Executive Summary

Title: DEVELOPING METHODS TO MONITOR AND ANALYZE EUTROPHICATION OF LAKE TANA FROM REMOTE SENSING IMAGERY USING MACHINE LEARNING– UPPER BLUE NILE, ETHIOPIA

As a result of population increase and agricultural intensification, eutrophication of fresh water bodies in the globe particularly in Africa is a challenge. Lake Tana is an internationally, nationally and locally important lake located in the tropical highlands of Ethiopia. It is the largest fresh water body in Ethiopia and source of Blue Nile. As the population is increasing, the dominant land uses becomes agricultural by removing forests in the hillside and drying wetlands. Expansion of water hyacinth becomes a recent phenomenon. One of the drawbacks has been not able to do proper monitoring of its water quality status spatially and temporally that likely indicates decision makers on the status of its quality.

Remote sensing technology can provide an affordable method to monitor water quality fresh water bodies such as Lake Tana. Hence, this study intends to use remote sensing data from LandSat-8 and use the potential of Machine Learning (ML) and Artificial Neural Network (ANN) to determine the water quality and trophic status of the Lake Tana. Seven trophic level indicators; secchi disk depth (SDD), Turbidity, Total Suspended Solid (TSS), Total Nitrogen (TN), Total Phosphorous (TP), Chlorophyll-A (chl_a) and Trophic Level Index (TLI) were selected to be retrieved from remote sensing using ensembles of ML and ANN algorithms.

For this study, previously collected secondary data and new additional data will be collected from 148 sampling points on the lake in two time steps. During the field observation, secchi disk depth, total suspended solid, and turbidity test will be conducted. At the same time water samples will be brought to laboratory to analyze total Nitrogen, total phosphorous and chlorophyll-A. The Trophic Level Index (TLI) will be calculated from the observed water quality parameters.

From remote sensing, 10 spectral bands will be selected and 19 derived vegetation and water indexes (NDWI, AWI, MNWDI, SR, SRWC, and WRI) will be used to retrieve the seven trophic level indicators. From available ML and ANN, **Genetic Algorithm (GA) – Artificial neural Network (ANN), Particle Swarm Optimized (PSO) – RBF, Particle Swarm Optimized (PSO) - Support Vector Machine (SVM), Particle Swarm Optimization (PSO) - Artificial Neural Networks (ANN), Random Forest Regression (RFR), XGBoost Regression, AdaBoost Regression (ABR) and Gradient Boosting Regression (GBR)** will be used to select the best performing algorithms. At the end, the algorithms will be helpful to retrieve important water quality parameters from RS imagery and determine the eutrophic level. The results from this work have an important implication to other lakes in Africa.