

MODELING OF RUNOFF POLLUTION LOAD IN A DATA SCARCE SITUATION USING SWAT, SONDU WATERSHED, LAKE VICTORIA BASIN

***CHERUIYOT C.K.¹ AND MUHANDIKI V.S.²**

¹Graduate School of Engineering, Nagoya University, Nagoya, 464-8601, Japan

²Program for Leading Graduate Schools, Nagoya University, Nagoya, 464-8601, Japan

Abstract

Sondu watershed is one of the major rivers flowing to Winam Gulf on the Kenyan side of Lake Victoria. The lake is a source of fish and freshwater supply and also provides routes for transportation. Lake deterioration is being driven by excessive pollution load: sediments (TSS) and nutrients (TN & TP). Previous assessments of run-off load in the basin used simple unit load methods because of lack of data and there is still need for comprehensive information on lake pollution and deterioration. Monitoring of weather and water quality data in the watershed is still low. Incorporation of Geographical Information System (GIS) and remote sensing based tools are options to improve data assembly and pollution load assessment. This study used Soil Water Assessment Tool (SWAT) to simulate temporal-spatial distribution of surface water runoff (river flow), sediment and nutrient generation in Sondu watershed, and to identify soil erosion and nutrient source hot spots. Annual sediment generation to the lake is 80,000 t/yr composed of mainly silt while TN & TP are 3,388 t/yr & 312 t/yr, respectively. Sediment and nutrients are mainly generated from agricultural crop areas at downstream, central (Sondu) and upstream West (lower Kisii/Nyamira) area of the watershed while the high water runoff yielding areas are upstream (Kericho/Kisii/Kericho) areas. The nutrients peak periods lagged behind sediments' by one to two months on average. Consideration of afforestation and adoption of best management practices such as filters on crop areas would be useful in reduction of pollution load to the lake and land conservation.

Key Words: Lake Victoria, Sondu watershed, Pollution load, SWAT model

Introduction

Lake Victoria in East Africa is a great resource to basin and downstream countries and global economies at large. The lake is a source of fish, freshwater supply and also provides routes for transportation. The average depth of the lake is 40 m (Muyodi *et al.*, 2010) and there are six main rivers on the Kenyan side of the basin that flow to the lake. The unique shape of the lake has some of its parts protruding inland such as

Winam Gulf on the Kenyan side of Lake Victoria. The gulf is long, slender, shallow, semi-enclosed and bottlenecked at the point of connection to the main lake (Cheruiyot and Muhandiki, 2014a). It has a surface area of 1,400 km², an average depth of 10 m (Calamari *et al.*, 1995) and connects to the main lake at Rusinga channel (Figure 1). The catchment of Winam Gulf is on the Kenyan side of Lake Victoria basin.

The gulf and by extension the lake and its ecosystem show evidence of dramatic changes with infestation by water hyacinth being one of the major concerns in the lake in recent years. Threats facing the lake have adversely affected efforts to improve livelihoods of the population dependent on the lake (Kayombo and Jorgensen, 2005). Pollution load parameters of concern include total nitrogen (TN), total phosphorous (TP) and total suspended solids-sediments (TSS).

There have been efforts to estimate pollution load in Lake Victoria in the past (Calamari *et al.*, 1995, COWI, 2002, LVEMP, 2005, Scheren, 2005 and 2003 and Scheren *et al.*, 2000 and 1995). The estimation has always been hampered by scarce data and limited deployment of spatial-temporal related technologies. Geographical Information System (GIS) and remote sensing are among suggested options by the studies to improve the estimation of pollution load. Use of GIS tools in Lake Victoria basin has not found much practice, but few studies have been undertaken. For example, Jayakrishnan *et al.* (2005) applied Soil Water Assessment Tool (SWAT) to model Sondu river watershed in Lake Victoria basin. The study assessed the impact on change in land cover driven by adoption of modern technology for smallholder dairy industry (Cheruiyot and Muhandiki, 2014b). Kimwaga *et al.* (2011) modeled non-point source pollution in Simiyu watershed in Tanzania using the SWAT model. The study used the model to quantify sediment and nutrient from two land cover scenarios. From the literature reviewed, Kimwaga *et al.* (2011) was the only study which aimed to predict and quantify non-point source pollution (sediments and nutrients) going into Lake Victoria using the SWAT model. Although the studies were faced with the

problem of data scarcity due to detailed model data requirements, more similar studies are needed in Lake Victoria. The SWAT model is a physically based and continuous time model that operates on daily/sub-daily time step and operates/runs in GIS frame work.

Sondu River which has a surface area is 3,508 km² is one of the main river watersheds draining its water to Winam Gulf and is the third largest by volume among the Kenyan rivers draining into Lake Victoria (Figure 1). It is a source of water supply and powers the 60 MW Sondu-Miriu hydropower station downstream. It supports a human population of about 1.1 million (LVEMP, 2005) with dense population in lower Kisii and Nyamira regions while its elevation range is 1,134-2,930 m above sea level. The watershed receives average annual rainfall of 1,000-2,500 mm, with upstream parts (Kericho, Molo, Kisii) receiving higher rainfall than downstream areas. There is a wide variation in average daily temperatures which range from 10°C to 31°C. According to LVEMP (2005), high temperatures are experienced downstream and around the lakeshore. Major land covers in the watershed are agriculture (67 %) and forest (29 %) which include the Mau forest complex. Forested areas are mainly upstream while the dominant agriculture enjoys even distribution across the watershed. The mau forest is the largest indigenous mountain forest in East Africa and also important ecosystem resource for not only Sondu watershed but for many other neighboring Kenyan rivers. The forest had earlier been encroached for human settlement but it is now under rehabilitation. Changes in forest cover and other aspects of the watershed would have adverse impact on development at both the watershed and within the lake.

The aim of this study was to simulate river flow, sediment and nutrient load in Sondu watershed using the SWAT model and its complementary programs; to assess

the spatial-temporal distribution of sources of the sediments and nutrients; and to identify soil erosion and nutrients depletion prone zones.

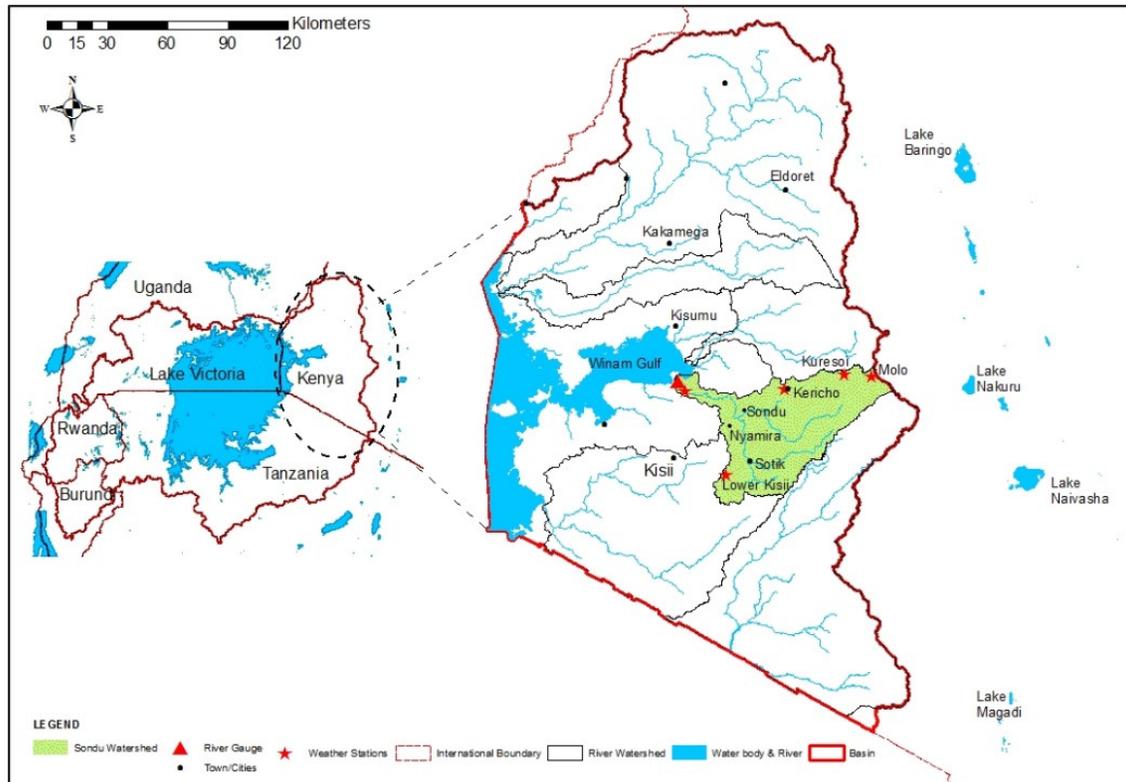


Figure 1: Sondu watershed in Lake Victoria basin

Materials and Methods

Remote sensing and field observed data were used to set up the SWAT model. Remote sensing data were sourced from relevant international institutions while observed river flow and limited water quality data (sediments – TSS, nitrogen – TN and phosphorous – TP) were collected from Kenyan institutions. Surface runoff, nutrient and sediment modeling parameters were selected based on their sensitivity and used to set up the model. The model run (2000 to 2010) was calibrated over the period 2005 to 2007 and validated using the year 2010 observed data. The model performance was evaluated based on model inbuilt statistics.

DEM, Land Use and Soil Data

Digital Elevation Model (DEM) represents elevation of the watershed. In this study, the Shuttle Radar Topography Mission (SRTM C-BAND) data from the National Aeronautics and Space Administration (NASA) homepage were used. The data were processed in Arc GIS 9.3 before importing into SWAT model at resolution of 90 m. Globcover Ver 2.3 (Land-use data of 2009) from European Space Agency (ESA) were sourced for this study. The resolution of the data is 300 m. The data were originally classified in UN Land Cover Classification System (UNLCCS) and were therefore reclassified

to SWAT model categories guided by accompanying literature describing the land cover types.

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 as per accompanying documentation does not have all soil parameters as required in SWAT model. In particular the missing parameters were (HYDGRP, SOL_K, SOL_ALB, and USLE_K). The Soil Plant Atmosphere Water (SPAW) software was used to get SOL_K (saturated hydraulic conductivity). The software calculates SOL_K by inputting relevant soil parameters while Universal Soil Loss Equation (USLE) was used to calculate USLE-K (soil erodibility K factor). The parameter is a product of soil erodibility factors of coarse-sand, clay-silt and organic-carbon soils. The HYDRGP (soil hydrologic group) and SOL_ALB (moist soil albedo) parameters were determined by

professional discretion guided by study on soil characteristics as described in SWAT documentation.

Weather Data

Precipitation, temperature, solar radiation, humidity, and wind speed are the five weather parameters required in the model. These data were collected from responsible institutions in Kenya (Kenya Meteorological Department - KMD and Water Resources Management Authority – WRMA) for five weather stations. The data period of interest was from 1990 to 2010; however, not all data days were available (Table 1). Weather data input were done on daily frequency. Statistics for the precipitation were calculated using pcpSTAT program (Stefan Liersch, 2003) while for the other parameters such as average maximum and minimum temperature were calculated in excel spreadsheets. In this study, inbuilt weather generator was used to simulate solar radiation and data gaps of the missing weather parameters.

Table 1 Daily weather data: Percentage of no data days (1990–2010)

Parameter	Station	Missing data days
Rainfall	Kisumu	0.4 %
	Kericho	0.8 %
	Molo	30.2 %
	Kuresoi	1.9 %
	Kisii	1.9 %
Temperature	Kisumu	8.5 %
	Kericho	11.9 %
	Kisii	21.2 %
Relative humidity	Kericho	15.6 %
	Kisumu	82.3 %
Wind speed	Kisumu	94.9 %

Observed river flow and nutrients (TN and TP)

Daily observed stream flow data and limited observed nutrients were sourced from Water Resources Management Authority (WRMA) in Kenya. The monthly averages were deduced from the daily observed values. Available data within a month were used to calculate the average for that month. Modeling in SWAT using monthly data is preferred especially if daily

data have significant gaps. Stream flow had a better temporal coverage than other variables (Table 2).

Table 2 Percentage of available monthly calibration and validation data

Period	Stream flow	Sediments	TN	TP
2005-2007	40 %	34 %	20 %	23 %
2010	75 %	23 %	23 %	23 %

SWAT modeling process

The SWAT model was set up in the conventional three step procedure: watershed partition, information input and simulation. Watershed was first set up in ArcSWAT and partitioned into 25 sub-watersheds. Partitioning the watershed into several sub-watersheds and referencing them spatially inform better analysis of runoff pollution load. Second, Hydrologic Response Units (HRUs) were created. The HRU forms a unit of analysis and is a product of land use, soil and slope data. Finally, weather data were input into the model and model was run over the period 2000 to 2010 for both daily and monthly time periods.

Sensitivity, calibration and validation

The arcSWAT 2009 was used to run sensitivity analysis for the readily inbuilt available 42 stream flow, sediment and nutrient parameters. Sensitivity analysis was run using both observed daily data and without observed data. The model was calibrated over the period 2005 to 2007 while validation period was the year 2010. The choice of the periods was based on level of data available, especially observed river flow, sediments and nutrients. The calculated monthly values were used for calibration and validation. The scarce weather and water quality data (Table 1 & 2) resulted in limited latitude on the length of calibration and validation periods.

The SWAT Calibration and Uncertainty Program (SWAT-CUP) was used for the calibration and validation of the model. Calibration was done step by step, one variable at a time. The SWAT-

CUP is a complementary program to arcSWAT and was preferred for calibration/validation process because it is faster as compared to inbuilt automated calibration in arcSWAT. Sequential Uncertainty Fitting Ver 2 (SUFI2) optimization algorithm option was used for analysis of the model.

Model performance

The four model performance factors (p-factor, r-factor, coefficient of determination- R^2 and Nash-Sutcliff efficiency-NS) were used to guide the calibration and validation processes. The p-factor is the percentage of data bracketed by the 95% prediction uncertainty (95PPU) while the r-factor is the average thickness of the 95PPU band divided by standard deviation of observed data. On the other hand, the NS and R^2 define goodness of fit between observations and final simulation (Karim, 2012). According to Moriasi *et al.* (2007), model simulation can be judged as satisfactory if $NS > 0.5$ and acceptable if NS is 0-1 for stream flow while Santhi *et al.* (2001) recommends $R^2 > 0.5$ as acceptable. Harmel *et al.* (2006) argue that uncertainty of measured data, which is characterized by measurement conditions, missing data days etc, be considered in model evaluation. That is to say, model evaluation for high certainty data should be stricter but relaxed in a case of low certainty data.

Results and Discussion

Several iterations (85) were run to arrive at reasonable simulation values for the four variables (stream flow, sediments, TN, TP). The main land covers are

agriculture and forest (67 % and 29 % respectively). The forested areas are spatially distributed in the watershed. However, the main forest is the Mau forest, located upstream and to the South East of the watershed (Figure 1).

Sensitivity analysis

The Cn2 (initial SCS curve number for moisture condition II), Spcon (linear parameter for calculating the maximum amount of sediment that can be reentrained

during channel sediment routing), Nperco (nitrogen percolation coefficient) and Biomix (biological mixing efficiency) parameters were the most sensitive parameters. However, less sensitive parameters such as SOL_AWC (available water capacity of the soil layer) were useful in calibrating peak river flows. The observed variables had different sets of top sensitive parameters but Cn2 consistently ranked among the top in the sets (Table 3).

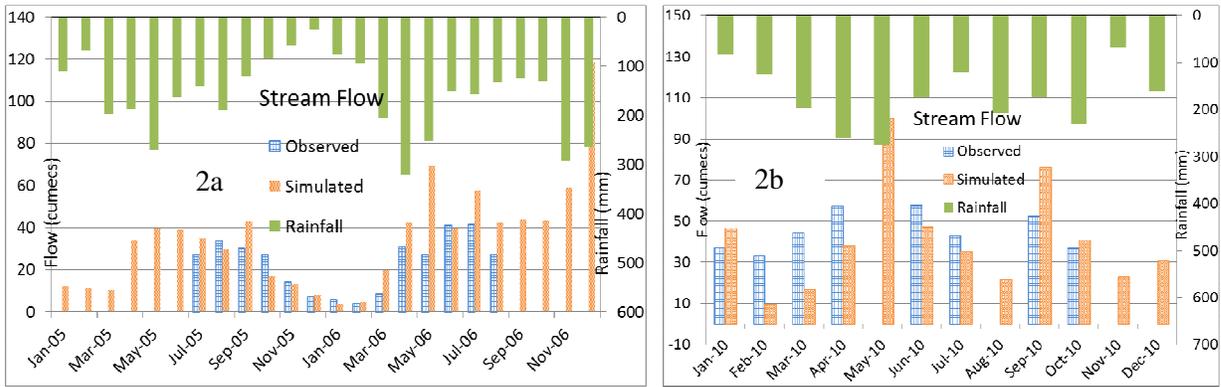
Table 3: Parameter Sensitivity derived using Observed Variables (Sensitivity decreases down the Table)

Rank	River Flow	Sediment	TN	TP
1	Cn2	Spcon	Nperco	Biomix
2	Alpha_Bf	Ch_K2	Cn2	Surlag
3	Rchrg_Dp	Ch_N2	Blai	Usle_P
4	Ch_K2	Cn2	Biomix	Canmx
5	Ch_N2	Spexp	Rchrg_Dp	Cn2
6	Esco	Alpha_Bf	Usle_P	Ch_K2

Stream flow

Initial iteration was run using the five stream flow most sensitive parameters, but subsequent improvement led to replacement and addition of other parameters. Eleven parameters were used for the final calibration and validation iterations. The peak and low flows were simulated mainly by varying Alpha_Bf (baseflow alpha factor-days) and SOL_AWC parameters. The model simulated stream flow with a p-factor of 0.86 and r-factor of 2.0 over the calibration period and 0.2 and 1.1 respectively for validation period. The calibration simulation had R² and NS indices of 0.6 and 0.46 while the values for validation were 0.3 and -4.45 respectively. The validation period did not perform as well as calibration period mainly because of scarce data in the period that led to limited

room for calibration. However the validation iteration captured the low and high flow seasons reasonably well and insynchrony with rainfall pattern (Figures 2a and b). The simulations compared favorably with similar studies on Lake Victoria such as Kimwaga *et al.* (2011) and Opere and Okello (2011) that attained R² value of 0.24. In the period 2005-2007, the high rainfall period was March-May while the high stream flow was from May-July (Figures 2a and b). The simulated annual average flow over the calibration period (2005-2007) was 35 m³/s whereas LVEMP (2005) recorded an average of 42 m³/s although the time periods are different. The LVEMP (2005) finding is based on 2003 flow measurement which was a relatively wetter year and this could explain the difference.



Figures 2a & b Observed and simulated stream flow and rainfall in the calibration period (2a) and validation period (2b)

The year 2006 had relatively higher water yield out of the watershed to Lake Victoria while 2005 had the least during the calibration period. The temporal precipitation had a similar trend. The high water yielding areas are watersheds in the upstream-North (Kericho region), upstream-South and South East (upper Sotik/Kuresoi regions) of the watershed (Figures 3a, b and c). These high water yielding areas not only receive higher precipitation but also have higher slope and the two characteristics mainly explain the water yield. Some high water yielding areas are mainly covered by agriculture (Sotik/Kericho) and some parts are mainly forested (Kuresoi) (Figures 3a, b, c and d). This was the same finding by Betrie *et al.* (2011) and Opere and Okello (2011) that agricultural areas are significant contributors of water runoff yields.

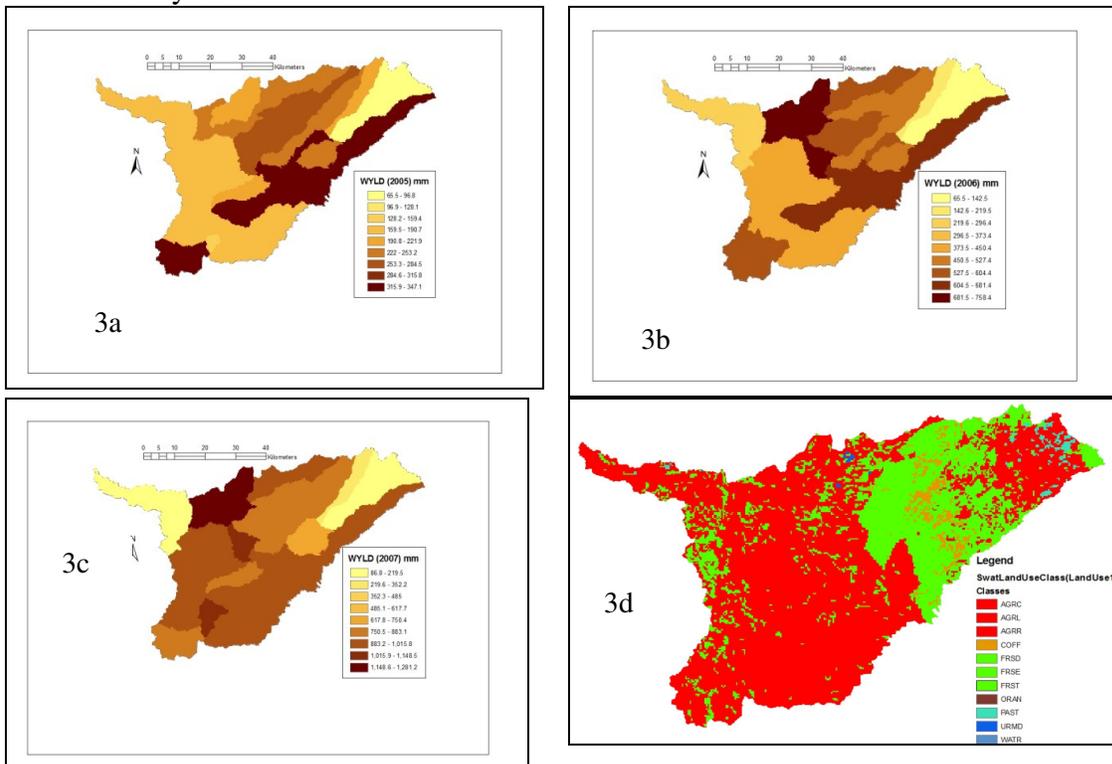


Fig. 3 a, b, c & d Water yield in the years 2005 (3a), 2006 (3b), 2007 (3c) and Land cover, 2009 (3d)

Sediment

The p-factor and r-factor indices for the final calibration iteration were 0.33 and 1.14 while for validation period were 1.0 and 2.0 respectively. Using the two indices, the model performance aggregately performed better over the validation period than calibration although periods are characterized by low observed data (Table 2). The R² and NS calibration indices were 0.2 and 0.07 while for validation they were 0.5 and 0.2 respectively. Based on Moriasi *et al.* (2007) and Santhi *et al.* (2001), the model performance was considered acceptable and within the observations of Harmel *et al.* (2006) in a case of low observed data. High sediment yield periods were Feb-April and Nov-Jan and they directly correlate with high rainfall seasons (Figures 4a and b).

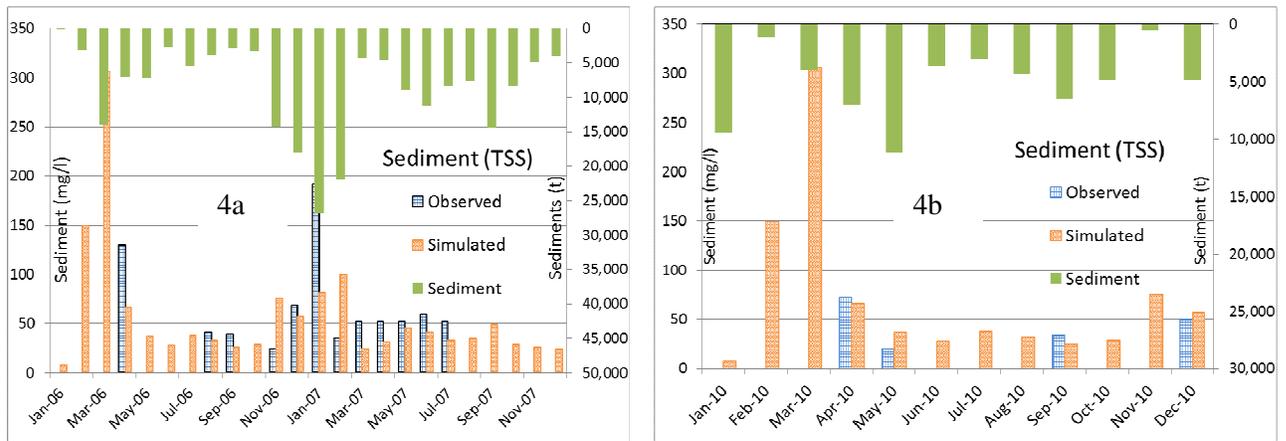


Fig. 4 a & b Observed and simulated sediment concentration and sediment load in the calibration period (4a) and validation period (4b)

Over the period 2005-2007, unlike the water yield, the year 2007 had high sediment yield while 2005 had the least. The upstream-South (upper Sotik) part of the watershed has high sediment yield relative to other parts (Fig. 5a, b & c). This could be attributed to slope terrain and agricultural activities in the zones. However, not all sediment yields from upstream watersheds end up in the outlet watershed because some are deposited in the river channels. This is to say, high upstream generated sediments are not necessarily responsible for serious threat to the lake or developments downstream such as dams. The sediment high yielding areas

are mainly agricultural areas (Fig. 5a, b, c, & d). The average simulated annual sediment load (80,000 t/yr) over the calibration period (2005-2007) was lower as compared to estimates by LVEMP (2005) of 145,000 t/yr. The year (2003) of data collection by LVEMP was wetter and higher sediment yield is expected. Generally sediment calibration process is challenging, as has been reported by most SWAT users, and this was compounded by minimal observed data for this particular study. Modeling of sediments are surrounded by hard-to-calibrate multiple factors and range of estimates are bound to be wide.

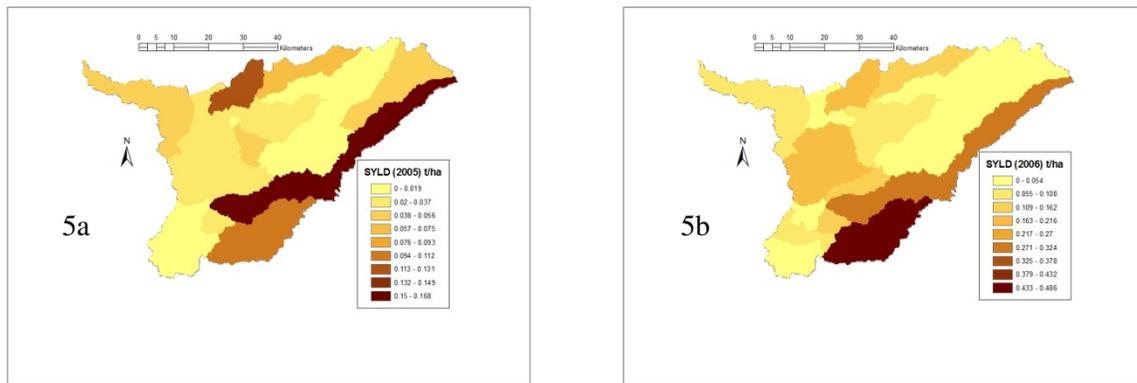


Fig. 5a & b Sediment Yield in the Year 2005 (5a) and 2006 (5b)

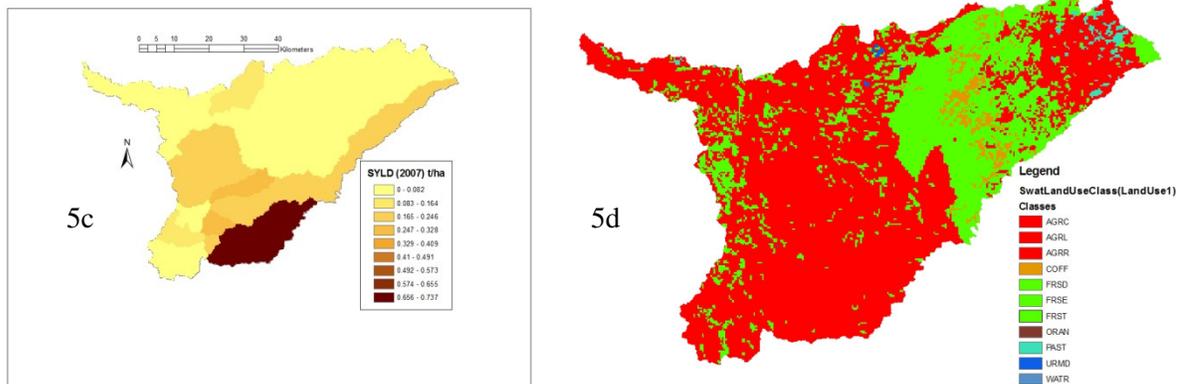


Fig. 5c & d Sediment Yield in the Year 2007 (5c) and Land Cover in the Year 2009 (5d)

Nutrient Load (TN and TP)

Better p-factor index for nutrients over the calibration period (TN: 0.71 & TP: 0.5) was attained as compared to sediment calibration. The coefficients of determination (R^2) were also better (TN: 0.67 & TP: 0.21). The model performance was considered reasonable (Harmel *et al.*, 2006). Monitoring of nutrients in Sondu watershed is low and observed data collected from the field were scarce. Monitored data over the calibration period were lower; nevertheless the available data were used for model development (Table 2).

April-May and Oct-Dec are high nutrient (TN & TP) yield seasons with

exception of Nov-Dec of 2005 and 2010 (Fig. 6a & b, 8a & b) which were low rainfall seasons (Fig. 2a & b). The downstream and central (Sotik region) to upstream-West (lower Kisii/Nyamira region) of the watershed are high nutrient yielding zones (Fig. 7a, b & c; Fig. 9a, b & c). The regions are agricultural areas while the lower Kisii/Nayamira region is not only covered by agriculture but it is also densely populated. The two land use characteristic explains the relatively high nutrients generation from these areas.

The simulated annual average TN and TP over the calibration period (2005-2007) was 3,388 t/yr and 312 t/yr respectively. The estimates by LVEMP (2005 & 2002),

from the field observations done in the year 2001 and 2003, found annual average range of TN & TP as 1,374-1,821 t/yr and 183-318 t/yr respectively. The TP findings compared favorably with this study while TN was on a higher side. Similar to

sediments, modeling of nutrients in an environment of scarce data as is the case in Sondu watershed, aggregate loads need to be assessed on the background of higher uncertainty of observed data.

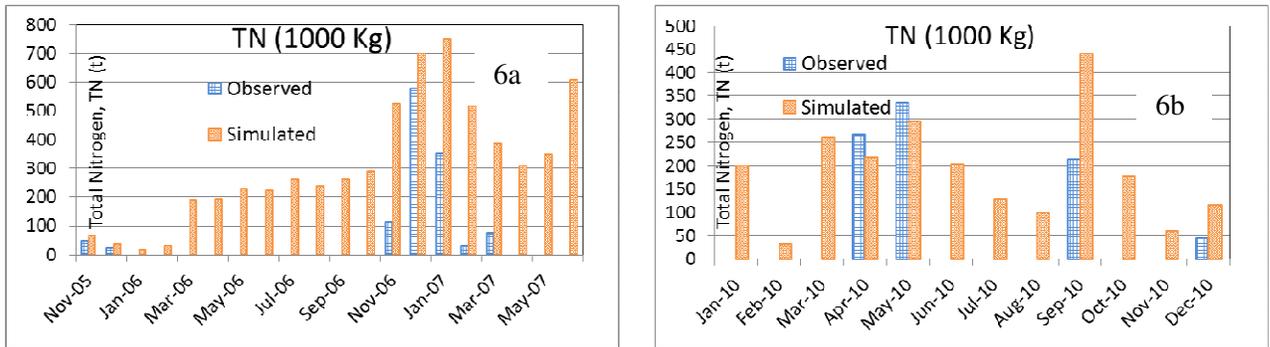


Fig. 6a & b Observed and simulated total nitrogen (TN) in the calibration period (6a) and validation period (6b)

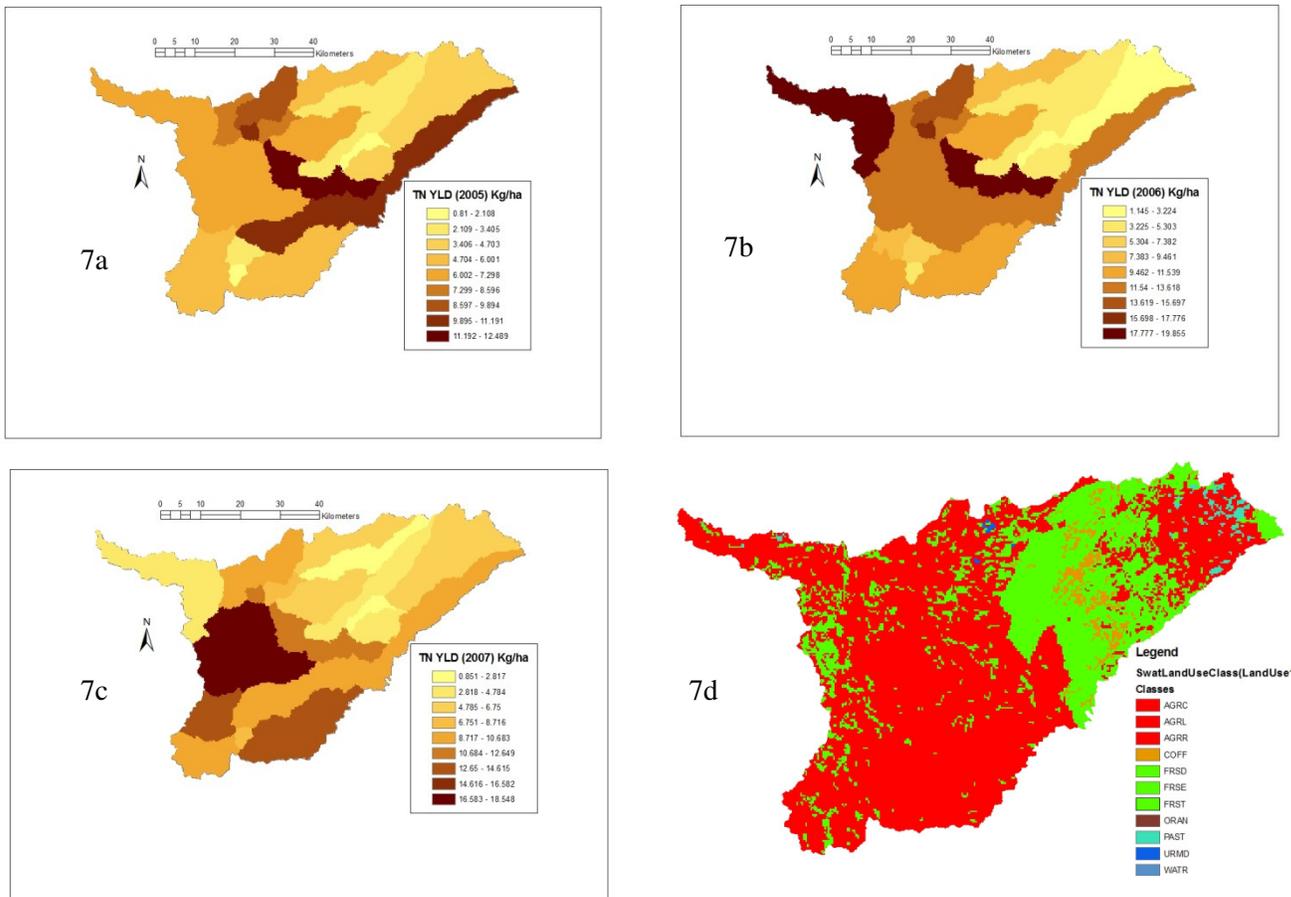


Fig. 7a, b, c & d Total nitrogen yield in the year 2005 (7a), 2006 (7b), 2007 (7c) and land cover in the year 2009 (7d)

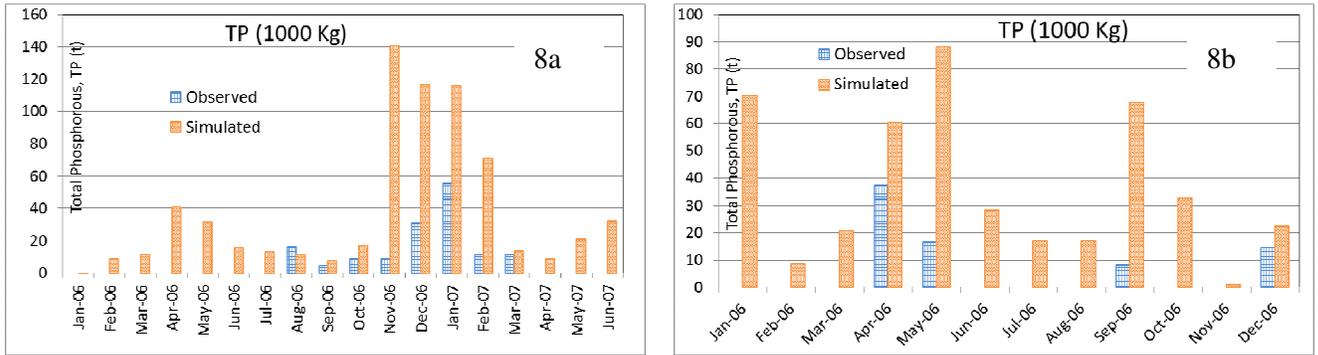


Fig. 8a & b Observed and simulated total nitrogen in the calibration period (8a) and validation period (8b)

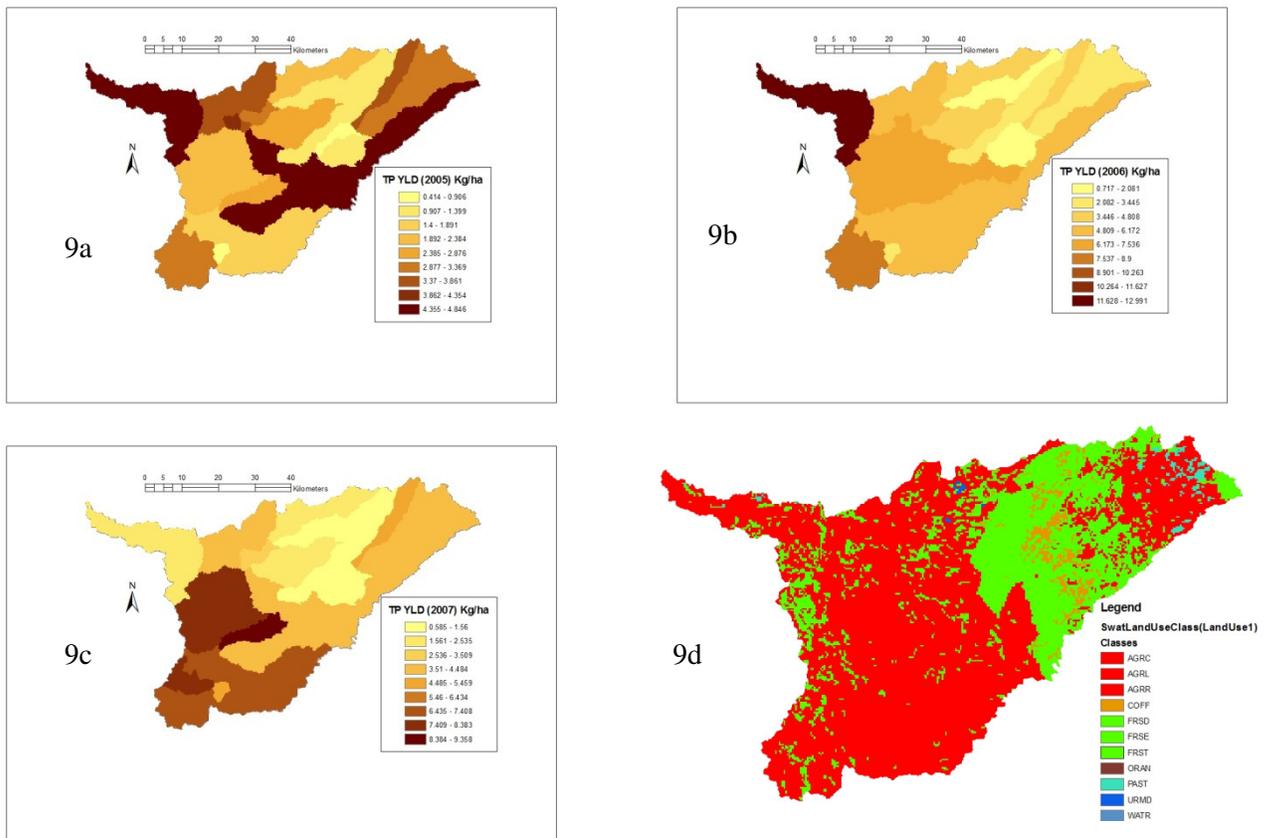


Fig. 9 a, b, c & d Total phosphorous yield in the year 2005 (9a), 2006 (9b), 2007 (9c) and land cover in the year 2009 (9d)

Conclusions

Annual average aggregate simulated sediment runoff load to Lake Victoria from Sondu watershed is 80,000 t/yr and is composed of mainly silt, annual stream

flow is 35 m³/s while total nitrogen and total phosphorous loads are 3,388 t/yr & 312 t/yr respectively. The pollution load estimates are reasonably within range of findings by LVEMP (2002 & 2005). Model

performance was reasonable in consideration of the circumstances of low observed data and guiding literature on evaluation of model performance. Modeling of stream flow was comprehensive, and had a lot of flexibility in parameter calibration because it had better data coverage. Although the model simulation of sediments and nutrients did not perfectly fit into observed values, it captured the patterns of low and high seasons.

The high water yielding areas are steep, slopy and agricultural crop dominant areas in the upstream (Kuresoi/Kisii/Kericho regions) of the watershed. Sediments and nutrients are relatively highly generated from agricultural crop areas at downstream (Sonde region), central (Sotik region) and upstream West side (Lower Kisii/Nyamira region) of the watershed. However, the upstream West of the watershed, characterized by dense human population, is a relatively high yielding zone for nutrients and not for sediments. Also, downstream area is not relatively high water runoff yielding but it ranks higher in sediment yield.

High sediment yield periods were Feb-April and Nov-Jan and directly correlate with high rainfall seasons. April-May and Oct-Dec are high nutrient yield seasons with exception of Nov-Dec of 2005 and 2010 which were low rainfall seasons. Feb-April is land preparation and crop planting season in the watershed and this partly explains the peak in sediment and nutrients. The sediment peaked while nutrient closely lagged behind by one to two months on average. River flow temporal variation was in synchrony with that of sediments.

Use of GIS tools in Lake Victoria basin provides a platform to model watersheds with improved temporal-spatial resolutions. This study should be considered as part of

initial attempts to simulate runoff pollution load in the watershed. Assessment of the performance of the model had to take into consideration the limitation of low availability of observed data. Continuous refinement of parameters and increased data collection to improve modeling process is recommended.

References

- Betrie, G.D., Mohamed, Y.A., Van Griensven, A. and Srinivasan, R., (2011). Sediment management modelling in the Blue Nile basin using SWAT model. *Hydrol. Earth Syst. Sci.*, 15: 807-818.
- 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. and Muhandiki, V.M., (2014a). Estimation of nutrient export coefficients in Lake Victoria basin. *Journal of Water and Environment Technology*, 12(3): 231-244.
- Cheruiyot, C.K., and Muhandiki, V.M. (2014b). Review of estimation of pollution load to Lake Victoria. *European Scientific Journal*, 10(5): 24-37.
- COWI Consulting Engineers (2002). The Integrated water quality/limnology study. LVEMP, Part II Technical Report, East African Community, Arusha, Tanzania.
- Harmel, R.M., Cooper, R.J., Slade, R.M., Haney, R.L. and Arnold, J.G. (2006). Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Trans. ASAE.*, 49 (3): 689-701.

- Jayakrishnan, R., Srinivasan, R., Santhi, C. and Arnold, J.G. (2005). Advances in the application of the SWAT model for water resources management. *Hydrological Processes*, 19: 749-762
- Karim, A.C. (2012). SWAT-CUP User's Guide. Ver. 5.1.4. Eawag: Swiss Federal Institute of Aquatic Science and Technology.
- Kayombo, S. and Jorgensen, S.E. (2005). Lake Victoria: experience and lessons Learned Brief. In: Lake Basin Management Initiative: Experience and Lessons Learned Briefs (on CD). pp. 431-446. International Lake Environment Committee Foundation (ILEC), Kusatsu, Japan.
- Kimwaga, R.J., Mashauri, D.A., Bukirwa, F., Banadda, N., Wali, U.G. and Nhapi, I. (2011). Modeling of non-point source pollution around Lake Victoria using SWAT model: A case of Simiyu catchment Tanzania. *The Open Environmental Engineering Journal*, 4: 112-123.
- Lake Victoria Environmental Management Project (LVEMP). (2005). Lake Victoria environment report on water quality and ecosystems status: Winam Gulf and river basins in Kenya, ministry of water and irrigation, Kisumu, Kenya.
- Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Binger, R.L., Harmel, R.D., Veith, T.L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* 50(3): 885-900.
- Muyodi, F.J., Bugenyi, F.W.B. and Hecky, R.E. (2010). Experiences and lessons learned from interventions in Lake Victoria Basin: The Lake Victoria Environmental Management Project. *Lakes and Reservoirs: Research and Management* 15: 77-88.
- Opere, A.O. and Okello, B.N. (2011). Hydrologic analysis for river Nyando using SWAT. *Hydrol. Earth Syst. Sci. Discuss*, 8: 1765-1797.
- Santhi, C., Arnold, J.G. and Williams, J.R., Dugas, W.A., Srinivasan, R., Hauck, L.M. (2001). Validation of the SWAT model on a large river basin with point and non-point sources. *Journal of American Water Resources Assoc.*, 37(5): 1169-1188.
- 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, Netherlands.
- Scheren, P.A.G.M., 2005. Integrated water pollution assessment in data and resource poor situations. In: Proceedings of the 11th World Lakes Conf. (2), 315-320, Nairobi, Kenya. Oct 31-Nov 4.
- 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 Africa Research and Development*, 25: 129-143.
- Scheren, P.A.G.M., Zanting, H.A., Lemmens, A.M.C. (2010). 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.
- Stefan Liersch. (2003). The Program pcpSTAT, Berlin.