Assessment of Pollution Load on the Kenyan Catchment of Lake Victoria Basin using GIS Tools

Charles Kipkoech Cheruiyot

1. Introduction

Lake Victoria is a freshwater lake in East Africa and has a large surface to basin area ratio of 1:3 i.e. 68,800 km^2 to 194,000 km^2 respectively. It is the second largest freshwater lake by surface area in the world and the largest in Africa. The lake is an economic zone to the three riparian countries, namely, Kenya, Tanzania and Uganda and also a lifeline source of water supply to dry downstream countries. The only outlet from the lake, River Nile, flows down all the way to Egypt. Intensive natural and human activities compounded by ever growing population, poor livelihoods and less investment in sanitation; have accelerated environmental degradation through deforestation, siltation, fishing malpractices, wetland destruction and direct disposal of sewage into the lake. Lake deterioration is being driven by excessive pollution load: sediments and nutrients (total nitrogen - TN & total phosphorous - TP).

Estimation of pollution load to Lake Victoria has been carried out by several studies in the past. Estimation of pollution load has always been hampered by scarcity of data which adversely affects the accuracy and reliability of results. The methods used borrowed nutrient export coefficients (UAL) to estimate pollution load. The borrowed coefficients were not adjusted to fit local conditions because of lack of relevant data and information. There is need to develop criteria of adjusting borrowed coefficients and or estimating local coefficients based on observed water quality and quantity data. Simulation of hydrology, sediment and nutrients as well as watershed management plans provides useful insights to watershed or lake manager especially on amount of pollution load and effectiveness of various watershed interventions. The studies have different estimates of pollution load which makes it difficult to determine which estimates are reliable and accurate. There is need to review and consider possible methods based on GIS modelling tools for future improvement. The Lake Victoria basin is geographically enormous and incorporation of remote sensing and GIS mapping technology has potential to improve reliability and accuracy of the estimates. Figure 1 is a map of Lake Victoria basin on the Kenyan side.

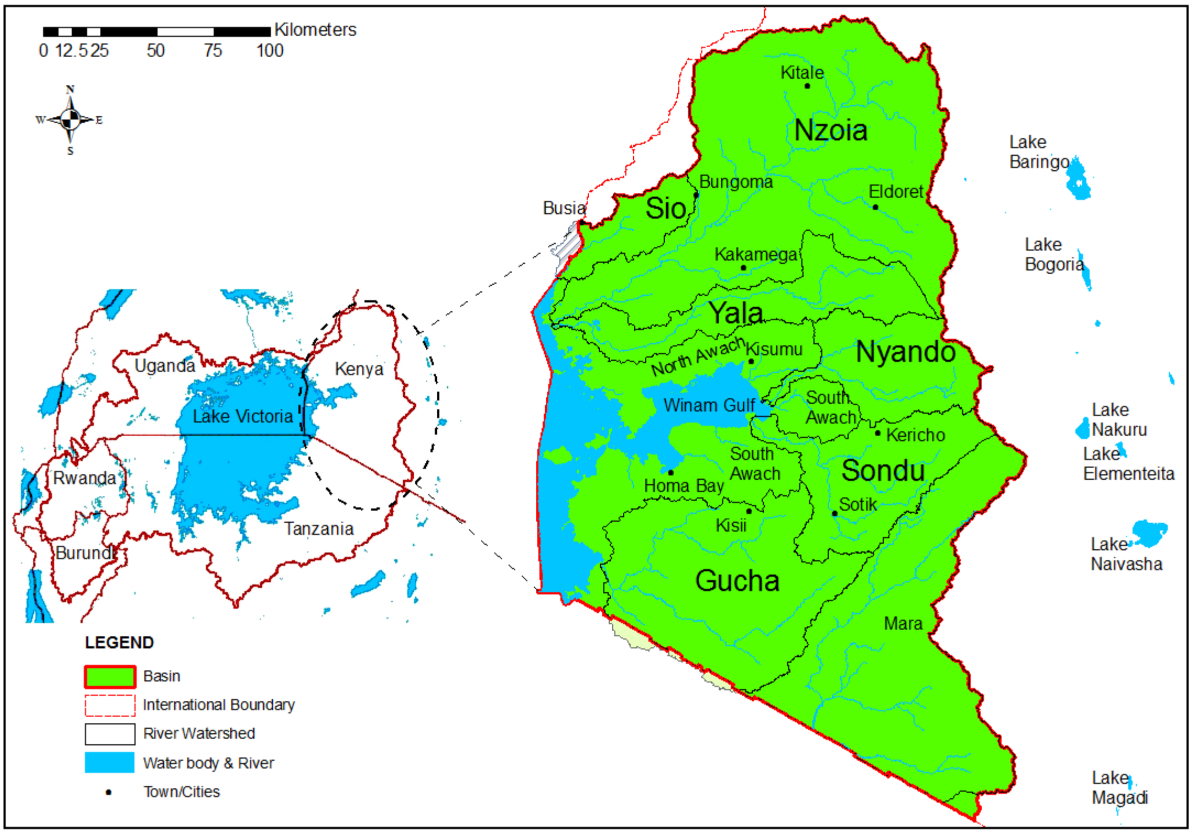


Figure 1. Map of Kenyan river watersheds in Lake Victoria basin

2. Methods of Estimation of Pollution Load

Pollution sources are usually conveniently classified into point and non-point sources. Municipal and industrial sources are classified as point sources while land runoff and atmospheric deposition as non-point sources. However, differences emerge regarding the methods of quantification of pollution load depending on data availability and quality.

Advancement in GIS and remote sensing technologies continue to complement research in water engineering. It bridges the gap in data scarcity and strengthens analysis component of research.

A sample of GIS and non GIS based models which have been used to estimate pollution load are described below.

2.1 Non GIS Models

Constant concentration model
This is a spreadsheet based model. Assumptions are made to simplify and fit into capabilities of spreadsheet. For example, a specified nutrient concentration in a river is assumed to be constant throughout the month or year i.e. it is not continuous. Total annual nutrient load is derived by

multiplying the annual river discharge with annual river concentration. Data input required include: river (stream) flow, river water quality, rainfall water quality and area of water surface. The model has limitation in simulation of water quality parameters which are rainfall driven events and continuously changing in concentration.

CMSS and Bayesian models
Catchment Management Support System (CMSS) is a simple unit area load model. CMSS estimates pollution load (TN & TP) from land runoff and allows for natural reduction (attenuation). Natural reduction is expressed as a function of river length, river channel depth and catchment area. The needed input data include: land cover types and areas of river watershed, length of river channel, slope of river, depth of river channel, and generation rates of land cover (Unit Area Load - UAL).

The CMSS model can be used in a Bayesian framework as done by Broad and Corkey (2011) in Tasmara, Southern Island State of Australia. Bayesian approach allows incorporation of uncertainty in the estimates in a natural way. All data are considered simultaneously and, in this respect, uncertainty is propagated through the model. In a Bayesian framework, CMSS can be used to calculate generation rates when you have observed pollution parameters. The river watershed is sub divided into sub catchments and CMSS model is applied to determine land cover generation rates (a case where you have all input parameters listed above with exception of generation rates).

2.2 GIS Models

AGNPS model
Agricultural Non-Point Source (AGNPS) model was developed by United States Department of Agriculture – Agricultural Research Service (USDA – ARS) (1980). The model estimates runoff load for a single storm event or for a continuous simulation (Young et al., 1987). The modified universal soil loss equation (MUSLE) is used to predict soil erosion and unit hydrograph is used to simulate hydrology flow. Input data required include: digital elevation information, soils, land cover and rainfall data. The model calculates runoff water quality of single rainfall event in a watershed and application area is limited to about 200 Km^2 hence its limitation (Aisha, 2005).

AnnAGNPS model
Annualized Agricultural Non-Point Source (AnnAGNPS) model was also developed by USDA – ARS (1990). It simulates runoff, sediment and nutrient loads from watersheds and evaluates conservation programs. The model is applied on in level of small watersheds in which the watershed can be delineated to accommodate land use and soil variation and conservation practices and remaining computationally feasible (Yongpin et al., 2011). The model routes the loads for a single day event and point sources are limited to constant loading. The model is limited by absence of pesticides consideration of mass balance calculations.

SWAT model
Soil Water and Assessment Tool (SWAT) is a GIS interface model. SWAT is a continuous model and operates on a daily frequency. It simulates watershed hydrology, sediment and nutrients transport. SWAT is a comprehensive hydrological model with capability to analyze

land and water management for agriculture and water quality. The main three modeling steps are: partitioning watershed and input information and simulation Hydrology in land phase and water or routing phase. Required data are: elevation information, soils, land cover and weather data (rainfall, humidity, wind speed, temperature, etc.).

The runoff hydrology is based on Curve Numbers (CN) by United States Soil Conservation Service (SCS). The model could be used to assess several watershed phenomena, for example to assess the impact of land cover change on a lake. Input data required are: elevation information, soils, land cover and weather data (rainfall, humidity, wind speed, temperature, land management practices, etc.).

3.0 Estimation of Pollution Load in Lake Victoria

There are several studies and projects which have been conducted in Lake Victoria in the recent past to estimate pollution load. They include: Calamari et al. (1995), Scheren et al. (1995, 2000), COWI (2002), LVEMP (2005), Scheren (2003, 2005).

A review of the studies shows that different methods have been used to estimate pollution load. Scarce and scattered data is consistently pointed out as a major limitation in the studies. The methods of estimating point and non-point pollution load can be broadly classified into:

1) Use of field data approach; and
2) Rapid assessment approach.

Field data approach uses locally measured data to estimate point and non-point pollution load in an estimation framework. For example, use of measured water quality and quantity or UAL to estimate runoff load. On the other hand, rapid assessment is applicable where data are scarce. For example, UAL from other basins (with similar characteristics with concerned study area) which is considered applicable is used to estimate runoff load.

Based on above summary of the studies, there is a need to have representative unit loads to effectively use rapid assessment methods to estimate municipal and industrial load. The main baseline data are urban population with their corresponding sanitation system, industrial production, etc. More important is to establish locally improved and applicable per capita unit load for municipal load and unit load per unit product for industrial load. The challenges experienced by past studies were pegged on the availability of these data. Population census and sanitation data are collected by national and local governments. For example, Kenya conducts population census once in 10 years. Industries are themselves better placed to provide their data if they do keep the relevant pollution records. In the studies, industrial data are the scarcest and surrounded by much uncertainty.

The alternative to rapid assessment is use of field data of municipal and industrial load; measurement of water quality and quantity of wastewater generated in a controlled system. For this alternative, data of wastewater treatment facilities for both industries and municipal are continuously monitored. Where wastewater does not flow through a control system it is accounted as non-point load. Estimates derived from wastewater treatment plants may be matched with number of persons using the plant to define the per capita load. The same may be done for industries to determine the waste load per unit of production. Treatment plants for industries, households and storm water should be separate for it to be successful. Also, data collection should be done over a reasonably long period of time enough to cover extreme cases of load flow fluctuations. Periods of storm runoff and peak production are examples of extreme cases. The information from locally measured data will be useful for future rapid assessment methods.

The two methods of estimating pollution load have been applied in Lake Victoria. They are complementary in use for lake management. The trend in the past studies is that rapid assessment was initially used because of scarce data. As more measured data are becoming available, it should be incorporated in the estimation process

Although the past studies provide useful information within the existing constraints, there is a lot of uncertainty in the accuracy and reliability of the estimated pollution loads. The Lake Victoria basin is geographically enormous and incorporation of remote sensing and GIS mapping technology will improve reliability and accuracy of the estimates. The ability to use GIS technology to collect data and predict various scenarios of land use in the quantification of pollutants should reduce the number of errors made when less-exact methods are used. Remote sensing and GIS relate spatial and temporal geographical relationships and reinforce weaknesses noted in the past studies and stand to improve the estimation of pollution load to Lake Victoria especially runoff load.

In summary, the efforts in the past studies on Lake Victoria were mainly hampered by lack of data and hence the choice of the simplistic methods used. Lacking data include management practices in the basin, water flow, water quality, municipal and industrial effluent generation, etc. Continuous time models such as Soil and Water Assessment Tool (SWAT) have not taken root in Lake Victoria. Use of SWAT to simulate pollution load requires diverse data, among, river water quality and quantity and weather data and thus its use has always been limited.

4.0 Case Studies

The description and findings of studies done by the author on estimation of pollution load on the Kenyan side of Lake Victoria basin are elaborated below. The approach was guided by challenges experienced by previous studies.

4.1 Estimation of Nutrient Export Coefficients (UAL) on Kenyan Catchment

Past studies estimated the runoff load using borrowed nutrient export coefficients from other regions. Borrowed export coefficients were not necessarily modified to match the attributes of the local area. The study estimated the nutrient export coefficients for three land uses using river runoff data measured at river mouths of watersheds on the Kenyan side of Lake Victoria. Measured nutrients at river mouths were distributed back to the watersheds using a model equation. Factors that influence the export of nutrients were also assessed and incorporated in the model. Land use areas and rainfall-runoff coefficient values were used as main variables to explain runoff load. Two sets of data were used, one to set up the model and the other to validate the model.

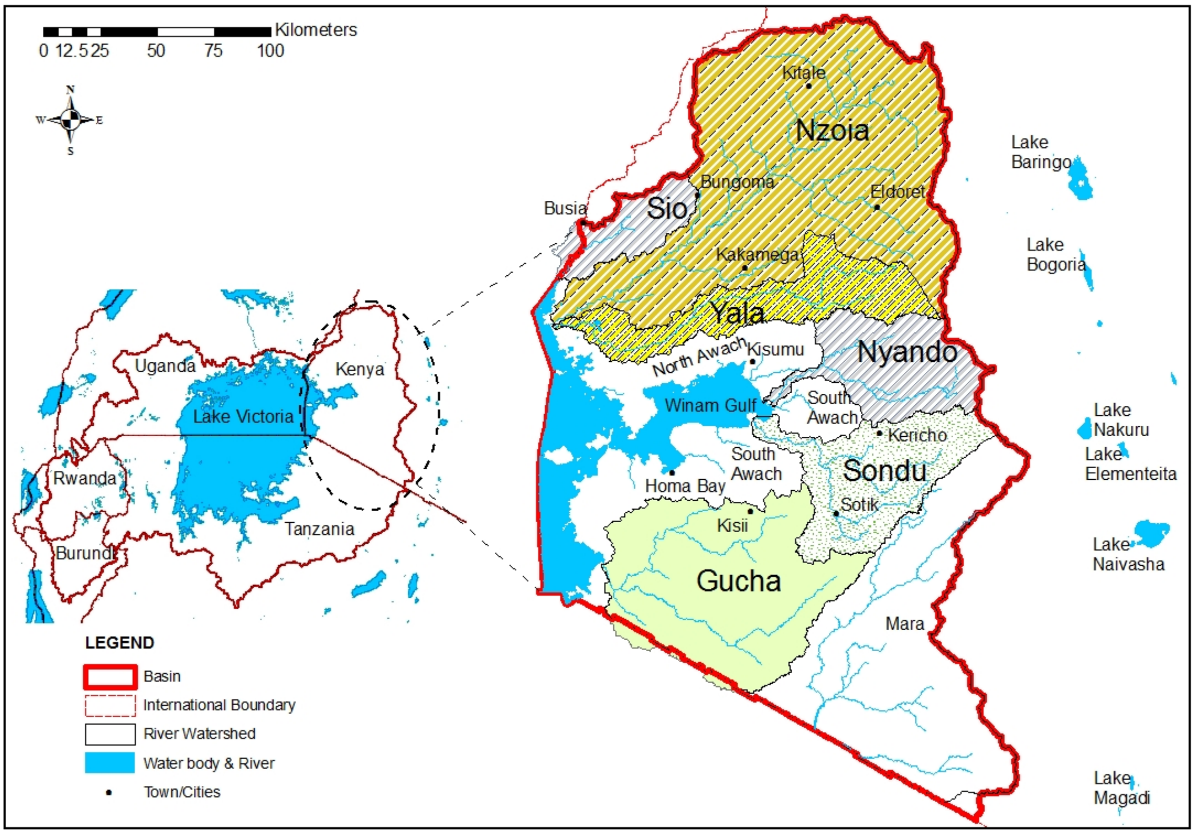
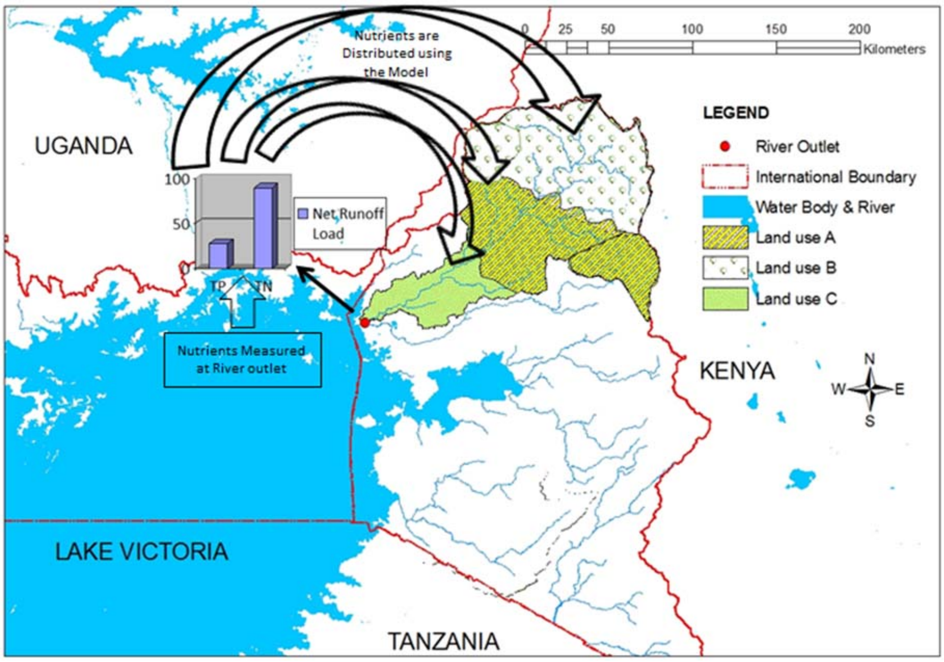
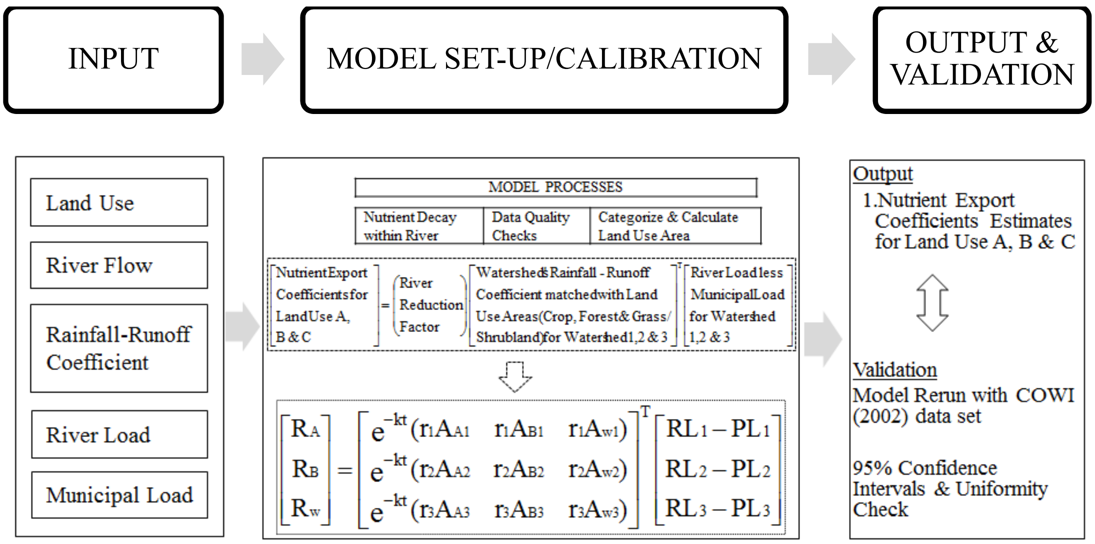


Figure 2. A map of the study area – Kenyan side of Lake Victoria basin

Conceptual approach
The conceptual approach of the study is illustrated in Figure 2 and Figure 3.



Figue 2. An illustration of distribution of nutrients measured at river mouths across various land uses



$$\begin{bmatrix} R_A \\ R_B \\ R_W \end{bmatrix} = \begin{bmatrix} e^{-kt}(r_1 A_{A1} & r_1 A_{B1} & r_1 A_{w1}) \\ e^{-kt}(r_2 A_{A2} & r_2 A_{B2} & r_2 A_{w2}) \\ e^{-kt}(r_3 A_{A3} & r_3 A_{B3} & r_3 A_{w3}) \end{bmatrix}^T \begin{bmatrix} RL_1 - PL_1 \\ RL_2 - PL_2 \\ RL_3 - PL_3 \end{bmatrix}$$

Figure 3. Model framework

Model description and calibration

Mean slope, flow length and area parameters were derived using ArcGIS software. The model equation explains the measured nutrients at the river mouth using land use and rainfall-runoff coefficient with consideration of the reduction of nutrients within the river system. It estimates nutrient export coefficients which are the only unknowns in the model. The model and its components are described below.

Runoff pollution load for each watershed was expressed as illustrated by Equation 1 using parameters described above and with consideration to municipal (point) load discharged upstream

$$\mathrm{RL} - \mathrm{PL} = \mathrm{r}\left(A_c\,\mathrm{R_c} + A_{gs}\,\mathrm{R_{gs}} + A_f\,\mathrm{R}_f\right)\exp(\text{-kt}) \qquad (1)$$

where RL is the measured load at river mouth (t/year); PL is the estimated point load generated in the watershed (t/year); A_i represent areas of respective three land uses in the watershed (c for cropland, gs for grassland and f for forest) (Km^2); R_i is the nutrient export coefficient (t Km^{-2}/year); r is the relative watershed rainfall-runoff coefficient (dimensionless); and $\exp(\text{-kt})$ represents nutrient reduction within the river system (dimensionless).

Cheruiyot & Muhandiki (2014) has more details on description of the model.

Findings

Table 1 which summarizes the estimates of export coefficients shows that vegetation/grassland/shrubland generates more nutrients per unit area annually while cropland generates the least with respect to both TN and TP. The land use also has relatively wider range. However, cropland is the main source of nutrients in terms of aggregate load due to its dominant coverage. The high coverage of land use under cropland in the catchment is explained by dominant tea, maize and sugarcane plantations in the study area. These are the main livelihood activities of the resident population.

Table 1. Estimates of nutrient export coefficients (Kg/ha/yr).

Nutrient	Statistics		Cropland	Forest	Vegetation/ grassland/ shrubland
TN		Minimum	0.643	3.123	12.862
		Mean	1.412	14.426	27.800
		Maximum	2.048	29.625	45.880
		Stdv*	0.543	10.902	33.032
	95% Confidence Interval	Margin of Error	0.238	4.778	5.232
		Low	1.174	9.648	22.569
		High	1.650	19.204	33.032
TP		Minimum	0.185	0.045	2.639
		Mean	0.257	1.958	5.611
		Maximum	0.296	5.778	11.423
		Stdv*	0.062	3.308	5.033
	95% Confidence Interval	Margin of Error	0.027	1.450	2.206
		Low	0.230	0.508	3.405
		High	0.284	3.408	7.817

*Standard Deviation

Conclusions

Runoff nutrient export is mainly influenced by the watershed's environmental attributes (land use, rainfall and soil characteristic, land slope, drainage density, etc.). Land use and rainfall characteristics are easily measureable and their relationship with nutrient generation is not complex. Rainfall-runoff coefficient influences the export of nutrients to the extent of their linear correlation. The relatively high river nutrient concentration with low rainfall-runoff depth of Nyando watershed suggests that driving factors other than land use and rainfall-runoff coefficient include loose soil characteristics. However, positive solutions for nutrient export coefficient demonstrated that land use and rainfall-runoff coefficient have significant influence and are usually available and useful variables to explain runoff load.

The estimated nutrient export coefficients are sufficient for large-scale rapid assessment of pollution load for a situation such as that of Lake Victoria where borrowed export coefficients are often used due to data scarcity. The estimated export coefficients represent the average values and not exact values due to spatial and inherent nature of environmental attributes across the catchment. The usage of estimated export coefficients elsewhere is subject to adjustment relative to rainfall-runoff coefficient of the base watershed. Rainfall-runoff coefficient is an appropriate variable to adjust the nutrient export coefficients from other areas to fit local conditions. Catchment assessment at small-scale level would give more precise results as it would reduce the variance of environmental characteristics. Further investigation of the influence of factors other than land use and rainfall-runoff coefficient will be considered in the successive studies. A model such as SWAT that captures all the characteristics holistically would be better.

4.2 SWAT Model Applied on Sondu Watershed to Assess Watershed Management Plans

During high precipitation seasons, Sondu River regularly bursts its banks downstream and causes flooding at Nyakach and Rachuonyo North districts. Excessive sedimentation not only leads to lake pollution, high operation costs of desilting irrigation channels and dredging of hydropower dams and reduced river capacity but also loss of lives and livelihood and human displacement.

Watershed management plans refer to assessment of impacts on stream flow, soil and nutrient loss by various management practices/plans. Simulation of hydrology and soil loss and impacts of watershed management practices provides useful insights to watershed or lake manager especially on mitigation of adverse impacts. Assessment of management practices done at watershed outlet gives a clearer understanding of temporal dynamics while at sub watershed level provides information on spatial distribution. A simulation of temporal-spatial characteristics of a watershed aids its management through identification of hot spot areas and time periods and resource needs which are useful for informed decision making.
The study assessed impacts of watershed management plans on hydrology and soil loss in Sondu watershed.

Conceptual approach

A step by step procedure of development of SWAT model is available online and in tutorial accompanying the software. The study used the Shuttle Radar Topography Mission (SRTM C-BAND) 90 m resolution elevation data from the U.S Geological Survey (USGS). The 300 m resolution - Globcover Ver. 2.3 (land-use data of 2005-2006) were sourced from European Space Agency (ESA). Harmonized World Soil Database (HWSD) Ver 1.2 (Soil Data, 2007) was sourced from International Institute for Applied Systems Analysis (IIASA). The HWSD data description in the accompanying documentation does not have all soil parameters as required in SWAT model.

The weather data were obtained from responsible institutions in Kenya (Kenya Meteorological Department - KMD) for five weather stations with exception of solar radiation (Kisumu, Kericho, Molo, Kuresoi, Kisii) (Fig. 4.4). The data period of interest was from 1990 to 2010. Observed daily river (stream) flow data and limited observed daily suspended solids concentration were sourced from Water Resources Management Authority (WRMA) in Kenya.

Watershed management plans

Watershed management plans represent interventions aimed at reducing the soil loss and environmental degradation. There are several choices of interventions in literature; however, the options should be based on local conditions and how realistic they are to implement in the watershed. In the study three management options were assessed and were considered realistic and reasonably comparable and as compared with similar studies in literature.

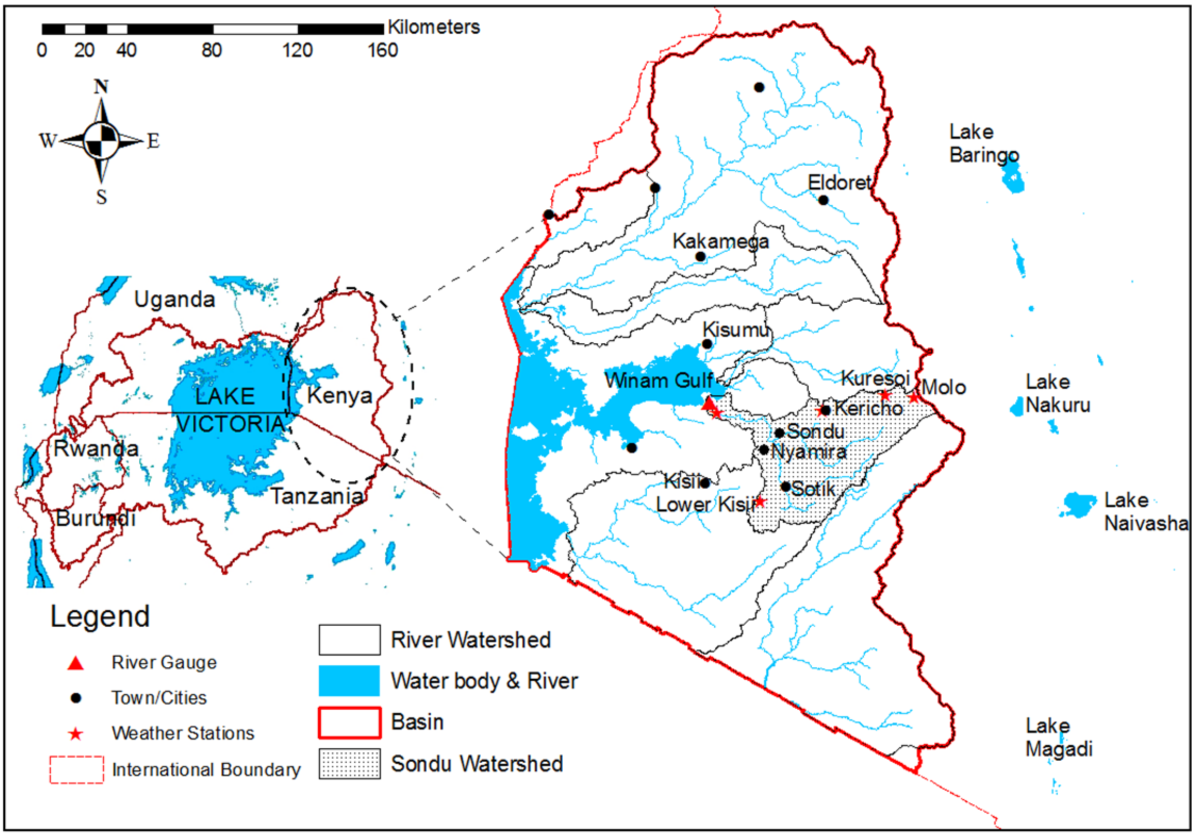


Figure 5. Map of Sondu watershed

Findings

The main land covers in Sondu watershed are agriculture (68 %) and forest (31 %). The forested areas are spatially distributed in the watershed and biased to upstream. In spite of scarce data, suspended solids simulation fitted into observed values with reasonable model performance. Soil yield ranking by sub basins of the baseline management option does not show a clear pattern with respect to upstream or downstream location but it has correlation with land cover distribution with agriculture driving up the soil yield while forest slowed down soil loss. Both filter and reforestation plans would be more effective in wetter months of the year. The simulation showed that months of April-May and November-December which are beginning of high rainfall seasons had high soil loss reduction rates for reforestation and filter plans. The reforestation plan consistently ranked higher with respect to soil yield reduction in all months of the year as monitored at basin outlet. Reforestation was relatively effective in reducing soil loss at most upstream sub basins while filters had more impact at most downstream sub basins of the watershed. Cheruiyot (2015) has more details on the findings of the study.

5. Conclusion

Although the past studies provide useful information within the existing constraints, there is a lot of uncertainty in the accuracy and reliability of the estimated pollution loads. The Lake Victoria basin is geographically enormous and incorporation of remote sensing and GIS mapping technology will improve reliability and accuracy of the estimates. The ability to use GIS technology to collect data and predict various scenarios of land use in the quantification of pollutants should reduce the number of errors made when less-exact methods are used. Remote sensing and GIS relate spatial and temporal geographical relationships and reinforce weaknesses noted in the past studies and stand to improve the estimation of pollution load to Lake Victoria especially runoff load.

In summary, the efforts in the past studies on Lake Victoria were mainly hampered by lack of data and hence the choice of the simplistic methods used. Lacking data include management practices in the basin, water flow, water quality, municipal and industrial effluent generation, etc. Continuous time models such as Soil and Water Assessment Tool (SWAT) have not taken root in Lake Victoria. Use of SWAT to simulate pollution load requires diverse data, among, river water quality and quantity and weather data and thus its use has always been limited.

References

Aisha, A. (2005): *Modeling of Non-Point Source Pollution, Experiences from River basin Mun, Thailand***.** Master's Thesis, Graduate School of Engineering, Asian Institute of Technology, Thailand.

Broad, S. T. and Corkrey, R. (2011): Estimating Annual Generation Rates of Total P and Total N for Different Land Uses in Tasmania, Australia. *Journal of Environmental Management,* 92. 1609 - 1617.

Calamari, D., Akech, M. O. and Ochumba, P. B. O. (1995): Pollution of Winam Gulf, Lake Victoria, Kenya: A Case Study for Preliminary Risk Assessment. *Lakes & Reservoirs: Research and Management*, 1. 89 - 106.

Cheruiyot, C. K. (2015): *Assessment of Pollution Load on the Kenyan Catchment of Lake Victoria Basin using GIS Tools*. PhD Thesis, Nagoya University, Nagoya, Japan.

Cheruiyot C. K. and Muhandiki V. M. (2014): Estimation of Nutrient Export Coefficients in lake Victoria Basin. *Journal of Water and Environment Technology.* 12 (3), 231 - 244.

Scheren, P. A. G. M., Zanting, H. A. and Lemmens, A. M. C. (2000): Estimation of Water Pollution Sources in Lake Victoria, East Africa: Application and Elaboration of the Rapid Assessment Methodology. *Journal of Environmental Management,* 58, 235 - 248.

Scheren, P. A. G. M. (2003): *Integrated Water Pollution Assessment in Data and Resource Poor Situations*: *Lake Victoria and Gulf Guinea Case Studies.* PhD Thesis, Technische University Eindhoven, the Netherlands.

Scheren, P. A. G. M. (2005): Integrated Water Pollution Assessment in Data and Resource Poor Situations, Proc. of 11th World Lakes Conference, Vol 2. October, 31 – November, 4, Nairobi, Kenya.

Scheren, P. A. G. M, Bosboom, J. C., Njau K. N. and Lemmens, A. M. (1995): Assessment of Water Pollution in the Catchment Area of Lake Victoria, Tanzania. *Journal of Eastern African Research and Development*, 25. 129 - 143.

Scheren, P. A. G. M., Zanting, H. A. and Lemmens, A. M. C. (2000): Estimation of Water Pollution Sources in Lake Victoria, East Africa: Application and Elaboration of the Rapid Assessment Methodology. *Journal of Environmental Management,* 58. 235 - 248.

USDA – ARS (1980): Agricultural Non-Point Source Pollution Model, National Sedimentation Laboratory, USA.

Yongping, Y., Megan, H., Mehaffey, Ricardo, D. L., Ronald, L., Bingner, R. B., Caroline, E., Michael, A. J. (2011): AnnAGNPS Model Application for Nitrogen Loading Assessment for the Future Midwest Landscape Study. *Water Journal,* 3, 196 – 216.

Young, R. A., Onstad, C. A., Bosch, D. D., Anderson, W. P. (1987): AGNPS: Agricultural Non-Point Source Model. A Large Watershed Analysis Tool, Conservation Research Report 35, Agricultural Research Service, United States Department of Agriculture, Washington, D. C.