inputs from sub-basins to tropical lakes. We integrated existing approaches into a seasonal model that accounts for seasonality in human activities, meteorology and hydrology, and we applied the model to the sub-basins of a representative tropical lake: Lake Tana, Ethiopia. The model quantifies the river export of DIN by season, source and sub-basin and accounts for open defecation to land as a diffuse source of N in rivers. Seasonality parameters were calibrated, and model outputs were validated against measured nitrogen loads in the main river outlets. The calibrated model showed good agreement with the measured nitrogen loads at the outflow of the main rivers. The model distinguishes four seasons: rainy (July–September), post-rainy (October–December), dry (January–March) and pre-rainy (April–June). The river export of DIN to Lake Tana was about 9 kton in 2017 and showed spatial and temporal variability: It was highest in the rainy and lowest in the dry seasons. Diffuse sources from agriculture were important contributors of DIN to rivers in 2017, and animal manure was the dominant source in all seasons. Our seasonal sub-basins and rivers model provides opportunities to identify the main nutrient sources to the lake and to formulate effective water quality management options. An example is nutrient application level that correspond to the crop needs in the sub-basins. Furthermore, our model can be used to analyse future trends and serves as an example for other large tropical lakes experiencing eutrophication.

### 3.1 Introduction

Eutrophication occurs in many aquatic systems worldwide, and the problems associated with it such as blooms of harmful algae often have negative consequences for the health of humans and ecosystems alike. Eutrophication in lakes is caused by excessive loads of nutrients delivered by rivers (Hecky, 1993; Lürling and van Oosterhout, 2013). Nitrogen (N) is one of the nutrients driving eutrophication. Nitrogen in rivers can result from diffuse and point sources. Diffuse sources include the use of animal manure, human waste (numerous locations of open defecation in the study basin) and synthetic fertilisers on land; biological $N_2$ fixation; atmospheric N deposition and erosion of organic N in the top soil. Point sources of N in rivers often include effluents from sewage systems via pipes. In some regions in the world, direct discharges of animal manure to rivers can also be a point source of water pollution (Strokal et al., 2016).

In tropical lakes, N is often the limiting element, thereby governing primary production of phytoplankton (Conley et al., 2009; Jeppesen et al., 2005; Lewis Jr, 1996; Wondie et al., 2007). This is also the case for Lake Tane as the measured ratio DIN:DIP is around eight in most cases. Therefore, we can conclude it is N- limited (Ptacnik et al., 2010). Lake Tana is a representative example of a eutrophicated lake,  located in the tropical environment in Ethiopia Lake Tana is part of the Blue Nile basin and is an important source of water for human activities. Agriculture has been developing rapidly over recent years in Ethiopia and makes a major contribution to the local economy. Socioeconomic development has stimulated agriculture in the Lake Tana basin, which has thus been identified as a major economic corridor by the Ethiopian government (Stave et al., 2017). This growth of agriculture has sparked the release of nutrients to surface and ground waters, leading to severe water pollution (Goshu and Aynalem, 2017; Goshu et al., 2010a; Selassie, 2017). Low sanitation conditions, high population pressure and degradation of land in the Lake Tana basin (Goshu and Aynalem, 2017; Yitaferu, 2007b) further jeopardise surface water quality in the Lake Tana basin.

Nutrient loading of rivers and lakes is a seasonal event specifically in countries with a highly variable precipitation. Resulting eutrophication effects importantly depend on seasonal cycles of N. Thus, a better quantification of the seasonality in N export to the lake is essential to understand and manage the timing and impact of eutrophication. Another important aspect is the seasonality of human activities (e.g., crop planting periods) and climate (e.g., temperature and hydrology) that influence N export to the lake. Furthermore, for modelling shifts in lake ecosystems, a seasonal N loading has added value compared to an average annual N loading, as it better reflects the seasonality in the effects and the feed-backs (Janssen et al. (2019a).

However, the seasonal river export of N in tropical lakes, such as Lake Tana, is still not well understood. This holds especially for the seasonal river export of N by source taking into account spatial variability (e.g., sub-basins). This hampers the formulation of effective management options (de Klein and Koelmans, 2011). Control and early warning systems related to eutrophication need source-, sub-basin– and season-specific data on N loads from sub-basins to rivers and lakes. Such information (monitoring data on N loads) is limited for large tropical lakes such as Lake Tana, especially for ungauged sub-basins. Some studies observed concentrations of N in rivers draining into Lake Tana (Goshu et al., 2017; Wondie et al., 2007), but they focused on a few rivers only and did not provide a systematic overview of N loadings into the lake. They did not address the sources of N in the lake by season. Therefore, the seasonality of eutrophication in the Lake Tana basin remains unknown.

There is a need for modelling tools to quantify the seasonal river export of N to lakes taking into account the seasonality in human activities on the land, climate, hydrology and their spatial variability. Moreover, the models should be able to attribute the sources of N on the scale of sub-basins. Such modelling tools hardly exist for drainage areas of lakes. Different models exist and have been applied to different basins in the world to quantify river export of nutrients (Douglas-Mankin et al., 2010; Schwarz et al., 2006; Mayorga et al., 2010). However, those models are often coarse for lakes and most of them are annual. Others are more detailed, however require a lot of input data implying that they cannot be used in data-poor regions. There are two exceptions. First, a sub-basin model has recently been developed (Strokal et al., 2016) and successfully applied to a few lakes in China (Yang et al., 2019, Li et al., 2019, Wang et al., 2019). This MARINA model (Model to Assess River Inputs of Nutrients to lakes) considers the spatial variability in human activities and hydrology (sub-basins) for river export of nutrients to lakes. However, this model is annual and does not consider the seasonality. Second, a seasonal modelling approach exists for global rivers: Global NEWS-DIN (Nutrient Export from WaterSheds, McCrackin et al., 2014). This modelling approach is for large basins, however, without considering the spatial variability among basins.

The aim of this study is to assess the seasonality and sources of dissolved inorganic N (DIN) from sub-basins to tropical lakes using Lake Tana as a case study. We focus on the river export of DIN in a spatially explicit manner, by merging the sub-basin scale approach of Strokal et al. (2016) with the seasonal approach of McCrackin et al. (2014). This results in a seasonal model for sub-basins and rivers discharging to tropical lakes such as Lake Tana. We implement the seasonal model to examine N exports in Lake Tana sub-basins and we validate the results with measured data in the same area.

## 3.2 Methodology

### 3.2.1 Study area

The Tana basin has a total drainage area of 16,500 km$^2$ (Fig. 3-1). More than six medium- to large-sized tributary rivers and more than 40 ephemeral streams drain into Lake Tana. Among the rivers, the Gilgel Abay, Dirma, Gumara, Gelda, Rib and Megech rivers contribute more than 90% of the inflow to the lake (Sirak, 2008). Lake Tana is the source of the Blue Nile, which is the only surface outflow. Lake Tana is a shallow lake, with a maximum depth of 14 m and mean depth of 8 m. This tropical lake is non-stratifying with a mean elevation of 1,800 m above sea level. The amount of rainfall is at its maximum during July and August, when it reaches 250–330 mm per month. Mean annual rainfall is nearly 1280 mm (Abebe and Minale, 2017). The rainy season (July–September) receives about two-thirds of the annual rainfall, while the dry season receives 2%; the pre-rainy season (April–June) receives 25% and the post-rainy season 8% of annual rainfall. The average seasonal air temperature reaches its maximum of 21.1 ºC in the pre-rainy season and its minimum (18.4 ºC) in the rainy season, and it shows a large diurnal but small seasonal change.

Lake Tana is the largest lake in Ethiopia, accounting for 50% of the fresh-water resources of the country (Vijverberg et al, 2009). It has a surface area of 3,111 km$^2$, 28.4 km$^3$ in volume and a maximum length of 90 km and width of 65 km. There are four administrative zones and eight districts in the Lake Tana basin. Bahir Dar city is a state capital. The population of the basin was projected to be 4.5 million in 2015 (CSA, 2007), with a population density of 228 persons per km$^2$ in 2007 (Anteneh, 2017), and 70% of the basin is agricultural land (Abebe and Minale, 2017). In this study, we distinguished 20 sub-basins draining into Lake Tana (Fig. 3-1).

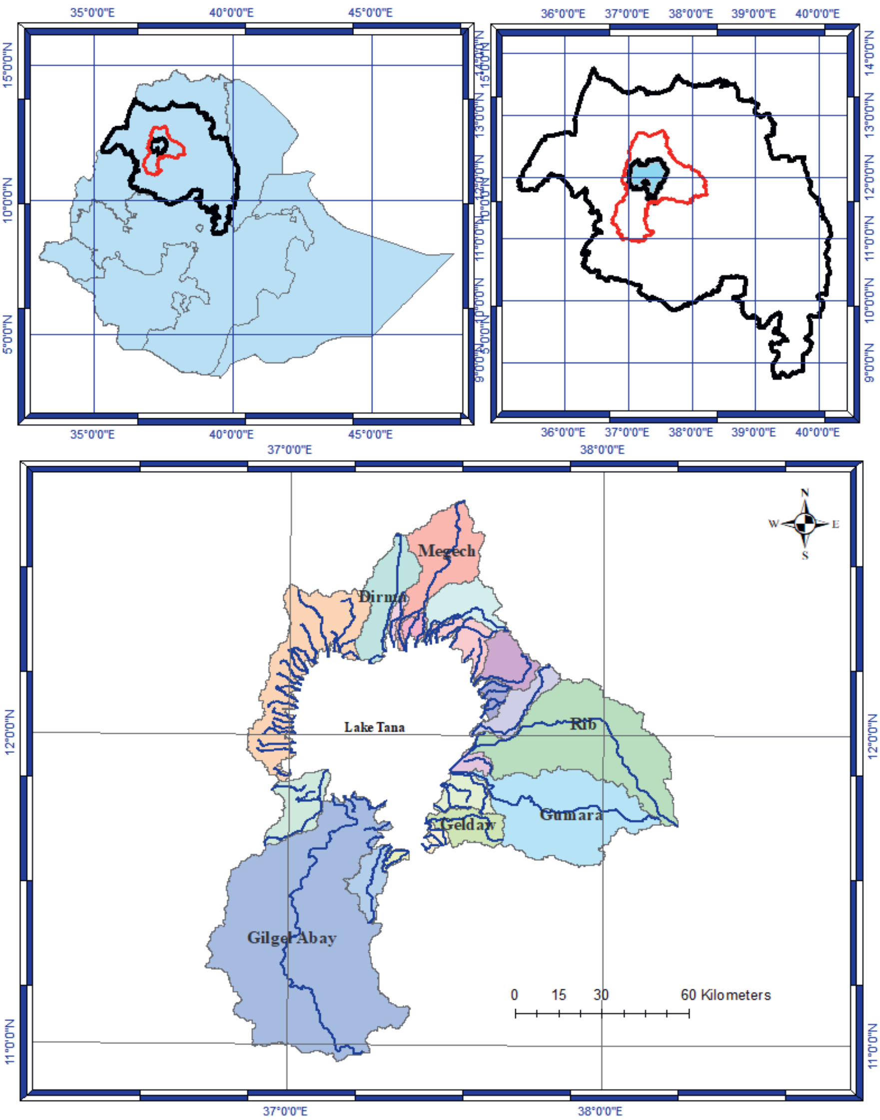



Fig. 3-1–Ethiopia and the Amhara region (upper left); the Amhara region, the Lake Tana basin and Lake Tana (upper right); the drainage area of the Lake Tana basin and the sub-basins draining into the lake (lower panel).

### 3.2.2 Model description

## General approach

We developed a seasonal sub-basins and rivers model, setup for the year 2017 by combining two existing modelling approaches with modifications for our study area. The focus of the approach are the sub-basins draining to a lake, without including the lake itself (Fig. 3-2). The two existing modelling approaches were the sub-basin scale modelling approach of the MARINA model (Strokal et al., 2016, Li et al., 2018, Yang et a., 2019, Wang et al., 2019) and the seasonal modelling approach of the NEWS-DIN(S) model (McCrackin et al., 2014). The MARINA model operates at the sub-basin scale on an annual basis for rivers exporting to lakes, while the seasonal NEWS-DIN(S) model operates at the basin scale for large rivers in the world. These models take a mass-balance approach to quantify nutrient inputs from land to rivers. This includes nutrient inputs to agricultural land (e.g., animal manure, synthetic fertilisers, deposition, fixation) and export from land via crop harvesting and grazing. The net nutrient inputs in soils (inputs minus export) are corrected for nutrient retentions and losses in soils, and the remainder enters rivers. Some nutrients enter rivers from sewage systems (e.g., from cities) and from natural areas. The human waste not collected in septic tanks and pit latrines but defecated to land was defined as open defecation and we categorized it as a diffuse source of nutrients. Human faeces and urine are defecated at many locations to the soil and eventually leaching N to surface waters in a diffuse way. Nutrients in rivers can be either lost or retained before reaching the river mouth. Retentions and losses of nutrients in rivers are estimated as functions of temperature, associated processes (e.g., denitrification), river damming and water consumption (Fig. 3-2).

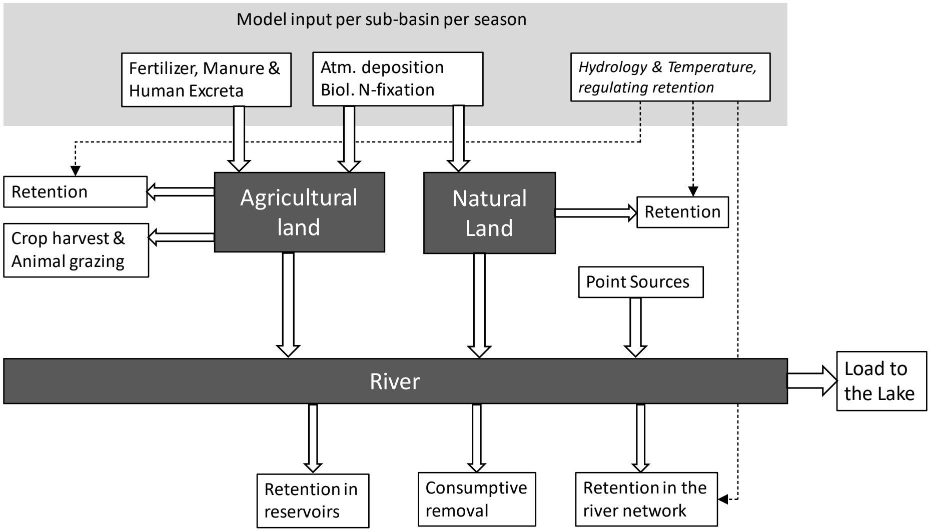



Fig. 3-2 – Conceptual diagram of the seasonal version of the MARINA-model for river export of dissolved inorganic nitrogen from sub-basins discharging into large tropical lakes such as Lake Tana. The seasonal version of the MARINA-model quantifies the river export of dissolved inorganic nitrogen to Lake Tana by season, source and sub-basin for the year 2017. Details are presented in Section 2 and Appendix 3.

In our study, we added seasonality into the sub-basin scale MARINA model based on the approach of McCrackin et al. (2014), resulting in a seasonal version of the MARINA model for tropical lakes such as Lake Tana. We made the following main modifications for our study area. First, we defined seasons according to the rainfall pattern (Table A.2.1), which differs from the approach of McCrackin et al. (2014) who based it on temperature. Second, we used mainly local sources of data for model inputs to represent the local situations of the Lake Tana basin. Third, we added open defecation as a new source of DIN pollution in rivers of Lake Tana. This is a diffuse source, because people defecate their waste on land and not directly into water. Finally, we recalibrated the model parameters influencing seasonality in the Lake Tana basin. Below, we define seasons (Section 2.2.2) and explain the seasonal model (Section 2.2.3 and A.3).

**Definition of Seasons**

In our study, we defined four seasons in this climate region by rainfall rather than by temperature. Based on the rainfall pattern, our defined seasons were pre-rainy: April, May and June (AMJ), rainy: July, August and September (JAS), post-rainy: October, November and December (OND) and dry: January, February and March (JFM).

**Modelling seasonal DIN export**

Our seasonal model estimates river export of DIN by season from sub-basins in three steps. First, inputs of N from diffuse and point sources to land and rivers were estimated. Second, the river export of DIN to the outlet of each sub-basin was estimated. Finally, the river export of DIN from sub-basin outlets to river mouths (the point at which DIN was discharged into Lake Tana) was estimated by quantifying retentions and losses of DIN within the river network.

The total load of DIN exported to the lake from each source was determined per sub-basin and per season. The overall equation to quantify annual river export over the four seasons to the river mouth $\mathrm{M_{DIN.y.j}}$ (kg $y^{-1}$) by source y from sub-basin j was as follows:

$$\mathrm{M_{DIN.y.j}} = \sum M_{DIN.y.j.S} \quad \textbf{(eq.1)}$$

Where $\mathrm{M_{DIN.y.j.S}}$ is the seasonal river export of DIN to the river mouth of Lake Tana by source y from sub-basin j in season S (kg $S^{-1}$).

$\mathrm{M_{DIN.y.j.S}}$ was estimated following the approach of the MARINA model (Strokal et al., 2016), but for seasons (modified from McCrackin et al., 2014) as follows:

$$\mathrm{M_{DIN.y.j.S}} = \mathrm{RS_{DIN.y.j.S}} \cdot \mathrm{FE_{riv.DIN.outlet.j.S}} \cdot \mathrm{FE_{riv.DIN.mouth.j.S}} \quad \textbf{(eq.2)}$$

Where $\mathrm{RS_{DIN.y.j.S}}$ was the input of DIN to rivers in sub-basin j from source y in season S (kg $S^{-1}$). The DIN in rivers results from both diffuse and point sources, the former of which include DIN that results from N inputs to land (see steps 1 and 2 for details below).

$FE_{riv.DIN.outlet.j.s}$ was the fraction of $RS_{DIN.y.j.S}$ that was exported to the outlet of sub-basin j from source y in season S (0–1).

$FE_{riv.DIN.mouth.j.S}$ was the fraction of $RS_{DINyj.S}.FE_{riv.DIN.outlet.j.S}$ that was exported to the river mouth of Lake Tana from source y and sub-basin j in season S (0–1, see details below).

We express the river export of DIN in terms of load to the Lake (kg $S^{-1}$) or yield (kg $km^{-2}$ of basin area $S^{-1}$).

**Step 1: Quantifying N inputs to land**

Nitrogen inputs to land include the use of animal manure, synthetic fertilisers, biological $N_2$ fixation and atmospheric N deposition on agricultural and natural areas and human waste (faeces + urine). Human waste is not used as fertiliser in the basin. We first estimated inputs of N to land from diffuse sources for 2017. We assumed that human waste (human faeces and urine) that was collected in pit latrines and septic tanks does not reach the surface waters (Van Drecht et al., 2009). To estimate N inputs to land, seasonal data were needed for the following aspects: hydrology, socio-economic development (population density, gross domestic product and sewage connection), sub-basin characteristics (land use and slope), agricultural inputs (fertiliser use and manure excretion, crops, livestock, etc.) and meteorological data (rainfall and temperature). Most model inputs were collected from local information, such as the Ethiopia Central Statistical Authority, Ethiopia Ministry of Water, Electric and Irrigation, Amhara Region Bureaus of Agriculture and Water, Amhara Design and Supervision Enterprise, Blue Nile Water Institute and Bahir Dar University and from zonal and districts Agricultural Offices in the Lake Tana basin. In the absence of locally available data, data for the Nile basin were used from the Global NEWS-2 model (Mayorga et al., 2010). To calculate N inputs to land and rivers at the sub-basin scale by season, model inputs were needed for sub-basins and seasons for 2017. For this, we used an area-weighted method to aggregate local (often district-level) data into the sub-basin scale data.

Annual values for N inputs to land for 2017 were available at the institutions listed above. To allocate annual N inputs to seasons, we took into account crop phenology and climate data (rainfall and temperature) in the natural and agriculture systems (Table A.4.5 and Table A.4.7). We allocated to seasons the following inputs for 2017: synthetic fertilisers, atmospheric N deposition and biological $N_2$ fixation. We assumed that N inputs to land from human waste (open defecation) and animal manure was the same among the four seasons (Appendix A.4.1 and A.4.2).

The seasonal distribution of the use of *synthetic fertilisers* was considered according to the approach of McCrackin et al. (2014). We took into account the planting area, application rate and application pattern of representative crops (see Appendix A.4 and Table A.5).

Annual values for *atmospheric N deposition* on agricultural and natural areas were available from the Global NEWS-2 Model. The annual values are distributed over seasons according to seasonal proportions of rainfall (McCrackin et al., 2014) (see Table A.4.5 and Table A.2.1).

Total seasonal *biological $N_2$ fixation* was calculated by summing the seasonal *biological $N_2$ fixation* by natural vegetation and seasonal fixation by agricultural area (see eq. A11- eq. A15). The seasonal *biological $N_2$ fixation* by natural vegetation was calculated from the annual *biological $N_2$ fixation* by natural vegetation with a formula that defines the relationship between air temperature and nitrogenase activity (Hijmans et al., 2005a) see eq. A.16). To estimate annual *biological $N_2$ fixation* in natural vegetation, we followed the spatial distribution of non-agricultural land use forms, and only fixation rates of tree cover areas, shrubs cover areas, wetland and lichen/mosses/spare vegetation were considered. The fixation rates of only these land use forms were taken from Global NEWS-2 data (Mayorga et al., 2010).

The seasonal *biological $N_2$ fixation* in agricultural area was the sum of *biological $N_2$ fixation* by non-legumes, rice and legumes. For non-symbiotic *biological $N_2$ fixation* (non-legumes) and grassland, a fixation rate of 5 kg $ha^{-1}$ $y^{-1}$, and for rice fields, a rate of 25 kg $ha^{-1}$ $y^{-1}$, as proposed by Smil (1999), was assumed. The seasonal fixation rates were estimated from annual rates by quantifying the annual rates as a function of air temperature based on approach of Hijmans et al. (2005b). The details can be found in Annex A.4. We calculated $N_2$ fixation by legumes as N in biomass harvested multiplied by two (for above- and below-ground biomass), following the approach of Bouwman et al. (2009). The total biological fixation of $N_2$ thus depends on the total production of legumes, as well as on the areas of grassland and cropland. The seasonal fixation rate was computed as a function of average seasonal temperature standardised by the nitrogenase activity.

The rates were calculated based on crop phenology (Houlton et al., 2008; Sacks et al., 2010) and local information. In addition to local information on crop calendar, the FAO crop data base (http://www.fao.org/agriculture/seed/cropcalendar/cropcal.do) was consulted to determine the planting and harvesting times of crops. Fixation by an individual crop was thus considered only for the period between planting and harvesting. To estimate *biological $N_2$ fixation* by crops, representative crop phenology was taken into account. We distinguished among rice, legume and non-legume growing areas. Since grassland was not fertilised, it was not part of the agricultural area in our model. We used crop production data compiled at a province level (CSA, 2016) and disaggregated into sub-basins based on an area-weighted method. The data contains crop type, number of holders, area covered, production and crop yield.

**Step 2: Quantifying DIN inputs to rivers from diffuse and point sources ($RS_{DIN.y.j.S}$)**

$RS_{DIN.y.j.S}$ was the sum of diffuse ($RSdif_{DIN.y.j.S}$) and point sources ($RSpnt_{DIN.y.j.S}$) in (kg $S^{-1}$).

- **Diffuse sources**

$RSdif_{DIN.y.j.S}$ for agricultural land was estimated following the approach of the MARINA model (Strokal et al., 2016), but for seasons, as follows:

$$RSdif_{DIN.y.j.S} = WSdif_{N.y.j.S} \cdot G_{N.j.S} \cdot FE_{WS.DIN.j.S} \qquad \textbf{(eq.3)}$$

Where $WSdif_{N.y.j,S}$ was N input to the land from diffuse source y in sub-basin j and season S (kg $S^{-1}$). Diffuse sources are specified in step 1; see step 1 how diffuse sources are derived for seasons and sub-basins.

$G_{N.j.S}$ was the fraction of N that remains in soils after crop uptake and animal grazing in sub-basin j and season S (0–1). This fraction was calculated following the MARINA model approach (Strokal et al., 2016). We use this fraction to estimate the amount of N that reaches rivers after correcting for N losses and retentions in the soils (e.g., denitrification losses as a function of runoff and temperature).

$FE_{WS.DIN.j.S}$ was the export fraction of N from diffuse sources entering the surface water as DIN in sub-basin j and season S (0–1). We followed the seasonal approach of McCrackin et al. (2014) to quantify $FE_{ws.DIN.j.S}$ by season. This approach includes temperature (see eq. 4) to account for the effect of temperature on N retention and losses from the soil (e.g., denitrification).

$FE_{ws.DIN.j.S}$ was estimated as follows (McCrackin et al., 2014):

$$FE_{ws.DIN.j.S} = FE_{ro.j.S} \cdot (1 - F_{temp.j.S}) \qquad \textbf{(eq.4)}$$

Where $FE_{ro.j.S}$ was the fraction of N entering the surface water as DIN in sub-basin j and season S, taking into account the influence of the surface runoff (0–1). $FE_{ro.j.S}$ was calculated as a function of seasonal natural runoff from land to streams in sub-basin j and season S (0–1) (eq. 5); $F_{temp.j.S}$ was the fraction of N retention in soils of sub-basin j and season S due to effects of temperature (0–1).

$F_{temp.j.S}$ was estimated using eq. 6, as follows (McCrackin et al., 2014):

$$FE_{ro.j.S} = b \cdot (Rnat_{j.S} * 4)^{a} \quad \textbf{(eq.5)}$$

$$F_{temp} = d(T_{j,S}/100)^{c} \quad \textbf{(eq.6)}$$

Where $Rnat_{j.S}$ was the seasonal runoff from land to streams in sub-basin j and season S (m $S^{-1}$). The natural runoff was multiplied by four, according to McCrackin et al. (2014), to ensure that $FE_{ro.j.S}$ was consistent with the annual fraction.

The a and b parameters (eq. 5) are used to determine the function of runoff (McCrackin et al., 2014). These parameters are recalibrated in for Lake Tana, because seasonality in our study area is driven more by rainfall rather than by temperature, as in the approach of McCrackin et al. (2014).

$T_{j.S}$ was the temperature for the sub-basin j in season S (°C), and the c and d constants (eq. 6) reflect the function of air temperature (McCrackin et al., 2014).

In eq. 6, $T_{j,S}$ is divided by 100 to fit the order of magnitude with $Rnat_{j.S}$ (McCrackin et al., 2014). $T_{j,S}$ for each sub-basin was calculated by taking the average observed temperature of the respective three months in the Lake Tana basin.

- **Point sources**

The point sources of DIN in rivers ($RSpnt_{DIN.y.j.S}$, kg $S^{-1}$) include effluents from sewage systems. $RSpnt_{DIN.y.j.S}$ was calculated based on Strokal et al. (2016, but for seasons, as follows:

$$\mathrm{RSpnt_{DIN.y.j.S}} = \mathrm{RSpnt_{N.y.j.S}} \cdot \mathrm{FEpnt_{DIN.y.j}} \quad \textbf{(eq.7)}$$

Where $RSpnt_{DIN.y.j,S}$ was DIN inputs to rivers from point source y in sub-basin j and season S (kg $S^{-1}$).

$RSpnt_{N.y.j.S}$ was the N input from point source y to rivers of sub-basin j in season S (kg $S^{-1}$).

$FEpnt_{DIN.y.j}$ was the export fraction of N from point source y in sub-basin j that was exported to rivers as DIN (0-1). $FEpnt_{DIN.y.j}$ was estimated using the approach of the MARINA model (Strokal et al., (2016).

In our model, $RSpnt_{N.y.j.S}$ was zero for 2017, as waste-water treatment plants and other point sources hardly exist in the Lake Tana basin.

**Step 3: Quantifying retentions and losses of DIN within the river network**

$FE_{riv.DIN.outlet.j.S}$ accounts for retentions and losses of DIN within the river network of sub-basins. This fraction was estimated based on Strokal et al. (2016 for sub-basins and on McCrackin et al. (2014) for seasons:

$$FE_{riv.DIN.outlet.j.S} = \left(1 - D_{DIN.j.S}\right) \cdot (1 - L_{DIN.j.S}) \cdot (1 - FQrem_{j.S}) \qquad \textbf{(eq.8)}$$

Where $D_{DIN,\,j.S}$ is the fraction of DIN retained in dammed reservoirs in sub-basin j and season S (0 -1), $L_{DIN.j.S}$ the fraction of DIN losses by denitrification in the river network of sub-basin j and season S (0 -1), and $FQrem_{j.S}$ the fraction of nutrients (generic for all nutrients) removed by water consumption in sub-basin j and season S (0-1).

The annual $D_{DIN.j.i}$ was estimated following Strokal et al. (2016). We assumed the same fractions for all seasons. $D_{DIN.j.i}$ was calculated as follows:

$$D_{DIN.j.i} = 0.8845 \times \left(\frac{h_{j.i}}{\Delta\tau_{R.j.i}}\right)^{-0.3677} \qquad \textbf{(eq.9)}$$

Where $h_i$ was the depth of reservoir i in sub-basin j (m), $\Delta\tau_{R.i}$ was the water residence time for reservoir i in sub-basin j (year) (data from Birhanu et al. (2014), and $L_{DIN.j.S}$ accounts for effects of temperature and was estimated using the approach of McCrackin et al. (2014), as follows:

$$L_{DIN.j.S} = \left(0.0605 \times \ln\left(Area_j\right) - 0.0443\right) \cdot Q_{10}^{\frac{\left(T_{j.S} - T_{average.j}\right)}{10}} \quad \textbf{(eq.10)}$$

Where $Area_j$ is the total area of sub-basin j ($km^2$) and $Q_{10}$ the air temperature coefficient that indicates the rate of changes in denitrification as a consequence of increasing the air temperature by 10 °C. The value of $Q_{10}$ is set to 2.54 in this study (Mineau et al., 2015). $T_{j.S}$ is the air temperature for season S in sub-basin j (°C) and $T_{average.j}$ the annual average temperature for sub-basin j (°C). The maximum value of $L_{DIN.j.S}$ is set at 0.65 to avoid extrapolation error (McCrackin et al., 2014).

$FQrem_{j.S}$ is calculated following the approach of Strokal et al. (2016) and McCrackin et al. (2014), as follows:

$$FQrem_j = 1 - Q_{act.j.S}/Q_{nat.j.S} \quad \textbf{(eq.11)}$$

Where $Q_{act.j.S}$ is the actual water discharge at the outlet of sub-basin j after water consumption in season S ($km^3\ S^{-1}$). $Q_{act.j.S}$ was collected from the Ethiopian Ministry of Water, Electric and Irrigation for the gauged sub-basins (daily data). The area of the gauged sub-basins comprises about 75% of the total Tana basin. Actual water discharge at the outlets of ungauged sub-basins, $Q_{act.j.S}$, were estimated by taking the actual water discharges of the adjacent or nearest (proximity analysis) gauged station and transferring this actual discharge to a new synthetic discharge for the ungauged sub-basin using the area and rainfall of ungauged basins based on the approach of (Yarahmadi, 2003) (see Appendix A5 for more detail). $Q_{nat.j.S}$ is the natural water discharge at the outlet of sub-basin j before water consumption in season S ($km^3\ S^{-1}$). $Q_{nat.j.S}$ is estimated by adding water consumption to $Q_{act.j.S}$ for each sub-basin and season. We accounted for water consumption for irrigation, animal watering, and surface water supply.

Industries and floriculture farms in the Lake Tana basin mainly use water either from deep wells and/or Lake Tana or by abstracting from rivers not included in the basin. Hence, water consumption by industries and floriculture farms is not accounted for in our model. We obtained the annual water consumption from local data (see Appendix A.4 for details). Seasonal water consumption is the distribution of annual water consumption based on the proportion and use of water in agriculture, industry and residents in seasons.

$FE_{riv.DIN.mouth.j.S}$ accounts for DIN retention and loss during transport from the outlets towards the river mouth (0–1, see eq. 8). $FE_{riv.DIN.mouth.j.S}$ is estimated as in Strokal et al. (2016). In our seasonal model, the distance between the hydrometric station (outlets) and the river mouths is reasonably short. Therefore, we assumed that all the DIN that reached the outlet will reach the river mouth, and $FE_{riv.DIN.mouth.j.S}$ was thus assumed to be 1 in our seasonal model.

**Monitoring water quality to estimate actual DIN loads**

Water samples were collected on a monthly basis in 2017 in 500 ml polyethylene bottles and were kept in ice and transported to the laboratory for immediate analysis (APHA-AWWA-WPCF., 1981). Chemical analysis of ammonium ($NH_4^+$ - N), nitrate ($NO_3^-$ - N) and nitrite ($NO_2^-$ - N) was done with colorimetry in accordance with the manufacturer instruction (Palintest transmittance – display photometer 8000). The DIN concentration (g $m^{-3}$) was calculated by summing $NH_4^+$ - N (g $m^{-3}$), $NO_3^-$ - N (g $m^{-3}$) and $NO_2^-$ - N (g $m^{-3}$). The DIN load to the lake (g $s^{-1}$) is a product of DIN concentration (g $m^{-3}$) and discharge ($m^3$ $s^{-1}$). Based on our data set, because discharge and concentration are independent and observed regularly, the average load to the lake can be calculated by multiplying the average discharge and the average concentration, and we thus preferred the straight-forward method (De Vries and Klavers, 1994), as follows:

Fout! Bladwijzer niet gedefinieerd.$\boldsymbol{L} = \frac{\sum_{i=1}^{m} Q_i.}{m} . \frac{\sum_{j=1}^{n} Cj.}{n}$ **(eq.12)**

Where,

L is the load of DIN (g $s^{-1}$).
$Q_i$ is the discharge at the gauging stations at time i ($m^3$ $s^{-1}$),
$C_j$ is the concentration at different time j (g $m^{-3}$),
m is the number of discharge measurements,
n is the number of concentration measurements.

**Model calibration and validation**

We modified the modelling approach of McCrackin et al. (2014) by re-calibrating the 'a' and 'b' model parameters that determine the runoff function (Table A.6.1 and eq. 3 above). We found the optimum values for parameters a and b that resulted in the minimum difference between modelled and observed values. To calibrate and validate the integrated model, we used different data sets of observed and modelled DIN loads and yields to the lake. For model calibration, we used the values of seasonal observed and modelled DIN at or near the mouths of the Gilgel Abay, Gelda, Gumara, Rib and Dirma rivers, all of which had DIN observed at or near the river mouth. We used seasonal DIN load and yield data of the Arno-Garno, Megech and Infranz rivers for model validation (See Fig. 1a-c for the location of the rivers and Appendix A.6.3 for observation data). We converted DIN to kg $km^{-2}$ $S^{-1}$ by dividing the observed tons of $S^{-1}$ of a sub-basin by its drainage area.

We assessed the model performance using $R^2_P$ (the Pearson's coefficient of determination, 0–1), $R^2_{NSE}$ (the Nash-Sutcliffe efficiency, -∞-1; (Nash and Sutcliffe, 1970) and RSR, > 0, according to Moriasi et al. (2007). $R^2_P$ shows the proportion of the variance in the observed data that is predictable from the modelled data, with higher values representing better performance. $R^2_{NSE}$ evaluates the fitness of the observed and modelled data with the 1:1 line, with higher values representing better performance. RSR is the ratio of root mean square error and standard deviation, with values close to zero showing very good model performance.

## 3.3 Results

### 3.3.1 Model calibration and validation

Model calibration results showed a good performance, that is, a high agreement between observed and modelled values of DIN load to the lake in ton S-1 (R2P = 0.90; R2NSE = 0.88; RSR = 0.31) and DIN yield in ton km-2 S-1 (R2P = 0.82; R2NSE = 0.80; RSR = 0.50; Fig. 3-3A, B). For validation the model also performed well against measured data for DIN loadings with R2P = 0.88, R2NSE = 0.81 and RSR = 0.41 ( Fig. 3-3C). As for DIN yields at the river mouth with observed values, goodness of fit metrics were R2P = 0.86, R2NSE = 0.69 and RSR = 0.50 (Fig. 3-3D). The total modelled DIN loading to the lake per season shows a good agreement with the observed loads to the lake (Fig. 3-3). The abovementioned indicators imply that model performance can be considered good (Moriasi et al., 2007).

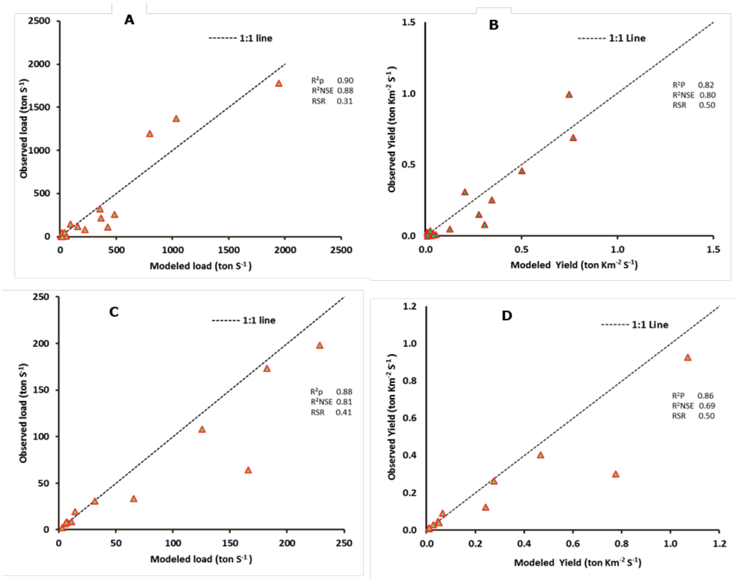



Fig. 3-3– Model calibration (panels A, B) and validation (panels C, D). Graphs show modelled versus observed DIN load (ton S$^{-1}$) and DIN yield (ton km $^{-2}$ S $^{-1}$) from different sub-basins per season. $R^2_P$, $R^2_{NSE}$ and RSR are the Pearson's coefficient of determination (0–1), the Nash-

Sutcliffe efficiency (0–1) and the root mean square error standardised by standard deviation (> 0), respectively.

### 3.3.2 Seasonal nitrogen inputs to land in the Lake Tana basin

The total annual N inputs to land in Lake Tana basin were estimated at 129 kton $y^{-1}$ in 2017 (Fig. 3-4). The N inputs to land (kg $km^{-2}$ $S^{-1}$) largely varied among seasons and sub-basins. The inputs were higher in the rainy season and lower in the dry season. The rainy season contributed 36% of the total annual N inputs to land, and the contributions of the post-rainy (21%), dry (21%) and pre-rainy (22%) seasons were lower (Fig. 3-5). The total annual N inputs to land varied between 0.1 and 51 kton $y^{-1}$ and between 232 and 2464 kg $km^{-2}$ $y^{-1}$ among sub-basins. The highest N inputs to land were delivered by the Gilgel Abay and the lowest by the Infranz-Bahir Dar sub-basin.

The source attribution differs slightly among the seasons. The seasonal N inputs to land for 2017 result from human excreta (open defecation), animal manure, synthetic fertilisers, biological $N_2$ fixation by crops and natural vegetation and atmospheric N deposition on agricultural and natural land (Table A.4.1; Appendix A.4). Among these diffuse sources, animal manure was the dominant contributor of N inputs to land in all seasons because animal grazing was assumed to occur at the same intensity in all seasons. Synthetic fertiliser was the second dominant N input to land, but only in the rainy season because it is only used during the rainy growing season (Table 3-1).

Table 3-1.Modelled N inputs to land by season and source for the Lake Tana basin in 2017 (kg km watershed area$^{-2}$ S$^{-1}$). Ranges indicate minimum and maximum values among sub-basins. *Over agricultural and natural land.

| Sources | Rainy | Post-rainy | Dry | Pre-rainy |
|---|---|---|---|---|
| Animal manure | 201–2,904 | 201–2,904 | 201–2,904 | 201–2,904 |
| Synthetic fertilisers | 260–2,873 | 0 | 0 | 0 |
| Human waste | 37–249 | 37–249 | 37–249 | 37–249 |
| Biological N2 fixation* | 100–334 | 135–237 | 11–162 | 31–181 |
| Atmospheric N deposition* | 131–196 | 10–24 | 2–6 | 43-77 |

### 3.3.3 Seasonal river export of DIN by sub-basin and source to Lake Tana

The model indicated that about 9 kton of DIN was exported by rivers to the lake in 2017 (Fig. 3-5). In terms of yield, this amount is 802 kg km$^{-2}$ of the total drainage area of the lake y$^{-1}$. The annual DIN yields from the sub-basins ranged from 232 to 2,464 kg km$^{-2}$ y $^{-1}$. The lowest river export of DIN was delivered by the Gemero Makesegnit and the highest by the Gelda sub-basin.

The six major rivers, namely, the Dirma, Gilgel Abay, Gumara, Gelda, Megech and Rib, together exported over two-thirds of total DIN to Lake Tana (Fig. 3-5). In the rainy season, about one-third of this export was from the Gilgel Abay sub-basin, 17% from the Gumara, 8% from the Rib, 6% from the Gelda and 6% from the Dirma. Small sub-basins such as the Infranz-Bahir Dar are estimated to export the smallest amounts of DIN to the lake.

River export of DIN to Lake Tana were highest in the rainy and lowest in the dry season. The export of DIN in the rainy season was estimated to be over half of the annual export to the lake. The DIN yields exported in the rainy season ranged from 141 to 1,393 kg km$^{-2}$ S$^{-1}$. The lowest amounts of DIN to the lake were exported from the Gemero-Makesegnit and the highest amounts from the Gelda sub-basin (Fig. 3-5). The Gilgel Abay sub-basin contributed 504 kg km$^{-2}$ S$^{-1}$. In the rainy season, the major sources of DIN inputs to the lake were animal manure

(49%), synthetic fertiliser (37%) and biological $N_2$ fixation (6%). Atmospheric N deposition (4%) had the lowest contribution (Fig. 3-4 A).

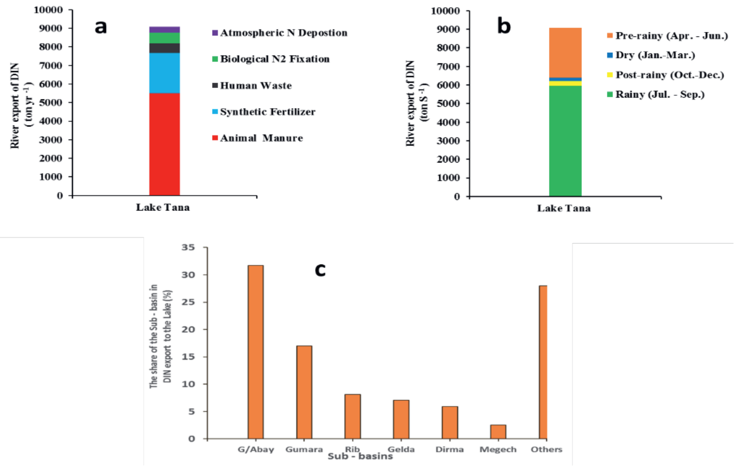



Fig. 3-4 – Total modelled annual river export of DIN by source (ton $y^{-1}$, a), season (ton $S^{-1}$, b) and the share of the sub-basins of the total annual river export of DIN (%, c) to Lake Tana in 2017. See Fig. 3-1for the location of the sub-basins.

The DIN load to the lake in the post-rainy season was by far lower than the DIN load in the rainy season (*Fig. 3-4*). The DIN load ranged from 0.5 to 93 ton $S^{-1}$ among sub-basins. Similar to the rainy season, the highest DIN load was exported by Gilgel Abay and the lowest by Infranz-Bahir Dar. By contrast, the highest DIN yield (53 kg $km^{-2}$ $S^{-1}$) was exported by Gelda and the lowest by Gemero-Maksegnet (5 kg $km^{-2}$ $S^{-1}$) (*Fig. 3-5*). In the dry season, Bahir Dar-Infranz exported 0.27 ton of DIN, which is much lower compared to the Gilgel Abay sub-basin (43 ton $S^{-1}$). This sub-basin exported the largest amount of DIN to the lake in the dry season (43 ton $S^{-1}$). In terms of yield, the river export of DIN ranged from 6 (Infranz _Bahir Dar) to 40 (Gelda) (kg $km^{-2}$ $S^{-1}$) among the sub-basins in the dry season (*Fig. 3-5*).

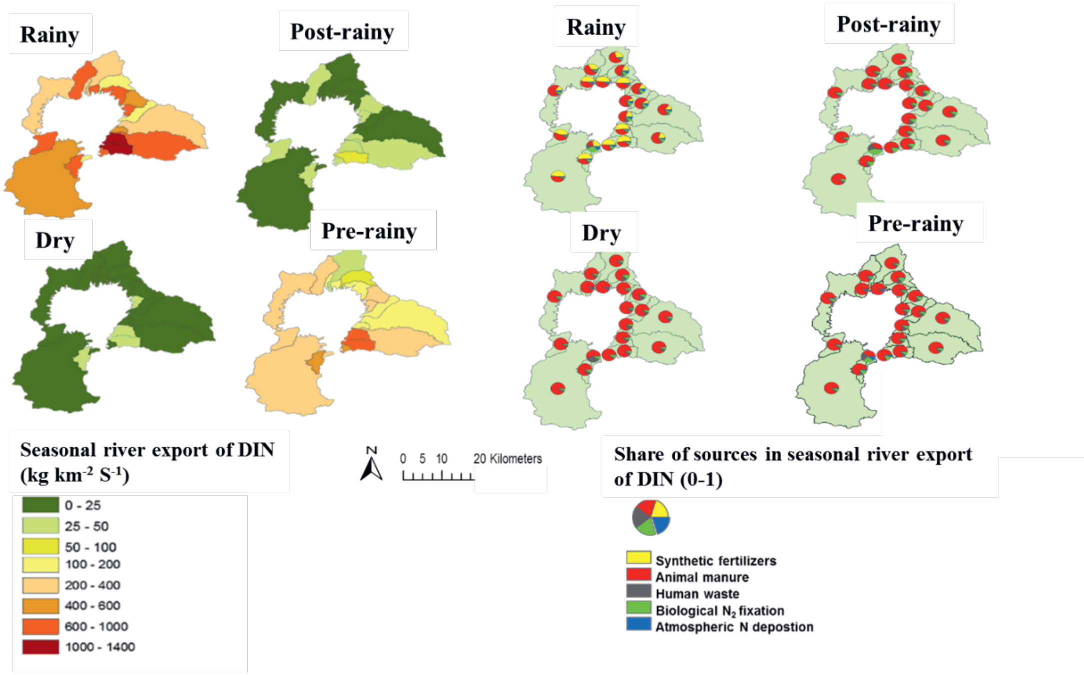



Fig. 3-5 – Modelled seasonal river export of dissolved inorganic nitrogen (DIN) by sub-basin to Lake Tana (kg km$^{-2}$ S$^{-1}$) and the share of the sources in the seasonal river export of DIN by sub-basin (0–1) in 2017.

The pre-rainy season was the second dominant contributor to the annual DIN load after the rainy season ( Table 3-1). In this season, 30 % of the annual DIN load was estimated to be exported to the lake. The DIN export ranged from 5.7 to 798 ton S$^{-1}$ among sub-basins. The lowest export was by the Megech-Dirma and Gilgel Abay sub-basins. Animal manure was estimated to be a dominant source of DIN in the lake (82%), and human waste was the second dominant source of DIN, followed by biological $N_2$ fixation (7%) and atmospheric N deposition (3 %).

## 3.4 Discussion

### 3.4.1 Strengths and weaknesses of the seasonal modelling approach

Our model was applied for 2017, which we consider to be a representative year because we found no significant difference in rainfall and temperature between 2017 and the long-term annual average (1952 - 2015) (Wilcoxon signed-ranked test, p=0.61 for rainfall, p=0.18 for temperature). Furthermore, fertilizer application amounts (the main contributor of DIN) in the basin are fairly constant over the years as they are based on fixed rates (blanket recommendations). From this we assume that 2017 can be seen as a representative year in terms of DIN sources. We also systematically collected observed data for all months of the year 2017, giving us confidence in representativeness of the observed values for different seasons. Although data were only available for one year, this enabled us to calibrate and validate our seasonal model. We realise that more years would be better, but after calibrating the seasonal parameters (Section 2.4), model outputs were generally in good agreement with observed data (e.g., $R^2_{NSE}$> 0.65, see Fig. 3-3). Nevertheless, our newly developed seasonal approach has some uncertainties that are largely associated with model inputs and parameters. We used Sentinel 2 land cover 2016 imagery to calculate areas of different land cover. Sentinel 2 (20 m–by–20 m resolution) land cover imagery has an overall area weighted accuracy of 65% +/- 1 % (Lesiv et al., 2017) but could nevertheless contribute to the uncertainty of our model. The use of global scale data, which has a 0.5° by 0.5° resolution, to local sub-basins of Lake Tana might also introduce some uncertainties. Nevertheless, validation results showed a good performance of the model to quantify seasonal river export of DIN to Lake Tana. Therefore, we consider the model uncertainties to not affect the main messages of our study.

We developed a seasonal version of the MARINA model for large tropical lake sub-basins and rivers by combining the sub-basin modelling approach (Strokal et al., 2016) with the seasonal modelling approach (Chen et al., 2019; McCrackin et al., 2014) and we did this with Lake Tana basin for the first time. The sub-basin scale MARINA model has been widely applied to other lakes, such as the Dianchi (Li et al., 2018), Taihu (Wang et al., 2019) and Guanting (Yang et al., 2019) lakes. These studies all validated the modelling approach against observations. Furthermore, Strokal et al. (2016) tested the sensitivity of model outputs to changes in model inputs and parameters for the sub-basin scale MARINA model. They found that the DIN export by rivers was generally sensitive to changes in animal manure, synthetic fertiliser and hydrology. McCrackin et al. (2014) also validated the seasonal modelling approach against observations for large rivers in the world. The sensitivity analyses of McCrackin et al. (2014) revealed that river export of DIN is also generally sensitive to changes in runoff, particularly in summer. The results of the sensitivity analysis for the sub-basin (from Strokal et al., 2016) and seasonal (from McCrackin et al., 2014) modelling approaches give us insights into the model inputs and parameters that need attention in implementing these approaches in our study area. We therefore used local information to derive model inputs and parameters that reflect best the local situations of the Lake Tana basin (see Section 2).

We argue that our seasonal model is suitable for assessing DIN export from sub-basins to Lake Tana. In general, the main advantage of our approach is that it can provide a relatively accurate way to link water quality with human activities, by quantifying the contribution of separate sources seasonally. Preferably the input is from local data, but if not available, global data can be used as well. Furthermore, the approach enables to estimate DIN export from ungauged sub-basins. The disadvantage is that an accurate assessment needs regional calibration of the seasonal parameters. In addition, for local analysis of water quality (e.g., cities), more in-depth model evaluation is needed.

### 3.4.2 Seasonal patterns and sources of river export of DIN in the Lake Tana

Our study presents results of the seasonal and sub-basin analyses of DIN export to Lake Tana. Our model enables the identification of which season and sub-basin contributes most to lake pollution, and from which source (e.g., agriculture or city). This provides new insight into the seasonality in river export of DIN, taking into account the socio-economic drivers, human activities, hydrological and climate characteristics. This has not been done before for Lake Tana. Our seasonal model for Lake Tana opens an opportunity to explore solutions and future trends in DIN export by rivers. This is relevant for science and policymaking as we provide a new tool and new insights that can support the formulation of effective water pollution strategies.

Seasonal estimates of the river export of DIN to Lake Tana in Ethiopia are scarcely available. The available studies for the Lake Tana basin and other bodies of water in Ethiopia are mainly based on measurements of DIN concentrations in rivers, river mouths and different parts of the lake. However, the number of observations is very limited. This holds especially for DIN concentrations in rivers of ungauged basins. For these rivers, water discharges are also often unknown, making it difficult to estimate concentrations of DIN. Our model can help fill this gap.

Our estimates of river export of DIN to Lake Tana were far higher than the estimate for the Nile basin provided by Yassin et al. (2010). We estimated DIN yields that ranged from 232 to 2,464 kg km$^{-2}$ y$^{-1}$ among sub-basins of Lake Tana. Yasin et al. (2010) reported DIN yield of 41 to 60 kg km$^{-2}$y$^{-1}$ for the Nile basin, on average. Our higher estimates could be explained by higher population density (228 people per km$^{2}$; Anteneh, 2017) and a larger share of agricultural land (70% )(Abebe and Minale, 2017) of the Lake Tana basin compared to the Nile basin. We think this has to do partly with relatively high N inputs in agriculture. Open animal grazing and open defecation of human excreta likely contribute to higher loads. The Nile basin has a population density of 13 people per km$^2$ and an agricultural land share of 42% (Yasin et al., 2010). Moreover, the High Aswan Dam in the Nile basin has a high nutrient reduction potential, which could result in lower estimates of river export of DIN, as the DIN outlet point was chosen after the dam site (Yasin et al., 2010).

Our estimates of the DIN loads were generally lower than the estimates of the Yangtze and Pearl River basins and Baiyangdian Lake in China (Liu et al., 2008; Strokal et al., 2016; Yang et al., 2019). This is because our study area is smaller than the basins of those rivers. In undisturbed conditions, the DIN export from a temperate watershed is lower than the DIN export from a tropical watershed (Kosten et al., 2009). However, the lower DIN yields in the Lake Tana basin compared to the Chinese basins could be explained by less human activity (for instance, less intensive application of synthetic fertiliser) and by the tropical climate in the Lake Tana basin (warmer temperature resulting in more N retention). These could contribute to lower DIN export to Lake Tana compared to river export of DIN by Yangtze and Pearl rivers. However, the warmer climate in the Lake Tana basin compared to the Yangtze River basin could favour an increased rate of biological $N_2$ fixation, though biological fixation is not a major source of DIN to the Lake Tana.

We found the highest DIN export in the rainy season and the lowest in the dry season for the Lake Tana basin, most likely because runoff reaches a maximum in the rainy season and a minimum in the dry season. McCrackin et al. (2014) reported that the seasonal DIN export was positively related to runoff and negatively to temperature. Agriculture in the Tana basin is rain-fed, and most of the agricultural activities have been carried out in the rainy season. Therefore, in the rainy season compared to the dry season, extensive (about two-thirds of the annual rainfall compared to 2% in the dry season), relatively low air temperature (18.4 $^{o}$C in the rainy compared to 21.1 $^{o}$C in the pre-rainy) and increased human activity (rain-fed agriculture) likely explain the DIN seasonality in river export of DIN to Lake Tana.

We reported that animal manure and synthetic fertilisers were dominant sources of DIN in Lake Tana. This could be explained by the high livestock density and open-grazing system of animal production in the area. In the Lake Tana sub-basin, livestock production comprises a large portion of farming activities (Alemayehu and Tassew, 2017; Tassew and Seifu, 2007).

Synthetic fertilisers and animal manure together explained about half of the river export of DIN to the coastal waters of Africa (Yan et al., 2010). Liu et al. (2007) reported that synthetic fertilisers were the largest contributor to the basin of lake Dianchi in China. Wang et al. (2019) reported that diffuse sources contributed more than 90% of the total dissolved nitrogen (TDN). Strokal et al. (2016) reported that animal manure as a point source (direct discharges to rivers) was the dominant source DIN in Chinese rivers.

### 3.4.3 Implications for eutrophication management

Shore areas and river mouths in the north eastern part of Lake Tana have already experienced eutrophication in the form of algal blooms and extensive growth of water hyacinth (Goshu et al., 2017). This is a result of excess N and P from the sub-basins. Negative impacts of the eutrophication problems on ecosystem and public health have been reported (Goshu and Aynalem, 2017; Wondie et al., 2007). So far, attempts to manage eutrophication, for example, by controlling water hyacinth, have been unsuccessful, and controlling water hyacinth and cyanobacterial nuisance have remained key challenges.

Our model provides new insights into seasonal aspects of managing agricultural N inputs to rivers and thus to the lake. We demonstrate that the river export of DIN is highest in the rainy and lowest in the dry seasons. This implies that the risk of N losses during a high runoff season is higher than during a low runoff season. Therefore, farming practices should avoid N losses to rivers, for the rainy season. Conserving the wetlands along the shores may have a nutrient stripping role during high runoff. This will prevent N export to the lake. Since the major source of N is animal manure, increasing reuse possibilities, for instance, by using animal manure as an organic fertiliser is advisable. Precision fertilisation of crops with respect to N demand and timing may reduce N losses from agricultural fields to rivers and thus eutrophication in the lake.

Our study highlights the seasons (e.g. rainy) and sub-basins (e.g. Gilgel Abay, Gumara, Rib, Gelda, Dirma and Megech) where reduction strategies are needed to avoid future lake pollution. Furthermore, our model identifies the causes of the lake pollution (e.g. manure). This can help policymakers identify adequate reduction policies for Lake Tana.

## 3.5 Conclusions

We integrated the two existing modelling approaches into a seasonal model for river export of DIN for large tropical lakes. This resulted in a seasonal version of the MARINA model for sub-basins and rivers discharging to tropical lakes such as Lake Tana. We applied this model for 2017, during which annual river export of DIN to Lake Tana was about 9 kton. River export of DIN to Lake Tana showed spatial and temporal variability, being highest in the rainy and lowest in the dry seasons. For example, two-thirds of the total annual DIN export was exported by rivers in the rainy season, whereas 30%, 3% and 2% of the DIN was exported by rivers in the pre-rainy, post rainy and dry seasons, respectively. Diffuse sources from agriculture were important contributors of DIN in rivers. Animal manure was the dominant source in all seasons. Synthetic fertiliser was the second dominant source in the rainy season. Human waste on land was a substantial diffuse source of N in all seasons in the year 2017. Over two-thirds of the annual river export of DIN was delivered by six of the twenty rivers, namely, the Gilgel Abay, Dirma, Megech, Rib, Gumara and Gelda. The Gilgel Abay sub-basin alone exported about one-third of the annual river export of DIN to the lake.

Our study shows new insights into the seasonality in river export of DIN to Lake Tana. We show the importance of diffuse sources in the lake pollution. This holds especially for wet periods where pollution levels are higher. We also show the areas (sub-basins) where pollution control is needed to avoid further pollution of the lake. This information can contribute to formulate effective management options for Lake Tana that are area-, season- and source-specific. Furthermore, our seasonal model can be applied to analyses of future trends in the lake pollution. We provide a seasonal approach that might be applied to sub-basins and rivers draining into other large tropical lakes that experience similar environmental problems.

For future research, we suggest three main directions. First, we suggest to set up a new project to monitor concentrations of N in rivers draining into the lakes. This will help to apply the model to other years and validate it for those years. Second, we suggest to conduct sensitivity and uncertainty analyses to increase trust in the model performance especially for those years for which observations are limited. Third, we suggest to apply our model to sub-basins and rivers of other lakes for effective nutrient management. This includes also application of the model to the future. In this future analysis, options to reduce lake pollution can be explored. Our model can also be applied to other sub-basins and rivers of tropical lakes that have comparable characteristics of the sub-basins and rivers of Lake Tana basin (e.g. precipitation, seasons).

**Acknowledgments**

This work was supported by the Netherlands Universities Foundation for International Cooperation (NUFFIC) – Project NFP PhD Goraw Goshu [grant number 5160957055]. The authors would like to acknowledge the Research and Community Service Vice President Office, Bahir Dar University.

# 4

# Importance of Spatial Heterogeneity to Explore Effects of Nutrient Loading on the Ecological Status of Lake Tana

**Goraw Goshu,[a, b, 1] , Micha Veenendaal [a], J.J.M de Klein[a]**

[a] *Aquatic Ecology and Water Quality Management Group, Department of Environmental Sciences, Wageningen University &Research, P.O. Box, 47,6700AA Wageningen, The Netherlands*
[b] *College of Agriculture and Environmental Sciences and Blue Nile Water Institute, Bahir Dar University, P.O. Box 1701, Bahir Dar, Ethiopia*

This chapter is accepted in revised form for publication in Journal of Hydrology

**Abstract**

Understanding spatial variability of nutrient loading and transport in large lakes, and its effect on the eutrophication status are indispensable for effective and sustainable management of a lake and its basin. Yet the level and importance of spatial variability in large shallow tropical lakes and their eutrophication effect is not known, hindering the sustainable management. This modelling study is, therefore, aimed to analyse the importance of spatial heterogeneity in exploring the effects of nutrient loading on tropical shallow lakes, taking Lake Tana as a case study. We applied a novel method of coupling a 2D application of the flow model Duflow with the ecosystem model PC Lake+ (a zero-dimensional food-web model). We defined different impact zones of major tributaries with a tracer model and simulated ecological processes and food web relations for each impact zone. Furthermore, we defined critical nitrogen and phosphorus loads for each impact zone and the whole lake. Subsequently, we analysed the spatial variability of phytoplankton and aquatic vegetation dynamics among the impact zones, as well as the differences in critical loadings. The model results indicate different ecological conditions in the impact zones, and the importance of spatial heterogeneity of eutrophication which is reflected in different critical loadings. The model shows that the North Impact Zone and Northwest Impact Zone are mainly vegetation dominated, and Southeast Impact Zone and Southwest Impact Zone are phytoplankton dominated. Simulated chlorophyll-a concentrations ranged from 0 to 59 µg $l^{-1}$ and the coverage of the vegetation ranged from 1 to 127 g d.w.$m^{-2}$ .These numbers are in line with measurements and observations in the lake. We demonstrate that this model approach can help to identify spatial heterogeneity in hydrology and eutrophication, which should be taken into account as such lakes cannot be regarded as completely mixed systems. Our new approach will aid in setting targets and define measures for specific regions making the management of the lake basin more effective and efficient.

**Keywords**: Spatial heterogeneity, Bifurcation analysis, 2D Tracer model, Critical load, Alternative Stable states, Tropical Lake

## 4.1 Introduction

Eutrophication is one of the most widespread environmental problems of inland water bodies. It has undesired effects on ecosystem state and services, public health, and socioeconomic activities, and some of these environmental effects in lakes include risk of (harmful) algal blooms leading to decrease of water clarity, hypoxia and fish kills, altering biogeochemical processes, and disrupting the aquatic food web (Scheffer, 1998). Eutrophication in lakes is caused by excessive loads of nutrients delivered by rivers (Kane et al., 2014; Lürling and van Oosterhout, 2013). The extent of eutrophication could increase as a result of increased nutrient inputs from the intensification of agriculture and the increase of the human population.

The management and restoration of lakes are often based on reducing the nutrient input to below the critical level. Critical loads represent an ecological threshold at which there is an abrupt change in an ecosystem state, property, or phenomenon, or where small changes in an environmental driver produce large responses in the ecosystem (Scheffer and Carpenter, 2003; Groffman et al., 2006; Xu et al., 2015). It is clearly documented that in shallow lakes eutrophication likely leads to an abrupt shift of conditions, often with alternative stable states either with submerged vegetation or phytoplankton dominated (Scheffer, 1998; Janssen et al., 2014)

Critical nutrient inputs can be derived from aquatic ecosystem models. A number of available lake models, including the ecosystem model PCLake+, was reviewed by Janssen et al. (2019). They conclude that to study state transitions in lakes a mechanistic dynamic model is required that include the aquatic food web. In this way non-linear processes and the relevant feedbacks are accounted for. PCLake meets the above criteria (Janse et al., 2010) and is widely used to explore the transition between alternative stable states and estimating critical nutrient loads. Therefore, we selected this model for our study.

Many aquatic models often assume lake are completely mixed systems. Natural lakes (especially large lakes) have clear spatial differences, both for hydrology as for chemistry. This is due to factors like complex morphology, the different depth, spatially distributed inflow of water and nutrients. In shallow lakes these are mainly horizontal differences. Whereas in deep lakes also vertical gradients exist due to stratification. All lakes may have spatial variability in the system as shown by different lake characteristics and connectivity (Fragoso Jr et al., 2008; Janssen et al., 2017; Liu and Qiu, 2007; Rahm and Danielsson, 2007; Yu et al., 2008). Consequently, we hypothesize that there are different zones in large shallow lakes based on lake hydrology and distributed inflows of rivers ('impact zones'). These impact zones have their own characteristics and can have different critical nutrient loads, which can differ from the critical load when we assume a completely mixed lake.

The spatial heterogeneity of eutrophication has implications for lake management. The effective and sustainable management of eutrophication and the success of lake restoration benefits from a clear understanding of spatial heterogeneity in nutrient loading and hydrology (Janssen et al., 2019a). In this way the effort of reducing nutrient loads in specific subbasins can be focused on the lake parts that are most vulnerable Applying the same management approach to a whole heterogenous lake might not be as efficient as defining more spatially different reduction targets.

The effects of spatial heterogeneity on the different biological components of an aquatic ecosystem have been studied by different authors (Ding et al., 2015; Pringle, 1990; Rychtecký and Znachor, 2011; Soares et al., 2012; Xu et al., 2020) and includes seasonal succession of phytoplankton, macrophyte growth, invasive alien species, and lake turbidity. These studies have shown that hydrology (different residence time) influences spatial heterogeneity, and this is important for a deeper understanding of ecological processes in freshwater lakes. It is also a mechanism maintaining the species diversity of phytoplankton communities.

The modelling of spatial heterogeneity of eutrophication in tropical lakes has not been adequately assessed to date. Most applications of ecological models regard lakes as one complete mixed system. This approach disregards spatial differences, e.g., rivers that discharge on the lake will only affect the lake part near the river mouth. And the littoral zones are generally shallower than the pelagic zones, resulting in different conditions for vegetation growth. In this paper, we modelled spatial differences in Lake Tana, to understand the spatial heterogeneity of eutrophication in a large tropical lake which might facilitate the design of more regional or locally targeted measures.

The main research questions of this study are:

1. What are the different impact zones of the major tributary rivers of Lake Tana? 2. Is there spatial variability of eutrophication (expressed in nutrient and chlorophyll concentrations) among the impact zones and Lake Tana? 3. What are the critical N and P loads for each impact zone? To address these questions, we use a novel modelling approach. First, we define different impact zones of major tributaries in Lake Tana by using a two-dimensional application of the flow and quality model Duflow (Clemmens et al., 1993). Secondly, we determine critical nutrient loads for these impact zones and Lake Tana separately with the ecosystem model PCLake+.

## 4.2 Methodology

### 4.2.1 Study area

Lake Tana is a shallow (maximum 14, average 8 m deep), non-stratifying and large tropical lake found at an altitude of 1800 m a s l. It is the largest lake in Ethiopia, which accounts for 50% of the surface water volume of the freshwater resource of the country. It has a surface area of ca.3111 km$^2$, 28.4 km$^3$ volume, and has a maximum length of 90 km and width of 65 km. The Tana basin has a total drainage area of 16,500 km$^2$ (Fig. 4-1). More than six medium- to large-sized tributary rivers and more than 40 ephemeral streams drain into Lake Tana. Among the rivers, the Gilgel Abay, Dirma, Gumara, Gelda, Rib, and Megech rivers contribute more than 90% of the inflow to the lake (Sirak, 2008). Lake Tana is the source of the Blue Nile, which is the only natural surface outflow. However, there is inter-basin transfer and abstraction of water from the lake. The interbasin transfer has relevance from the water and nutrient balances perspective as the transfer is from Tana basin which consists of Lake Tana, and its tributaries.

Rainfall is maximum during July and August at 250–330 mm per month. Mean annual rainfall is nearly 1280 mm (Abebe and Minale, 2017). The rainy season (July–September) receives about two-thirds of the annual rainfall, while the dry season receives 2%; the pre-rainy season (April–June) receives 25% and the post-rainy season 8% of annual rainfall. The average seasonal air temperature reaches its maximum of 21.1 $^{o}$C in the pre-rainy season and its minimum (18.4 $^{o}$C) in the rainy season, with a large diurnal but small seasonal change.

The population of the basin was projected to be 4.5 million in 2015 (CSA 2007), with a population density of 228 persons per km$^2$ in 2007 (Anteneh, 2017), and 70% of the basin is agricultural land (Abebe and Minale, 2017). In this study, we mainly focused on the largest sub-basins: Gilgel Abay, Dirma, Megech, Gumara, Gelda, Rib, Arno, and Garno (Fig. 4-1).

Lake Tana is an example of a large tropical lake that probably cannot be considered completely mixed. This is caused by multiple inflowing rivers with different discharges and nutrient concentrations. We can roughly conclude that the lake is not a homogenous system. We observe spatial variations in phytoplankton, nutrient concentrations, and vegetation conditions. These variations may result in spatial patterning of critical nutrient loads in Lake Tana. The measured and secondary data about the current trophic status of Lake Tana show spatial variability among the shore, open water, and river mouths of the lake. Most of the river mouths and a significantly smaller portion of the lakeshore area show eutrophic conditions, and most of the lake exhibited more oligotrophic conditions (Goraw Goshu, unpublished).

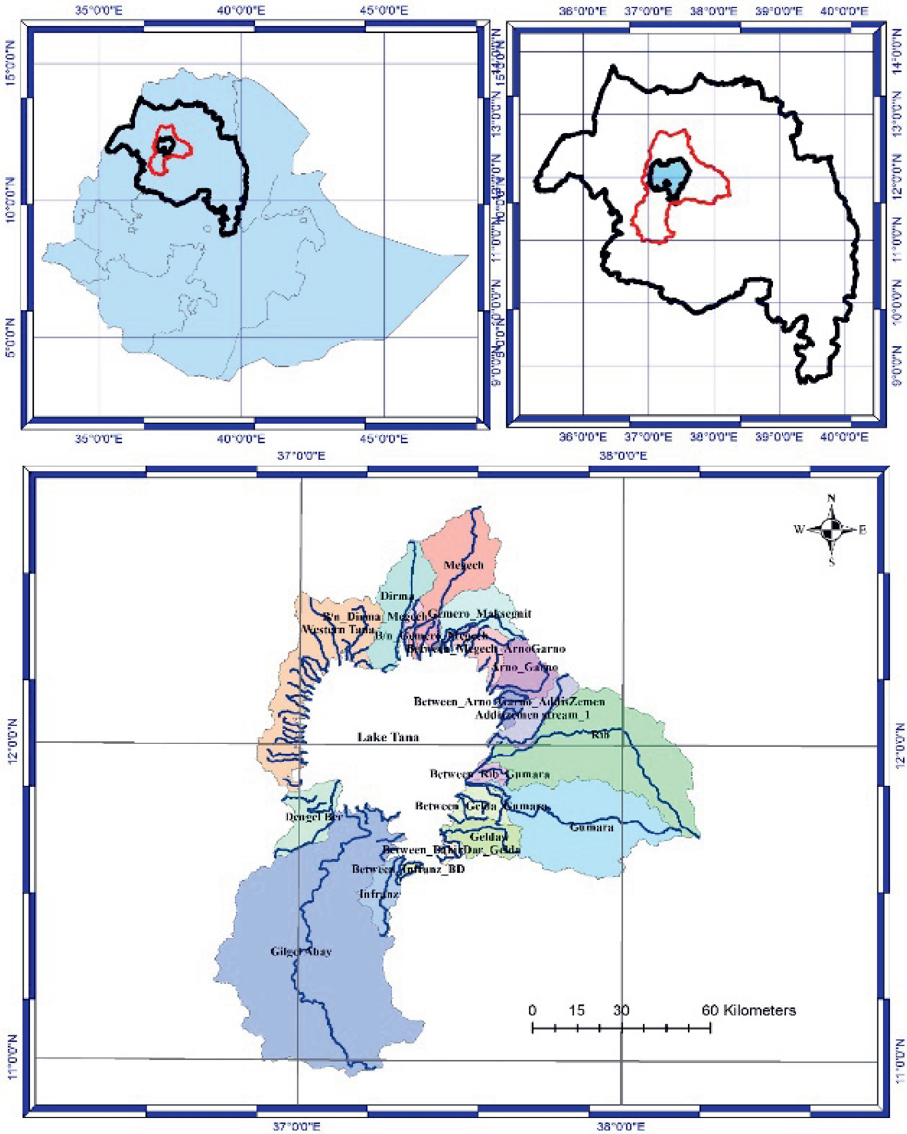



Fig. 4-1 –Ethiopia and the Amhara region (upper left); the Amhara region, the Lake Tana basin and Lake Tana (upper right); the drainage area of the Lake Tana basin and the sub-basins draining into the lake (lower panel).

### 4.2.2 Research approach

We developed a method to facilitate the analysis of spatial variability of eutrophication in a heterogenous lake (summarized in Fig. 4-2). We combined two modelling approaches with modifications for our study area based on the models Duflow (Clemmens et al., 1993) and PCLake+ (Janse et al. 2010). We used a two-dimensional application of the flow and water quality model Duflow, to perform unsteady flow computations and to simulate tracer distribution in the lake to define so-called impact zones. Thereafter, we ran the PCLake+ model for each impact zone separately and for Lake Tana as one mixed system to simulate ecosystem conditions (e.g. nutrient concentrations, phytoplankton, vegetation) and to determine critical nutrient loads for ecosystem transitions (Janse et al., 2010).

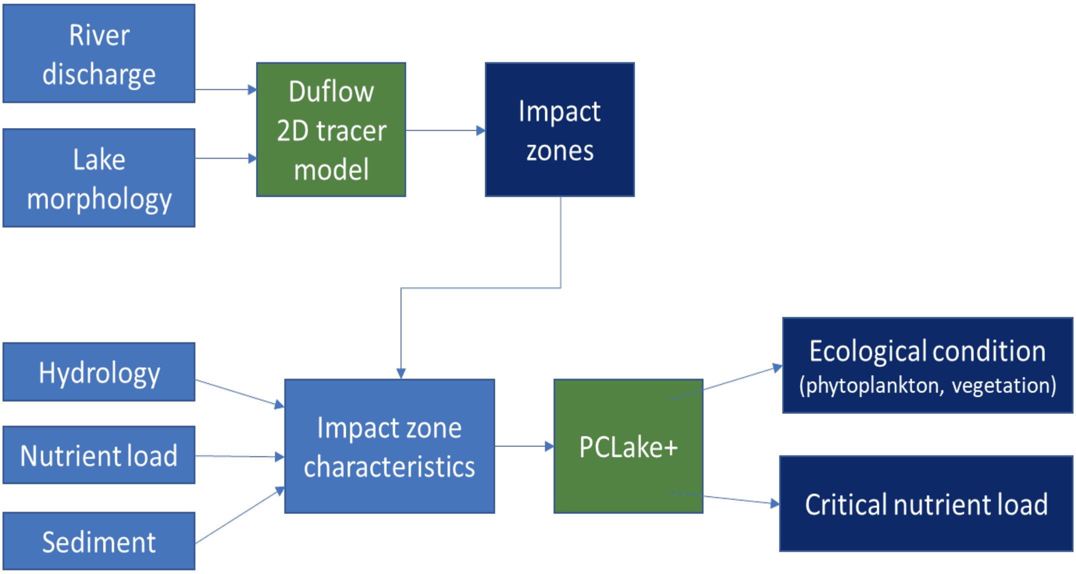



Fig. 4-2 –Conceptual diagram of model coupling of Duflow and PCLake+ (blue: input, green: model, dark blue: result)

### 4.2.3 Modelling Lake Tana with a tracer model in Duflow

**Network schematization**

Duflow is designed as a 1D unsteady flow and quality model. For this study, we set up a 2D application by constructing a dense network of sections in multiple directions to be able to simulate water flow and solute transport in the lake. Duflow solves the transport equation (advection-dispersion), and the user can add specific physical and (bio)chemical processes. We defined a simple tracer model with only transport and no processes. The first step in the configuration of the Duflow model is building the network. We classified the lake into four depth zones: 0-2 m, 2-6 m, 6-10 m, and 10-14 m. The 0-2 m deep zone is defined as shore (4%), 2-6 is littoral (12%), 6-10 is sublittoral (21%), and 10-14 m is classified as a pelagic zone (62%). We built the network on a geographical background by adding the bathymetric map of Lake Tana (Wale et al., 2008). The length of the vertical and horizontal cross-sections (internode distance) ranges from 0.5 km mostly in the shore area to 5 km in offshore zones. We defined the width and length of the sections in such a way that the sum of the section areas corresponded with the total lake area.

**Flow Model**

After network schematization, we defined initial and boundary conditions for the hydrology, and configured the calculation. We defined the boundary conditions, i.e., inflow of a tributary river, at the edge of the lake system, based on discharge monitoring data that were aggregated to monthly discharges. Furthermore, we added water abstractions such as outflow to the hydropower plant (Tana-Beles) and the Blue Nile outlet as negative discharge values and the surplus (total rainfall - evaporation over the lake area ) as negative or positive depending on the net value. We configured the model with a calculation time step of one hour and a one-day output step. Once the flow calculation was verified, we added the quality model.

To explore the spatial distribution of solutes coming from separate inflows, we added a multiple tracer quality model. The model definition was without chemical processes to simulate the transport of a conservative substance. We added four different tracers in a concentration of 100 mg $l^{-1}$ in the main inflowing rivers These are the gauged rivers; no tracer concentration was added to other tributary rivers. Depending on the hydraulic conditions, the tracers are spread over the lake and possibly transported to abstractions including the Blue Nile weir and Tana-Beles outflow.

**Delineation of the 'impact zones' of major incoming rivers with the 2D tracer model**

The impact zone is the area of the lake that is under the specific influence of the feeding tributary river. We focused on the major tributary rivers, and we grouped the major tributary rivers into four zones: South-west (Gilgel Abay), South-east (Gelda, Gumara, Rib), North-east (Arno-Garno), and North (Dirma, Megech). Grouping of tributaries was done as some rivers are close to each other and difficult to separate.

To delineate impact zones of the rivers in the lake, we used two indicators from the tracer model: the concentration of the four tracers in all sections of the lake and the fraction of one tracer from the sum of all tracer concentrations (**eq 1**):

Fout! Bladwijzer niet gedefinieerd.$$\boldsymbol{Fract_i} = \frac{\mathbf{100} * \boldsymbol{Tracer_i}}{\boldsymbol{Tracer_1} + \boldsymbol{Tracer_2} + \boldsymbol{Tracer_3} + \boldsymbol{Tracer_4}} \qquad \textbf{(eq.1)}$$

where, $Fract_i$ is the fraction of tracer 1, 2, 3, or 4, and $Tracer_i$ is the concentration of tracer 1, 2, 3, or 4.

For delineation of the impact zones, we used the criteria that $Fract_i$ >90% and $Tracer_i$ >10 mg $l^{-1}$

### 4.2.4 PCLake+ model configuration for the Lake Tana impact zones

The modelling of N, P critical loads to Lake Tana with PCLake+ encompassed the hydraulic and nutrient loads from major tributaries to the lake. Only the major tributary rivers which are gauged were accounted for. PCLake+ is a zero-dimensional ecosystem model that disregards spatial heterogeneity. The PCLake+ model is a complex ecological model used to assess the nutritional status of stratified and non-stratified freshwater lakes (Janssen et al., 2019b). The PCLake+ model includes biological (phytoplankton, submerged vegetation, and simplified food web) and non-biological (transparency and nutrients) modules (Janse et al., 2010). The model is built in a sequence from primary producers to top predators. N and P cycles in the water column and sediments combine with higher levels in the food web (Janse et al., 2010). The Marsh area is a PCLake+ module, involving a simplified growth model of *Phragmites australis*, an emergent species of macrophytes, combined with vegetative processes in the sediment and the water column of the marsh area (Sollie and Verhoeven, 2008). This area with emergent plants is connected to the open water by a flow of water between them (Janssen et al., 2019b). There are four types of lake-specific settings to parameterize PCLake+: specific lake characteristics, nutrients input, lake hydrology, and sediment type. The specific impact zone characteristics mean depth, impact zone area and main wind direction were computed in this study. We used Arc GIS to calculate the fetch, length of the impact zone in the main wind direction. The mean depth of each impact zone was estimated in ArcGIS using a weighted average technique from the bathymetric map of Lake Tana. Marsh area, as fraction of the lake area was retrieved by taking into account all wetland areas within the catchment of Lake Tana (Aynalem et al., 2017). The mean fetch of each impact zone was calculated by the square root of the surface area. Sediment characteristics are parameterized using the data from previous studies. The Lake Tana sediment composition, texture class, is predominately clay (67 %), sand (17 %), and silt (16%). The average organic matter content of the sediment is 16 g $kg^{-1}$, and the sediment available phosphorus was 19 mg $kg^{-1}$ (Kebedew et al., 2020). .

We set up the models for the year 2017. We quantified the monthly N and P loads from the six major sub-basins of Lake Tana from discharge and nutrient data. The dynamic simulation of the aquatic ecosystem condition in each impact zone consisted of two parts. First, the current condition was simulated, for which the model was initialized and run for 50 years to reach equilibrium (Janssen et al., 2019a). The results in the last 2 years of simulation were taken as final output and chlorophyll-a concentrations were compared with the measurements of 2016 and 2017.

Secondly, a bifurcation analysis was done for each impact zone separately and for Lake Tana as one mixed system. For this, PCLake+ is run sequentially with increasing and decreasing nutrient inputs, to find the critical nutrient loads for a shift from phytoplankton dominance to macrophyte dominance and vice versa. Each individual run for the bifurcation analysis was done over a period of 30 years, and the average of the May-September period in the last 2 years of simulation was taken as the result. In this study, the ratio between the external nutrient loads (N and P load) was measured in 2017. We took this as a reference for the bifurcation analysis with a constant N:P ratio. The critical nutrient loads were used to identify the required reductions of external nutrient loading to restore the clear stable state in the lake. Or, in case the clear state is present, it is used to identify the maximum loads before it switches to the turbid state (Janse et al, 2008).

### 4.2.5 Hydrology and water quality

**Water Balance of Lake Tana**

The water balance is needed to calculate the hydraulic load of the impact zones and the whole lake as input for PCLake+. The water balance was validated with literature data before implementing it. The water balance was calculated for the year 2017, as 2017 is the most complete year for the discharge data to and from the lake and nutrient concentrations (needed for calculating nutrient fluxes). The monthly water balance of Lake Tana was calculated using equation 2.

$$\frac{\Delta S}{\Delta T} = P + Qgauged + Qungauged - E_0 - Q_{HP} - Q_{Dam} \qquad \textbf{(eq.2)}$$

Where,

| | |
|---|---|
| $\frac{\Delta S}{\Delta T}$ | Storage change (closure term) [$L^3\ T^{-1}$] , |
| $P$ | Precipitation over Lake Tana [$L^3\ T^{-1}$] |
| $Qgauged$ | Surface water discharge from gauge reading near lake inlet [$L^3\ T^{-1}$] |
| $Qungauged$ | Surface water discharge from ungauged towards Lake Tana [$L^3\ T^{-1}$] |
| $E_0$ | Penman open water evaporation [$L^3\ T^{-1}$] |
| $Q_{HP}$ | Surface water discharge towards Hydropower plant Belles [$L^3\ T^{-1}$] |
| $Q_{Dam}$ | Surface water discharge towards Blue Nile river [$L^3\ T^{-1}$] |

The surface-water flow to Lake Tana includes six major sub-basins (*Table 4-1*) which are gauged and 14 sub-basins which are not gauged. Groundwater flow within the Lake Tana Basin takes place in a heterogeneous aquifer system with the groundwater flow converging to Lake Tana, however, groundwater flow also takes place from the Lake Tana Basin towards the adjacent Beles River basin (Mamo et al., 2016; Nigate et al., 2017). Furthermore, as concluded in Mamo et al. (2016 ), estimated groundwater input to the lake was found to be of minor importance and therefore groundwater flow was excluded from the water balance for this research. The detailed computation of the water balance components is described in the supplementary materials.

*Table 4-1.* Sub-basins of the Lake Tana basin and the percentage of the surface area of the specific sub-basin within the total Lake Tana Basin. Source Goshu et al., 2020

| Sub-basin | Surface area sub-basin ($km^2$) | Percentage surface area of the total basin (%) | Gauged (yes/no) |
|---|---|---|---|
| Gilgel Abay | 3866 | 34.4 | Yes |
| Arno-Garno | 269 | 2.4 | Yes |
| Dirma | 461 | 4.1 | Yes |
| Gelda | 261 | 2.3 | Yes |
| Gumara | 1376 | 12.2 | Yes |
| Megech | 659 | 5.9 | Yes |
| Rib | 1727 | 15.4 | Yes |

**Hydraulic loadings**

The hydraulic loadings of tributary rivers in each impact zone and the Lake Tana were computed by dividing the sum of the discharges of the tributary rivers in the impact zone by the area of the impact zone, estimated based on the tracer model study (eq.3). Water discharge data of a tributary for months in 2017 were obtained from the Ethiopian Ministry of Water, Electric and Irrigation.

Fout! Bladwijzer niet gedefinieerd. $$\boldsymbol{HLR} = \frac{\sum_{i=1}^{m} Q_i}{A} \qquad \textbf{(eq.3)}$$

Where,

HLR is the hydraulic loading rate (m $s^{-1}$),

$Q_i$ is the discharge at the gauging stations at a site I ($m^3$ $s^{-1}$),

A is the area of the impact zone ($m^2$),

m is the number of discharge measurements.

**Nutrient loadings**

The water sampling and analysis of the different nitrogen and phosphorus species are described in detail in Goshu et al. (2020). The DIN or DIP load to the impact zone and Lake Tana (g $s^{-1}$) is a product of DIN or DIP concentration (g $m^{-3}$) and discharge ($m^3$ $s^{-1}$)(see eq.4). Based on our data set, because daily discharge and monthly nutrient concentrations were available for the year 2017, the average load to the impact zone was calculated by multiplying the average discharge and the average concentration (De Vries and Klavers, 1994), providing a formula as follows :

Fout! Bladwijzer niet gedefinieerd.$$\boldsymbol{L} = \frac{\sum_{i=1}^{m} Q_i.}{m} . \frac{\sum_{j=1}^{n} Cj.}{n} \qquad \textbf{(eq.4)}$$

Where,

L is a load of DIN or DIP (g $s^{-1}$),
$Q_i$ is the discharge at the gauging stations at the time I ($m^3$ $s^{-1}$),
$C_j$ is the concentration at different times j (g $m^{-3}$),
m is the number of discharge measurements; n is the number of concentration measurements.

Chlorophyll-a was measured as an indicator for the phytoplankton concentration. During three periods (August 2016, December 2016, March 2017) water samples for chlorophyll-a analysis were taken at 143 locations at Lake Tana in a spatial grid of 5 km by 5km, and chlorophyll-ac samples were also collected from the river monthly for 2017. The chlorophyll-a concentration was determined using the acetone extraction procedure (Wetzel and Likens, 1991). From the analysis, a map of the chlorophyll-a over the lake was constructed using the Inverse Distance Weighted (IDW) interpolation technique in ArcGIS.

To classify the current condition in Lake Tana, the trophic status of the lake and river mouths can be defined at different levels (oligotrophic to hypertrophic). The indicators used for this classification are total phosphorus (μg $l^{-1}$) and Chlorophyll-a (μg $l^{-1}$) (Cunha et al. (2013). The trophic status index of Cunha et al (2013) is specifically developed for tropical and subtropical lakes.

## 4.3 Results

### 4.3.1 Current ecological condition of the lake based on measurements

**Current status of eutrophication**

The trophic status of Lake Tana clearly showed spatial variability among the shore, littoral, pelagic, and river mouth stations of the lake. Most of the river mouths and a small part of the lakeshore and littoral areas showed eutrophic conditions (Cunha et al. (2013), especially after the rainy season (Jun-Sep). The majority of the lake exhibited oligotrophic conditions based on chlorophyll-a concentration (see Fig. 4-3).

### 4.3.2 Chlorophyll-a in littoral, pelagic, and river mouths of Lake Tana

The chlorophyll-a concentration in Lake Tana shows high spatial and temporal variability and ranged from <0.5 to 191 µg $l^{-1}$ (Fig. 4-3). Generally, the chlorophyll-a concentrations are highest in the river mouths (average 87 µg $l^{-1}$) and lowest in the pelagic zone of the lake (average 14 µg $l^{-1}$) except for some inshore and littoral areas where temporally low chlorophyll-a concentrations were found. This might be caused by the presence of water hyacinth infestation and its ecological effect of depleting nutrients and light blockage (Gezie et al., 2018).

The three sampling periods (August 2016, December 2017, March 2017) show clear seasonal variability in chlorophyll-a concentrations. During the rainy season (August) more than 85% of Lake Tana had chlorophyll-a concentrations below 4 µg $l^{-1}$. In December 2016 (post rainy season), the chlorophyll-a concentration ranged from <0.5 µg $l^{-1}$ to 148 µg $l^{-1}$, and more than half of the lake (56 %) had chlorophyll-a concentration between 4 and 27 µg $l^{-1}$. Finally, in March 2017 (dry season), the chlorophyll-a concentrations ranged from < 0,5 µg $l^{-1}$ to 191 µg $l^{-1}$, with 59 % of the lake had concentrations above 20 µg $l^{-1}$.

Spatial and temporal variations of chlorophyll-a concentrations are observed also in the river mouths of Dirma, Gilgel Abay, Gelda, Gumara, and Rib (Table 4-2).

The chlorophyll-a concentrations in the river mouths were generally higher than in the lake (22.8 - 351.5 μg $l^{-1}$) . Spatially, the minimum chlorophyll-a concentration was noted in the G/Abay River mouth, and the maximum was measured in the Rib River mouth. Nevertheless, there were no significant spatial and temporal differences in chlorophyll-a concentrations among the river mouths ($P > 0.05$, Repeated Measures ANOVA, n = 60).

Based on the monthly measurements in 2017, the minimum chlorophyll-a was noted in the rainy season and the maximum was noted in post rainy and pre rainy seasons for most of the river mouths.

*Table 4-2.*The minimum and maximum Chl-a concentration (μg $l^{-1}$) of the river mouths of major tributary rivers of Lake Tana in 2016/2017 (n=12).

| Statistic | Dirma mouth | Gilgel Abay mouth | Gelda mouth | Gumara mouth | Rib mouth |
|---|---|---|---|---|---|
| Min | 27 | 1 | 15 | 8 | 23 |
| Max | 94 | 78 | 259 | 185 | 352 |


| Statistic | Dirma mouth | Gilgel Abay mouth | Gelda mouth | Gumara mouth | Rib mouth |
|---|---|---|---|---|---|
| Min | 27 | 1 | 15 | 8 | 23 |
| Max | 94 | 78 | 259 | 185 | 352 |

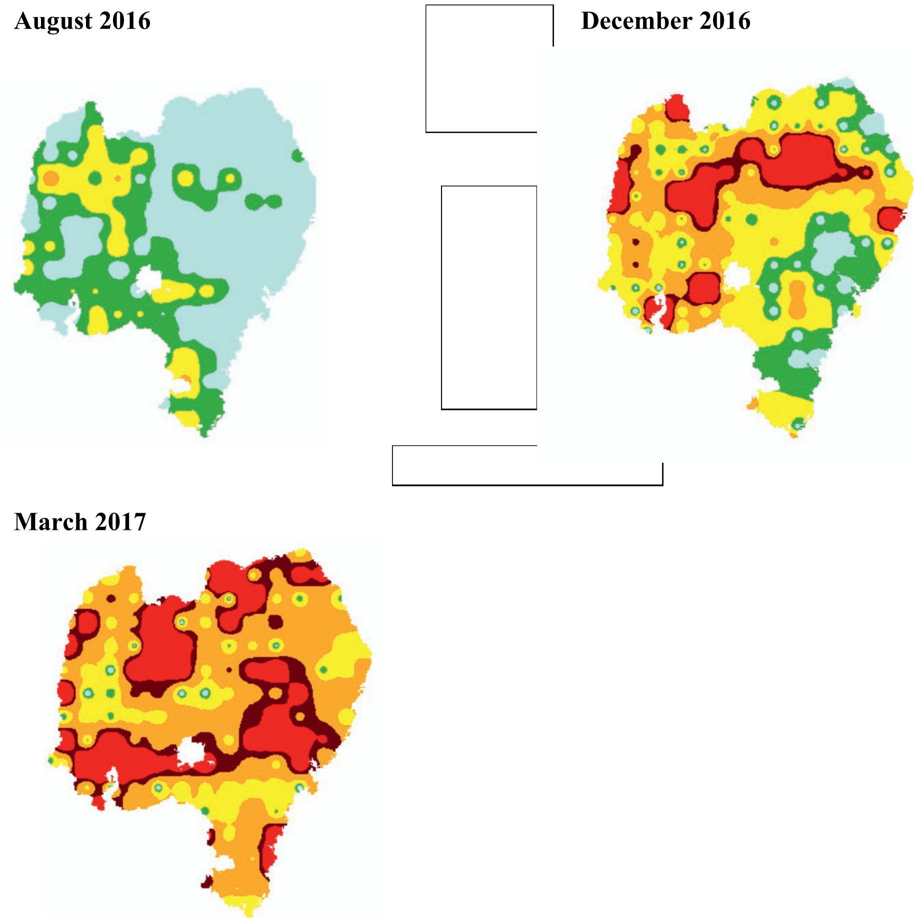



Fig. 4-3 –Spatial Maps of measured Chl-a (μgl$^{-1}$) for littoral, pelagic, and river mouths of Lake Tana for months of the year 2016/2017

### 4.3.3 Delineation of the 'impact zones' of major incoming rivers

The contours of the impact zones were defined by the tracer study for May, September, and December of the year 2017 (Fig. 4-4). We observed a direct relationship between the river basin area and the impact zone area. The larger the river basin area and the subsequent river discharge, the larger the impact zone. The southwest impact zone has the largest area in all months study except December. The northeast impact zone has the smallest area in all months of the model study. Generally, the impact zones have the smallest area in the dry season (May), and the largest in the post rainy season (December). Surprisingly, the impact zones are mainly confined in the shore and littoral parts of the lake and do not cover the pelagic part (Fig. 4-4).

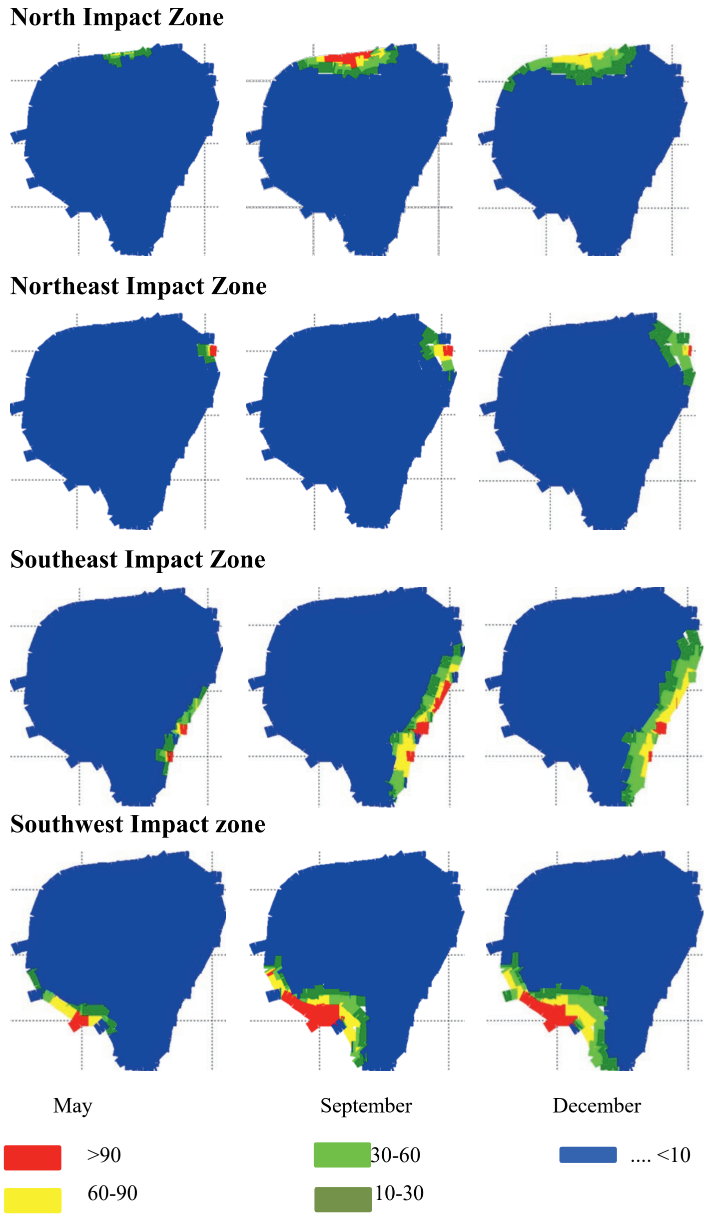



Fig. 4-4 – Spatial coverage of the different impact zones in Lake Tana defined by the tracer study, in May, Sep, and Dec of the year 2017. (colours indicated tracer concentration (mg $l^{-1}$); initial concentration was 100 mg $l^{-1}$).

Based on the impact zones delineation we computed the different characteristics of the separate zones needed for the PCLake+ modelling. These include the fetch, marsh area and mean depth (*Table 4-3*), and the monthly N and P load (*Table 4-5*).

*Table 4-3.* Impact zones defined by tracer study, and their area in May, Sep., and Dec. of the year 2017; including mean fetch, marsh area, and mean depth. North Impact Zone-Dirma and Megech; North East Impact Zone-Arno-Garno; South East Impact Zone-Rib, Gumara and Gelda, and South West Impact Zone-GilgelAbay

| **Impact zone** | **May** | **Sep** | **Dec** |
|---|---|---|---|
| Area (km$^2$) | | | |
| North | 43.6 | 213 | 308 |
| North East | 30 | 88 | 150 |
| South East | 102 | 362 | 490 |
| South West | 134 | 365 | 475 |
| Fetch (km) | | | |
| North | 7 | 15 | 18 |
| North East | 5 | 9 | 12 |
| South East | 10 | 19 | 22 |
| South West | 12 | 19 | 22 |
| Marsh Area (km2) | | | |
| North | 0.10 | 0.50 | 0.72 |
| North East | 0.07 | 0.21 | 0.35 |
| South East | 0.24 | 0.84 | 1.14 |
| South West | 0.31 | 0.85 | 1.11 |
| Mean depth (m) | | | |
| North | 2.00 | 2.79 | 2.39 |
| North East | 0.21 | 2.00 | 1.6 |
| South East | 2.64 | 4.43 | 4.05 |
| South West | 1.63 | 3.41 | 3.03 |

### 4.3.4 Hydrology and water balance

The water balance study of Lake Tana shows the water level is fluctuating a lot. The water level drops during dry and pre rainy seasons and increases in the rainy season (Fig. 4-5, *Table 4-4*). This water level change is well considered in the Duflow modelling. Fig. 4-5 displays the seasonal variation in water balance of Lake Tana for 2017. Table 4-4 gives the calculated annual water balance terms of Lake Tana. The main inflow of water to Lake Tana is coming from the gauged sub-basins. Furthermore, evaporation exceeds precipitation on an annual basis (*Table 4-4*). The lake exhibits water storage from May until September, with the largest positive storage for July and August. Despite the outflow flowing into the hydropower plant Tana Beles, and the monitored part of the Blue Nile, there is a large cumulative closure term in the water balance of Lake Tana at the end of the year, which should be compensated by an additional loss term from the lake. This large closure term might be because the largest part of the outflow towards the Blue Nile River is not monitored.

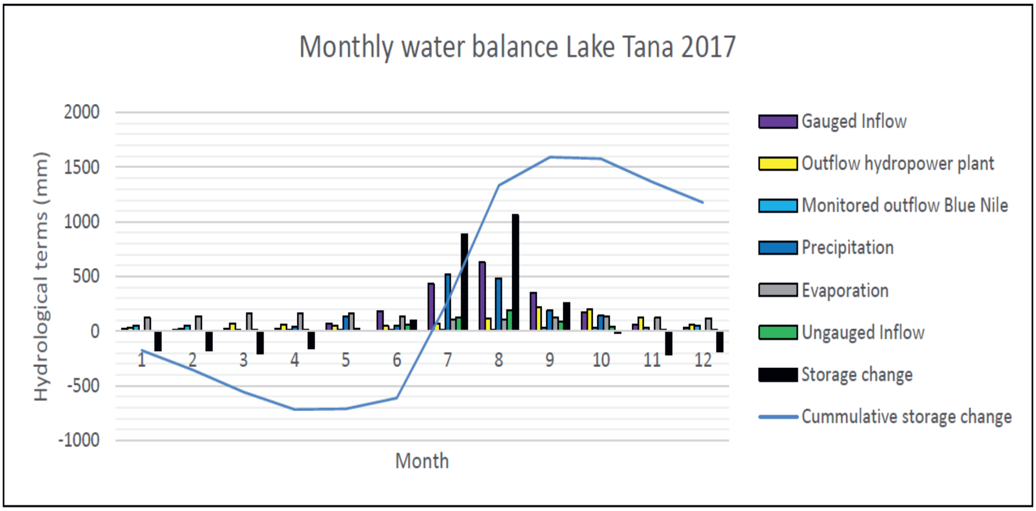



Fig. 4-5–Monthly water balance of Lake Tana for the year 2017

*Table 4-4.* Calculated water balance of Lake Tana for the year 2017

| Water balance components | mm/year | Mm³/year |
|---|---|---|
| Qgauged | +2002 | +6229 |
| $Q_{HP}$ | +1064 | -3311 |
| $Q_{Dam}$ | -319 | -994 |
| P | +1561 | +4857 |
| $E_0$ | -1584 | -4928 |
| Qunguaged | +582 | +1810 |
| $\frac{\Delta S}{\Delta T}$ | +1178 | +3664 |

### 4.3.5 Simulation of ecological indicators and critical nutrient loads of the four impact zones and Lake Tana

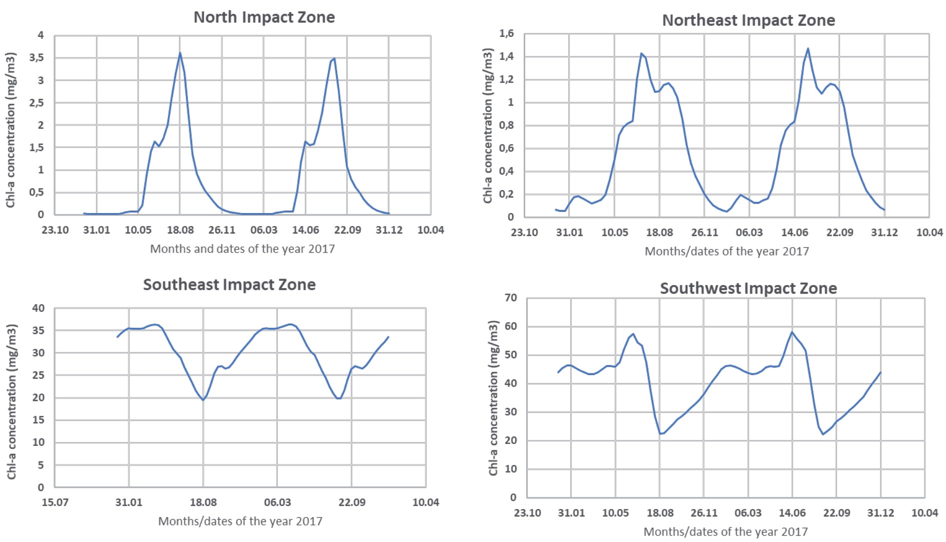
The simulation of phytoplankton (expressed as chlorophyll-a) showed that chlorophyll-a was more dominant in the Southeast and Southwest zones (maximum chlorophyll-a 40 and 70 $\mu g$ $l^{-1}$), than the North and Northeast zones (maximum chlorophyll-a 3,5 and 1.6 $\mu g$ $l^{-1}$)(see Fig.6a). Contrastingly, vegetation was more dominant in the North and Northeast impact zones (maximum vegetation = 135 and 245 g d.w. $m^{-2}$), than the Southeast and Southwest impact zones (maximum vegetation = 0.2 and 0.2 g d.w. $m^{-2}$) (Fig. 4-6b). The simulation results of TN, TP, and Secchi depth are shown in Annex 2.

The bifurcation analyses of the four impact zones and Lake Tana as one mixed system are presented in Fig. 4-7and Fig. 4-8 .They show clear critical loadings, except for Northeast zone , where a linear increase of chlorophyll with increasing P-load is found. The North and Southwest impact zones display hysteresis (2 critical loadings with alternative stable states). This bifurcation run was done with a changing P-load throughout the year, coupled to the N-load with an TN/TP ratio of 2.33 based on the incoming nutrient fluxes. This analysis gives an impression of the effect of a combined reduction of N- and P- loads to the impact zones. The average measured P-load to the impact zones were 0.0017, 0.0031, 0.0108, and 0.0146 (g $m^{-2}$

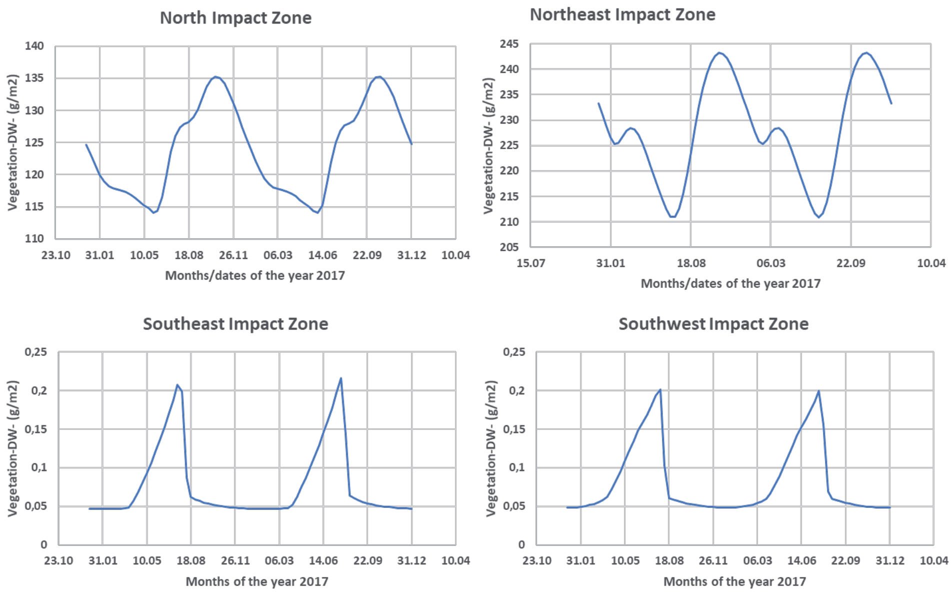
d$^{-1}$) and the average N-load to the impact zones is 0.0075, 0.0226, 0.0109, and 0.0240 (g m$^{-2}$ d$^{-1}$) (*Table 4-5*).

*Table 4-5*. N and P loads of the different impact zones for the months of the year 2017

| | **North Zone** | | **Northeast Zone** | | **South East Zone** | | **South west Zone** | |
|---|---|---|---|---|---|---|---|---|
| | Nload | Pload | Nload | Pload | Nload | Pload | Nload | Pload |
| Months | g m$^{-2}$ d$^{-1}$ | g m$^{-2}$ d$^{-1}$ | g m$^{-2}$ d$^{-1}$ | g m$^{-2}$ d$^{-1}$ | g m$^{-2}$ d$^{-1}$ | g m$^{-2}$ d$^{-1}$ | g m$^{-2}$ d$^{-1}$ | g m$^{-2}$ d$^{-1}$ |
| Jan | 0.0001 | 0.0000 | 0.0002 | 0.0002 | 0.0003 | 7.14E-05 | 0.0008 | 0.0003 |
| Feb | 0.0024 | 0.0002 | 0.0035 | 0.0027 | 0.0009 | 4.10E-04 | 0.0062 | 0.0037 |
| Mar | 0.0001 | 0.0001 | 0.0015 | 0.0006 | 0.0028 | 2.52E-03 | 0.0017 | 0.0012 |
| Apr | 0.0012 | 0.0004 | 0.0000 | 0.0003 | 0.0025 | 6.07E-03 | 0.0017 | 0.0012 |
| May | 0.0026 | 0.0002 | 0.0100 | 0.0010 | 0.0054 | 1.02E-02 | 0.0311 | 0.0046 |
| Jun | 0.0250 | 0.0100 | 0.0195 | 0.0008 | 0.0143 | 1.20E-02 | 0.1009 | 0.0752 |
| July | 0.0211 | 0.0080 | 0.0985 | 0.0040 | 0.0241 | 1.25E-02 | 0.0439 | 0.0365 |
| Aug | 0.0271 | 0.0150 | 0.1138 | 0.0060 | 0.0433 | 2.09E-02 | 0.0352 | 0.0000 |
| Sep | 0.0083 | 0.0018 | 0.0161 | 0.0033 | 0.0299 | 6.42E-02 | 0.0426 | 0.0345 |
| Oct | 0.0015 | 0.0012 | 0.0057 | 0.0008 | 0.0063 | 6.39E-04 | 0.0200 | 0.0177 |
| Nov | 0.0005 | 0.0001 | 0.0012 | 0.0003 | 0.0004 | 4.08E-04 | 0.0024 | 0.0001 |
| Dec | 0.0002 | 0.0000 | 0.0006 | 0.0001 | 0.0003 | 2.02E-04 | 0.0011 | 0.0003 |
| Ave. | 0.008 | 0.003 | 0.023 | 0.002 | 0.011 | 0.011 | 0.024 | 0.015 |




(a)




(b)

Fig. 4-6 – Simulation results of chlorophyll-a (mg $m^{-3}$, panel a) and vegetation (g dry weight $m^{-2}$, panel b) of the four impact zones in Lake Tana in 2017, during the last 2 years of the simulation period.

The bifurcation analysis (Fig. 4-7 and Fig. 4-8) shows that for the two impact zones that are in the turbid state (high chlorophyll-a) the external nutrient loading should be reduced considerable (to around 10% of the current load) to switch back to the clear vegetation state. In contract, the current nutrient loads for the two impact zones that exhibit vegetation dominance (clear state) are predicted to be below the critical values.

*Table 4-6.* The current average P-loads, reduction/increase factors (Turbid to clear and Clear to Turbid), and critical loads (Turbid to Clear and Clear to Turbid) of the separate impact zones and Lake Tana.

| | current average P-load (mgP $m^{-2}$ $d^{-1}$) | multiplier Turbid to Clear (-) | multiplier Clear to Turbid (-) | critical load Turbid to Clear (mgP $m^{-2}$ $d^{-1}$) | critical load Clear to Turbid (mgP $m^{-2}$ $d^{-1}$) |
|---|---|---|---|---|---|
| **North Zone** | 3.0 | 2.5 | 9.5 | 7.5 | 28.5 |
| **Northeast Zone** | 2.0 | 0.05 | 0.1 | 0.1 | 0.2 |
| **South East Zone** | 11 | undetermined | undetermined | | |
| **Impact zone 4** | 15 | 0.2 | 0.7 | 3 | 10.5 |
| **South west Zone** | 3.9 | 0.05 | 0.2 | 0.20 | 0.78 |

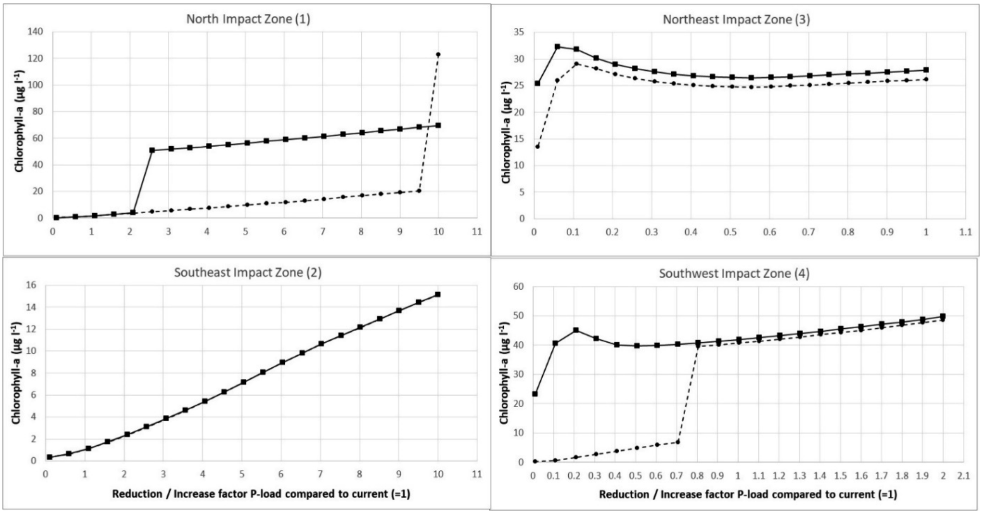



Fig. 4-7–Bifurcation analysis results of the four impact zones in Lake Tana. On the x-axes are the reduction factor (<1) or increase factor (>1) of the P-load; on the y-axes average chlorophyll

concentration in the last 2 years (of 30-years run) in the months May – September. Dotted line is transition from clear to turbid state, solid line is transition from turbid to clear state.

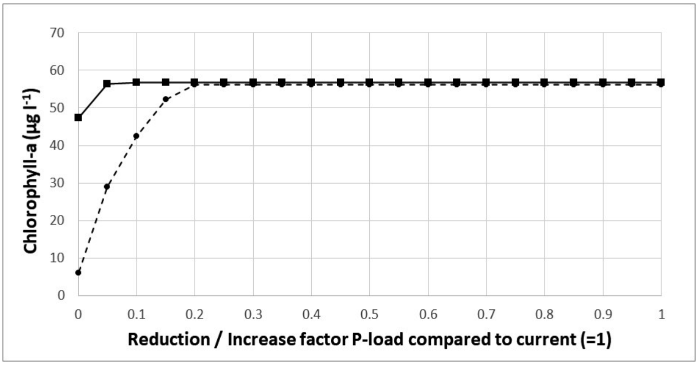



Fig. 4-8 – Bifurcation analysis for the whole Lake Tana when modelled as one mixed system. On the x-axes the reduction factor (<1) or increase factor (>1) of the P-load; on the y-axes average chlorophyll concentration in the last 2 years (of 30-years run) in the months May – September. Dotted line is transition from clear to turbid state, solid line is transition from turbid to clear state.

### 4.4 Discussion

## 4.4.1 Spatial heterogeneity

With the approach presented here we show that spatial variability can be an important factor in analysing the ecological condition and determining critical loads in a large lake. The variability in the lake conditions is strongly related to the locations and magnitude of the in flowing rivers. The results for the separate impact zones show clear differences, among each other and when compared to the situation that the lake is modelled as one mixed system. Nevertheless, we acknowledge the uncertainties that are involved in application of the PCLake+ model when only very limited measured data are available.

There is high vegetation and low chlorophyll-a concentrations in North and Northeast impact zones, and the other way around in Southeast and Southwest impact zones. This finding is consistent with the reports of Dersseh et al. (2019) and Wondie et al. (2012). However, the simulated chlorophyll-a concentrations are deviating to some extent from the observed chlorophyll-a concentrations. Still, the general pattern is in line with the observations, especially if we consider the chlorophyll-a measurements at the river mouths. The pattern of higher concentrations in March 2017 compared with August 2016 is clear and represents the actual observations. However, the absolute chlorophyll-a concentrations contain uncertainties. Possible reason could be that we ran PCLake+ with default settings and not specifically calibrated for tropical lakes.

Our finding of spatial variability of eutrophication in Lake Tana is also reported in other tropical and sub-tropical large lakes of the world. Lake Taihu (Southeast China) is a good example of a large shallow lake with high spatial variation; macrophytes are established at the shores and in the bays, whereas they are absent in the lake's centre due to strong wind forces (Janssen et al., 2014; Xu et al., 2015b; Zhao et al., 2013). The study by Soares et al. (2012) reported that the tropical Funil Reservoir (Brazil) demonstrated spatial heterogeneity that could affect the occurrence and distribution of algal blooms along the reservoir, the fluvial, intermediate, and lentic compartments. This supports the conclusion that high spatial heterogeneity in macrophytes and phytoplankton abundance suggests variation in the responses to eutrophication within the lake. This variation results in spatial patterning of critical nutrient loads in Lake Tana.

### 4.4.2 Critical loads and bifurcation analysis

The critical nutrient loads for the impact zones and for the lake are summarized in Table 4-6.

These are derived from the average current loads (*Table 4-5*) and the reduction factors (multipliers) from Fig. 4-7 and Fig. 4-8. Given the variability of the current monthly loads an average annual value has only limited meaning. However, the critical loads estimated here enable a comparison with other studies in lakes where critical loads were estimated.

From Table 4-6, it is evident that current nutrient loads to Southeast and Southwest impact zones are higher than the estimated critical loads to these impact zones. On the other hand, we found a different ecological condition in North and Northeast impact zone, where the phytoplankton is less dominating.

Lake Tana is a large tropical lake with a mean depth of 8 m, but the impact zones are generally only 2 m deep. This is comparable to shallow Dutch and Chinese lakes where PCLake+ was applied. In Janse et al. (2008), critical P-loadings were determined that can be considered representative for Dutch shallow lakes. It was observed that for a default lake (see exact settings in Janse, et al., 2008), critical bifurcation points of the lake, turning from a turbid to clear, and clear to the turbid state were 0.9 mgP $m^{-2}$ $d^{-1}$ and 3 mgP $m^{-2}$ $d^{-1}$ respectively. The general numbers for Lake Tana are similar (Table 4-6) , but the range is larger (0.1 to 28.5 mgP $m^{-2}$ $d^{-1}$), possibly due to the spatial heterogeneity of the impact zones with different characteristics. Studies with PCLake+ for deeper lakes, such as described in (Li et al., 2019) for Lake Dianchi with an average depth of 4.9 m, showed critical loadings of 0.34 mgP $m^{-2}$ $d^{-1}$ and 0.38 mgP $m^{-2}$ $d^{-1}$. Furthermore, lakes with larger depths will have lower critical P-loadings because less submerged macrophytes will be present, due to less favourable light conditions in deeper lakes (Janse, 2005; Janse et al., 2008). This causes more nutrients available for the growth of phytoplankton causing more turbid conditions at lower nutrient loadings.

### 4.4.3 Model uncertainties

PCLake+ is originally parameterized for temperate shallow lakes. and not specifically for shallow tropical lakes. This was done based on a dataset of around 40 lakes (Janse et al. 2010). Such an intensive calibration is not yet done for tropical lakes, but the ecological and chemical concepts are generic so we believe applications in other regions can be justified. This is supported by studies where the model in its present form has been applied successfully in several tropical and subtropical lakes (Janssen et al 2016; Li et al. 2019; Kong et al 2017;

Fragoso et al. 2008). Still, the results should be considered with care. Lake Tana is one of the first attempt of applying PCLake+ to African lakes, and further studies in other lakes will gain more insight. We acknowledge that the derived critical loads for Lake Tana contain uncertainties. Comparison with critical loads in other regions has limited value, but it shows that the values for Lake Tana can be plausible. Future monitoring and modelling in tropical lakes can contribute to reducing the uncertainties. Nevertheless, we believe that the general conclusion of importance of spatial heterogeneity in large shallow lakes is justified here.

### 4.4.4 Implications

Studies conducted in China and Brazil have shown the importance of spatial heterogeneity of shallow lakes (Janssen et al, 2017; Fragoso et al., 2008). For ecological modelling as well as defining nutrient management options large tropical lakes like Lake Tana are traditionally treated as completely mixed systems. Because of this, the traditional approach of nutrient management has been only limited successful. Our research in this lake has clearly shown that importance of spatial heterogeneity in terms of eutrophication effects and critical nutrient loads.

Our spatial modelling provides new insights in exploring the effects of eutrophication in a large shallow tropical lake. We demonstrate that Lake Tana is a heterogeneous lake and the conditions that drive this are likely to be existing also in other lakes. The modelled impact zones clearly show spatial differences in N and P critical loadings, implying that each zone has a different nutrient reduction target. Therefore, in impact zones like the Northeastern and western parts of Lake Tana, priority should be given to nutrient reduction measures. Lake managers could aim to reduce the present nutrient load to below the lowest critical nutrient load in northeast and western impact zones. The major N and P sources to Lake Tana are animal manure, synthetic fertilizers, and human waste (Goraw et al., 2020). Therefore, nutrient reduction measures from the above sources should be practiced in sub-basins of prioritized impact zones of Lake Tana.

## 4.5 Conclusions

This study showed that there are clear distinct impact zones in Lake Tana, related to the main tributaries, with different characteristics. Both the measurements and the modelling with PCLake+ show that the current ecological status has spatial variability over the lake and among the impact zones. Critical nutrient loadings vary among the impact zones, implying that each zone has a different reduction target. For Lake Tana in North Impact Zone and Northeast Impact Zone, the current level of P load is below the critical loading. In the other two zones, the current loading is above the critical loading, so in the North-eastern and Western part of the lake, the river loading should be decreased substantially. The critical loadings of N and P are quite different when we take Lake Tana as one mixed system. Our findings demonstrate the importance of accounting for spatial heterogeneity of eutrophication to determine sub-catchment critical nutrient loads in a relatively large lake. We acknowledge that the derived critical loads for Lake Tana contain uncertainties. And comparison with critical loads in other regions has limited value, but it shows that the values for Lake Tana can be plausible. Future monitoring and modelling in tropical lakes can contribute to reducing the uncertainties. Nevertheless, we believe that the general conclusion of importance of spatial heterogeneity in large shallow lakes is justified here.

**Acknowledgments**

This work was supported by the Netherlands Universities Foundation for International Cooperation (NUFFIC) Project NFP Ph.D. Goraw Goshu [grant number: 5160957055]. Finally, we thank the Blue Nile Water Institute, the Research and Community Service Vice President Office for their financial and logistic support. We thank Aron Ateka for providing us chlorophyll-a data.

# 5

# Performance of Faecal Indicators of Bacteria, Microbial Source Tracking, and Pollution Risk Mapping in Highland Tropical Water

**Goraw Goshu,[a, b, 1] , A.A. Koelmans [a], J.J.M de Klein[a]**

[a] *Aquatic Ecology and Water Quality Management Group, Department of Environmental Sciences, Wageningen University &Research, P.O. Box, 47,6700AA Wageningen, The Netherlands*

[b] *College of Agriculture and Environmental Sciences and Blue Nile Water Institute, Bahir Dar University, P.O. Box 1701, Bahir Dar, Ethiopia*

[c]

This chapter is published in revised form as:

Goshu G, Koelmans AA, de Klein JJ. Performance of faecal indicator bacteria, microbial source tracking, and pollution risk mapping in tropical water. Environmental Pollution. 2021 May 1;276:116693. ISSN 0269-7491, https://doi.org/10.1016/j.envpol.2021.116693.

**Abstract**

Water pollution by pathogenic microorganisms causes public and ecosystem health risks. Faecal Indicator Bacteria (FIB) were used for the assessment of faecal pollution and possible water quality deterioration. There is growing evidence that FIB used in temperate regions is not adequate and reliable to detect faecal pollution in tropical regions, such as the Lake Tana basin. This study aims to evaluate the performance of FIB, including total coli forms (TC), *Escherichia coli* (EC), Enterococci (IEC), and *Clostridium perfringens* (CP). In addition to FIB, for the purpose of microbial source tracking (MST), a ruminant associated molecular marker was applied at different water types and altitudes and map bacteria pollution risks based on consensus FIB. The performances of indicators were evaluated at 22sites, from different water types and altitudes, sampled monthly from June to December 2017, and risk maps were developed for all sub-basins of Lake Tana using consensus FIB. Physicochemical measurements were also taken to characterize study sites. The physicochemical data indicated a diverse range of aquatic habitats and pollution patterns that formed the basic framework for testing the performance of selected FIB. The performance of FIB varied with altitude and water type. *E.coli* and CP results indicated a consensus picture of faecal pollution in most of the investigated environments. Markers associated with ruminants (BacR) were identified in more than three-fourths of the sites, indicating the presence of mainly ruminant-associated faecal pollution. We found moderate to high levels of faecal pollution in most sub-basins, with highest levels in the rainy season. Based on the consensus parameters, a pollution risk map for sub-basins of Lake Tana was developed, including the un-gauged sub-basins. Our research can aid to improvements to water quality testing and reduce risk to the general population from in stream bacteria.

## 5.1 Introduction

Water pollution by pathogenic microorganisms from human and animal waste causes public and ecosystem health risks. Management of this problem depends on identifying the right indicators and knowing which sources of faecal matter is the cause. In temperate regions, faecal coli form such as presumptive EC*, CP*, and *Enterococci* are used as indicator bacteria for the assessment of faecal pollution and possible water quality deterioration in various freshwater sources (APHA-AWWA-WPCF., 1981; Byamukama et al., 2005; Toranzos et al., 1997). For a long time, faecal coliforms have been used as a standard indicator of recent faecal pollution and are accepted under most conditions in temperate freshwaters. These microbial indicators have been used in tropical countries as well (Byamukama et al., 2005).The maximum allowable contaminant levels (MCL) established for many temperate areas have been accepted without question by tropical nations, even though source water quality in most tropical areas differs from that of temperate areas in three major ways: 1) physicochemical 2) biological and 3) socio economic factors(Hazen and Toranzos, 1990). Furthermore, there is a growing body of evidence that the underlying assumptions of the assays being used were not valid in tropical climate(Byamukama et al., 2005; Desmarais et al., 2002; Espinosa et al., 2009; Reischer et al., 2013; Sinigalliano et al., 2010).

Studies in tropical freshwater have shown that high proportions of faecal coliform-positive isolates may be of non-faecal origin (Scott et al., 2002). Furthermore, some studies have reported that presumptive EC can become a normal inhabitant of tropical waters, as reported for pristine environments in some tropical waters (Ahmed et al., 2008; Byappanahalli and Fujioka, 1998; Hazen, 1988; Rivera et al., 1988). This apparent unreliability of traditional FIB in tropical conditions should lead to the development of alternative pollution indicators.

In addition to the development of alternative, easy to perform, and affordable FIB, tracking the source of faecal pollution is important for proper management of drinking water sources at a watershed scale. To this end, various genetic faecal markers have been developed for temperate waters. There are numerous and diverse assay tools that target faecal pollution from different animal sources to discriminate between pollution sources, including BacR = ruminant-associated faecal pollution, PigIIBac = pig-associated, and Bach HUM or HF 183 Taqman = human-associated faecal pollution (Mayer et al., 2018; Reischer et al., 2013). Recently developed qPCR assays as well as standard faecal indicators have been widely applied in temperate regions.

However, their performance has not been adequately evaluated in a tropical country like Ethiopia where there is a high range of altitude that varies from -100 m.a.s.l.to 4533m.a.s.l.and varied water types and socioeconomic contexts. Moreover, neither standard FIB nor molecular markers (BacR) we re-evaluated at different altitudes (low, mid, and highlands) and water sources (surface and ground water) in tropical countries like Ethiopia. An urgent need exists to evaluate existing methods and parameters to come up with affordable, easy - to - perform and reliable techniques for faecal contamination monitoring in tropical regions.

Pollution risk mapping using a selected indicators is important for the overall water quality management in a lake basin. Benefits of pollution risk mapping include easy identification of areas that need intervention, management of potential sources of contamination, and the integration of drinking water protection activities with other environmental programs at the state, zone, district, and local levels. However, the pollution risk map for the Lake Tana basin both for gauged and un-gauged sub-basins remains unknown.

Therefore, this study aims to evaluate the performance of FIB including total coli forms (TC), presumptive *E.coli* (EC), intestinal enterococci (IEC), and presumptive *Clostridia perfringens* spores (CP), and to determine ruminant-associated faecal pollution using qPCR assay (BacR) at different water types located at different altitudes (1100-3835 m.a.s.l.in a highland tropical country. Based on the results of the FIB evaluation and enumeration, a map was created showing the degree of pollution of sub-basins of the Lake Tana basin.

## 5.2 Materials and Methods

### 5.2.1 The Study Sites in General

The study includes two sets of sites. The first set, which consists of 22 sites, was used to evaluate the performance of FIB and then to select a ‘’consensus’’ parameter (see section 2.6). The second set consists of 20 sub-basins (7 gauged and 13 un-gauged) where a map showing the degree of pollution of various waters in the Lake Tana basin was set up based on the results of the FIB enumeration.

### 5.2.2 Study sites for performance evaluation and microbial source tracking

The performance study of FIB was mainly conducted in the Lake Tana watershed and partly in the Chokie Mountains and Chifenchifit watersheds in the northern part of Ethiopia (see *Table 5-1*). A total of 22 sampling sites -12 groundwater and 10 surface water - were studied in rainy and post-rainy seasons from June to December 2017. The examined locations were from different water types, pollution categories, and altitudes. The sampling sites were selected based on altitude, sources of water, and presumptive pollution category as well as an assumption that there is a direct relationship between population and livestock density and level of pollution. Three altitude categories, low, mid, and highland were sampled. In this study, the lowland was operationally defined as an area having an altitude of ≤ 1100 m.a.s.l., mid land (>1100, and < 2300 m.a.s.l.), and highland (> 2500 m.a.s.l.). The surface waters were streams, and the ground waters were wells. The examined groundwater sampling sites were located in Bahir Dar City and peri-urban areas (see *Table 5-1*). The highly polluted groundwater sites were located in the most interior part of the city, close to the Kidane Mihret Church. This was a slum area and densely populated. Sites with intermediate pollution were located close to the Bata Church. Wells with intermediate pollution were found in an area where there was low population density, and the wells were better protected by a top metal cover and a circular metal case on the periphery to hold the surrounding soil. They are located around the Bata church. The last group of wells is located on the outskirts of Bahir Dar City around Robbit Bata and Wondata Michael. Among all well sites, these wells were the best protected and they were found in an area where the population density was low. These wells are protected with concrete structures used for drinking purposes.

The ten sampling sites from surface water were located at different altitudes that ranged from 1100 – 3850 m.a.s.l. Two sites of Chifenchifit stream represented the lowland, four sites of Infranz and Gudo-Bahir streams represented midland, and four sites of Awisha and Chokie Streams represented highland.

The Chifenchifit sampling sites were located in the Blue Nile George, close to the Hidase Bridge at an altitude of approx. 1100m.a.s.l. The up-stream site was a drinking water source. The portion of the watershed around the upstream site was not highly populated. It is an agricultural region heavily degraded by abundant run off from the uplands. The downstream study site was located closer to the Hidase Bridge. Human activities such as bathing, and car washing are often practiced at this site. The stream joins the Blue Nile River at Hidase Bridge about 1000 m downstream of the downstream site.

Infranz stream is located ca.15 km north-west of Bahir Dar. The major source of Infranz flow is the overland flow of the springs which are drinking water sources for Bahir Dar city. It is a perennial stream flowing down about 17 kms to Lake Tana. The stream passes through an extensive wetland, part of the Infranz Wetland which has an estimated area of 526 ha in 2011(Sewnet, 2015). Agriculture is the basis of the livelihoods of almost all the people in the watershed. It is mixed farming, where the rural people depend on both crop and livestock production.The livestock species kept in the rural areas of Lake Tana basin include cattle, sheep, goats, equines, and chickens (Alemayehu and Tassew, 2017).The population of the rural Lake Tana basin is about 32,000 people, with a density of approx. 124 people/km$^2$ (Csa, 2007). The upstream and downstream sites have a distance of 5 km.

Awhisha Stream is located close to Lake Zengena (2500 m.a.s.l.). It was located in an agricultural watershed and many people and livestock depend on it for drinking. The water point of the stream was not protected and one can expect an influence from humans and livestock. The upstream and downstream sites are separated by about 1km.

The Chokie Mountains watershed is one of the highest places in northern Ethiopia located some 50 km to the north of Debre Markos. The Chokie Mountains are known as the ''Water tower'' of the Blue Nile basin, the source of more than 23 major rivers and 273 small streams that flow to the Blue Nile. A mixed farming system is practiced in the form of traditional agriculture in the Chokie Mountains that contribute to severe soil erosion and natural resource degradation (Simane et al., 2013). The upstream Chokie Mountains site is located at the highest tip of the watershed (see Table 5-1) and the downstream study site is located about 1 km downstream of the upstream Chokie Mountains site.

*Table 5-1.* GPS readings of sampling sites used for the evaluation of FIB, altitude category, water type, presumptive pollution category and different uses of the source water. Abbreviations: S -stream, G-ground water wells, I- irrigation, C- cattle watering, D- drinking, and W-washing. Llu-lowland upstream, Lld-lowland downstream, Mld-midland downstream, Mlu-mid land upstream, St2u-stream 2 upstream, St2d-stream 2 downstream, Hlzu-highland Zengena upstream, Hlzd-highland Zengena downstream, Hlu-highland upstream, Hld-highland downstream, Gwm-groundwater middle city, Gwi-groundwater inner city, and Gwo-groundwater outer city.

| Sampling Site code | Name of a sampling site | Geographic latitude and longitude | | Altitude Category | Water Type | Presumptive Pollution category | Uses of the water |
|---|---|---|---|---|---|---|---|
| Llu | Chifenchifit upstream | 411064 | 1114600 | Lowland | S | Low | D,W,C |
| Lld | Chifenchifit downstream | 411234,9 | 1114309 | ‘’ | S | High | - |
| Mld | Infranz downstream | 313494 | 1285297 | Midland | S | ‘’ | D, I, C |
| Mlu | Infranz upstream | 311266 | 1282125 | ‘’ | S | Low | D, I, C |
| St2u | GudoBahir Upstream | 322657,8 | 1281582 | ‘’ | S | ‘’ | I, C |
| St2d | GudoBahir downstream | 324958,4 | 1280174 | ‘’ | S | High | - |
| Hlzu | Awhisha upstream | 277356 | 1206429 | Highland | S | Low | D, I, C, W |
| Hlzd | Awhisha down stream | 276941 | 1206598 | ‘’ | S | High | I, C, W |
| Hlu | Chokie upstream | 372573 | 1176467 | ‘’ | S | Low | D, C, W |
| Hld | Chokie downstream | 372573 | 1176467 | ‘’ | S | High | I, C, W |
| Gwm1 | Bata Church | 321614,4 | 1284519 | Midland | G | Medium | I, C |
| Gwm2 | ‘’ | 321507 | 1284311 | ‘’ | G | ‘’ | I, C |
| Gwm3 | ‘’ | 321516 | 321516 | ‘’ | G | ‘’ | I, C |
| Gwm4 | ‘’ | 321491,3 | 1284203 | ‘’ | G | ‘’ | I, C |
| Gwi1 | Kidanmihret Church | 322450,6 | 1281903 | ‘’ | G | High | C, W |
| Gwi2 | ‘’ | 322447,5 | 1281899 | ‘’ | G | ‘’ | C, W |
| Gwi3 | ‘’ | 324311,3 | 1280930 | ‘’ | G | ‘’ | C, W |
| Gwi4 | ‘’ | 324319,1 | 1282322 | ‘’ | G | ‘’ | C, W |
| Gwo1 | Robit area | 332052,1 | 1292130 | ‘’ | G | Low | D, I, C, W |
| Gwo2 | ‘’ | 332052,1 | 1292136 | ‘’ | W | ‘’ | D, I, C, W |
| Gwo3 | ‘’ | 323985,5 | 1268804 | ‘’ | W | ‘’ | D, I, C, W |
| Gwo4 | ‘’ | 324206,3 | 1268726 | ‘’ | W | ‘’ | D, I, C, W |

### 5.2.3 Study sites for pollution risk mapping

Pollution risk maps were developed for the 20 sub-basins of Lake Tana for the rainy (July-September) and post-rainy (October- December) seasons. Six of the twenty sub-basins are gauged river sub-basins in the Lake Tana basin. Lake Tana has a total drainage area of 16,500 $km^2$ (Fig. 5-1). More than six medium- to large-sized tributary rivers and more than 40 ephemeral streams drain into Lake Tana. Among the rivers, the Gilgel Abay, Dirma, Gumara, Gelda, Rib, and Megech Rivers contribute more than 90% of the inflow to the lake (Sirak, 2008). Lake Tana is the source of the Blue Nile and the Blue Nile is the only surface outflow from the lake. Lake Tana is relatively shallow, with a maximum depth of 14 m and a mean depth of 8 m. This tropical lake is non-stratifying with a mean elevation of 1,800 m.a.s.l. The amount of rainfall is at its maximum during July and August when it reaches 250–330 mm per month. The mean annual rainfall is nearly 1280 mm (Abebe and Minale, 2017). The rainy season (July–September) receives about two-thirds of the annual rainfall, while the dry season receives 2%; the pre-rainy season (April–June) receives 25% and the post-rainy season 8% of annual rainfall (Goshu et al., 2020). The average seasonal air temperature reaches its maximum of 21.1$^{o}$C in the pre-rainy season and its minimum (18.4$^{o}$C) in the rainy season, and it shows a large diurnal but small seasonal change.

Lake Tana is the largest lake in Ethiopia, accounting for 50% of the fresh-water resources of the country (Vijverberg et al, 2009). It has a surface area of 3,111 $km^2$, 28.4 $km^3$ in volume, and a maximum length of 90 km, and a width of 65 km. The population of the basin was projected to be 4.5 million in 2015 (CSA, 2007), with a population density of 228 persons per $km^2$ in 2007 (Anteneh, 2017), and 70% of the basin is agricultural land (Abebe and Minale, 2017).

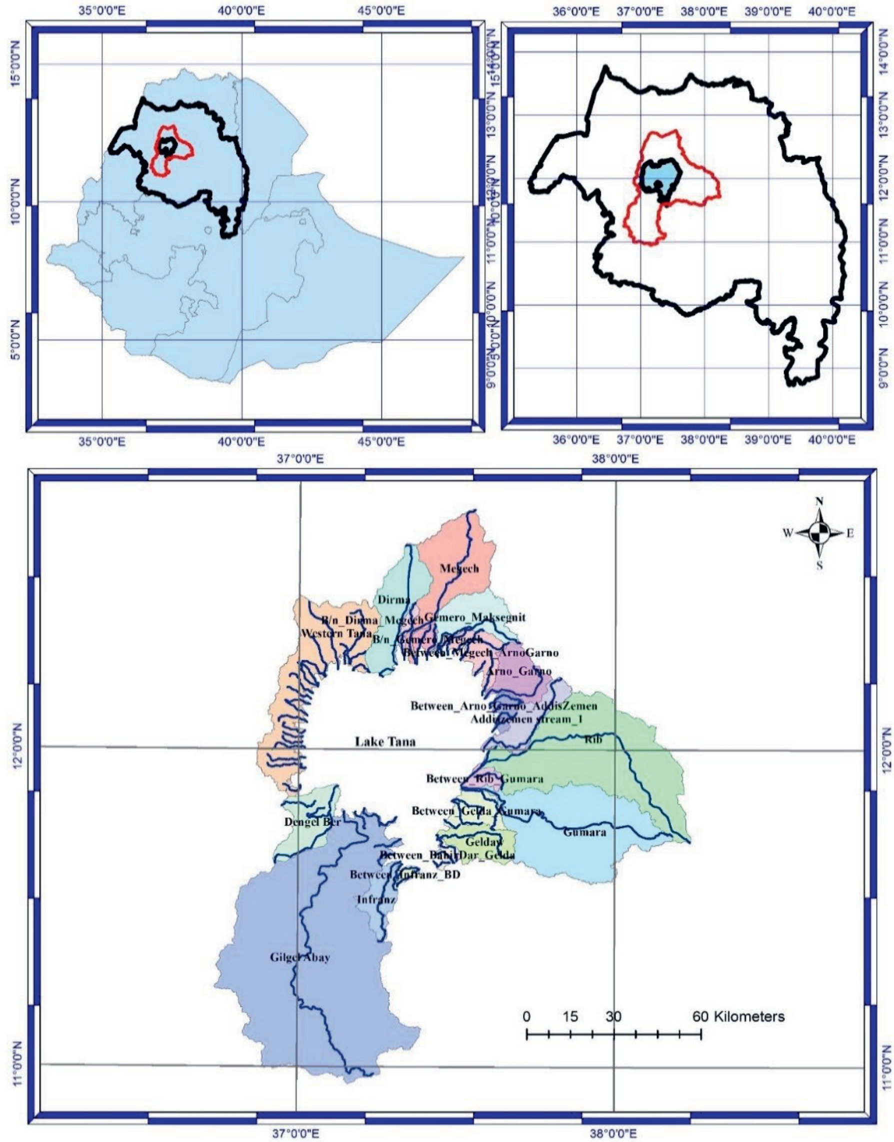



Fig. 5-1 – a) Ethiopia and Amhara Region – b) Amhara Region, Lake Tana Basin, and Lake Tana – c) the drainage area of the Lake Tana Basin. Map of the Lake Tana Basin shows sub-basins draining into the lake. Source: (Goshu et al., 2020).

### 5.2.4 Sampling

A total of 241water samples (192 for performance and 49 for mapping) were collected from wells, streams, and rivers. Surface water samples were collected using sterile Kimax Kimble glass bottles with butyl rubber stoppers from a 30 cm-deep flowing section of a stream. The groundwater samples from wells equipped with a pump were collected after 2 minutes of water flushing. In the remaining wells provided with ropes and buckets, water samples were collected directly from the bucket after washing. The sample bottles were immediately kept in the dark in a cooling box and transported to the Food and Chemical Engineering Laboratory of Bahir Dar University and samples were analysed within 6 hours of taking the first sample. Water samples for physicochemical analysis were collected in 500 ml polyethylene bottles (Thomas Scientific), kept in an icebox, and transported to the laboratory for immediate analysis within six hours (APHA-AWWA-WPCF., 1981).

### 5.2.5 Physicochemical Characteristics

For basic characterization of the investigated habitats, physicochemical variables were measured: dissolved oxygen, pH, water temperature, electrical conductivity, and nitrogen. In-situ measurements of electrical conductivity (Con), pH, dissolved oxygen (DO), temperature (T), total dissolved solids (TDS) were done with a YSI Pro Plus multi parameter meter, Ohio, USA. Ammonia, nitrite, and nitrate were determined photo metrically at the water quality laboratory of the School of Civil and Water Resource Engineering, Bahir Dar University, using Palin test Transmittance-display photometer 8000, North East, UK, following the manufacturer instruction.

### 5.2.6 Determination of Standard Faecal Indicator Bacteria

#### Presumptive detection, enumeration, and confirmatory tests

Presumptive simultaneous detection of TC and EC was done by membrane filtration technique using Chromo Cult Coli form Agar (CCA). *CP* spores were detected by solid TSC agar, IEC were detected by Slanetz and Bartley agar media.

CCA media were amended by the addition of Cefsulodin (5mg/l; Sigma, Vienna, Austria), and CCP by Clostridium perfringens supplement (0.4g/ l).

The pink and blue colonies on CCA agar were classified as TC and EC respectively. Black colonies growing on TSC plates were classified as CP. Pink to red colonies was classified as IEC. All agar media was from Merck.

About 20 well-isolated presumptive CP colonies from TSC were randomly selected and sub-cultured in the selective media and characterized by the rapid ID 32 A biochemical test (Bio Merieux Sa, Francais-I, France) according to the manufacturers' instructions.

#### Sampling and DNA extraction of water samples

Water samples were filtered over 0.2 µm polycarbonate filters (Millipore, Bedford, MA) and stored on dry ice at -80°C during transportation to TU Vienna for further processing.

DNA extraction was performed using bead-beating and phenol/chloroform as described previously in (Griffiths et al., 2000; Mayer et al., 2018; Reischer et al., 2008). In brief, cell lysis was achieved by the addition of CTAB buffer and glass beads in a Fast Prep 24 bench-top homogenizer for cell lysis (MP Biomedical Inc., Irvine, CA) at a speed setting of 6 $m.s^{-1}$ for 30s. Polycarbonate filters were completely dissolved at this step and the DNA was subsequently purified by washing procedures. Precipitation of the DNA was achieved by the addition of isopropanol. The extracted

DNA was eluted in 10 m MTRIS buffer (pH 8.0) and stored at -80°C until further analysis. Before the qPCR assays, the DNA concentration of the sample DNA extracts was determined using the Quanti Fluor® dsDNA Kit (Promega, USA) according to the manufacturer's instructions. Fluorescence readings were taken on a Glomax Micro-plate Reader (Promega, USA).

### 5.2.7 MST Genetic Marker Detection

In addition to the applied host-associated faecal genetic 16S-rRNA-gene marker targeting ruminant-associated faecal pollution (BacR; Reischer et al. 2006), a general *Bacteroidetes* marker, AllBac (Layton et al., 2006), was used as quality control to assess the ability to amplify DNA in the samples. The assay was applied as a duplex qPCR including an internal amplification control (IAC, non-competitive) using the *ntb2* gene from *Tabaco nicotianum* L.(Anderson et al., 2011). All samples were measured in duplicate in at least two 4-fold DNA dilution steps and the results were compared. Samples with two matching All Bac concentrations (i.e., the ratio [concentration 1:16·4]/ [concentration1:4] was between 0.5 and 2) in the 1:4 and 1:16 dilutions were judged free of PCR inhibiting substances in the 1:4 dilution. In addition to the dilution method to identify possible PCR inhibition, results from the IAC were evaluated. Samples were judged of PCR inhibition if the threshold cycle (Ct value) of the IAC assay in a sample was shifted towards higher Ct values by more than one cycle in comparison to the negative control. Applying these rules, in neither of the samples, inhibition of the qPCR reaction could be detected. Controls furthermore included no-template controls as well as filtration and DNA extraction blanks.

All sample DNAs in the qPCR assay to detect ruminant-associated faecal pollution (BacR (Reischer et al., 2006) were measured in duplicate. Quality assessment of qPCR data was done as previously described (Mayer et al., 2018; Reischer et al., 2006; Reischer et al., 2011). In brief, the reaction efficiency of all qPCR runs ranged from 95-105%. All negative controls and no-template controls were consistently negative (i.e., fluorescence never exceeded the threshold) and samples with replicate standard deviations of a Ct-value >1 in the 4-fold DNA extract dilutions we re-judged not quantifiable and were not considered for further analysis. qPCR standard dilutions ranging from $10^0$ to $10^6$ targets per reaction were used in a linear regression model for calculation of the qPCR calibration curve. Results are reported as marker equivalents per filtered water volume (ME $vol^{-1}$) as previously described by Reischer *et al.* (2006). Filtration volume, the use of 2.5-µl of undiluted DNA extract in the qPCR and the minimal theoretically detectable marker concentration per reaction (1 copy) defines the threshold of detection (Reischer *et al.* 2006). The threshold of detection of the herein presented data set was 107 copies.

All qPCR assays were run on a Rotor-Gene Q thermocycler (Qiagen Inc.) in a total reaction volume of 15µl with 2.5 µl sample DNA dilution (1:4). The respective reaction mixtures were composed of 7.5 µl Rotor-Gene Multiplex PCR master mix (Qiagen Inc.), 2.5 µl sample DNA dilution, and 400 ng $\mu l^{-1}$ bovine serum albumin, while the originally published primer and probe concentrations were maintained. Cycling parameters for All Bac including the IAC were 5 min at 95 °C for denaturation and 45 cycles of 30 s at 95 °C followed by 45 s at 60 °C and cycling parameters for BacR were 5 min at 95 °C for denaturation and 45 cycles of 15 s at 95 °C followed by 60 s at 60 °C.

### 5.2.8 Evaluation of FIB and Pollution Risk Mapping

Regarding pollution risk maps, gauged and un-gauged sub-basins of the Lake Tana basin were the areas of interest. First, we evaluated different FIB at different altitudes and water types, and then selected the best FIB among the tested ones, which we termed the ''consensus'' picture. Consensus FIB were selected based on their performance which is defined as the discrimination efficacy of the tested FIB at different altitudes and water types of different levels of presumptive pollution.

Better performing FIB has a higher discrimination ability. The discrimination ability is measured as a ratio of the number of pairs of compared sites of presumptively high and low pollution

$$Bac_{conung.j.S} = \frac{TLUD_{ung.j.s}}{TLU_{Dg.j}} \cdot \frac{PNLD_{ung.j}}{PNLD_{g.j}} \cdot \frac{Fews_{ung.j.S}}{Fews_{g.j.S}} \cdot \frac{Feriv_{ung.j.S}}{Feriv_{g.j.S}} \cdot Bac_{cong.j.S} \qquad \text{(eq.1)}$$

categories that have significant statistical test outcomes ( $p \leq 0{,}05$) to the total number of pairs of compared sites. Those FIB which have better discrimination efficacy were selected as consensus FIB. Then, we established pollution classes. Pollution classes were established between the highest achievable level in faecal material and raw sewage and the natural background level of faecal pollution of surface water and groundwater habitats (expected levels < 1 CFU of the consensus indicator per 100 ml;(Kavka et al., 2006). For the classification of the investigated sites into respective pollution classes, ninety percentile log values of consensus FIB was used to develop the pollution map based on a statistical conservative estimate (Kavka et al., 2006). To estimate the ninety percentile bacteria concentration for the un-gauged sub-basins, we used the approach of Yarahmadi (2003and Hofstra and Vermeulen (2016but modified as the following formula:

Where,

$Bac_{con\ ung.j.S}$is the bacterial concentration in un-gauged sub-basin j and season S (CFU/100 ml $S^{-1}$)

; $Bac_{con\ g.j.S}$isthe bacterial concentration in gauged sub-basin j and season S (CFU/100 ml $S^{-1}$); $TLUD_{ung.j.s}$isthe tropical livestock unit density of un-gauged sub-basin j and season (no.km$^{-2}$S$^{-1}$); $TLUD_{g.j.s}$isthe tropical livestock unit density of gauged sub-basin j and season S (no.km$^{-2}$S$^{-1}$); $PNLD_{ung.j}$is the population density with no toilet facility of un-gauged sub-basin j and season S (no.km$^{-2}$S$^{-1}$); $PNLD_{g.j}$isthe population density with no toilet facility of gauged sub-basin j and season S (no.km$^{-2}$S$^{-1}$);Fews$_{ung.j.S}$isthe N watershed export fraction of un-gauged sub-basin j and season S (S$^{-1}$); Fews$_{g.j.S}$isthe N watershed export fraction of gauged sub-basin j and season S (S$^{-1}$); Feriv$_{ung.j.S}$isthe DIN river export fraction of un-gauged sub-basin j and season S (S$^{-1}$); Feriv$_{g.j.S}$ is the DIN river export fraction of gauged sub-basin j and season S (S$^{-1}$)

### 5.2.9 Data Analysis

We compared the mean differences of bacterial counts of each indicator bacteria for a pair of sites in the surface water and site classes in groundwater sampling sites. For statistical analysis, the software package IBM SPSS Statistics 25was used. Non-parametric tests (i.e., Kruskal-Wallis H test) were carried out. We also run a non-parametric Monte Carlo test, using 10,000 re-samplings to evaluate the sensitivity of the indicators to differentiate between presumptive high- and low-influence sites. A general significance level of ($p \leq 0, 05$) was set for all the tests. To calculate a non-parametric-based but standardized indication of variability (as in analogy to the parametric coefficient of variation), we calculated the inter quartile range and divided it by the median. The bacteriological concentration ratios of medians (BCRM)of various indicators were calculated by dividing the median concentration of an indicator bacteria of a high influence site by the median concentration of an indicator bacteria of a low influence site to further analyse the detected differences in discrimination ability of the indicators.

## 5.3 Results

### 5.3.1 Physical and Chemical Characteristics of the Sites

The results of the physicochemical characteristics of the examined water habitats used to study the performance of the FIB are presented in *Table 5-2*. There is a significant difference within the examined surface as well as groundwater habitats in terms of most of the physicochemical parameters presented in *Table 5-2* except nitrite for surface water and ammonia for groundwater habitats($p<0,05$, n = 63 for surface water and n= 60 for groundwater sites, Kruskal-Wallis H test). This result confirmed the presence of a range of differing habitats and pollution patterns in the investigated water systems that build the basic framework for having different categories of faecal pollution and testing the performance of indicators.

*Table 5-2*. Physicochemical Characteristics (Mean value ± standard deviation) of the examined water habitats; Abbreviations: Ll- Surface water from lowland, Ml-Surface water from midland, Hl- Surface water from high land, Gwi-Groundwater from the inner part of the city, Gwm-Groundwater from the middle part of the city, Gwo- Groundwater from the outer part of the city, T- Temperature, DO - Dissolved oxygen, Cond. - Conductivity, $NO_3$ - Nitrate, $NO_2$ - Nitrite, $NH_3$ - Ammonia, TDS - Total dissolved solids

| | Surface water | | | | | | | Ground water | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Ll | | Ml | | Hl | | | Gwi | | Gwm | | Gwo | | |
| Parameter | Mean | SD | Mean | SD | Mean | SD | p value | Mean | SD | Mean | SD | Mean | SD | P value |
| T (°C) | 20,6 | 2,2 | 19,6 | 2,3 | 14 | 5 | p<0,05 | 21,7 | 1,5 | 22,2 | 0,9 | 23,3 | 0,4 | p<0,05 |
| DO (mg/ l) | 8,3 | 2,2 | 4,6 | 2,4 | 8,5 | 2,7 | p<0,05 | 1,6 | 0,5 | 2,5 | 0,7 | 3,1 | 1 | p<0,05 |
| Cond.( µs/cm) | 913 | 312 | 387 | 201 | 131 | 74 | p<0,05 | 999 | 259 | 186 | 99 | 424 | 377 | p<0,05 |
| pH (-) | 8,4 | 1 | 7,7 | 1 | 8,5 | 1 | p<0,05 | 6,8 | 0,3 | 6,3 | 0,2 | 6,7 | 0,4 | p<0,05 |
| $NO_3$ (mg / l) | 2,52 | 2,41 | 0,88 | 0,86 | 0,48 | 0,21 | p<0,05 | 29,5 | 21,5 | 3 | 2,4 | 2,7 | 2,6 | p<0,05 |
| $NO_2$ (mg / l) | 0,02 | 0,02 | 0,07 | 0,23 | 0 | 0,01 | p>0,05 | 0,4 | 0,6 | 0,1 | 0,1 | 0,1 | 0,2 | p<0,05 |
| $NH_3$ (mg / l) | 0,06 | 0,04 | 4,72 | 5,65 | 0,08 | 0,07 | p<0,05 | 3,7 | 5,1 | 2,4 | 4,2 | 4,4 | 5,6 | p>0,05 |
| TDS (g / l) | 0,65 | 0,22 | 0,23 | 0,1 | 0,1 | 0,05 | p<0,05 | 0,5 | 0,1 | 0,1 | 0,1 | 0,2 | 0,1 | p<0,05 |

### 5.3.2 General Level of Occurrence and Abundance of FIB

The general level of occurrence of indicators used for the detection of faecal pollution of the examined surface and groundwater habitats is shown in Fig.5-2. Total Coli forms (TC), presumptive*, and CP* were detected in 97, 91, and 93% of all sampling sites throughout the sampling period respectively. In surface water study sites, TC, EC, and CP were detected in 100, 96, and 95% of the sampling sites. In groundwater study sites, among the tested indicators (TC, CP, and EC), TC was detected in 95%, CP in 90%, and EC was detected in 85% of all the sampling sites. Among the detected multiple microbial indicators, TC resulted in the highest value with a concentration of log 3.39 CFU per 100 ml (n = 123; Fig.2A, B, C) in a pooled data. It had also the highest concentration on all surface as well as groundwater sampling sites (Fig 5-2 A,B,C). The maximum concentrations of the other indicators were in the order of presumptive (log 3.21 CFU per 100 ml), IEC (log 3.28 CFU per 100 ml), and *CP* (log 2.38 CFU per 100 ml). In groundwater study sites, among the tested indicators (TC, CP, and EC), TC was detected in 95%, CP in 90%, and EC was detected in 85% of all the sampling sites. Among the detected multiple microbial indicators, TC resulted in the highest value with a concentration of log 3.39 CFU per 100 ml (n = 123; Fig. 5-2 A,B,C) in a pooled data. It had also the highest concentration on all surface and groundwater sampling sites (Fig. 5-2 A,B,C). The maximum concentrations of the other indicators were in the order of presumptive (log 3.21 CFU per 100 ml), IEC (log 3.28CFU per 100 ml), and *CP* (log 2.38 CFU per 100 ml).

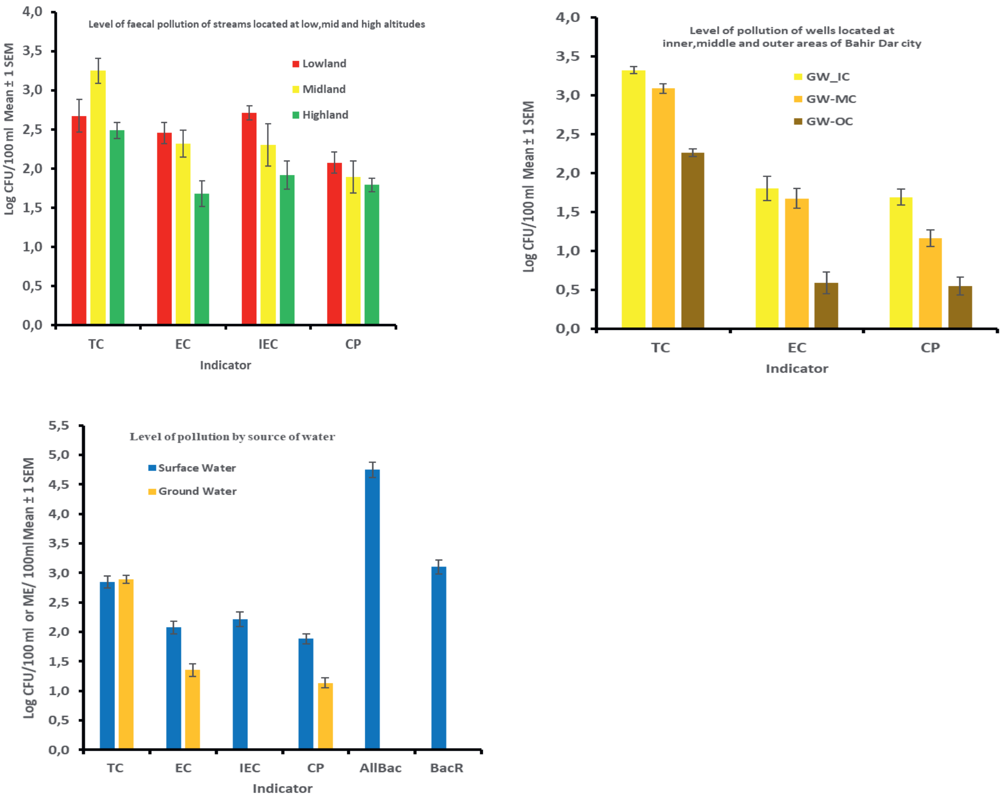



Fig. 5-2 – The general level of occurrence of faecal indicators - a) in three altitude categories of Lake Tana Basin, Ethiopia; highland (≥ 2500m.a.s.l.); midland ( >1100 and < 2500m.a.s.l.) and lowland (≤ 1100 m.a.s.l.; left panel) - b) in three presumptive pollution categories of the groundwater sampling sites (right panel). The FIB included in the study were TC, EC, IEC, CP, and their concentrations were in colony-forming units (CFU) per 100 ml expressed as Mean. The error bars indicate ± 1 Standard Error of the mean (n= 14 -20; 14 per altitude category of the surface water and 20 per pollution category of the groundwater). GW-IC-groundwater inner city, GW-MC-groundwater middle city, and GW-OC-groundwater outer city.

### 5.3.3 Performance of FIB

The performance of FIB (discrimination efficacy) was evaluated based on eight corresponding low to very high influence sites and site classes for comparisons (*Table 5-3*A and B). Based on these comparisons, the various FIB showed varying levels of discrimination efficacy and sensitivity. Presumptive and CP showed equal performance and ability to significantly discriminate between paired sites [63% (5/8)], TC 50%, and IEC 20% (*Table 5-3*A). In surface water study sites, EC and CP were able to discriminate between40% of the sites and TC and IEC significantly discriminate between20% of the sites. In groundwater systems, EC, CP, and TC were able to significantly discriminate between100% of the sites (3/3).

The bacterial concentration ratio of the medians (BCRM), which is the ratio of the median of higher pollution category to the median of lower pollution category, indicated the sensitivity of faecal indicators. The higher value of BCRM meant higher sensitivity and the BCRM value of EC ranged from 1.55 to 18.83, CP 1.28-14.96, and TC 0.66 -11.88 in pooled samples. In surface water samples, the BCRM of EC ranged from 1.55 -3.69, CP 1.28-3, and IEC 1.12- 47. In groundwater systems, the BCRM values were relatively higher compared to surface water systems and the BCRM of EC ranged from 2.33-18.83, CP 3.34 -14.96, and TC 1.78-11.88 (*Table 5-3*B).

*Table 5-3*. Pair wise habitat comparison based on bacteriological parameters of presumptively high and low influence sites for surface water systems (streams; a) and groundwater systems (wells; b). The bacteriological parameters were TC, EC, and CP in log CFU/100 ml. See Fig. 5-1 for the location of the examined habitats and *Table 5-1* for the abbreviations of sampling sites.

| **a)** | **Statistical test outcome of the examined surface water systems** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Comparison Type | Compared pair | **EC** | | **CP** | | **IEC** | | **TC** | |
| | | P | BCRM | P | BCRM | P | BCRM | P | BCRM |
| Site by site | Hlu-Hld | ,031 | 4 | ,041 | 3 | ,042 | 4,25 | ,250 | 3,33 |
| | Hlzu-Hlzd | ,569 | 1,8 | ,470 | 3 | ,982 | 1,12 | ,361 | 0,4 |
| | Mlu-Mld | ,179 | 3,69 | ,476 | 1,8 | ,506 | 47 | ,482 | 0,66 |
| | Llu-Lld | ,306 | 1,58 | ,787 | 1,4 | ,265 | 1,76 | ,944 | 0,66 |
| | St2u-St2d | ,047 | 1,55 | ,049 | 1,28 | ,113 | - | ,029 | 1,32 |

| b) | **Statistical test outcome of the examined groundwater systems** | | | | | | |
|---|---|---|---|---|---|---|---|
| | | **EC** | | **CP** | | **TC** | |
| Comparison type | Compared pair | P | BCRM | P | BCRM | P | BCRM |
| Site class by site class | Gwi-Gwm | ,05 | 2,33 | ,005 | 4,47 | ,001 | 1,78 |
| | Gwi-Gwo | ,001 | 18,83 | ,001 | 14,96 | ,002 | 11,88 |
| | Gwm-Gwo | ,000 | 14,12 | ,000 | 3,34 | ,000 | 6,68 |

The result of the correlation analysis of faecal indicators is presented in *Table 5-4*. The result revealed that there is generally a significant correlation p<=0, 01 among most of the indicators (*Table 5-4*). Total coli forms strongly correlated with EC, CP, and IEC (r = 0,356 – 0,528, n =123) in pooled samples. EC correlated significantly with IEC and CP in pooled samples (r = 0,601 - 0,647, n= 123; see Table 4). CP revealed a significant correlation with TC and EC only. The correlation among the indicators is more pronounced at highland than mid and lowlands located study sites (r = 0,624 – 0,848, n =28).TC showed a very strong correlation with EC and CP only at highland sites (r = 0,6 –0,624). IEC showed a significant correlation with EC only in pooled samples (r=0,848). The correlation coefficient among indicators in mid and lowland study sites is generally very low except EC with TC and IEC correlations. All the evaluated faecal indicators showed a general trend towards presumptive high-influenced sites that were typically more polluted than presumptive low influenced sites.

*Table 5-4*. Pearson correlation coefficients (r) between the various microbial parameters (TC, EC, IEC and CP) of samples collected from surface and groundwater sources distributed over three altitude categories: low-, mid- and highlands. Asterisks indicate * significant (0, 05) and ** highly significant (0, 01)

| Pooled | TC | EC | IEC | CP |
|---|---|---|---|---|
| TC | | ,528** | ,356* | ,413** |
| EC | ,528** | | ,647** | ,601** |
| IEC | ,356* | ,647** | | ,167 |
| CP | ,413** | ,601** | ,167 | |
| Highland | | | | |
| TC | | ,624** | ,244 | ,600** |
| EC | ,624** | | ,848** | ,660** |
| IEC | ,244 | ,848** | | -,277 |
| CP | ,600** | ,660** | -,277 | |
| Midland | | | | |
| TC | | ,352 | ,195 | -,158 |
| EC | ,352 | | ,227 | -,124 |
| IEC | ,195 | ,227 | | ,266 |
| CP | -,158 | -,124 | ,266 | |
| Low land | | | | |
| TC | | ,755* | ,238 | ,450 |
| EC | ,755* | | ,594 | ,384 |
| IEC | ,238 | ,594 | | -,133 |
| CP | ,450 | ,384 | -,133 | |

The results of the coefficient of variation expressed as non-parametric based variation (NBV) values are summarized and presented in *Table 5-5*A and B. There is no significant difference in NBV values among the values of all faecal indicators in surface water sampling sites (p>0, 05, Kruskal-Wallis H test, n= 7with 8 sampling sites). In groundwater systems, the NBV values of TC were significantly different from EC and CP NBV values. However, there remained no significant difference between NBV values of EC and CP (p>0, 05, Mann-Whitney U test, n= 7 with 12 sampling sites).

*Table 5-5*. Bacteriological characteristics of faecal indicators for surface water (a) and groundwater (b) The concentrations of faecal indicator bacteria for TC, EC, and CP (log CFU/100 ml). NBV-Nonparametric based variation; M-Median; R-Range. See *Table 5-1* for the abbreviations of the sites.

| a) | TC | | | EC | | | IEC | | | CP | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sampling sites | M | R | NBV | M | R | NBV | M | R | NBV | M | R | NBV |
| Hld | 2,7 | 1,7-2,9 | 0,2 | 1,8 | ND -2,5 | 0,6 | 2,2 | ND - 2,5 | 0,3 | 1,8 | 1,6 -2,4 | 0,4 |
| Hlu | 2,2 | 1,6 -2,6 | 0,3 | 1,3 | ND -2,5 | 0,5 | 1,6 | ND - 2,3 | 1,0 | 1,3 | 1- 2,4 | 0,6 |
| Hlzd | 2,6 | 2,6 - 3,4 | 0,3 | 2,0 | 1,5-2,6 | 0,3 | 2,3 | 1,0 - 3,3 | 0,8 | 2,2 | 1,3- 2,3 | 0,4 |
| Hlzu | 2,8 | 2,7 - 2,9 | - | 1,7 | 1,5 - 3,3 | 0,5 | 2,2 | 1,0 - 3,3 | 0,8 | 1,7 | 1,3- 2,4 | 0,4 |
| Lld | 2,9 | 2,0 - 3,3 | 0,7 | 2,6 | 2,1 - 3 | 0,4 | 2,9 | 2,6 - 3,1 | 0,2 | 2,1 | 1,3- 2,7 | 0,5 |
| Llu | 2,8 | 1,6 - 3,2 | 0,5 | 2,4 | 1,6 - 3 | 0,3 | 2,7 | 2 - 3,1 | 0,4 | 2,2 | 1,6- 2,7 | 0,4 |
| Mld | 2,7 | 2,6 - 2,9 | 0,1 | 2,4 | 2,1 - 2,5 | 0,2 | 2,7 | 2,2 - 3 | 0,2 | 2,1 | 1,3- 2,8 | 0,4 |
| Mlu | 2,9 | 2,4 - 3,1 | 0,2 | 1,8 | 1,3 - 2,6 | 0,5 | 1,9 | ND - 3,1 | 1,1 | 1,0 | ND - 2,8 | 1,4 |

| b) | TC | | | EC | | | CP | | |
|---|---|---|---|---|---|---|---|---|---|
| Sampling sites | M | R | NBV | M | R | NBV | M | R | NBV |
| Gwm | 3,1 | 2,6 - 3,6 | 0,1 | 1,8 | ND -2,4 | 0,3 | 1,1 | 0,5 -2,2 | 0,6 |
| Gwi | 3,3 | 3 - 3,7 | 0,02 | 1,75 | ND -2,8 | 0,6 | 1,7 | 0,8 -2,5 | 0,38 |
| Gwo | 2,2 | 2-2,7 | 0,19 | 0,5 | ND -2,1 | 2,35 | 0,45 | ND - 1,6 | 1,78 |

### 5.3.4 Microbial Source Tracking (MST)

Markers associated with ruminants (BacR) were detected at all sites except MLU at least once during the study period from June to December 2017. The percentage occurrence of BacR ranged from ND(Mlu) to 86% (Hld). The overall percentage occurrence of BacR in pooled sites is 37%. The BacR concentration ranges from NDto 3.95 log ME/100 ml. The greatest concentration of the BacR marker was detected at Mld (log 6.95 ME/100 ml).

*Table 5-6*. Distribution and detection frequency of microbial source tracking markers by PCR and range of concentrations by q PCR for host-associated markers. Comparisons of FIB abundance (%) and concentration (log CFU/100 ml) for different streams and the MST study (BacR)are also presented. See *Table 5-1* for the abbreviations of the sites. See Fig. 5-2 for the abbreviations of FIB. Ruminant Associated Bacteroides - BacR nd-non detected

| **Number of positive samples (%)** | | | | | | **BacR** | | **FIB** | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **qPCR assay** | **FIB** | | | | **log ME/100 ml** | | **(Log CFU/100 ml)** | | |
| Site (Sampling events) | BacR | EC | *CP* | *TC* | *IEC* | BacR | EC | *CP* | *TC* | *IEC* |
| Llu (7) | 2 (29%) | 7(100%) | 6(86%) | 7(100%) | 7(100%) | ND-2,97 | 1,61-3,02 | 1,61-2,67 | 1,61-3,18 | 2-3,09 |
| Lld (7) | 3(43%) | 6(86%) | 7(100%) | 7(100%) | 7(100%) | ND-3,37 | 2,15-3,02 | 1,32-2,7 | 1,96-3,3 | 2,56-3,08 |
| Mld(6) | 3(50%) | 5(83%) | 6(100%) | 5(83%) | 6(100%) | ND-3,95 | 2,08-2,53 | 1,32-2,79 | 2,6-2,89 | 2,23-3 |
| Mlu (6) | ND | 6(100%) | 5(83%) | 5(83%) | 6(100%) | ND | 1,32-2,58 | ND-2,76 | 2,4-3,06 | ND-3,11 |
| Hlzu (7) | 2(29%) | 6(86%) | 7(100%) | 6(86%) | 7(100%) | ND-2,93 | 1,5-2,3 | 1,3-2,4 | 2,7-2,9 | 1-3,28 |
| Hlzd (7) | 3(43%) | 6(86%) | 7(100%) | 6(86%) | 7(100%) | ND-3,52 | 1,5-2,6 | 1,3-2,3 | 2,6-3,4 | 1-3,3 |
| Hlu (7) | 1(15%) | 6(86%) | 7(100%) | 6(86%) | 7(100%) | ND-3,48 | ND-2,55 | 1,04-2,4 | 1,61-2,57 | ND-2,3 |
| Hld (7) | 6(86%) | 7(100%) | 7(100%) | 7(100%) | 7(100%) | ND -3,82 | ND-2,52 | 1,61-2,45 | 1,71-2,9 | ND -2,49 |

### 5.3.5 Pollution Risk Map

**Sites used for performance study**

About one-third of the sites (Hlu, Hlzu, Gwm-2, Gwm-3, Gwm-1, Gwi-2 Gwi-4 Gwo-1 Gwo-2 Gwo-3 Gwo-4 St-1 St-2 ) were classified into the same pollution class by the consensus FIB, presumptive EC, and CP approaches (*Table 5-7*,*Table 5-8*, and Fig. 5-3). Sampling sites having equal or greater than log 5.5 EC and log 3.7 CP concentrations per 100 ml would have been classified as *excessively polluted* but none of the sites qualified for this category. Log 3.5 - log 4.4 and log 4.5 - log 5.4 concentrations per 100 ml were classified as *very highly* and *highly polluted*, respectively, based on EC. None of our study sites fell into this pollution level except the Gudo-Bahir stream sites. Sites having log 90% percentile presumptive EC values in the range of log 2.5 and log 3.4 CFU per 100 ml were deemed *moderately polluted* and one-third of the study sites were in this pollution category. Forty percent of the sites had fallen under a *low level* and only two sites, both ground water, were under *very low* levels of pollution (Tables 5-7, 5-8, and Fig. 5-3).

*Table 5-7*.Dual classification scheme of faecal pollution levels based on *E.coli* (Kavka et al., 2006), and *Clostridium perfringens* spore concentrations(this work).

| Pollution class | EC (Log CFU /100 ml ) | *CP* (Log CFU/100 ml) | | Pollution class |
|---|---|---|---|---|
| Excessive | > 5.5 | > 3.7 | 1 | |
| Very high | 4.5 – 5.4 | 3.03 – 3.69 | 2 | |
| High | 3.5 – 4.4 | 2.36 – 3.02 | 3 | |
| Moderate | 2.5 – 3.4 | 1.69 – 2.35 | 4 | |
| Low | 1.5 – 2.4 | 1.02 – 1.68 | 5 | |
| Very low | up to 1.4 | up to 1.01 | 6 | |

*Table 5-8*. Ninety percentile of log EC and log CP for all sampling sites (n = 4 -7 per sampling sites), classified into the respective level of pollution based on the pollution classes of *Table 5-7*. (See *Table 5-1* for the abbreviations of the sites.)

| **Site** | **EC Log CFU/ 100 ml** | **CP Log CFU/100 ml** | **Pollution class EC** | **Pollution class CP** |
|---|---|---|---|---|
| Hld | 2.3 | 2.4 | 5 | 3 |
| Hlu | 2.1 | 2.3 | 5 | 4 |
| Hlzd | 2.3 | 2.3 | 5 | 4 |
| Hlzu | 2.7 | 2.2 | 4 | 4 |
| Lld | 3.0 | 2.6 | 4 | 3 |
| Llu | 2.7 | 2.6 | 4 | 3 |
| Mld | 2.5 | 2.5 | 4 | 3 |
| Mlu | 2.5 | 2.7 | 4 | 3 |
| Gwm-1 | 1.4 | 1.1 | 6 | 5 |
| Gwm-2 | 2.1 | 1.1 | 5 | 5 |
| Gwm-3 | 2.2 | 1.4 | 5 | 5 |
| Gwm-4 | 2.1 | 2.2 | 5 | 4 |
| Gwm-1 | 2.6 | 2.3 | 4 | 4 |
| Gwi-2 | 2.1 | 1.5 | 5 | 5 |
| Gwi-3 | 2.3 | 2.3 | 5 | 4 |
| Gwi-4 | 2.6 | 2.2 | 4 | 4 |
| Gwo-1 | 0.0 | 0.6 | 6 | 6 |
| Gwo-2 | 0.5 | 0.3 | 6 | 6 |
| Gwo-3 | 1.9 | 1.6 | 5 | 5 |
| Gwo-4 | 0.5 | 0.8 | 6 | 6 |
| St-1 | 3.7 | 2.4 | 3 | 3 |
| St-2 | 3.5 | 2.7 | 3 | 3 |

## General level of Pollution and risk map

*E.coli* was detected in all sampled rivers and the average EC concentration in pooled sites ranged from log 2.34 to log 3.14CFU/100ml, the minimum average EC concentration in Gumara and the maximum in Rib, respectively. In the rainy season, the average EC concentration ranged from log 2.46 CFU/100 ml to log 3.14CFU/100 ml, the minimum noted in Garno and the maximum noted in Dirma. In the post-rainy season, the average EC concentration ranged from log 2.34 CFU/100 ml to log 2.57CFU/100 ml, the minimum noted in Gumara, and the maximum noted in Rib River (see Fig. 5-3).

*CP* was detected in all sampled rivers. The average *CP* concentration in pooled sites ranged from log 1.76 to log 2.83CFU/100 ml, the minimum average *CP* concentration in Gilgel Abay and the maximum noted in Gumara. In the rainy season, the average *CP* concentration ranged from log 1.76 CFU/100 ml to log 2.83 CFU/100 ml, the minimum noted in G/Abay and the maximum noted in Gumara. In the post rainy season, the average *CP* concentration ranged from log 2.34 CFU/100 ml to log 2.57CFU/100 ml, the minimum noted in Gumara, and the maximum noted in Rib River (see Fig. 5-3).

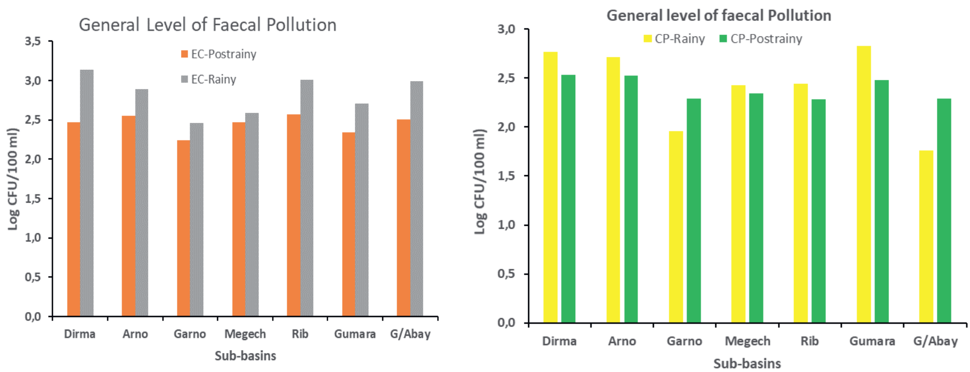



Fig. 5-3 – The general level of occurrence of FIB for average EC concentration (a; log CFU/100 ml) and average *CP* (b; log CFU/100 ml) in major tributary rivers of the Lake Tana basin (see Fig. 5-1 for the location of the rivers) in rainy and post rainy season of the year 2017.

*Table 5-9*.The sub-basins of Lake Tana, tropical livestock unit density (TLUD; no.km$^{-2}$S$^{-1}$), population density without latrine coverage (PNLD; no.km$^{-2}$S$^{-1}$), watershed export fraction(FEWS), and river export fraction for different seasons of the sub-basins of Lake Tana (Goshu et al., 2020).

| | | | Rainy | | Post-rainy | |
|---|---|---|---|---|---|---|
| Sub-basin | TLUD | PNLD | FE riv (0-1) | FEWS (0-1) | FE riv (0-1) | FEWS (0-1) |
| A/Zemen stream | 90 | 162 | 0.65 | 0.15 | 0.43 | 0.02 |
| Arno_Garno | 125 | 154 | 0.77 | 0.28 | 0.65 | 0.03 |
| Between_Arno_Garno_AddisZemen | 113 | 150 | 0.81 | 0.30 | 0.69 | 0.03 |
| Between_BahirDar_Gelda | 53 | 30 | 0.84 | 0.39 | 0.79 | 0.04 |
| Between_Gelda_Gumara | 105 | 142 | 0.77 | 0.35 | 0.74 | 0.03 |
| Between_Megech_ArnoGarno | 149 | 145 | 0.79 | 0.20 | 0.71 | 0.01 |
| Between_Rib_Gumara | 86 | 203 | 0.78 | 0.27 | 0.66 | 0.02 |
| Dengelber | 156 | 117 | 0.72 | 0.18 | 0.63 | 0.02 |
| Dirma | 142 | 183 | 0.73 | 0.25 | 0.54 | 0.02 |
| Gilgel Abay | 134 | 136 | 0.31 | 0.31 | 0.28 | 0.03 |
| Gelda | 102 | 138 | 0.77 | 0.39 | 0.75 | 0.04 |
| Gemero_Makesegnit | 149 | 146 | 0.72 | 0.11 | 0.59 | 0.01 |
| Gumara | 118 | 170 | 0.65 | 0.37 | 0.60 | 0.03 |
| Infranz | 68 | 118 | 0.78 | 0.31 | 0.72 | 0.04 |
| Megech | 167 | 179 | 0.32 | 0.20 | 0.28 | 0.01 |
| Rib | 119 | 174 | 0.62 | 0.16 | 0.44 | 0.01 |
| WesternTana | 114 | 114 | 0.67 | 0.18 | 0.55 | 0.02 |
| Between Infranz and Bahir Dar | 106 | 106 | 0.66 | 0.37 | 0.85 | 0.03 |
| Between Gumara and Megech | 182 | 182 | 0.30 | 0.24 | 0.71 | 0.01 |
| Between Megech and Dirma | 182 | 182 | 0.59 | 0.27 | 0.60 | 0.02 |

In the rainy season, nearly two-thirds of the Lake Tana basin was moderately polluted. These include major tributary rivers such as Gilgel Abay, Gumara, Rib, Megech, and Dirma sub-basins.

Twenty percent of the basin was highly polluted and 20% of the basin had very low to low levels of pollution (See Fig. 5-4). In the post-rainy season, nearly half of the sites were within the low level of pollution, 35% were moderately polluted, and 10% were highly polluted as indicated by 90% EC concentration (see Fig. 5-4).

In the rainy season, based on ninety percentile CP concentration, 10% of the basin was classified as low to very low, 35% moderate, 35% high, 15% very high, and 5% excessively polluted. In the post-rainy season, based on ninety percentile CP concentration, 10% of the basin was classified as low to very low,30% moderate,40% high, 10% very high, and 5% excessively polluted (see Fig. 5-4).

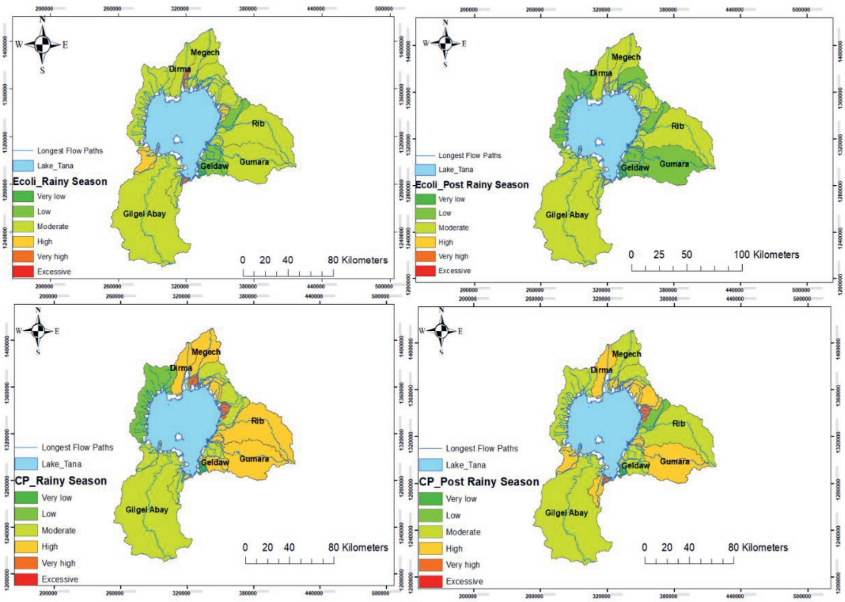



Fig. 5-4 – Rainy and post-rainy season bacteria pollution risk maps of the Lake Tana basin classified into respective pollution levels based on the pollution classes of Kavka et al. (2006) (see *Table 5-7*). Bacteria pollution numbers for the un-gauged sub-basins are estimated with Equation 1.

## 5.4 Discussion

### 5.4.1 Physicochemical Characteristics of the Examined Water Habitats

The result of the physicochemical characterization of the examined water habitats indicated the presence of a wide range of different aquatic habitats and pollution patterns. The physicochemical characteristics of surface and groundwater sites support our presumptive pollution gradient approach: the pollution levels of the sampling sites follow the levels of anthropogenic activity in the surrounding areas. This provided a firm basis for the site selection (Byamukama et al., 2005).

The characteristics of the groundwater sampling sites showed a chemical pollution gradient from the inner city where there is a relatively high population density and more anthropogenic activities, to the middle, and then to the outskirts of the city where there is relatively less human density. However, groundwater sites also receive diffuse pollutants from livestock and humans from the watershed. The nonpoint source pollution from humans is due to many points of open defecation in the watershed. The inner-city groundwater sites showed significantly higher values of temperature, electrical conductivity, ammonia, and total dissolved solids than the middle and outer city stations. This might be explained by the high organic pollution of the sites. Accordingly, the dissolved oxygen concentration was significantly lower in the inner city than the middle and outer city groundwater sampling sites because organic waste pollution consumes oxygen as it degrades.

The surface water sampling sites also showed significant variations across most physicochemical characteristics. Lowland sampling sites had higher values of temperature, conductivity, nitrate, nitrite, ammonia, and total dissolved solids than mid and high altitudes. Dissolved oxygen and pH values of lowland streams were lower than high altitude but higher than mid altitude located sites. The highland streams had the lowest temperature, highest dissolved oxygen, and were slightly more alkaline than other streams.

In general, our physicochemical results for ground and surface water are in line with other studies (Goshu et al., 2010a; Kebede et al., 2020; Zhang et al., 1996) insofar as higher values of nutrients and organic pollution were noted in water bodies where there were high anthropogenic activities. (Pawar et al. 1998) also reported low dissolved oxygen concentration in water as the result of high amount of organic matter because of the influence of increased anthropogenic activities.

Though we did not observe a clear concentration gradient for the physicochemical characteristics of the surface water as we did for the groundwater stations, the downstream sites had relatively higher values than the upstream side of the same stream in most altitude categories. This could be because of intermittent pollution inputs within the watershed as a result of open defecation and open grazing magnified by watershed degradation such as soil erosion (Goshu and Aynalem, 2017; Moges et al., 2016; Selassie, 2017; Sewnet, 2015; Simane et al., 2013).This result confirms the presence of a range of differing habitats, and pollution pattern in the investigated water systems that build the basic framework for having different categories of faecal pollution and testing the performance of indicators.

### 5.4.2 General Level of Bacteria Occurrence and Abundance

The general level of occurrence and abundance of FIB shows more or less the same gradient as that of the physicochemical characteristics of the examined water habitats. This further supports the basic framework to test the performance of the indicators at different altitudes and source waters to come up with affordable, easy-to-perform, and practicable indicators that help to map the pollution risks. The general level of occurrence and abundance of TC, EC, and CP in most surface water sampling sites was higher than in groundwater sampling sites. In most cases, the sampling sites from the lowland area had a relatively higher level of occurrence and abundance than mid and high-altitude streams. EC, IEC, and CP had a relatively higher pooled mean abundance in surface and groundwater habitats compared to other indicators. This was expected as it was observed in other similar studies in Ethiopia (Goshu and Aynalem, 2017; Kebede et al., 2020) and East Africa countries (Byamukama et al., 2005). The observed higher TC levels at all sites were thus probably because of naturally occurring populations in the sampled habitats without being highly indicative of faecal pollution (Byamukama et al., 2005; Goshu et al., 2010a).

In surface water sampling sites, the occurrence and abundance of most indicators followed an altitude gradient wherein lowland sampling sites had the highest occurrence and abundance of all FIB than mid and highland sampling sites and midland sampling sites showed higher occurrence and abundance than highland sampling sites, except for TC. Highland sites showed the least occurrence and detection of all indicators (See Fig. 5-2, A-C).

A study conducted in mid-altitude areas of the Lake Tana Basin by Goshu et al. (2010a reported a comparable level of abundance, but slightly higher values (2 logs) of TC(log 6.3),EC (log 6.1), and CP (log4.1). Kebede et al. (2020) also reported a maximum EC concentration of log 3.88 and an IEC concentration of log (4.04) in the Awisha River, a mid-altitude river in the central highland of Ethiopia. Byamukama et al. (2005 studied several water bodies around Lake Victoria, Uganda, and reported comparable numbers but higher maximum median concentrations for TC (log 6.5) and EC (log 4.3).

### 5.4.3 Performance of FIB

Results indicated that presumptive EC and CP spores' determination showed better discrimination efficacy than other indicators as they were able to significantly discriminate about two-thirds of the sampling sites in pooled samples of surface and groundwater sources located at all altitudes. Moreover, EC and CP revealed higher values of BCRM compared to other indicators.

This finding is in agreement with Byamukama et al. (2005) who tested the performance of EC and CP in tropical waters of Uganda (1180 m.a.s.l.). However, our study evaluated the performance of indicators including very high altitudes, as high as 3850 m.a.s.l. and different sources of water. Edberg et al. (2000 and Odonkor and Ampofo (2013 reported that EC was the best indicator of the bacteriological quality of water in most examined habitats. *CP* seems to be the choice in published literature for measuring past or intermittent pollution and in circumstances where resistance to disinfectants and environmental stress was highest (Cabelli ,1977).

Second to EC and CP, TC and IEC showed useful efficacy to discriminate between water bodies in high and low pollution categories.

The discrimination efficacy of EC and CP was even stronger in groundwater habitats than surface water habitats, presumably due to the strong pollution gradient in the groundwater habitat. This was further supported by comparatively higher values of BCRM in groundwater than in surface water. All the evaluated indicators in the groundwater that included EC, CP, and TC demonstrated strong discrimination efficacy.

The discrimination efficacy of indicators in surface water habitats was not as good as that of groundwater habitats, possibly because surface water habitats are prone to intermittent faecal pollution. Nonetheless, discrimination efficacy varied with altitude, and we noted that the efficacy increased with increasing altitude. The lowland sites have larger watershed areas, are hotter, and have higher population densities of humans and animals than the mid and highlands. Interestingly, the BCRM of EC and CP in surface water sites followed the altitude gradient with higher values for highlands and lower values for lowlands. In most cases, the sensitivity of the indicators increased from lowland to highland.

The sensitivity of EC, CP, and TC in detecting low and high pollution categories follows the presumptive pollution gradient in groundwater sites that were located in the same altitude category. This was further confirmed by higher values of BCRM in the inner city than the middle and outer parts of the city. We noted lower values of BCRM in the outer city groundwater sampling sites.

Although all applied microbial parameters and indicators showed a very high correlation among surface water samples, only EC and CP also revealed a high correlation in the groundwater samples. EC and CP thereby resulted in most of the investigated environments in a consensus picture of faecal pollution and a pollution risk map could be created for those environments. As these two independent parameters of faecal pollution, with different ecological potential and persistence in the environment (Medema et al., 1997), revealed high to very high correlation in both ground and surface water results strongly suggest that these two indicators seemed to work in the studied area. In conclusion, results agree with the findings of Byamukama et al. (2005 for the Ugandan environment and support the use of EC and CP spores as a couple of complementary faecal indicator parameters in high altitude tropical countries of Africa.

### 5.4.4 Microbial Source Tracking

This study indicates that the BacR marker can nearly always be detected in all study sites except midland upstream. Reischer et al. (2013 and Chase et al. (2012 also reported that BacR has strong host adaptation and broad distribution within the targeted bacterial subpopulation, and represents obligatory symbionts in the ruminant digestive system, thus making it an ideal MST target.

In our study, we also detected FIB (TC, EC, CP, and IEC) in most of the sampling sites.

The inability of BacR to be detected in the study sites where we have a clear detection of FIB suggests a need for more qPCR assays to be applied in higher spatial and temporal resolutions along with other MST assays.

### 5.4.5 Pollution Risk Map

Pollution risk mapping as applied by Kavka and Poetsch (2006) was often based on one parameter, most frequently on presumptive EC quantification. However, as it could be demonstrated in this work, the combined use of presumptive enumeration and CP spore determination proved an effective means for faecal pollution monitoring for the Ethiopian situation. For the creation of a faecal pollution map of the investigated study sites, a dual faecal mapping system was thus developed and applied. Some of the groundwater sampling sites, especially located in the outskirt of the city were closed wells and we expect remote contamination, and the environment was expected to be of low oxygen or close to an anaerobic environment. In such circumstances, the combined use of CP with EC is advisable as the CP method only detected spores from CP species, which are strictly anaerobic and exclusively of faecal origin (Ellis 1989).

Based on this mapping approach, in pooled samples, more than 60% of the sites were classified into the same pollution class (Fig. 5-3). In surface water sampling sites, only 30% of the sites were in the same pollution category.

In groundwater sampling sites, three-fourth of the sites were classified into the same pollution class. The very clear pollution gradient in the ground water sites is neatly captured by the consensus FIB. This was especially interesting as EC are derived gram-negative, non-spore-forming, and only facultative anaerobic Entero bacteria cease whereas the CP method only detected spores from *CP*, which are strictly anaerobic and only exclusively of faecal origin (Ellis 1989).

The investigated habits were used for different purposes. Most of the surface water (streams) was mainly used for animal watering, irrigation, washing clothes, bathing, cleaning utensils. In rural areas, it was not possible to imagine that they could also be used for drinking purposes by the rural community. The groundwater especially found in the inner part of the city was rarely used for washing clothes; those in the middle city were used for washing clothes and agricultural activities in their backyards and rarely for drinking. However, one-third of the groundwater sampling sites (four out of twenty) was located in the outskirts of Bahir Dar city and was used for drinking by the local people. We suspected that pollution was infiltrating most of the sampling sites we investigated except wells in the outskirts of the city which had non-detectable to very low levels of pollution. The level of pollution of sampling sites was considerably higher than the WHO drinking water standard. According to WHO drinking water standards, EC and CP should be absent in a water sample of 100 ml(santé et al., 2004).Coli forms were common in the environment and were generally not harmful. However, the presence of EC and CP bacteria in drinking water was usually a result of point and nonpoint source pollution and this is a serious public health concern (Abdelzaher et al., 2010; Polo et al., 1998; Wright et al., 2004).

### 5.4.6 The General Level of Pollution and Risk Map of the Lake Tana Basin

The level of faecal pollution of tributary rivers and the pollution risk map of the associated sub-basins, including the un-gauged ones, remain unknown or not done though about 4.5 million people directly or indirectly depend on the surface and groundwater resource for agriculture, industry, recreation, ecosystem, and drinking in Lake Tana basin.

The results from this study indicate the remarkably high impact potential of anthropogenic faecal sources on receiving rivers of the Lake Tana basin. More than 80% of the gauged rivers' sub-basins have a moderate to very high level of pollution. The level of pollution is more pronounced in the rainy season than the post-rainy season since the rainy season in the study area receives two-thirds of the annual rainfall and the runoff washes untreated agricultural and domestic wastes into the receiving rivers. The estimation of faecal pollution in un-gauged basins based on measurements of the level of faecal pollution in the adjacent rivers, population density without latrine coverage, livestock density, watershed nitrogen export fraction, and nitrogen river export fraction (Goshu et al., 2020) all indicated that more than three-fourths of the un-gauged sub-basins have moderate to very high levels of pollution.

This finding is in line with the findings of (Abera et al., 2017; Abera et al., 2014; Ewnetu et al., 2014; Goshu and Aynalem, 2017; Goshu et al., 2010a; Goshu et al., 2017; Mengesha et al., 2004; Wondie, 2009). The moderate to very high levels of pollution in the Lake Tana basin could be attributed to low coverage of sanitation facilities and high livestock densities with open grazing systems (Alemayehu and Tassew, 2017; Mengesha et al., 2004) aggravated by land degradation in the basin (Selassie, 2017; Yitaferu, 2007b).

## 5.5 Conclusion

The performance of FIB to differentiate between high and low pollution categories (discrimination efficacy) varied with altitude and source water in highland tropical water. The FIB showed better efficacy in groundwater systems than surface water systems because of nonpoint intermittent pollution loads to surface water systems due to high anthropogenic influence in the watersheds. Among the indicators, presumptive EC and *CP* spores showed the better performance to differentiate high and low pollution categories in groundwater than in surface waters. Most indicators showed better discrimination efficacy to differentiate high and low pollution categories at highland sites (3835 m.a.s.l.) than mid (1800 m.a.s.l.) and lowland sites (1100 m.a.s.l.). Although all applied microbial parameters and indicators showed a very high correlation in the surface water samples, only EC and CP also revealed a high correlation for the groundwater environment. Presumptive EC cell enumeration and CP spore determination for faecal pollution monitoring performed well in the high-altitude tropical country of Ethiopia. Based on this couple of faecal indicators a pollution risk map could be created for the investigated environment. EC and CP thereby resulted in most of the investigated environments in a consensus picture of faecal pollution. Most of the sub-basins of Lake Tana are found to be moderate to highly polluted and the level of pollution is higher in the rainy season than in the post-rainy season. Markers associated with ruminants (BacR) are identified in more than three-fourths of the sites, indicating the presence of mainly ruminant-associated faecal pollution.

Further study on the performance of molecular markers should be conducted at different source waters and altitudes with higher spatial and temporal resolutions as this study only represented a discrete region of the Ethiopian environment.

# 6

# General Discussion and Conclusion

# Chapter 6. General Discussion and Conclusion

This chapter includes a general discussion and conclusion of the thesis. First, I give a general discussion, then I summarize the main findings and novelties of the thesis and reflect on the modelling and experimental methods used in this thesis. Finally, I draw a lesson and present a future outlook.

## 6.1 General Discussion

Food security and unemployment are huge challenges facing the Ethiopian government. To solve these problems and promote social and economic growth, the Ethiopian government formulated an agricultural-based development strategy in the mid-1990s, namely, after the Agricultural Development-Led Industrialization (ADLI), it began to carry out agricultural transformation. Agricultural development-led industrialization believes that agriculture is the main stimulus for creating more production, employment, and income for people, and it also serves as a springboard for the development of other economic sectors (Keeley and Scoones, 2000). The Ethiopian government has formulated an agricultural intensification strategy, which is a technology-based, supply-driven intensification strategy that includes better supply and promotion of agricultural inputs and fertilizers (Kassa, 2005).

The Lake Tana Basin has been identified as an economic corridor by the Ethiopian government. To this end, numerous plans for water resources, agricultural and industrial developments (short, medium, and long-term plans) have been planned, and some are being implemented. Irrigation development plans for the basin include dams and diversion projects that span many tributaries, as well as plans to pump water directly from Lake Tana. The basin flower farms are in different stages of operation. The largest farm, Tana Flora (with a total area of 124 hectares) performs at full capacity taking water directly from Lake Tana. The farm is still being expanded to include more flower-growing areas. The Robit flower farm along the road to the Gonder and Meshenti areas is at the infrastructure development stage and hopes to be operational as soon as possible. The last two farms are not far from Lake Tana, and they must use deep wells as water sources in their yards to grow flowers. More integrated horticulture and floriculture farms are in pipeline in the western part of Lake Tana.

As a result, the agricultural and industrial intensification, urbanization, low sanitation coverage, and open grazing system in the Tana Lake Basin have produced a large amount of fertilizers, pesticides, nutrients, and bacteria wasted on the surface and groundwater systems. Extensive degradation of the watershed exacerbates this situation, manifested by sediment loading, loss of soil fertility and decline of land productivity, the disappearance of wetlands, drying of rivers, and seasonal floods in the plains. Population pressure in the basin (5.48 million people, with an annual growth rate of 2.8%) has exacerbated these problems. The above-mentioned challenges, coupled with the challenges of recently implemented plans and the plans to be implemented make the Lake Tana Basin more susceptible to watershed degradation. Eutrophication has become one of the environmental problems in the Tana Lake Basin. This is evidenced by the presence of invasive alien species in the northern and north-western parts of the Lake and the presence of algal blooms in the river mouths and lakeside areas. The problem of water hyacinth is a prevailing problem though a daunting annual campaigns to manually remove weed. Lake Tana was once a poorly nutritious lake, but its nutritional status is gradually changing, especially since the estuary has received more load from the sub-basin. Of course, due to the increasing concentrations of phosphorus and nitrogen, water resources all over the world are under pressure, which harms public and ecosystem health, and various economic sectors. The increase in nutrient concentration is caused by point and diffuse sources. Because of the importance of Lake Tana in supporting intensive irrigated agriculture now and in the future, it serves as a source of hydroelectric power (the water source of Tana Beles and GERD hydroelectric power plants that generate 460 and 6000 MW respectively), fisheries and tourism industry, bird habitats and biodiversity resources, and the Lake Tana basin environment is worth protecting. However, for example, lakes are receiving more and more waste; point and non-point sources of nutrients and bacteria and show signs of deteriorating water quality.

## 6.2 The Research Objectives, Main Findings and Novelty of my Thesis

The overall research objective is, therefore, to understand better river export of nutrients to Lake Tana by sub-basin, season, and source and to analyse the importance of the spatial heterogeneity of eutrophication, the performance evaluation of faecal pollution indicators, and the risk mapping of faecal pollution. Four research sub-objectives are formulated to achieve the overall objective ( Fig. 6-1). These are:1) overview water quality of Lake Tana Basin, and also presenting some of my works, 2) develop a seasonal, source, and the sub-basin scale model to assess N inputs to Lake Tana as influenced by human activities, climate, and hydrology in the basin. This was done by developing a model following the sub-basin approach of the MARINA model and the seasonal approach of NEWS 2-DIN-S model approach, and applying the model to analyze N inputs and their sources, 3) assess the spatial variability of eutrophication in tropical shallow lakes taking Lake Tana (Ethiopia) as a case show, 4) evaluate the performance of Faecal Indicator Bacteria (FIB) including total coliforms (TC), presumptive *E.coli* (EC), intestinal enterococci (IEC), and presumptive Clostridia perfringens spores (CP), and to determine ruminant-associated faecal pollution using qPCR assay(BacR) at different water types and altitudes (1100-3835 m.a.s.l) in a highland tropical country, Ethiopia.

The above objectives have been addressed in Chapters 2-5 of this thesis. The novelties of Chapters 2-5 are discussed below with the focus on the export of nutrient and bacteria pollution from sub-basins to the lake and the effects of nutrient pollution on the ecosystem of the lake. Fig 6.1 summarizes the main findings of this thesis.

Chapter 2 is unique in that it systematically reviews the available data on the water quality of Lake Tana Basin. The Tana Basin is identified as an important economic corridor due to the huge potential of its water resources for social and economic development. Following this decision, intensification of agriculture in the basin is expected and this development effort together with an expansion of industrialization and urbanization produce huge waste which affects the water quality of the Tana basin. Therefore, to know the current level of deterioration, understand the trend, and effectively and sustainably manage water resources in the basin, the water quality must be evaluated in the spatial and temporal dimensions. However, scientific information about Ethiopian water bodies in general and the Lake Tana basin, in particular, is scarce. The reason why there is no solid database of water quality is due to the lack of a sustainable monitoring plan. Hence, there is no detailed and systematic water quality characterization based on long-term and spatially representative data. The water quality data and research seem scattered, nonetheless, existing information on Lake Tana shows that nutrient and bacteria contamination, the nutritional status of the lake has gradually reached an alarming level to the extent that it poses mesotrophic and eutrophic conditions at the shore and river mouths of the lake. Faecal contamination and toxin-producing bacteria are detected in the lake, especially onshore areas and river mouths.

Water quality impairment is going on for years, which significantly affects the basin's potential for agriculture, industry, hydroelectric, ecosystem, water supply, and recreation sectors. The available data shows sedimentation, pollutants, wetland encroachment, and hydrological alterations are the main issues of water quality management in the basin.

Chapter 3 is novel and introduces a seasonal version of the MARINA model (a model for assessing nutrient input from rivers to lakes) of the lake and the rivers of the basin. We integrate two existing modelling approaches into a seasonal model for river export of DIN for large tropical lakes. This led to a seasonal version of the MARINA model for sub-basins and rivers that discharge into Lake Tana. We applied this model for the year 2017, during which time the export of DIN was approximately 9 kilotons to Lake Tana. River export of DIN to Lake Tana shows spatial and temporal variability, being highest in the rainy and lowest in the dry season. For example, two-thirds of the total DIN exports each year are exported through rivers during the rainy season, while 30%, 3%, and 2% of DIN are exported through rivers before, after, and dry seasons, respectively. The diffuse source from agriculture is the main contributor to DIN in the river. Animal manure is the main source in all seasons. Synthetic fertilizers are the second-largest source in the rainy season. In all seasons of 2017, human excrement on the earth is an important source of nitrogen. More than two-thirds of DIN's annual river exports come from 6 of the 20 rivers, namely Gilgel Abay, Dirma, Megech, Rib, Gumara, and Gelda. The Gilgel Abay sub-basin alone exports about a third of the annual DIN of river exports to lakes.

My research shows new insights into the seasonality of the DIN river export to Lake Tana. I demonstrate the importance of diffuse sources in lake pollution. This is especially suitable for wet periods when contamination levels are high. I also show areas (sub-basins) that need pollution control to avoid further pollution of the lake. This information helps develop effective management plans for the lake by sub-basin, season, and specific source. Additionally, my seasonal model can be used to analyze future trends in lake pollution. I provide a seasonal method that can be applied to sub-basins and rivers that flow into other large tropical lakes that face similar environment.

Chapter 4 applies a new method to couple the Duflow two-dimensional model (usually ID and calculate the unsteady flow) with the PC Lake model of the ecosystem (zero-dimensional ecosystem model), and define the different impact zones of the tributary rivers of Lake Tana, analyze the spatial variability of the eutrophication of Lake Tana, and determine the critical load of N and P. Understanding the spatial variability of eutrophication and its impact on lake characteristics is essential for the effective and sustainable management of lakes and their watersheds. This modelling study took Lake Tana as an example to analyze the importance of spatial heterogeneity in exploring the nutrient load of shallow tropical lakes. The model results indicate the importance of the different ecological conditions in the affected area and the spatial heterogeneity of eutrophication in shallow non-stratified lakes. This spatial variability of eutrophication is reflected in the critical load. The model indicates that vegetation in affected areas 1 and 2 is dominant, and vegetation in affected areas 3 and 4 is less dominant. The southern and eastern parts of the lake and the shore areas show eutrophication/turbidity conditions, as evidenced by the high concentration of Chl-a and low vegetation density. The model shows that the concentration of Chl-a varies from n. d to 58.72 $\mu gl^{-1}$, and the vegetation coverage ranges from 1.33 to 127.54 $\mu gl^{-1}$. I show how eutrophication and hydrological spatial heterogeneity can contribute to the management of large tropical lakes.

The novelty of Chapter 5 is that it provides the first Faecal Indicator Bacteria (FIB) performance evaluation and qPCR test, ruminant-related faecal pollution when used to distinguish high and low pollution categories that vary with altitude and water source. In tropical highland waters. among these indicators, the *Escherichia coli* and *Clostridium perfringens* spores perform better in distinguishing between high pollution and low pollution categories in groundwater than in surface water. Most indicators show better efficiency in distinguishing between high and low pollution categories of highland (3850 m.a.s.l), rather than medium (1800 m.a.s.l) and lowland (1100 m.a.s.l). Among the applied faecal pollution indicators, only EC and PC showed a very high correlation in surface water and groundwater samples. FIB shows better efficiency in the groundwater system than in the surface water system. This is due to the high degree of human influence on the basin resulting in the intermittent non-point pollution load of the surface water system. In most of the samples studied, EC and CP resulted in inconsistent images of Faecal contamination. In the high-altitude tropical countries of Ethiopia, presumptive EC cell counts and PC spore determination seem to be a good indicator for faecal contamination. Based on these double consensus Faecal indicators, a pollution risk map can be developed for the water body under study. It was found that most of the sub-basins of Lake Tana were moderate to highly polluted, and the pollution level in the rainy season was higher than that in the post-training season. Ruminant-related markers (BacR) have been identified in more than three-quarters of the locations, indicating the presence of faecal contamination mainly related to ruminants. My research can help improve water quality testing and reduce the risk of bacteria in streams to the general population. More research should be done on the performance of molecular markers of different water sources and altitudes with a higher spatial and temporal resolution, as this research only represents a discrete area of the Ethiopian environment.


Fig. 6-1– An overview of the main findings of this Ph.D. thesis.

## 6.3 Modelling nutrient losses from sub-basins to the Lake: Models of this thesis

Observed data on river export of nutrients from large and ungauged catchments to lakes are generally scarce in tropical regions, and most empirical studies have been performed in temperate zones(Ivens et al., 2011).This creates the problem that in many tropical lakes and rivers eutrophication is observed, but data on the causes of eutrophication are lacking, and this makes the development of early warning system, and sustainable management of eutrophication a daunting exercise and ineffective. In such data scarce regions, nutrient export models can be helpful in the formulation of policies aimed at reducing eutrophication.

Basin-scale nutrient export models are useful tools to understand the causes of nutrient exports from land to water systems, and to explore effective solutions. The MARINA family models can quantify flows of different forms of nutrients from land to coastal water areas by source and by sub-basin for current and future years. However, these models have not been applied to the sub-basins of lakes in Ethiopia while accounting for seasonal variation. The Duflow model which I upgraded to 2D application of the flow model helps to define the impact zones in lakes by rivers. This furthers helps to Understand spatial variability of nutrient loading and transport in large lakes, and its effect on the eutrophication status. Another group of models are aquatic ecosystem models (such as PCLake+), that can help to estimate critical nutrient loads. Critical nutrient loads reflect the amount of nutrients that the aquatic ecosystem can absorb. Linking basin-scale nutrient export models and aquatic ecosystem models for nitrogen and phosphorus can provide new insights in how to reduce external nutrient loads so that critical nutrient loads in lakes are not exceeded.

**MARINA model (Basin-scale nutrient export model)**

MARINA model (Model to Assess River Inputs of Nutrients to seAs) is a sub-basin version of the Global NEWS-2 (Nutrient Export from WaterSheds) model with an improved approach for nutrient losses from animal production and population for China. It has a higher spatial resolution than the Global NEWS-2 model. The model computes annual river export of different N and P species that include dissolved inorganic (DIN, DIP) and organic (DON, DOP) forms for 1970, 2000, and 2050 by sub-basin. The model computes river export of nutrients to Chinese seas as a function of human activities on land, sub-basin characteristics, and retentions of nutrients in rivers. The first version of the MARINA model family, MARINA 1.0, was developed and applied to six large river basins in China (Strokal et al., 2016a). Subsequently, extended versions were developed to quantify river export of nutrients to Chinese seas in more temporal detail (e.g., MARINA 1.1) (Chen et al., 2019a), or more spatial detail (e.g., MARINA 3.0) (Chen et al., 2019b) or from a future climate change perspective (e.g., MARINA 2.0) (Wang et al., 2020a).

**NEWS 2-DIN-S**

NEWS 2-DIN-S (McCrackin et al., 2014) is a seasonal version of the NEWS-DIN model which is the DIN module of the Global NEWS-2 model. NEWS.2-DIN-S operates at the basin scale for large rivers in the world. It is a spatially explicit model that analyses the seasonal flow of DIN to rivers.

I developed the seasonal version of the MARINA-model for river export of dissolved inorganic nitrogen from sub-basins discharging into a large tropical lake by merging the sub-basin scale approach of Strokal et al. (2016) with the seasonal approach of McCrackin et al. (2014) **(Chapter 3)**. This results in a seasonal model for sub-basins and rivers discharging to tropical lakes such as Lake Tana. We applied the seasonal model to assess N exports in Lake Tana sub-basins and we validate the results with measured data in the same area.

**PCLake+ (Aquatic ecosystem model)**

The PCLake+ model is a complex ecological model used to assess the nutritional status of stratified and non-stratified freshwater lakes around the world (Janssen et al., 2019b). PCLake+ is an updated version of the PCLake model was originally developed for shallow lakes (Janse, 1995). Both models are known for their ability to calculate critical nutrient loads in shallow lakes. The PCLake+ model includes biological (for example, phytoplankton, submerged vegetation, and simplified food web) and non-biological (for example, transparency and nutrients) modules within the aquatic food web framework (Janse et al., 2010). The food web starts with submerged plants and three groups of algae (i.e., green algae, diatoms, and cyanobacteria) with the lowest nutritional level, and ends with predatory fish at the top of the food web. N and P cycles in the upper layer of the water column and sediments combine with higher levels in the food web (Janse et al., 2010). The swamp is also a PCLake+ model module. This module involves a simplified growth model of *Phragmites australis*, an emergent species of macrophytes, combining with vegetative processes in the upper layer of sediments in the water column and the swamp area (Sollie et al., 2008). The swamp belt with emergent plants is connected to the open water by a flow of water between them (Janssen et al., 2019b).

**Duflow (2D Tracer modelling)**

It is a flow and quality model that computes the flow and quality. With this program, one can perform unsteady flow computations in networks of open watercourses. Duflow is also useful in simulating the transportation of substances in free surface flow and more complex water quality processes.

We combined two modelling approaches with modifications for our study area based on the models Duflow (Clemmens et al., 1993) and PCLake+ ( Janse et al. 2010). We used two-dimensional applications of the flow and quality model Duflow, to perform unsteady flow computations and to simulate tracer distribution in the lake to define the so-called impact zones. Thereafter, we ran the PCLake+ model for each impact zone separately, to simulate ecosystem conditions (e.g. nutrient concentrations, phytoplankton, vegetation) and to determine critical nutrient loads for ecosystem transitions (Janse et al., 2010).

### 6.1.1 Building trust in the used models

Building trust in the models can be done in many ways that include comparing model output with measurements (validation), comparing model inputs to independent datasets, comparing to other models, or by sensitivity analysis (Strokal et al., 2016).In the case of Lake Tana basin where there is lack of data except the PhD project data, we did validation, used reliable sources of data (mostly from government institutions), and reviewed the sensitivity analysis of the model conducted somewhere else in a similar environment. I mostly used the classical method of model validation which is measurement. We conducted a one-year field campaign to measure concentrations of different nitrogen, phosphorous and bacteria concentrations, and Chl-a data from the major tributary rivers of the Lake Tana Basin. The measurement was done every month in the year 2016/ 2017.

As part of validation, the performance of the models was assessed by computing the different model performance metrics. The model performance metrics include $R^2P$ (the Pearson's coefficient of determination, 0–1), $R^2NSE$ (the Nash-Sutcliffe efficiency, $-\infty$-1; (Nash and Sutcliffe, 1970) and RSR, > 0, according to Moriasi et al. (2007). $R^2P$ shows the proportion of the variance in the observed data that is predictable from the modeled data, with higher values representing better performance. $R^2NSE$ evaluates the fitness of the observed and modeled data with the 1:1 line, with higher values representing better performance. RSR is the ratio of root mean square error and standard deviation, with values close to zero showing very good model performance.

I modified the modelling approach of McCrackin et al. (2014) by re-calibrating the 'a' and 'b' model parameters that determine the runoff function. I found the optimum values for parameters a and b that resulted in the minimum difference between modeled and observed values. To calibrate and validate the integrated model, I used different data sets of observed and modeled DIN loads and yields to the lake. For model calibration, we used the monthly values of seasonal observed and modeled DIN at or near the mouths of the Gilgel Abay, Gelda, Gumara, Rib, and Dirma rivers, all of which had DIN observed at or near the river mouth. We used seasonal DIN load and yield data of the Arno-Garno, Megech, and Infranz rivers for model validation.

I validated the results of the PCLake+ model with the limited seasonal dataset, including Chl-a concentration and vegetation density variables. Due to a shortage of data, we did not use the coefficient of determination ($R^2$) and the mean relative absolute error (RE) to evaluate the model results by following the previous study (Janssen et al., 2017).

Validation results provide trust in modelling seasonal nutrient exports to Lake Tana, Tana Basin (Chapters 3 ) and quantifying the critical nutrient loads for Lake Tana (Chapter 4). I evaluated the model inputs of seasonal MARINA, Duflow, and PC Lake + models. The data sources are broadly reports of Central Statistical Agency of Ethiopia, Agriculture and Water and Development Bureaus and Offices, data from design and supervision enterprises, published works, and monitoring data during the fieldwork. I considered these input sources reliable because they reflect the ideal local conditions in Lake Tana Basin and Ethiopia and contain the most complete information that can be used in my modelling. For example, the MARINA model input data broadly needs land use, climate, and hydrology data.

The local information on irrigation demand was collected from the agricultural offices in the basin. The amount of DAP and Urea applied in each sub-basin is derived from districts' data on the amount of DAP and Urea utilization. To know the harvested biomass of legumes, we used crop production data compiled at a zonal level (CSA, 2016) and disaggregated it into respective sub-basins based on the area-weighted method. The data contains crop type, number of holders, the area covered, production, and yield. The % protein content of a legume (16%) was taken from the United States Department of Agriculture, Agricultural Research Service, the National Nutrient Database for Standard Reference Legacy Release (http://www.fao.org, retrieved on 14 May 2018). In addition to local information on the crop calendar, the FAO crop database (http://www.fao.org, retrieved on 14 May 2018) was consulted to know the planting and harvesting times of crops. We derived the $N_2$ fixation rate by natural vegetation from the database of the Global *NEWS*- 2 (Nutrient Export from Watersheds) model (Mayorga et al., 2010). The fixation rate we got from the Global *NEWS*- 2 is the annual fixation rate and it should be further disaggregated to seasons. The disaggregation of the annual rate to seasons is based on the seasonal air temperature as it affects the activity of the bacteria. However, there is no local data about the rates of deposition both in agricultural and natural areas and the only available data is annual atmospheric N deposition rates in agricultural and natural areas from the database of the Global *NEWS*-2 model (Mayorga et al., 2010). These annual rates, however, were disaggregated to seasonal rates. The disaggregation was done based on seasonal rates using seasonal rainfall patterns. The data on the number of Housing Units in urban and rural areas and their sanitation types for toilet facilities were taken from the Ethiopian National Household Census Report (CSA, 2007). Livestock data was collected from district agricultural offices that are located in the

Lake Tana Basin. First, we derived livestock population per district. Then, district information was converted to sub-basins.

The meteorological data such as air temperature and rainfall are collected from Meteorological Offices. We defined seasons based on rainfall. The monthly rainfall of a sub-basin was taken from the nearest meteorological stations in the Lake Tana basin.

The hydrology data set, actual water discharges at the outlets of the gauged sub-basins for seasons (Qact, $km^3 S^{-1}$) in 2017 were collected from the Ethiopian Ministry of Water, Electric, and Irrigation.

### 6.1.2 Strengths and Limitations of the Models Used in This Thesis

I applied three models in my PhD thesis. These are the  Seasonal MARINA (Goshu etal., 2020), Duflow(Clemmens et al., 1993) and PCLake+(Janse et al., 2010) models. I did coupling of models to better understand the situation in Lake Tana Basin. For instance, I coupled Duflow and PCLake+ models to define an impact zone in Lake Tana by tributary rivers and set critical thresholds of N and P for Lake Tana. This multi model approach is a strength of my thesis. The strength of each model is described briefly as follows.

**Strengths of Seasonal MARINA model (Goshu etal., 2020)**

The strong side of my modelling is that there was almost complete data for the year 2017. I set up the model for 2017. I assume that 2017 can be seen as a representative year in terms of DIN sources. In the first place, I have adequate data of all model inputs for the year 2017. Moreover, I found no significant difference in rainfall and temperature between 2017 and the long-term annual average (1952 - 2015). Furthermore, the fertilizer application in the Lake Tana basin is not based on the fertilizer requirement of the agricultural land. So, the rate of fertilizer application in the basin is constant over the years. I also systematically collected observed data for all months of the year 2017, giving us confidence in the representativeness of the observed values for different seasons. Although data were only available for one year, this enabled me to calibrate and validate my seasonal model. I realize that more years would be better, but after calibrating the seasonal parameters (Chapter 3), Seasonal MARINA model outputs were generally in good agreement with observed data (e.g., $R^2$NSE> 0.65, see Chapter 3).

The seasonal MARINA model accounts for the seasonal patterns of river export of DIN to lakes. The seasonal MARINA model is novel in quantifying seasonal river export of N to Lake Tana. The seasonal MARINA model accounts for seasonal variability in human activities (e.g., agriculture) and natural processes (e.g., atmospheric N deposition, biological N2 fixation) on land, climate (e.g., air temperature), and hydrology (e.g., runoff and river discharges). For example, the annual N losses from land to rivers are distributed across seasons according to crop phenology, local farming practices, and hydro-climatic conditions (e.g., runoff and air temperature) (Chapter 3, Chen et al., 2019a; McCrackin et al., 2014; Yang et al., 2021). Retentions and losses of N in rivers are quantified as a function of temperature and associated processes (e.g., denitrification), river damming and water consumption (e.g., irrigation) (Chapter 3; Chen et al., 2019a; McCrackin et al., 2014; Yang et al., 2021). The model results provide quantitative information on the seasonal patterns of nutrient export to lakes. The Seasonal MARINA model is a useful tool to provide seasonal information for decision-making to reduce eutrophication problems such as harmful algae blooms. This is because algal blooms often last for certain seasons or months (Davis et al. 2009; Tang et al. 2006a).

The Seasonal MARINA model quantifies river export of nutrients by source, form, and sub-basin. The seasonal MARINA model is strong in quantifying sources of nutrient export from land to lakes by nutrient form and sub-basin (Chapter 3, Chen et al., 2019a; Strokal et al., 2016a; Wang et al., 2020a). The model distinguishes between point and diffuse sources of nutrients in rivers draining into the lakes. Diffuse sources of nutrients in rivers include N and P from the use of synthetic fertilizers, animal manure, and human excretion on land, atmospheric N deposition, and biological N2 fixation. Point sources of nutrients in rivers include direct discharges of waste to rivers (without treatment) and sewage systems in rural and urban areas. The seasonal MARINA model distinguishes between dissolved inorganic (DIN and DIP) and dissolved organic (DON and DOP) N and P.

Different forms of nutrients in rivers result from different sources. Quantitative information on sources of different nutrient forms in rivers supports the formulation of effective policies to reduce nutrient pollution in the lakes. For example, DON and DOP in rivers are generally dominated by human waste and animal manure. In contrast, most DIN and DIP in rivers are often from the use of synthetic fertilizers and direct discharges of manure. The seasonal MARINA model quantifies nutrient exports to lakes by sub-basin (Chapter 3; Li et al., 2019b; Wang et al., 2019). It enables a better understanding of the origin (i.e., sub-basin) of the nutrient pollution in rivers and lakes. This can help identify critical areas for improving nutrient management practices.

**Strengths of the PCLake+(Janse et al., 2010)**

A large diversity of lake ecosystem models exists to study the water quality and eutrophication with different level of complexity and data requirement. There are very simple lake ecosystem models like model on alternative stable states that have few equations(Scheffer and van Nes, 2007) and very complex ecosystem models like Delft 3D-WAQ/ECO(Los, 2009) which help to understand water quality and eutrophication. There are also lake ecosystem models of intermediate complexity such as My lake (Saloranta and Andersen, 2007),and PCLake+(Janse et al., 2010). Compared with other models with an intermediate level of complexity, PCLake+ includes key lake food web components. Ecological feedbacks between these food web components, such as the interaction between macrophytes and algae, are key to critical nutrient loadings at which the ecosystem state shifts from a clear to turbid state or vice versa. As PCLake+ includes these key feedbacks within the food web, the model is well-suited to studying impacts of eutrophication and oligotrophication on lake ecology and water quality(Janse et al., 2010).

**Strengths of the Duflow(Clemmens et al., 1993)**

The Duflow computer program for unsteady open-channel flow is relatively inexpensive and easy to use. It has moderate complexity and moderate data requirement. Data entry and presentation of results are conveniently handled.

**Linking the seasonal MARINA model to estimate nineteen percentile bacterial concentration for the ungauged sub-basins**

In chapter 5, I used a novel method to estimate the ninety percentile log values of ‘’consensus’’ faecal indicator bacteria for the classification of the investigated sites into respective pollution classes. To estimate the nineteen-percentile bacteria concentration for the ungauged sub-basins, I used the output of the seasonal MARINA model, the N watershed export fraction of gauged sub-basin j and season S ($S^{-1}$)($Fews_{ung.j.S}$ ), and the DIN river export fraction of gauged sub-basin j and season S ($S^{-1}$)( $Fews_{g.j.S}$). The linking of the MARINA model to the estimation of the nineteen-percentile bacteria concentration for the ungauged sub-basins is a novel aspect of this thesis because it is a new application for the Lake Tana basin where there is a lack of a tool to estimate diffuse emissions of bacteria.

**Linking seasonal MARINA model to PCLake+ model**

The linking of the seasonal MARINA model and PCLake+ is a novel aspect of this thesis because it is a new application for Lake Tana that accounts for both N and P, helps to set critical nutrient thresholds and it provides new insights into the causes and effects of nutrient exports to Lake Tana. Such a model system is a useful tool to analyze the effects of nutrient export on water quality and aquatic ecosystems of lakes (Li et al., 2019b; Wang et al., 2019). This is done by quantifying the gaps between critical nutrient loads from PCLake+ and actual nutrient loads to the lakes from the seasonal MARINA model, especially in a data-poor region. Critical nutrient loads indicate the maximum nutrient loads that a water body or ecosystem could absorb without exceeding water quality standards (Landis, 2008). The exceeding of actual nutrient loads to water quality thresholds indicates high possibilities for water deterioration and thus for damaging the ecosystem services of the lakes.

**Linking the seasonal Duflow model to the PCLake+ model to define the so-called impact zones.**

The linking of the Duflow(Clemmens et al., 1993) and PCLake+( Janse et al. 2010) to define the so-called impact zones is the novel aspect of this thesis. Duflow is normally 1D, but I upgraded it to

2D, two-dimensional applications of the flow and quality model Duflow, to perform unsteady flow computations and to simulate tracer distribution in the lake. Thereafter, we ran the PCLake+ model for each impact zone separately, to simulate ecosystem conditions (e.g., nutrient concentrations, phytoplankton, vegetation) and to determine critical nutrient loads for ecosystem transitions.

**Limitations**

A multi-model approach is defined as one in which more than one model-each derived from a different perspective and utilizing correspondingly distinct reasoning and simulation strategies-are employed. A weakness of the multi modelling approaches is that multi-modelling approaches always have inconsistencies between the models that may affect the interpretation of the model results.

However, as more models are included in an ensemble, the obtained consistency of performance increases: for larger ensemble sizes, the spread between the best and the weakest performing model combination becomes considerably lower. The use of a large multi model ensemble should be a good choice for obtaining stable and reliable results and at the same time also allows for an estimation of the (model) uncertainty (Donat et al., 2010).

**The models of this thesis have the following limitations:**

Limitations of the MARINA model are related to the model's structure, inputs, parameters, and scenario assumptions. First, the MARINA model lumps processes within a sub-basin to quantify nutrient retentions in soils and rivers. The model uses parameter-based modelling approaches (Chapters 3,4 and 5) that may introduce uncertainties. Second, model inputs for the seasonal MARINA model were derived from many sources, such as statistics books, government reports, etc. To my knowledge, these data sources are the most complete and reliable in Ethiopia. However, in the absence of some local data, we used data from the national, regional or global data bases.

The default settings of the PC Lake+ model are Dutch shallow lakes. However, we ran PCLake+ with default settings and not specifically calibrated for tropical lakes, Lake Tana. Moreover, in the PCLake+ model, floating vegetation is not explicitly modeled.

Our newly developed seasonal approach has some uncertainties that are largely associated with model inputs and parameters. We used Sentinel 2 land cover 2016 imagery to calculate areas of different land cover. Sentinel 2 (20 m–by–20 m resolution) land cover imagery has an overall area-weighted accuracy of 65% +/- 1 % (Lesiv et al., 2017) but could nevertheless contribute to the uncertainty of our model. The use of global-scale data, which has a 0.5° by 0.5° resolution, to local sub-basins of Lake Tana might also introduce some uncertainties. Nevertheless, validation results showed a good performance of the model to quantify seasonal river export of DIN to Lake Tana. Therefore, we consider the model uncertainties to not affect the main messages of our study.

## 6.4 Comparison to Other Modelling Approaches

Three different types of models are used in this thesis. MARINA model is a basin-scale nutrient export model, Duflow model is useful in simulating the transportation of substances in free surface flow and more complex water quality processes. PCLake+ is a complex dynamic model accounting for nutrient cycles in the food web of lakes. Below, I compare the MARINA model to other basin-scale nutrient export models and PCLake+ to other aquatic ecosystem models (i.e., lake-scale), and Duflow to other flow and quality models.

**Basin-scale nutrient export models**

Models are needed to improve our understanding of the interaction between different processes in different environmental compartments (soils, groundwater, riparian zones, streams, rivers, lakes, reservoirs), and quantify the net result (Kroeze et al., 2012).There exist different models with different purposes; deterministic versus stochastic; static versus dynamic; empirical versus mechanistic, and lumped versus distributed.

In deterministic models, the output is completely dependent on input variables and parameters, the output is a single value. In stochastic models, there is intrinsic randomness, same set of input variables and parameters will lead to ensemble of varying outputs and the output can be represented by a distribution accounting for the uncertainty. Static models are time-invariant and calculate the equilibrium of a system. They represent the static components of a system and the relations between them. Dynamic models account for the time-dependent changes in the state of a system. Empirical models (functional, inductive) are based on the observed relation between inputs and outputs. The internal mechanics of a system (the processes) are not known, but the results are known. In mechanistic models (deductive), the internal mechanics of a system (the processes) are known and translated into rules and equations.

The lumped model considers individual sub-basins as a single unit, whereas the distributed model sub-divides each sub-basin in smaller cells. A lumped parameter model will consider the content of one equipment to be homogeneous; it will not consider radial or axial gradients in the fluid properties. A fully distributed model will consider the variation of fluid properties and interaction with its neighbouring elements in all three dimensions and over time (Kroeze et al., 2012).

Different types of nutrient export models serve different purposes. The lumped annual-scale models are useful for assessing past and future trends in nutrient export to coastal seas at annual temporal scale and the spatial scale of world regions and groups of river basins. They have also proven useful in assisting decision making by identifying likely causes of coastal water pollution in data-poor regions On the other hand, dynamic and distributed models may be the most appropriate types to improve our understanding of the various nutrient loading, retention, impacts and seasons (Kroeze et al., 2012).

Basin-scale nutrient export models are often applied to estimate nutrient losses from land to water systems such as rivers, lakes, and coastal waters. In Ethiopia, there is generally a lack of tools to model nutrient and bacteria export and to simulate the effect of these nutrients in the ecological condition of the Lake. Lake Tana basin is a data-poor region, and we hardly find applications of water quality models. Though SWAT was data-intensive, it was applied in the Lake Tana basin to spatial delineate soil erosion, assessing the implications of water harvesting intensification on upstream-downstream ecosystem service and hydrological and sediment yield modelling (Dile et al., 2013; Dile et al., 2016; Setegn et al., 2008; Setegn et al., 2009).

The SWAT calculates pollutant fluxes based on hydrological response units, and the outputs are further aggregated to basins (Dong et al., 2018; Wang et al., 2014c). Thus, this model applies to both large and small basins, especially for predominantly agricultural basins. Moreover, SWAT is a dynamic model, providing high-resolution results of multiple pollutants (e.g., total N and P, pesticides, heavy metals) at daily, monthly or seasonal scales. The spatial and temporal scales rely on the model inputs. However, the demand for data is high. Therefore, the SWAT model is often applied to past and present years. It is challenging to predict future trends of nutrient export from land to water systems. Moreover, the output of SWAT cannot explicitly provide the source attribution of nutrient loads to water systems.

Though the following water quality models are not intensively used in the Lake Tana basin, they are reviewed as presented below (see *Table 6-1*).

*Table 6-1.*Purpose, modelling approach, nutrient forms, scales, sources of nutrients, complexity and data requirement of the models used in the thesis and other nutrient export models

| **Models** | **Global NEWS** | **MARINA** | **NEWS 2_DIN-S** | **Duflow** | **PCLake+** | **SWAT (Soil and water assessment Tool)** | **SPARROW** |
|---|---|---|---|---|---|---|---|
| Purpose | Analyze nutrient flows from land to sea as affected by human activities | Analyze nutrient flows from land to sea as affected by human activities | Analyze nutrient flow from land to seas | Computes flow and quality | Analyze critical thresholds of N and P and spatial ecosystem bifurcation analysis in combination with Duflow | Analyze the quality and quantity of surface and ground waters as affected by the land management practices | Analyze nutrient, pesticide, |
| The main approach of nutrient cycling | Steady state | Steady-state | Steady-state | Unsteady** | Steady-state | Dynamic | Steady-state |
| Nutrient forms | TN, TP.DIP,DIN,DOP, DON,PN,PP,DSI | DIP, DIN,DON,DOP | DIN | TN,TP,DIN,DIP | TN, TP,DIP,DIN | TN, TP | TN, TP |
| Temporaal sales | Annual | Annual | Seasonal | Flexibel | Flexibel | Inter-annual | Annual |
| Spatial scales | Basin | Sub-basins | Basins | Basins, lake | Lake | Basins | WaterSheds (basin or sub-basin) |
| Sources of nutrients | Diffuse and Point source | Diffuse and Point | Diffuse and Point | Diffuse and Point | Diffuse and Point | Diffuse and Point | Diffuse and Point |
| Complexity | Moderate | Simple | Simple | Moderate | Moderate | Complex | Moderate |
| Data requirement | Extensive | Moderate | Moderate | Moderate | Moderate | Extensive | Extensive |
| References | (Mayorga et al., 2010) | (Strokal et al., 2016) | (McCrackin et al., 2014) | (Clemmens et al., 1993) | (Janse et al., 2010) | (Douglas-Mankin et al., 2010) | (Schwarz et al., 2006) |

*TN= total nitrogen. TP= total phosphorus. DIN= dissolved inorganic nitrogen. DON= dissolved organic nitrogen. DIP= dissolved inorganic phosphorus. DOP= dissolved organic phosphorus, PN= particulate nitrogen, PP= particulate phosphorous, DSI= dissolved silicon, ** A flow in which quantity of liquid flowing per second is not constant

## 6.5 Methods for Assessing Faecal Pollution in Lakes

My work is mainly limited to the performance evaluation of standard faecal indicator bacteria (SFIB) such as *E. coli, C. perfringens*, Enterococci, and qPCR assays across altitude ranges and different water sources. This is done by measuring the bacteria concentrations at different water bodies. However, it is still unfeasible to experimentally monitor their levels at the high spatiotemporal resolution often needed in real applications. Therefore, in addition to the determination of the concentration of SFIB, performance evaluation of SFIB and qPCR assays focusing on this limited number of indicator organisms, I combined direct SFIB measurements with the use of models. In this work, I combined the seasonal Marina model outputs such as $\mathrm{Fews}_{\mathrm{g.j.S}}$ which is the N watershed export fraction of gauged sub-basin j and season S ($S^{-1}$), and $\mathrm{Feriv}_{\mathrm{ung.j.S}}$ which is the DIN river export fraction of gauged sub-basin j and season S ($S^{-1}$) as a surrogate parameter to estimate the ninety percentile bacteria concentration for the un-gauged sub-basins, I used the approach of Yarahmadi (2003 and Hofstra and Vermeulen (2016 ) but modified it as described in detail in the annex of chapter 5.

Faecal coliforms such as *E. coli, C. perfringens*, Enterococci, genetic faecal marker (BacR = ruminant-associated Faecal pollution, PigIIBac = pig-associated, BachHUM or HF 183 Taqman = human-associated Faecal pollution) are used to indicate the quality of water in temperate regions (APHA, 1995; Toranzos and McFeters, 1997; Reischer et al.,2013). For a long time, faecal coliforms have been used as a standard indicator of recent faecal pollution and are acceptable under most conditions in temperate freshwater. These microbial parameters of water quality have been used in tropical countries too. The maximum allowable contaminant levels (MCL) established for many temperate areas have been accepted without question by tropical nations through source water quality in most tropical areas differs from that of temperate areas in three major ways: 1) by physical and chemical 2) by biological and 3) by social and economic factors (Hazen and Toranzos, 1990).

Despite this trend, there is a growing body of evidence that the underlying assumptions of the assays being used are not valid in the tropical climate. Studies in tropical freshwater have shown that a high proportion of faecal Coliforms-positive isolates may be of non-Faecal origin. Furthermore, some studies have reported that *E. coli* can become a normal inhabitant of tropical waters as reported for pristine environments in some tropical waters. Studies report that traditional faecal indicators are not performing well in tropical conditions and this apparent untrustworthiness of traditional Faecal pollution indicators should lead to an immediate development of alternative suggestions for pollution indicators.

Standardized laboratory procedures were followed for presumptive detection, enumeration, and confirmatory tests for faecal pollution indicator bacteria (E.coli: ISO 16649 2; intestinal Enterococci: ISO 7899-2; CP-spores, ISO 14189 and genetic markers as per the approach of Reischer etal.,2013 ). In addition to the traditional culture-based detection methods, modern molecular microbiological techniques like fluorescent antibody staining techniques, DNA probes, and genetic ribosomal 16sDNA characterization were done. For biochemical characterization for instance selected presumptive *E. coli and C. perfringens isolates*, API 20 E test strips, and Rapid id32(BioMerieux Sa, Marcy I' Etoile, France) were used respectively. Advanced molecular analysis of the faecal samples collected for genetic markers was done at Biotechnology and Molecular ecology laboratory, Technical University of Vienna, Austria.

There is a limitation in the application of standard faecal indicators and qPCR assays in the Lake Tana basin in a more spatial and temporal resolution. The SFIB, as well as the qPCR assay, should have been evaluated at different altitudes in Ethiopia. However, the sampling sites for the lowland were not having an altitude far below 1000-m asl. The study sites that represent low land are having an altitude close to 1000m asl. Our study did not get a good site from the low altitudes. The spatial representation of the evaluation of the indicators was not tested in a clear altitude category.

Only BacR assay was evaluated in three altitude categories and source waters, BacR was selected as the nitrogen export study showed that animal manure is the dominant source of Nitrogen. Further research on the evaluation of other qPCR assays should be conducted. In some of the qPCR analyses, the extraction efficiency was reported to be low and the possible explanation for this could be some of the samples are turbid and this turbidity may interfere with the DNA extraction efficiency.

The purposes and the standard laboratory procedures of the indicator bacteria is summarized and presented below in *Table 6-2*.

*Table 6-2.*Faecal pollution indication purpose of Faecal Pollution Indicators (FIB) and qPCR assays used in the thesis

| **Indicators** | **TC** | **FC** | **EC** | **IEC** | **CP** | **BACR** | **PigIIBac** | **BachHUM or HF 183 Taqman** |
|---|---|---|---|---|---|---|---|---|
| Purpose | Indicates Faecal pollution of animal and human origin<br>Indicates bacterial treatment efficiency in water treatment plants (drinking water) | A good Faecal indicator of the high probability of Faecal pollution in many environmental waters | Indicates recent Faecal contamination and the possible presence of intestinal pathogens | Indicates Faecal pollution | Indicates remote contamination with Faecal matter<br>Their presence in finished waters, therefore, suggests deficiencies in treatment filtration processes | Indicates the source of ruminant-associated Faecal pollution | Indicates the source of pig-associated Faecal pollution | Indicates human-associated Faecal pollution |
| Standardized laboratory procedures | ISO 9308-1:2014 | ISO, 1990b | ISO 16649 2 | ISO 7899-2 | ISO 14189 | Reischer etal.,2013 | Reischer etal.,2013 | Reischer etal.,2013 |
| TC= total coliforms; FC=Faecal coliforms; EC-E.coli; IEC= Intestinal enterococci; CP=Clostridium perfringens; BacR, a ruminant-associated Faecal marker; Pig-2-Bac, a pig-associated Faecal marker, and Bach HUM or HF 183 Taqman, a human-associated Faecal marker | | | | | | | | |

## 6.6 Lessons and Future Outlook

### 6.1.3 Lessons for Modelling Nutrient Losses to the Lake

Based on my study, I draw the following main lessons.

***Lesson 1: Downscaling annual inputs of nutrients to seasonal inputs supports a better understanding of temporal trends in river exports of nutrients to Lake Tana***

In this thesis, I developed a seasonal MARINA-Lake model to analyze seasonal patterns of river export of nutrients to Lake Tana (Chapter 2). Hydrology is an important factor affecting N transport from land to rivers (Chapter 3, Chen et al., 2019a; McCrackin et al., 2014). The seasonal MARINA model accounts for this factor by applying seasonal runoff and river discharges in modelling nutrient exports. This model accounts for seasonal variability in human activities (e.g., agriculture), natural processes (e.g., atmospheric N deposition, biological N2 fixation), and climate (e.g., air temperature). The net effect of anthropogenic factors, natural processes, climate, and hydrology provides new insights into the temporal trends in river exports of nutrients to Lake Tana especially for the regions with intensive agriculture and low retention during certain seasons. I applied the seasonal MARINA model to the 20 sub-basins of Lake Tana as illustrative example. I learned that downscaling annual inputs to seasons can help us to better understand the temporal trends of river export of nutrients to lake. This seasonal model can also be applied to other lakes in a similar condition experiencing nutrient pollution problems and lacking an understanding of the relevant temporal trends.

***Lesson 2: PCLake+ model development should consider the parameter setting when applied to a shallow, non-stratified lake in a tropical environment.***

The PCLake+ is developed for shallow, non-stratified, and temperate lakes. The PCLake+ model was used in my study area to simulate ecological conditions in Lake Tana basin, Ethiopia which is a tropical lake. The model generally performs well, however, parametrization of the default settings to the tropical environment may improve the model performance.

***Lesson 3: Linking MARINA and PCLake+ helps to better understand nitrogen pollution in Lake Tana Basin.***

I linked the seasonal MARINA and PCLake+ models to better understand the causes and effects of N pollution in the Lake Tana basin and Lake Tana (Chapters 3, 4, and 5). The seasonal MARINA model is a transparent model that provides temporally and spatially explicit outputs. These results help identify the causes of nutrient export to lakes and their origin (i.e., sub-basin) over time (Chapters 3,4,5). The PCLake+ model is known for its ability to estimate critical nutrient loads of lakes, which can help to assess the effects of nutrient exports from land on the water quality and ecosystem state of lakes (Chapter 5). Linking the seasonal MARINA model and PCLake+ models is also possible for other lakes in Ethiopia, especially in an area where there is a critical shortage of data.

***Lesson 4: A multi-model approach helps to better understand the spatial heterogeneity of eutrophication in large lakes.***

First, I used Duflow (2D, tracer modelling) to identify the so-called impact zones. Then I linked Duflow and PCLake+ to do bifurcation analysis and simulate the ecological condition of the Lake Tana. So, linking of Duflow and PC Lake + in other water bodies of Lake Tana and tropical lakes is important to define impact zones, analyze bifurcation and simulate ecological conditions of the Lakes and set critical thresholds of the N and P.

The spatial heterogeneity can be large and should be taken into account in (ecological) modelling. Lake Tana is the largest lake in Ethiopia that covers an area of about 3100 (km2), and the tradition of eutrophication management in the lake assumes that Lake Tana is a homogeneous system, even though one can see differences in the different parts of the lake, reflected by differences in turbidity and vegetation coverage. In this thesis, I applied a novel method of integrating Duflow and

PCLake+ and I verified that Lake Tana is a heterogeneous system, with spatial heterogeneity of nutrient loading and ecological conditions.

***Lesson 5: Faecal indicators developed for temperate areas cannot be used as they are, but need to be evaluated for tropical conditions before use***

The use of faecal coliforms *E. coli, C. perfringens,* Enterococci, and others as indicator bacteria for the assessment of faecal pollution and possible water quality deterioration in various freshwater sources is a widely accepted concept in temperate regions. However, these indicators have different discrimination efficacy when they are applied in tropical areas. I learned that discrimination efficacy is different among indicators and within one indicator, altitude and source water also affect their performance.

### 6.1.4 Outlook

***-Research*** [2].

For future research, I suggest three main directions. First, I propose a new project to monitor concentrations of N in rivers draining into the lakes. This helps to build trust and apply the model to other areas. Second, I suggest conducting sensitivity and uncertainty analyses to increase trust in the model performance especially for those years for which observations are limited. Third, I propose applying our model to sub-basins and rivers of other lakes in Ethiopia for effective nutrient management. This includes also the application of the model to the future, options to reduce lake pollution can be explored. Our model can also be applied to other sub-basins and rivers of tropical lakes that have comparable characteristics of the sub-basins and rivers of the Lake Tana Basin

[2] This part is partly based on Goshu G, Aynalem S, Damtie B, Stave K. Research Needs in the Lake Tana Basin Social-Ecological System. In Social and Ecological System Dynamics 2017 (pp. 631-646). Springer, Cham

(e.g., precipitation, seasons). Regarding the further evaluation of qPCR assays, more performance evaluation studies should be conducted in different agro ecologies of Ethiopia.

***- Recommendations for science aiming to improve model approaches***

Based on my results, the following recommendations can be made for future modelling studies. To model nutrient exports to lakes by the seasonal MARINA model (i.e., the basin-scale nutrient export model), the following steps are recommended. First, nutrient retention on wetlands and floodplains should be considered. Wetlands and flood plains in Lake Tana basin cover a substantial area and they are also important hotspots of biodiversity and Lake Tana is identified as important bird area because of the seasonal aggregation of birds and migratory species that reaches 217 species. The nutrient and bacteria input from birds is substantial but unfortunately this input was not taken into account in the previous modelling works. The model can be improved by including specific locations of temporary bird areas, livestock farms and by adding missing sources such as local industries that are not included in the current seasonal MARINA model. Second, the model does not explicitly account for nutrient pollution in groundwater in Ethiopia. Nitrate concentrations in groundwater frequently exceed the World Health Organization quality standard for drinking water of 50 mg $L^{-1}$. This happens especially in the flood plain areas and intensively managed agricultural regions in Eastern Lake Tana. It is important to analyze the causes of nutrient pollution in groundwater and to explore mitigation solutions by expanding the seasonal MARINA model with a ground-water module. Third, scenario analyses need to consider not only the reduction of nutrient pollution but also the social and economic feasibility when exploring possible options in future studies. In future studies, this could be done via participatory approaches that involve stakeholders in scenario development. Finally, to better validate the models, more experimental data is needed. I recommend that water quality be systematically monitored in rivers and lakes water.

To estimate critical nutrient loads by PCLake+ (i.e., the aquatic ecosystem model), the following steps are recommended. Firstly, the definition of the impact zones by the tributary rivers over the lake should be done, then the estimation of critical nutrient loads needs to account for spatial heterogeneity within the lake water and temporal variation of nutrient exports to the lake. Heterogeneity within lake water could lead to flexible load-response curves and consequently critical nutrient loads could vary throughout the lake (Janssen et al., 2017). The critical nutrient load may change across seasons due to the seasonal patterns of nutrient exports to the lake. Secondly, the future changes of climate and N/P mass ratio (gN/gP) could be accounted for calculating critical nutrient loads. Climate change can result in heat uptake by natural lakes (Vanderkelen et al., 2020) and observed surface temperatures of lakes have increased rapidly in recent decades (O'Reilly et al., 2015; Schneider and Hook, 2010). Therefore, the critical nutrient loads could decrease due to climate change (Mooij et al., 2007). Moreover, critical nutrient loads have been found to decrease with the increasing N/P mass ratio during oligo-eutrophication (Chang et al., 2020).

**- *Recommendations for improving nutrient management***

Clean water is necessary for human activities and for protecting biodiversity. It is also connected with one of the Sustainable Development Goals (SDGs) of the United Nations (UN) for water quality (UN, 2015a). It is SDG 6 “Clean water and sanitation”. In recent years, the Ethiopian government has made major efforts to keep water quality high. My thesis indicates that animal manure and synthetic fertilizers are the main causes of increased nutrient exports from land to lakes (Chapters 3). Therefore, Ethiopia should introduce a policy of “strengthening prevention and control of N and P pollution from point and diffuse sources”

In Chapter 3, river export of N to Lake Tana was highest in the rainy season and lowest in the pre-rainy season. So, effectively retaining the nutrient in the sub-watershed is thus needed to reduce lake pollution. In Chapters 3, river exports of nutrients to Lake Tana were estimated by sub-basin and by source. This helps water managers to locate key periods, areas, and sources, thereby implementing effective measures for future basin management.

In addition, the focus of pollution control in Ethiopia should shift from water quality control exclusively to include maintaining ecosystem health. Chapter 4 of this thesis is relevant in this respect. Chapter 4 identified critical nutrient loads by applying a complex ecological dynamic model (i.e., PCLake+) and integrating with Duflow model. Critical nutrient loads can be used as an indicator representing the ecosystem state of lakes, as it shows whether the lakes receive excessive nutrient loads that can cause eutrophication or the shift of ecosystem states (i.e., clear or turbid). By comparing the nutrient loads with the critical nutrient loads (derived from the PCLake+ model), the possible nutrient management in sub-basins of the lake was discussed. This linking model serves as a starting point to explore nutrient management strategies for sub-basins aiming to protect lake ecosystem health in the future. For future water resource development in the Lake Tana basin, faecal pollution risk mapping supports provides important information on where and when to give priority in Lake Tana Basin, from this end, Chapter 5 of the thesis is very important.

# Summary

The Ethiopian economy is projected to grow in the future (WEO, 2020). Agriculture and industry are the major sectors that contribute to the national economy. Agriculture supports more than 80 % of the national population but contribute to nutrient and faecal pollution in large lakes. The Lake Tana Basin is one of the main economic corridors of Ethiopia. This basin is not only an agricultural corridor, but it also hosts unique and fragile aquatic, wetland and flood plain ecosystems that provide social, economic, ecological and hydrological roles. Unfortunately, the development and consumption patterns in the Lake Tana basin outweighs the conservation and protection efforts and basin degradation is manifested in different forms such as invasive alien species and disease outbreaks. Nutrient and faecal pollution is increasing in Lake Tana. Nutrients such as nitrogen (N) and phosphorus (P) cause eutrophication issues in the lake.

Therefore, the overall research objective of my PhD study is to understand better river export of nutrients to Lake Tana by sub-basin, season, and source and to analyse the importance of the spatial heterogeneity of eutrophication, the performance evaluation of faecal pollution indicators, and the risk mapping of faecal pollution.

The four sub-objectives were formulated and addressed in the PhD thesis:

1. Perform a systematic overview of the water quality status of the Lake Tana basin through a literature review (Chapter 2).

2. Develop and apply a basin model for quantifying river export of N and P to Lake Tana by sub-basin, season, and source (Chapter 3).

3. Assess the spatial variability of eutrophication in tropical shallow lakes, taking Lake Tana as a case study and determining critical loads of N and P for different zones of the lake (Chapter 4).

4. Evaluate the performance of rapid and practical techniques to monitor faecal pollution in Lake Tana by testing them on a pollution gradient in low, mid-, and high altitudinal environments and water types (Chapter 5).

In **Chapter 2**, I justified my research objectives with a systematic overview of the available water quality data in the Lake Tana Basin. I presented my previous research works of water quality conducted at different times in the Lake Tana Basin. I identified knowledge gaps for Lake Tana associated with eutrophication and faecal pollution, and its sustainable management in the basin.

In **Chapter 3**, I focused on nutrient pollution issues in Lake Tana. I developed a seasonal version of the MARINA (water quality) model for Lake Tana. I did this by integrating the sub-basin modelling approach of MARINA with the seasonal modelling approach of NEWS 2-DIN-S (nutrient export model). My model quantifies seasonal river export of DIN (dissolved inorganic nitrogen) to Lake Tana by source and from twenty sub-basins. I made the four main modifications in my model compared to the existing MARINA model. First, I defined seasons according to the rainfall pattern. Second, I used mainly local sources of data for model inputs to represent the local situations of the Lake Tana basin. Third, I added open defecation as a new source of DIN pollution in the rivers of Lake Tana. This is a diffuse source because people defecate their waste on land and not directly into the water. Fourth, I recalibrated the model parameters influencing seasonality in the Lake Tana region. My model quantifies seasonal river export of DIN in three steps. First, inputs of N from diffuse sources to land were estimated, and inputs of N from point sources to rivers were also estimated. Second, the river export of DIN to the outlet of each sub-basin (river) was estimated. Third, the river export of DIN from sub-basin outlets to river mouths (the point at which DIN was discharged into Lake Tana) was estimated by quantifying retentions and losses of DIN within the river network.

In **Chapter 4**, I introduced a novel method that helps to understand the importance of spatial heterogeneity to explore the effects of nutrient loading on tropical shallow Lake, Lake Tana. It also helps to set critical loads of N and P. I linked the seasonal MARINA model from Chapter 3 by using the outputs of this model as an input to the new other two approaches. As a result, I combined three modelling approaches: Duflow, PCLake+, and seasonal MARINA model. I used a 2D dimensional flow and quality model, Duflow, to perform unsteady flow computations and to simulate the transportation of N and P in the Lake Tana ecosystem. I coupled the zero-dimensional PCLake+ model as a quality model to Duflow to simulate more complex water quality processes in Lake Tana. I used the monthly estimated N and P load from the MARINA model from Chapter 3.

In **Chapter 5**, I addressed faecal pollution and risk mapping in the Lake Tana basin (both ungauged and gauged sub-basins). I evaluated the performance of Faecal Indicator Bacteria (FIB) including total coliforms (TC), presumptive *E.coli* (EC), intestinal enterococci (IEC), and presumptive *Clostridia perfringens* spores (CP). I determined ruminant-associated faecal pollution using qPCR assay (BacR) at different water types located at different altitudes (1100 - 3835 m.a.s.l.) in a highland tropical country**.** Based on the results of the FIB evaluation and enumeration, a map was created showing the degree of pollution of sub-basins of the Lake Tana basin. To estimate the bacterial concentration especially for the ungauged basin, I based the seasonal DIN export model I developed (Chapter 3). DIN export fractions estimated by the seasonal MARINA model (Chapter 3) were used along with the population without latrine and animal densities data.

In **Chapter 6**, I discussed the outcomes of my thesis, presented main findings and outlook. I also formulated the main lessons that I learned in my study.

My study showed that nutrient and faecal pollution of water systems within the Lake Tana basin (especially river mouths and shore areas) have been increasing and this is caused by point and diffuse sources. The main findings of the PhD thesis per chapter are as follows:

**Main findings from Chapter 2:**

- Water quality information about Ethiopian water bodies is rare and the available ones are expeditious origin.
- The water quality data of surface, as well as groundwaters, is neither detailed nor systematic.
- There is no long-term and spatially representative data due to the absence of a sustainable monitoring program.
- The review indicates that the trophic status of Lake Tana has gradually changed from oligotrophic to mesotrophic due to nutrient loads.
- The review identified that sedimentation, eutrophication, faecal pollution, wetland encroachment, and hydrological alterations are the main issues of water quality management in Lake Tana basin.

**Main findings from Chapter 3:**

- We integrated existing models to capture seasonality in river export of nitrogen to tropical lakes.
- The model was applied to a representative tropical lake: Lake Tana.
- We modelled nutrient exports to Lake Tana, showing good agreement with measured loads.
- We found that river export of nitrogen to Lake Tana is the highest in rainy and the lowest in dry seasons.
- We found that animal manure is the dominant source of nitrogen in rivers in all seasons.

**Main findings from Chapter 4:**

- We defined impact zones by tributary rivers, using 2D tracer modelling
- We showed the importance of spatial heterogeneity of eutrophication in a shallow, non-stratifying lake
- Spatial variability of eutrophication is reflected in the critical loadings
- The critical N and P loads for each impact zone was defined
- Spatial heterogeneity can be large and should be taken into account when doing (ecological) modelling

**Main findings from Chapter 5:**

- E.coli and C. perfringens are identified as "consensus picture" of faecal pollution in tropical water.
- The performance of faecal indicator bacteria (FIB) varies with altitude and source water.
- Markers associated with ruminants (BacR) are identified in more than three-fourths of the sites.
- Moderate to high levels of faecal pollution were found in most sub-basins and the highest levels were in the rainy season.
- A bacterial pollution risk map is developed for sub-basins of Lake Tana including the un-gauged sub-basins.

I identified five main lessons that are learned in the course of the Ph.D. study. First, downscaling annual inputs of nutrients to seasonal inputs supports a better understanding of temporal trends in river exports of nutrients to Lake Tana. Second, PCLake+ model development should consider the parameter setting when applied to a shallow, non-stratified lake in a tropical environment.

Third, linking MARINA and PCLake+ helps to better understand nitrogen pollution in Lake Tana Basin. Fourth, a multi-model approach helps to better understand the spatial heterogeneity of eutrophication in large lakes. Fifth, faecal indicators developed for temperate areas cannot be used as they are but need to be evaluated for tropical conditions before use.

Further research outlooks were highlighted based on Goshu G, Aynalem S, Damtie B, Stave K. Research Needs in the Lake Tana Basin Social-Ecological System. In Social and Ecological System Dynamics 2017 (pp. 631-646). Springer, Cham.

The scientific significance is that I increase our understanding on the total N and P inputs, outputs, retentions, and the factors that control patterns of nutrient retentions in the Lake Tana basin. I show the critical N, P loads that trigger a trophic status change. I provide model tools that are useful to analyze seasonal nutrient loads and ecosystem effects. The application of the proposed linked seasonal MARINA -PCLake+ model approach is expected to set a new standard for eutrophication assessment in tropical areas. This study provides novel insights on the discrimination efficacy of the faecal pollution assays including performance characteristics of qPCR assays for microbial source tracking (MST) and risk categories of the water sources. Finally, the new findings concerning the main focal points of this thesis project plan, i.e., nutrient impacts and faecal pollution will be integrated into a synthesis providing recommendations for water quality management of Lake Tana on the system scale.

The societal relevance is that my outcomes can contribute to develop sustainable development strategies and support country policies and programs to address the issue of nutrient, faecal pollution, and improve water quality management in Ethiopia. This Ph.D. project aimed to provide scientifically based knowledge about the spatial and temporal variations of water quality and options of sustainable management. A workshop to present the scientific result and training on the new tools and methods was organized for appropriate stakeholders in collaboration with the responsible government institutions. The scientific results have been disseminated to users in the form of manual, proceedings, articles and policy briefs.

My Ph.D. thesis reveals innovative insights for effective environmental policies in Ethiopia. It shows the importance of manure, synthetic fertilizer, and human waste from diffuse sources in water pollution by nutrients and faecal matter. Managing these sources will likely reduce eutrophication and faecal pollution in the future. Furthermore, the implementation of advanced technologies is essential when dealing with urban pollution. Hopefully, my Ph.D. thesis is also useful for other tropical regions that have similar environmental problems to Lake Tana. The new seasonal MARINA model is rather simple and thus can be applied to other large, data-poor basins that may benefit from the allocation of effective management options. With this, I hope to contribute to increasing the availability of clean water and environment for the current and next generations, not only in Ethiopia but also in other world regions.

# List of publications

(During the study period)

## 2022

**Goraw Goshu**, Micha Veenendaal , J.J.M de Klein. Importance of spatial heterogeneity to explore effects of nutrient loading on the Ecological Status of Lake Tana. Accepted to Journal of Hydrology,

## 2021

**Goshu G**, Koelmans AA, de Klein JJ. Performance of faecal indicator bacteria, microbial source tracking, and pollution risk mapping in tropical water. Environmental Pollution. 2021 May 1;276:116693

## 2020

**Goshu G,** Strokal M, Kroeze C, Koelmans AA, de Klein JJ. Assessing seasonal nitrogen export to large tropical lakes. Science of the Total Environment. 2020 May 6:139199.

## 2017

**Goshu G**, Koelmans AA, de Klein JJ. Water quality of Lake Tana basin, Upper Blue Nile, Ethiopia. A review of available data. In Social and ecological system dynamics 2017 (pp. 127-141). Springer, Cham.

**Goshu, G**., Aynalem, S., Damtie, B. and Stave, K., 2017. Research Needs in the Lake Tana Basin Social-Ecological System. In Social and Ecological System Dynamics (pp. 631-646). Springer, Cham.

**Goshu G**, Aynalem S. Problem overview of the Lake Tana basin. In Social and Ecological System Dynamics 2017 (pp. 9-23). Springer, Cham.

Stave K, **Goshu G**, Aynalem S. Social and Ecological System Dynamics. Springer Verlag; 2017.

Mengistu AA, Aragaw C, Mengist M, **Goshu G**. The fish and the fisheries of Lake Tana. In Social and ecological system dynamics 2017 (pp. 157-177). Springer, Cham.

Aynalem S, **Goshu G**, Wondie A. Wetlands of the Lake Tana Watershed. In Social and Ecological System Dynamics 2017 (pp. 245-256). Springer, Cham.

Stave, K., Kopainsky, B., Anteneh, M., Mengistu, A. A., Sebhat, M. Y., Aynalem, S., ... **Goshu G** , & Beyene, B. S. (2017). Participatory System Dynamics Mapping for Collaboration and Socioecological Integration in the Lake Tana Region. In Social and Ecological System Dynamics (pp. 615-630). Springer, Cham.

# References

Abdelzaher AM, Wright ME, Ortega C, Solo-Gabriele HM, Miller G, Elmir S, et al. Presence of pathogens and indicator microbes at a non-point source subtropical recreational marine beach. Appl. Environ. Microbiol. 2010; 76: 724-732.

Abebe WB, Minale AS. Land Use and Watershed Management Practices in Lake Tana Basin. Social and Ecological System Dynamics. Springer, 2017, pp. 479-521.

Abera B, Bezabih B, Hailu D. Microbial quality of community drinking water supplies: A ten year (2004–2014) analyses in west Amhara, Ethiopia. Sustainability of Water Quality and Ecology 2017; 9: 22-26.

Abera B, Kibret M, Goshu G, Yimer M. Bacterial quality of drinking water sources and antimicrobial resistance profile of Enterobacteriaceae in Bahir Dar city, Ethiopia. Journal of water, sanitation and hygiene for development 2014; 4: 384-390.

Ahmed W, Goonetilleke A, Gardner T. Alternative indicators for detection and quantification of faecal pollution. Water 2008; 39: 46-49.

Alcamo J, Flörke M, Märker M. Future long-term changes in global water resources driven by socio-economic and climatic changes. Hydrological Sciences Journal 2007; 52: 247-275.

Alemayehu K, Tassew A. Challenges and Opportunities for Increased Farm Animal Productivity in the Lake Tana Sub-Basin. Social and Ecological System Dynamics. Springer, 2017, pp. 399-415.

Anderson A, Pietsch K, Zucker R, Mayr A, Müller-Hohe E, Messelhäusser U, et al. Validation of a duplex real-time PCR for the detection of Salmonella spp. in different food products. Food Analytical Methods 2011; 4: 259-267.

Anteneh M. Demographic Characteristics of the Lake Tana Basin. Social and Ecological System Dynamics. Springer, 2017, pp. 283-292.

APHA-AWWA-WPCF. Standard methods for the examination of water and wastewater: APHA American Public Health Association, 1981.

Association APH. American Environment Federation 1995. Standard methods for the examination of water and wastewater. 19th ed. Washington, DC.

Berhanu B, Seleshi Y, Melesse AM. Surface water and groundwater resources of Ethiopia: potentials and challenges of water resources development. Nile River Basin. Springer, 2014, pp. 97-117.

Birhanu K, Alamirew T, Dinka MO, Ayalew S, Aklog D. Optimizing reservoir operation policy using chance constraint nonlinear programming for Koga irrigation Dam, Ethiopia. Water resources management 2014; 28: 4957-4970.

Briggs D. Environmental pollution and the global burden of disease. British medical bulletin 2003; 68: 1-24.

Byamukama D, Mach RL, Kansiime F, Manafi M, Farnleitner AH. Discrimination efficacy of fecal pollution detection in different aquatic habitats of a high-altitude tropical country, using presumptive coliforms, Escherichia coli, and Clostridium perfringens spores. Appl. Environ. Microbiol. 2005; 71: 65-71.

Byappanahalli M, Fujioka R. Evidence that tropical soil environment can support the growth of Escherichia coli. Water Science and Technology 1998; 38: 171.

Cabelli V. Clostridium perfringens as a water quality indicator. Bacterial Indicators/Health Hazards Associated with Water. ASTM International, 1977.

Carpenter SR, Caraco NF, Correll DL, Howarth RW, Sharpley AN, Smith VH. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological applications 1998; 8: 559-568.

Chase E, Hunting J, Staley C, Harwood V. Microbial source tracking to identify human and ruminant sources of faecal pollution in an ephemeral F lorida river. Journal of applied microbiology 2012; 113: 1396-1406.

Chen X, Strokal M, Kroeze C, Ma L, Shen Z, Wu J, et al. Seasonality in river export of nitrogen: a modelling approach for the Yangtze River. Science of the Total Environment 2019; 671: 1282-1292.
Chislock MF, Doster E, Zitomer RA, Wilson AE. Eutrophication: causes, consequences, and controls in aquatic ecosystems. Nature Education Knowledge 2013; 4: 10.
Clemmens A, Holly Jr F, Schuurmans W. Description and evaluation of program: Duflow. Journal of irrigation and drainage engineering 1993; 119: 724-734.
Conley DJ, Paerl HW, Howarth RW, Boesch DF, Seitzinger SP, Havens KE, et al. Controlling eutrophication: nitrogen and phosphorus. American Association for the Advancement of Science, 2009.
Csa C. The 2007 population and housing census of Ethiopia. CSA Addis Ababa 2007.
Cunha DGF, do Carmo Calijuri M, Lamparelli MC. A trophic state index for tropical/subtropical reservoirs (TSItsr). Ecological Engineering 2013; 60: 126-134.
de Klein JJ, Koelmans AA. Quantifying seasonal export and retention of nutrients in West European lowland rivers at catchment scale. Hydrological Processes 2011; 25: 2102-2111.
De Vries A, Klavers H. Riverine fluxes of pollutants: monitoring strategy first, calculation methods second. European Water Management 1994; 4: 12-17.
Dersseh MG, Kibret AA, Tilahun SA, Worqlul AW, Moges MA, Dagnew DC, et al. Potential of water hyacinth infestation on Lake Tana, ethiopia: A prediction using a gis-based multi-criteria technique. Water 2019; 11: 1921.
Dersseh MG, Tilahun SA, Worqlul AW, Moges MA, Abebe WB, Mhiret DA, et al. Spatial and temporal dynamics of water hyacinth and its linkage with lake-level fluctuation: Lake Tana, a sub-humid region of the Ethiopian highlands. Water 2020; 12: 1435.
Desmarais TR, Solo-Gabriele HM, Palmer CJ. Influence of soil on fecal indicator organisms in a tidally influenced subtropical environment. Applied and environmental microbiology 2002; 68: 1165-1172.
Dile YT, Berndtsson R, Setegn SG. Hydrological response to climate change for gilgel abay river, in the Lake Tana basin-upper blue Nile basin of Ethiopia. PloS one 2013; 8: e79296.
Dile YT, Karlberg L, Daggupati P, Srinivasan R, Wiberg D, Rockström J. Assessing the implications of water harvesting intensification on upstream–downstream ecosystem services: A case study in the Lake Tana basin. Science of the Total Environment 2016; 542: 22-35.
Ding J, Cao J, Xu Q, Xi B, Su J, Gao R, et al. Spatial heterogeneity of lake eutrophication caused by physiogeographic conditions: An analysis of 143 lakes in China. Journal of Environmental Sciences 2015; 30: 140-147.
Donat MG, Leckebusch GC, Wild S, Ulbrich U. Benefits and limitations of regional multi-model ensembles for storm loss estimations. Climate Research 2010; 44: 211-225.
Douglas-Mankin K, Srinivasan R, Arnold J. Soil and Water Assessment Tool (SWAT) model: Current developments and applications. Transactions of the ASABE 2010; 53: 1423-1431.
Edberg S, Rice E, Karlin R, Allen M. Escherichia coli: the best biological drinking water indicator for public health protection. Journal of applied microbiology 2000; 88: 106S-116S.
Espinosa AC, Arias CF, Sánchez-Colón S, Mazari-Hiriart M. Comparative study of enteric viruses, coliphages and indicator bacteria for evaluating water quality in a tropical high-altitude system. Environmental Health 2009; 8: 49.
Ewnetu DA, Bitew BD, Chercos DH. Determination of surface water quality status and identifying potential pollution sources of Lake Tana: particular emphasis on the Lake Boundary of Bahirdar City, Amhara Region, North West Ethiopia, 2013. 2014.
Fragoso Jr CR, Marques DMM, Collischonn W, Tucci CE, van Nes EH. Modelling spatial heterogeneity of phytoplankton in Lake Mangueira, a large shallow subtropical lake in South Brazil. Ecological modelling 2008; 219: 125-137.

Frumin G, Gildeeva I. Eutrophication of water bodies—A global environmental problem. Russian Journal of General Chemistry 2014; 84: 2483-2488.

Galloway JN, Dentener FJ, Capone DG, Boyer EW, Howarth RW, Seitzinger SP, et al. Nitrogen cycles: past, present, and future. Biogeochemistry 2004; 70: 153-226.

Gleick PH. Water in crisis: paths to sustainable water use. Ecological applications 1998; 8: 571-579.

Goshu G, Aynalem S. Problem Overview of the Lake Tana Basin. Social and Ecological System Dynamics. Springer, 2017, pp. 9-23.

Goshu G, Byamukama D, Manafi M, Kirschner AK, Farnleitner AH. A pilot study on anthropogenic faecal pollution impact in Bahir Dar Gulf of Lake Tana, Northern Ethiopia. Ecohydrology & Hydrobiology 2010a; 10: 271-279.

Goshu G, Byamukama D, Manafi M, Kirschner AK, Farnleitner AH. A pilot study on anthropogenic faecal pollution impact in Bahir Dar Gulf of Lake Tana, Northern Ethiopia. Ecohydrology & Hydrobiology 2010b; 10: 271-279.

Goshu G, Koelmans A, de Klein J. Water quality of Lake Tana basin, Upper Blue Nile, Ethiopia. A review of available data. Social and ecological system dynamics. Springer, 2017, pp. 127-141.

Goshu G, Koelmans A, de Klein J. Performance of faecal indicator bacteria, microbial source tracking, and pollution risk mapping in tropical water. Environmental Pollution 2021; 276: 116693.

Goshu G, Strokal M, Kroeze C, Koelmans A, de Klein J. Assessing seasonal nitrogen export to large tropical lakes. Science of the Total Environment 2020: 139199.

Goshu G, Tewabe D, Adugna BT. Spatial and temporal distribution of commercially important fish species of Lake Tana, Ethiopia. Ecohydrology & Hydrobiology 2010c; 10: 231-240.

Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ. Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA-and rRNA-based microbial community composition. Appl. Environ. Microbiol. 2000; 66: 5488-5491.

Groffman PM, Baron JS, Blett T, Gold AJ, Goodman I, Gunderson LH, et al. Ecological thresholds: the key to successful environmental management or an important concept with no practical application? Ecosystems 2006; 9: 1-13.

Haregeweyn N, Fikadu G, Tsunekawa A, Tsubo M, Meshesha DT. The dynamics of urban expansion and its impacts on land use/land cover change and small-scale farmers living near the urban fringe: A case study of Bahir Dar, Ethiopia. Landscape and urban planning 2012; 106: 149-157.

Hazen TC. Fecal coliforms as indicators in tropical waters: a review. Toxicity assessment 1988; 3: 461-477.

Hazen TC, Toranzos GA. Tropical source water. Drinking Water Microbiology. Springer, 1990, pp. 32-53.

Hecky R. The eutrophication of lake Victoria. Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen 1993; 25: 39-48.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 2005a; 25: 1965-1978.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology: A Journal of the Royal Meteorological Society 2005b; 25: 1965-1978.

Hofstra N. Quantifying the impact of climate change on enteric waterborne pathogen concentrations in surface water. Current Opinion in Environmental Sustainability 2011; 3: 471-479.

Hofstra N, Vermeulen LC. Impacts of population growth, urbanisation and sanitation changes on global human Cryptosporidium emissions to surface water. International journal of hygiene and environmental health 2016; 219: 599-605.

Houlton BZ, Wang Y-P, Vitousek PM, Field CB. A unifying framework for dinitrogen fixation in the terrestrial biosphere. Nature 2008; 454: 327.

Ivens WP, Tysmans DJ, Kroeze C, Löhr AJ, van Wijnen J. Modelling global N2O emissions from aquatic systems. Current opinion in environmental sustainability 2011; 3: 350-358.

Janse J, Scheffer M, Lijklema L, Van Liere L, Sloot J, Mooij W. Estimating the critical phosphorus loading of shallow lakes with the ecosystem model PCLake: sensitivity, calibration and uncertainty. Ecological Modelling 2010; 221: 654-665.

Janse JH. Model studies on the eutrophication of shallow lakes and ditches: Wageningen University and Research, 2005.

Janse JH, Domis LNDS, Scheffer M, Lijklema L, Van Liere L, Klinge M, et al. Critical phosphorus loading of different types of shallow lakes and the consequences for management estimated with the ecosystem model PCLake. Limnologica 2008; 38: 203-219.

Janssen AB, de Jager VC, Janse JH, Kong X, Liu S, Ye Q, et al. Spatial identification of critical nutrient loads of large shallow lakes: implications for Lake Taihu (China). Water research 2017; 119: 276-287.

Janssen AB, Teurlincx S, Beusen AH, Huijbregts MA, Rost J, Schipper AM, et al. PCLake+: A process-based ecological model to assess the trophic state of stratified and non-stratified freshwater lakes worldwide. Ecological modelling 2019a; 396: 23-32.

Janssen AB, van Wijk D, van Gerven LP, Bakker ES, Brederveld RJ, DeAngelis DL, et al. Success of lake restoration depends on spatial aspects of nutrient loading and hydrology. Science of the Total Environment 2019b; 679: 248-259.

Jeppesen E, Søndergaard M, Jensen JP, Havens KE, Anneville O, Carvalho L, et al. Lake responses to reduced nutrient loading–an analysis of contemporary long-term data from 35 case studies. Freshwater Biology 2005; 50: 1747-1771.

Kavka GG, Kasimir GD, Farnleitner AH. Microbiological water quality of the River Danube (km 2581-km 15): Longitudinal variation of pollution as determined by standard parameters: na, 2006.

Kebede G, Mushi D, Linke RB, Dereje O, Lakew A, Hayes DS, et al. Macroinvertebrate indices versus microbial fecal pollution characteristics for water quality monitoring reveals contrasting results for an Ethiopian river. Ecological Indicators 2020; 108: 105733.

Kosten S, Huszar VL, Mazzeo N, Scheffer M, da SL Sternberg L, Jeppesen E. Lake and watershed characteristics rather than climate influence nutrient limitation in shallow lakes. Ecological applications 2009; 19: 1791-1804.

Kroeze C, Bouwman L, Seitzinger S. Modelling global nutrient export from watersheds. Current Opinion in Environmental Sustainability 2012; 4: 195-202.

Langergraber G, Muellegger E. Ecological Sanitation—a way to solve global sanitation problems? Environment international 2005; 31: 433-444.

Layton A, McKay L, Williams D, Garrett V, Gentry R, Sayler G. Development of Bacteroides 16S rRNA gene TaqMan-based real-time PCR assays for estimation of total, human, and bovine fecal pollution in water. Appl. Environ. Microbiol. 2006; 72: 4214-4224.

Lesiv M, Fritz S, McCallum I, Tsendbazar N, Herold M, Pekel J-F, et al. Evaluation of ESA CCI prototype land cover map at 20m. 2017.

Lewis Jr WM. Tropical lakes: how latitude makes a difference. Perspectives in tropical limnology 1996; 4364.

Liu C, Watanabe M, Wang Q. Changes in nitrogen budgets and nitrogen use efficiency in the agroecosystems of the Changjiang River basin between 1980 and 2000. Nutrient Cycling in Agroecosystems 2008; 80: 19-37.

Liu W, Qiu R. Water eutrophication in China and the combating strategies. Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental & Clean Technology 2007; 82: 781-786.

Los FJ. Eco-hydrodynamic modelling of primary production in coastal waters and lakes using BLOOM. Vol 1: Ios Press, 2009.

Lürling M, van Oosterhout F. Controlling eutrophication by combined bloom precipitation and sediment phosphorus inactivation. Water research 2013; 47: 6527-6537.

Mamo S, Ayenew T, Berehanu B, Kebede S. Hydrology of the Lake Tana Basin, Ethiopia: Implication to Groundwater-Surface Waters Interaction. J. Environ. Earth Sci 2016; 6: 54-66.

Mayer RE, Reischer GH, Ixenmaier SK, Derx J, Blaschke AP, Ebdon JE, et al. Global distribution of human-associated fecal genetic markers in reference samples from six continents. Environmental science & technology 2018; 52: 5076-5084.

Mayorga E, Seitzinger SP, Harrison JA, Dumont E, Beusen AH, Bouwman A, et al. Global nutrient export from WaterSheds 2 (NEWS 2): model development and implementation. Environmental Modelling & Software 2010; 25: 837-853.

McCrackin ML, Harrison JA, Compton JE. Factors influencing export of dissolved inorganic nitrogen by major rivers: A new, seasonal, spatially explicit, global model. Global Biogeochemical Cycles 2014; 28: 269-285.

Medema G, Bahar M, Schets F. Survival of Cryptosporidium parvum, Escherichia coli, faecal enterococci and Clostridium perfringens in river water: influence of temperature and autochthonous microorganisms. Water Science and Technology 1997; 35: 249.

Mengesha A, Wubshet M, Gelaw B. A survey of bacteriological quality of drinking water in North Gondar. The Ethiopian Journal of Health Development (EJHD) 2004; 18.

Mineau MM, Wollheim WM, Stewart RJ. An index to characterize the spatial distribution of land use within watersheds and implications for river network nutrient removal and export. Geophysical Research Letters 2015; 42: 6688-6695.

Moges MA, Tilahun SA, Ayana EK, Moges MM, Gabye N, Giri S, et al. Non-point source pollution of dissolved phosphorus in the Ethiopian highlands: The Awramba watershed near Lake Tana. CLEAN–Soil, Air, Water 2016; 44: 703-709.

Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE 2007; 50: 885-900.

Nash JE, Sutcliffe JV. River flow forecasting through conceptual models part I—A discussion of principles. Journal of hydrology 1970; 10: 282-290.

Nigate F, Ayenew T, Belete W, Walraevens K. Overview of the Hydrogeology and Groundwater Occurrence in the Lake Tana Basin, Upper Blue Nile River Basin. Social and ecological system dynamics. Springer, 2017, pp. 77-91.

Odonkor ST, Ampofo JK. Escherichia coli as an indicator of bacteriological quality of water: an overview. Microbiology research 2013; 4: e2-e2.

Pardo LH, Fenn ME, Goodale CL, Geiser LH, Driscoll CT, Allen EB, et al. Effects of nitrogen deposition and empirical nitrogen critical loads for ecoregions of the United States. Ecological Applications 2011; 21: 3049-3082.

Pawar N, Pondhe G, Patil S. Groundwater pollution due to sugar-mill effluent, at Sonai, Maharashtra, India. Environmental Geology 1998; 34: 151-158.

Polo F, Figueras M, Inza I, Sala J, Fleisher JM, Guarro J. Relationship between presence of Salmonella and indicators of faecal pollution in aquatic habitats. FEMS Microbiology Letters 1998; 160: 253-256.

Pringle CM. Nutrient spatial heterogeneity: effects on community structure, physiognomy, and diversity of stream algae. Ecology 1990; 71: 905-920.

Rahm L, Danielsson Å. Spatial heterogeneity of nutrients in the Baltic Proper, Baltic Sea. Estuarine, Coastal and Shelf Science 2007; 73: 268-278.

Reischer G, Haider J, Sommer R, Stadler H, Keiblinger K, Hornek R, et al. Quantitative microbial faecal source tracking with sampling guided by hydrological catchment dynamics. Environmental microbiology 2008; 10: 2598-2608.

Reischer GH, Ebdon JE, Bauer JM, Schuster N, Ahmed W, Astrom J, et al. Performance characteristics of qPCR assays targeting human- and ruminant-associated bacteroidetes for microbial source tracking across sixteen countries on six continents. Environ Sci Technol 2013; 47: 8548-56.

Reischer GH, Kasper DC, Steinborn R, Mach RL, Farnleitner AH. Quantitative PCR method for sensitive detection of ruminant fecal pollution in freshwater and evaluation of this method in alpine karstic regions. Appl Environ Microbiol 2006; 72: 5610-4.

Reischer GH, Kollanur D, Vierheilig J, Wehrspaun C, Mach RL, Sommer R, et al. Hypothesis-driven approach for the identification of fecal pollution sources in water resources. Environmental science & technology 2011; 45: 4038-4045.

Ritter KS, Paul Sibley, Ken Hall, Patricia Keen, Gevan Mattu, Beth Linton, Len. Sources, pathways, and relative risks of contaminants in surface water and groundwater: a perspective prepared for the Walkerton inquiry. Journal of Toxicology and Environmental Health Part A 2002; 65: 1-142.

Rivera SC, Hazen TC, Toranzos GA. Isolation of fecal coliforms from pristine sites in a tropical rain forest. Appl. Environ. Microbiol. 1988; 54: 513-517.

Rychtecký P, Znachor P. Spatial heterogeneity and seasonal succession of phytoplankton along the longitudinal gradient in a eutrophic reservoir. Hydrobiologia 2011; 663: 175-186.

Sacks WJ, Deryng D, Foley JA, Ramankutty N. Crop planting dates: an analysis of global patterns. Global Ecology and Biogeography 2010; 19: 607-620.

Saloranta TM, Andersen T. MyLake—A multi-year lake simulation model code suitable for uncertainty and sensitivity analysis simulations. Ecological modelling 2007; 207: 45-60.

santé Omdl, Organization WH, programme W--W, Staff WHO, Zdrowia ŚO, WHO. Guidelines for drinking-water quality. Vol 1: World Health Organization, 2004.

Scheffer M, van Nes EH. Shallow lakes theory revisited: various alternative regimes driven by climate, nutrients, depth and lake size. Shallow lakes in a changing world. Springer, 2007, pp. 455-466.

Schwarz G, Hoos AB, Alexander R, Smith R. The SPARROW surface water-quality model: theory, application and user documentation. US geological survey techniques and methods report, book 2006; 6: 248.

Scott TM, Rose JB, Jenkins TM, Farrah SR, Lukasik J. Microbial source tracking: current methodology and future directions. Applied and environmental microbiology 2002; 68: 5796-5803.

Seitzinger SP, Mayorga E, Bouwman AF, Kroeze C, Beusen AHW, Billen G, et al. Global river nutrient export: A scenario analysis of past and future trends. Global Biogeochemical Cycles 2010; 24: n/a-n/a.

Selassie YG. Problems, Efforts and Future Directions of Natural Resources Management in Western Amhara Region of the Blue Nile Basin, Ethiopia. Social and Ecological System Dynamics. Springer, 2017, pp. 597-613.

Setegn SG, Srinivasan R, Dargahi B. Hydrological modelling in the Lake Tana Basin, Ethiopia using SWAT model. The Open Hydrology Journal 2008; 2.

Setegn SG, Srinivasan R, Dargahi B, Melesse AM. Spatial delineation of soil erosion vulnerability in the Lake Tana Basin, Ethiopia. Hydrological Processes: An International Journal 2009; 23: 3738-3750.

Sewnet A. Land Use/Cover Change at Infraz Watershed, Northwestren Ethiopia. Journal of Landscape Ecology 2015; 8: 69-83.

Simane B, Zaitchik BF, Ozdogan M. Agroecosystem analysis of the Choke Mountain watersheds, Ethiopia. Sustainability 2013; 5: 592-616.

Sinigalliano CD, Fleisher JM, Gidley ML, Solo-Gabriele HM, Shibata T, Plano LR, et al. Traditional and molecular analyses for fecal indicator bacteria in non-point source subtropical recreational marine waters. Water research 2010; 44: 3763-3772.

Sirak T. Watershed Modelling of Lake Tana Basin Using SWAT. Arbaminch, Ethiopia 2008.

Smil V. Nitrogen in crop production: An account of global flows. Global biogeochemical cycles 1999; 13: 647-662.

Smith VH. Eutrophication of freshwater and coastal marine ecosystems a global problem. Environmental Science and Pollution Research 2003; 10: 126-139.

Soares MCS, Marinho MM, Azevedo SM, Branco CW, Huszar VL. Eutrophication and retention time affecting spatial heterogeneity in a tropical reservoir. Limnologica 2012; 42: 197-203.

Stave K, Goshu G, Aynalem S. Social and Ecological System Dynamics: Springer, 2017.

Strokal M, Kroeze C, Wang M, Bai Z, Ma L. The MARINA model (Model to Assess River Inputs of Nutrients to seAs): Model description and results for China. Sci Total Environ 2016; 562: 869-888.

Tassew A, Seifu E. Production, handling, traditional processing practices and quality of milk in Bahir Dar milk shed Area, Ethiopia. Haramaya University, 2007.

Toranzos G, McFeters GA, Hurst C, Knudsen G, McInerney M, Stetzenbach L, et al. Detection of indicator microorganisms in environmental freshwaters and drinking waters. Manual of environmental microbiology. 1997: 184-194.

Van Drecht G, Bouwman A, Harrison J, Knoop J. Global nitrogen and phosphate in urban wastewater for the period 1970 to 2050. Global Biogeochemical Cycles 2009; 23.

Verhougstraete MP, Martin SL, Kendall AD, Hyndman DW, Rose JB. Linking fecal bacteria in rivers to landscape, geochemical, and hydrologic factors and sources at the basin scale. Proceedings of the National Academy of Sciences 2015; 112: 10419-10424.

Wale A, Rientjes T, Dost R, Gieske A. Hydrological Balance of Lake Tana Upper Blue Nile Basin, Ethiopia. ITC The Netherlands, 2008.

Wang M, Strokal M, Burek P, Kroeze C, Ma L, Janssen AB. Excess nutrient loads to Lake Taihu: Opportunities for nutrient reduction. Science of the Total Environment 2019; 664: 865-873.

Wang Y, Law R, Pak B. A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. Biogeosciences 2010; 7: 2261-2282.

Wetzel RG, Likens GE. Primary productivity of phytoplankton. Limnological Analyses. Springer, 1991, pp. 207-226.

Wondie A, Mengistu S, Vijverberg J, Dejen E. Seasonal variation in primary production of a large high altitude tropical lake (Lake Tana, Ethiopia): effects of nutrient availability and water transparency. Aquatic Ecology 2007; 41: 195-207.

Wondie A, Seid A, Molla E, Goshu G, Gkidan W, Shibabaw A, et al. Preliminary assessment of water hyacinth (Eichornia crassipes) in Lake Tana. Proceedings of the National Workshop (Biological Society of Ethiopia), Addis Ababa, Ethiopia, 2012.

Wondie TA. The impact of urban storm water runoff and domestic waste effluent on water quality of Lake Tana and local groundwater near the city of Bahir Dar, Ethiopia. Cornell University Ithaca, NY, 2009.

Wright J, Gundry S, Conroy R. Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. Tropical medicine & international health 2004; 9: 106-117.

Xu H, Paerl H, Qin B, Zhu G, Hall N, Wu Y. Determining critical nutrient thresholds needed to control harmful cyanobacterial blooms in eutrophic Lake Taihu, China. Environmental science & technology 2015; 49: 1051-1059.

Xu T, Yang T, Xiong M. Time scales of external loading and spatial heterogeneity in nutrients-chlorophyll a response: Implication on eutrophication control in a large shallow lake. Ecological Engineering 2020; 142: 105636.

Yan W, Mayorga E, Li X, Seitzinger SP, Bouwman A. Increasing anthropogenic nitrogen inputs and riverine DIN exports from the Changjiang River basin under changing human pressures. Global Biogeochemical Cycles 2010; 24.

Yang J, Strokal M, Kroeze C, Wang M, Wang J, Wu Y, et al. Nutrient losses to surface waters in Hai He basin: A case study of Guanting reservoir and Baiyangdian lake. Agricultural water management 2019; 213: 62-75.

Yarahmadi J. The integration of satellite images, GIS and CROPWAT model to investigation of water balance in irrigated area. ITC, 2003.

Yasin JA, Kroeze C, Mayorga E. Nutrients export by rivers to the coastal waters of Africa: Past and future trends. Global Biogeochemical Cycles 2010; 24: n/a-n/a.
Yitaferu B. Land degradation and options for sustainable land management in the Lake Tana Basin (LTB), Amhara Region, Ethiopia. University of Bern, 2007a.
Yitaferu B. Land Degradation and options for sustainable land management in the Lake Tana Basin (LTB), Amhara region, Ethiopia. 2007b.
Yu Y, Yan Q, Feng W. Spatiotemporal heterogeneity of plankton communities in Lake Donghu, China, as revealed by PCR–denaturing gradient gel electrophoresis and its relation to biotic and abiotic factors. FEMS Microbiology Ecology 2008; 63: 328-337.
Zhang W, Tian Z, Zhang N, Li X. Nitrate pollution of groundwater in northern China. Agriculture, Ecosystems & Environment 1996; 59: 223-231.
Zimmerman JB, Mihelcic JR, Smith, James. Global stressors on water quality and quantity. ACS Publications, 2008.

# Supplementary Materials

Supplementary materials contain additional information to the following chapters of the phD thesis

- Chapter 3 (Published as Goraw et al.,2020)
- Chapter 4 (is accepted as Goraw et al.,)
- Chapter 5 (Published as Goraw et al.,2021)

The text, figures and tables of the supplementary materials from published articles and the articles under review have been adjusted to the phD Theis format The adjusted phD version of the supplementary materials is available up on request (goraw.yemer@wur.nl). The published versions of the supplementary materials are available online with the published articles.

# Acknowledgments

First and foremost, I would like to thank God, the Almighty, for giving me the strength, patience and endurance to successfully complete my PhD. I would like to express my special gratitude and thanks to my promotors Prof. Dr. Koelman Bart and Prof. Dr. Carolien Kroeze. I have no words to express your(her) support. You have taught me a great deal of things, from critical thinking to careful writing of my papers. I really enjoyed the freedom you provided me to express my idea. Your critical constructive comments and feedbacks were a considerable help in refining my scientific skills. I am thankful for the opportunities you have provided me to develop myself in other aspects like supervisions, trainings and symposiums. It was a great privilege to work under your guidance. It would have been impossible for me to reach this stage without the unreserved support of my co-promotors. I would like to thank my co-promotor, Dr. JJM de Klein, from the aquatic ecology and water quality management group and Dr Maryna Strokal from the waters systems and Global change chair group. Jeroen, you helped me better understand the real science of modelling and nutrient management. I am grateful for your commitment to travel all the way to Ethiopia and visit my sampling sites, even at times of ongoing conflict and violence in Ethiopia. The discussions we held on to the sites and afterwards about my research progress, model validation and calibration techniques were very critical and essential as they gave me the spirit and encouragement to excel in my paper writing. I also wish to thank him for helping me get to know the Dutch scientific and social cultures. Thanks a lot for organizing the celebration of the publication of our paper 2. I would like to extend my sincere gratitude to Jeroen and his wife for their hospitality. I am also pleased to say thank you to my other co-promotor, Dr. Strokal Maryna, from the water systems and global change group, Wageningen university. I am extremely grateful that she helped me comprehend her model, MARINA model, Global NEWS and NEWS-DIN(S).

I managed to develop the seasonal version to the Lake Tana basin by integrating the two approaches. It was challenging for me to understand the model in a short span of time and apply it to the Lake Tana basin. I am extremely grateful for your precious support and insightful suggestions throughout my PhD research work. I would also like to thank Prof. Andreas F. and Dr. Rita L. from Technical University of Vienna, Austria for your hospitality in taking care of me and for organizing me a training on the Microbial Source Tracking (MST).

I thank very much my wife, Banchiamlak (Mastewal) Worku for your love and care you have provided to me during my PhD study. Maste it was impossible to finish my PhD without your great support. I understand that taking care of twins ( 06 months when I started my PhD), and managing the whole family is very tough. I never forget the hard times you have spent taking care of all the kids in my absence for the PhD. I really appreciate very much your commitment you have paid for the success of my PhD. You are a great gift of GOD to me !! I thank very much all my children Haileamlak G., Hana G., F/Markos G., Dawit G. Addis A., and Masresha A. for your motivation, care and love you have provided me during my PhD study. The phone conversions, daily family virtual prayer ceremony we had during the peak times of COVID 19 was very much inspiring and relived my stress, sitting alone and not doing much. I thank all the family members for your unreserved effort to organize that.

I sincerely thank and my brother Melkamu G. and my sisters, Wubete, Tirngo, Tarik, Amele, and Banchi Gizie, and my father- in- law Worku G/Selassie and my mother- in-law Melkmarim Walelgne for their invaluable emotional support, heart-warming kindness and encouragement. I duly acknowledge my brothers-in-law, Yonas A., Assefa M. and Melese S. and my nephews Yihnew D. and Biniyam G., who have shown me their love, care and kindness despite the long distance between us.

I also want to extend my appreciation to my colleagues for their unfailing encouragement, inspiration and intellectual guidance: Dr. Shimeles A., Wubneh B., Chalchew A., Lakachew Y., Dr. Mintesinot A., Teshome D., and Prof Yihnew G/Selassie at BDU and Yeshiwas, Dr Belay T., Dr. Sileshi G., Dr. Kalkidan, Dr. Lemessa, Mrs Hana R., Dr. Abebe C., Zerihun A,, Dr Samuael T., Markos, Penagos, Nancy, Frits, Annett and Chen at Wageningen University.

I want to acknowledge the financial support received from NUFFIC and BDU-RCS/BNWI. I am beyond grateful to Gerold W. for his financial support and encouragement throughout my PhD. Gerold words are not enough to express my thanks to you and your organization.

I greatly appreciate my field and lab assistants for their meaningful contribution towards my PhD: Tewodaj G., Masresha B., Hassen M., and Belay G. I am also grateful to lab analysts, Daniel and Birkutait, for their enormous support. And of course, my thanks go out to my drivers and boat-skippers who were helpful numerous times: Chale, Manale, Cherie, Wossen, Woretaw, Teshome and Mengie.

Now I am happy to say: My PhD thesis is finished, and new challenges are starting!

# About the author

Goraw Goshu was born in Bichena, Gojjam, Ethiopia. After high school, Goraw studied at Addis Ababa University (1993 - 1997) and received a BSc degree in Biology in 1997. After that, he studied his MSc at UNESCO-IHE (2005-2007) and received an MSc degree in Environmental Science in 2007, the Netherlands. During his master study, Goraw focused on environmental pollution, mainly evaluation of faecal pollution indicators, and water quality assessment of surface and ground water systems in rural and peri urban areas of Ethiopia. For this he did a lot of laboratory and field work. After his BSc graduation, Goraw has continued working as a teacher, agricultural expert and researcher for more than 10 years at different institutions such as bureaus of education, agriculture and Amhara Regional Agricultural Research Institute. In 2010, Goraw worked as Environmental/Water quality specialist for the World Bank financed project: -Tana Beles Integrated Water Resources Development Project. Since January 2011, Goraw has joined Bahir Dar University and he teaches, conducts research and community services. Goraw was founder of the Blue Nile Water Research Institute, and he directs the research institute for more than four years. In November 2015, he received a PhD position at the Aquatic Ecology and Water Quality Management group (AEW) group of Wageningen University & Research. He has also supervisors from the Water Systems and Global Change group of Wageningen University & Research. His PhD thesis is about analyzing water quality and its impacts towards a sustainable basin management: Modelling and experimental approaches as affected by human activities. He focuses on modelling nutrient and faecal pollution that causes surface and ground water problems. For this he developed a seasonal version of an integrated

*MARINA* model, in short for Model to Assess River Inputs of Nutrient to seAs. This model aims at identifying causes of nutrient pollution and exploring solutions to reduce this pollution in a spatially explicit way. He further models the importance of spatial heterogeneity to explore effects of nutrient loadings in a large tropical Lake, Lake Tana. Goraw also evaluates and applies new microbial source tracking (MST) bioassays to map the pollution risk map of surface and ground water resources. During his PhD period Goraw was active in education and supervision activities. Goraw contributed largely to extending the network with Austria on faecal pollution management. He has also attended and presented oral presentations at various international conferences and symposia and. Now, Goraw is looking forward to new challenges in his scientific career.

# Awards

- ✓ In 2005, Goraw received a NUFFIC scholarship for an MSc study at UNESCO-IHE in The Netherlands.
- ✓ In 2015, Goraw received a NUFFIC scholarship for an PhD study at Wageningen

University & Research in The Netherlands

# SENSE Diploma

Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment

D I P L O M A

*for specialised PhD training*

The Netherlands research school for the
Socio-Economic and Natural Sciences of the Environment
(SENSE) declares that

***Goraw Goshu Yemer***

born on 7th October 1974 in Bichena, Ethiopia

has successfully fulfilled all requirements of the
educational PhD programme of SENSE.

Wageningen, 18th October 2022

| | |
|---|---|
| Chair of the SENSE board | The SENSE Director |
| Prof. dr. Martin Wassen | Prof. Philipp Pattberg |

*The SENSE Research School has been accredited by the Royal Netherlands Academy of Arts and Sciences (KNAW)*

K O N I N K L I J K E N E D E R L A N D S E
A K A D E M I E V A N W E T E N S C H A P P E N

The SENSE Research School declares that **Goraw Goshu Yemer** has successfully fulfilled all requirements of the educational PhD programme of SENSE with a work load of 49.7 EC, including the following activities:

**SENSE PhD Courses**

- Environmental research in context (2016)
- Research in context activity: 'Initiating and organizing joint book inauguration ceremony on 'Social and ecological systems dynamics: characteristics, trends, and integration in the Lake Tana Basin, Ethiopia (2019)
- Model training for Scenario analysis (2018)
- Robustness of aquatic ecosystems in the face of climate change (2017)
- The art of modelling (2019)

**Other PhD and Advanced MSc Courses**

- Quantitative Microbial Source Tracking-an introduction, Technical University of Vienna (2016)
- Project and time management, Wageningen Graduate Schools (2016)
- Advanced GIS and RS, Bahir Dar University (2019)
- Introduction to Water, Sanitation, and Hygiene, UNICEF-online (2016)
- Water quality, Wageningen University (2016)

**Management and Didactic Skills Training**

- Editor of the book- Stave K, Goshu G, Aynalem S. Social and Ecological System Dynamics. Springer (2017)
- Organized symposium 'Nile Water Resource Development towards River-led Continental Integration'' (2017)
- Teaching in the MSc course 'Advanced Water Quality Management' (2017)
- Teaching in the MSc course 'Water and Waste Water Quality Analysis' (2017)

**Oral Presentations**

- *Modelling seasonal river export of nutrients to Lake Tana, upper Blue Nile. International conference on the Nile and GERD. 20-21 August 2020, Florida, United States of America*

SENSE coordinator PhD education

Dr. ir. Peter Vermeulen

The research described in this thesis was financially supported by the Netherlands Foundation for International Cooperation (NUFFIC).

Financial support from Wageningen University, Aquatic Ecology and Water Quality Management Chair Group for printing this thesis is gratefully acknowledged.