CALIBRATION OF HYDROLOGICAL STREAMFLOW MODELING USING MODIS

Manithaphone Mahaxay[1], Wanchai ARUNPRAPARUT[2], Yongyut TRISURAT[2], Nipon TANGTHAM[3]
[1]UNESCO Bangkok Office, 920 Sukhumvit Rd, Bangkok 10110, Thailand
Email: manithaphone@gmail.com
[2]Faculty of Forestry, Kasetsart University, Bangkok 10900, Thailand
Email: fforwca@ku.ac.th, and fforyyt@ku.ac.th
[3]Forest Research Center, Faculty of Forestry, Kasetsart University, Bangkok 10900, Thailand
Email: ffornpt@ku.ac.th

KEY WORDS: Land Use and Land Cover (LULC); MODIS; SWAT; Modeling; Streamflow

ABSTRACT: Land use and land cover play an important role in mathematical hydrological modeling. As in a number of countries around the world, available land use and land cover are not always updated to reflect the most current situation. In this regard, the objectives of this study were to investigate the potential capability of moderate resolution satellite imagery such as MODIS, acquired in 2010 for updated land use and land cover. This issue was illustrated through the application of the most current land use and land cover as one of the data inputs of the SWAT model in the Tonlé Sap Lake Basin, a sub-basin of the Mekong River. The streamflow was tested using moderate resolution land use and land cover of 500 meters. The result showed good agreement between observed and simulated values for both monthly and daily streamflow. The statistical evaluation results at a monitoring station for model calibration and validation showed that the R^2 for daily and monthly values range from 0.76 to 0.88 and 0.86 to 0.89 respectively (Table 2), whereas the Nash-Sutcliffe efficiency daily and monthly values range between 0.75 to 0.85 and 0.76 to 0.87 respectively. The simulation result demonstrates that land use and land cover at moderate resolution, based on MODIS imagery, holds considerable potential as an effective water quantity modeling tool. An additional level of confidence is provided by the notion that the methods described here could be applied in similar watershed conditions.

INTRODUCTION

Land use and land cover (LULC) datasets, which are important in a watershed for hydrological and environmental modeling, require accurate LULC datasets to parameterize the physical system being simulated (Burian *et al*., 2002). It is important that land-cover data be based on the most current data available, since the land-cover changes over time (Chen *et al*., 2005). In watersheds, where LULC change takes place over the modeling period, using a single land-use geospatial data is not a true representation of the watershed condition (Pai and Saraswat, 2011). The LULC data are one of the essential inputs for the Soil and Water Assessment Tool (SWAT) model to which this research was applied.

SWAT is considered one of the most suitable physically-based models for simulating hydrological condition and is one of the most widely used watershed-scale water-quality models in the world. Nearly 600 peer-reviewed SWAT-related journal articles have been published and hundreds more have been published in conference proceedings and other formats (Gassman *et al*., 2010). Rossi *et al.* (2009) pointed out that SWAT can potentially be used as an effective water quantity tool within Mekong basin. In which, SWAT model has been setup to simulate streamflow in each Mekong sub-basin (Mainuddin *et al*., 2010). In the Mekong Sub-basin the SWAT model has been calibrated using the most up-to-date available land use data of 2003 generated from Landsat image against available streamflow data for the period 1985-2000 (Johnston *et al*., 2003). The SWAT simulation result provided daily estimates of flow for 138 sub-basins covering entire the Lower Mekong basin except the delta south of Phnom Penh (Rossi *et al*., 2009). However, whether using simple or complex models, an accurate LULC dataset with an appropriate spatial or temporal resolution and level of detail is paramount for reliable predictions (Huang, 2013).
Landsat imagery is widely used to produce high resolution LULC data covering large river watershed. Although high resolution satellite imagery data can be extremely useful for LULC change detection and monitoring efforts, it can be difficult to obtain an image over the entire study area during a particular timeframe. In other words, only rarely is it possible to generate more than one scene of high resolution satellite imagery in a day. The revisit characteristics of the satellites, as well as the presence of cloud cover, can limit the availability of data (WRP, 1994). In addition, spatial data, including land use, are usually expensive to obtain. This paper explores alternatives aimed at overcoming the limitations of LULC for hydrological modeling. To achieve the overall goal of the research, the status of LULC in 2010 was mapped out using both GIS analysis and remote sensing data such as MODIS with 500 m resolution. The principle objective of this study is to assess whether free-data-MODIS can be effectively applied as an input for hydrological modeling. It is expected that the results of this study will contribute useful hydrologic information regarding the possibility of moderate-resolution of LULC data for large river watershed assessments.

STUDY AREA

Tonlé Sap Lake Basin is located in the northwest of Cambodia, between approximately latitudes 102° 15' to 105° 50'E and longitudes 11° 40' to 14° 28'N. The Tonlé Sap Lake Basin is a sub-catchment of the Mekong basin. The total drainage area of Tonlé Sap Lake Basin is approximately 85,786 km^2, including a permanent lake area of around 2,350 km^2. That is approximately 10.8% of the total area of the Mekong basin (Mekong River Commission, 2003). The majority of the catchment is located in Cambodia and only 5% is in Thailand (Figure 1). Ground altitudes range from 1 m to 1,500 m above sea level. About one third of the area is covered by forests that consist of a mixture of deciduous trees. There are agricultural areas and numerous small settlements as well.

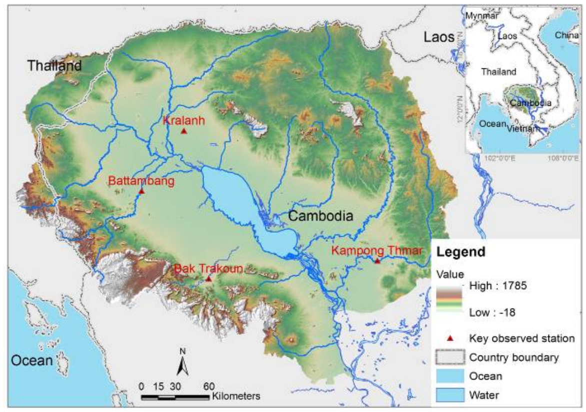

Figure 1 Tonlé Sap watershed.

MATERIALS AND METHODS

Materials

Data: Time series of 16-day composite MODIS imagery of MOD09A1 with 500 m resolution was acquired for LULC classification and mapping to temporal scenarios of spatial LULC of 2010. The other spatial data used are soil map of 50 m resolution based on FAO\UNESCO (1988) classification system up to level three category and DEM data of 50 m resolution. The other hydro-climatological quantities have been used from available gauges over the study area. Figure 2 shows sets of required spatial data for SWAT hydrological modeling.

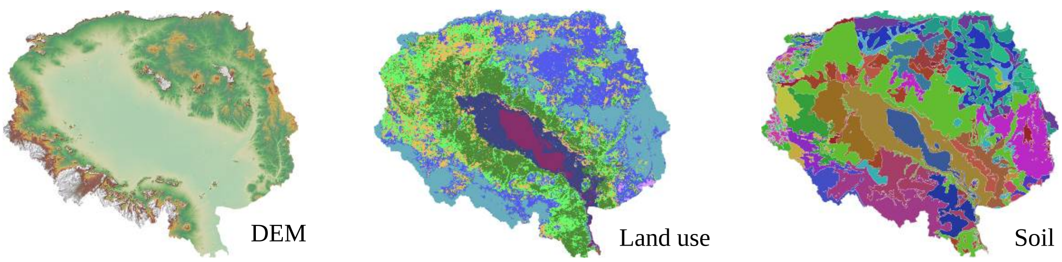

Figure 2 Spatial data for SWAT model input.

Software: Most of the data preparation and analysis in this research was carried out using ArcGIS 10.1. Some specific image processing operations were executed using the ERDAS Imagine software Version 8.0 (ERDAS Imagine is a remote sensing application designed for geospatial applications). Other types of software employed are Idrisi 15.0 for modeling of LULC; ArcSWAT Version 2012.10_1.7 for streamflow modeling and MRT (MODIS Reprojection Tool) for MODIS reprojection and transformation.

Methods

Description of SWAT model: SWAT is a model developed by the United States Department for Agriculture, Agricultural Research Service (USDA-ARS). It consists of a river-basin scale model developed to quantify the impact of land management practices in large, complex catchments (Neitsch *et al.*, 2001). The main components of SWAT include hydrology, weather, sedimentation, soil temperature, crop growth, nutrients, pesticides and agricultural management. The model can be used to predict impacts of land management practices on water, sediment and agricultural chemicals in catchments (Neitsch *et al.*, 2002; Chaplot, 2004). The SWAT model simulates hydrology as a two-component system, composed of land hydrology and channel hydrology. The land portion of the hydrologic cycle is based on a water mass balance. Soil water balance is the primary consideration by the model in each hydrological response unit (HRU), which Arnold *et al.* (1998) represent as follow:

$$SW_t = SW + \sum_{i=1}^{t} (R_i - Q_i - ET_i - P_i - QR_i) \quad (1)$$

Where *SW* is the soil water content; *i* is time in days for the simulation period *t*; and *R*, *Q*, *ET*, *P* and *QR* respectively are the daily precipitation, runoff, evapotranspiration, percolation and return flow.

Description of SWAT model: LULC data used for this hydrological modeling were derived from satellite MODIS imagery. The LULC classification of 2010 LULC was carried out using supervised classification and every training site was carefully selected. Post-classification was performed based on existing land use map of 2003 generated from Landsat, DEM and ground survey. Accuracy assessment was also executed based on those field surveys and existing land use data. Overall classification accuracy was greater than 80%. To make LULC data useable for SWAT, ArcSWAT interface requires a table linking the values represented to LULC types already defined in the model. Hence, the look-up table that converts the LULC classification codes to SWAT land cover/plant codes was created manually in "ASCII .txt" format. Table 1 represents a look-up table for LULC categories conversion. The soil units were also translated into SWAT user soil database. ArcSWAT creates the hydrologic response unit by combining Digital Elevation Model (DEM), soil and slope. Once DEM, land use and land cover, and soil data have been overlaid, the hydrological Response Units (HRUs) were generated.

Table 1 Look-up table for the land use database use in SWAT

Land use and land cover class		Land use class No.	SWAT Database
Forest land	Evergreen	1	FRSE
	Deciduous	2	FRSD
	Plantation	3	PLAN
	Shrubland	4	SHRB
Crop land	Upland	5	AGRL
	Lowland paddy	6	PDDY
Others	Wetland	7	WETL
	Built-up land	8	URBN
	Water (rivers, lakes)	9	WATR

Rainfall data from 31 stations with time-series data from 1980 to 2008 were used as input data in SWAT. Additional rainfall data related to 2009 and 2010 were compensated by Global Weather Data for SWAT at http://globalweather.tamu.edu/. When all inputs were successfully entered, simulation was activated. Sensitivity analysis was carried out for help in determining the sensitivity of parameters by comparing variances in output caused by variability in the inputs. It also facilitates the selection of important and influential parameters for a model calibration by indicating the parameters that display higher sensitivity in output due to input variability. Streamflow simulations were calibrated for each LULC data annually, while the other spatial data are fixed. Overall procedure of SWAT application in this research is shown in Figure 3.

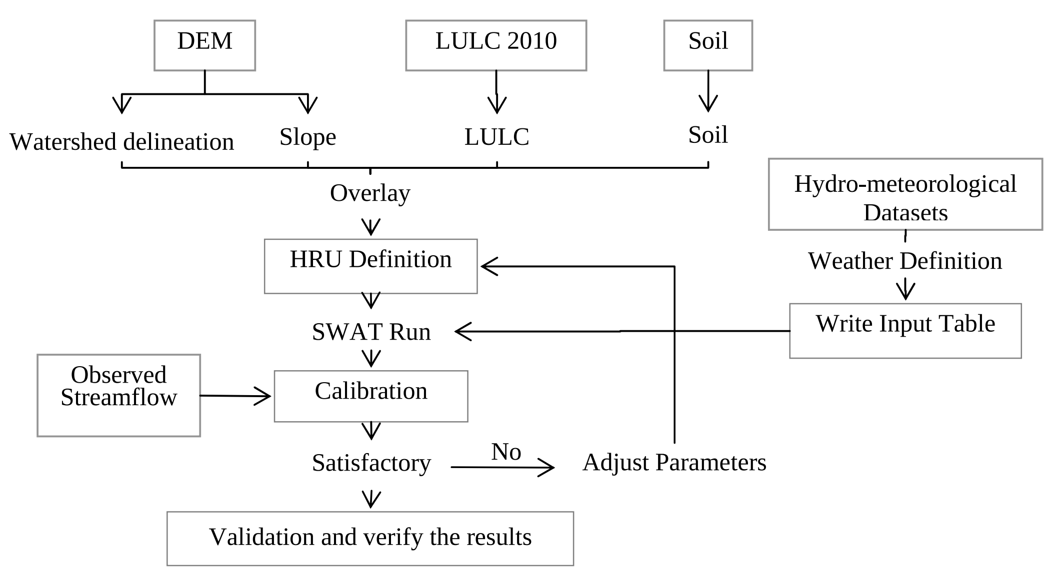

Figure 3 Flow chart of streamflow modeling process.

The streamflow was run at the outlet of selected hydrological stations at daily and monthly time steps for the period January through December 2010. Calibration was performed on the 1997 to 2009 years, while the years from 1980 to 1996 were used for model warm-up period. To verify the results, the performance of the model in simulating streamflow was evaluated using Nash–Sutcliffe efficiency (*ENS* or *NSE or* E_{NS}) and the coefficient of determination (R^2) (Eisenhauer, 2003). The Nash-Sutcliffe statistic is a measure of how well the observed variance is simulated (Nash and Sutcliffe, 1970). The equations used were as follows:

$$E_{NS} = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2} \qquad (2)$$

Where O_i and P_i are the observed and simulated data, respectively; $\overline{O}$ is the average of the observed data and n is the total number of data records.

$$R^2 = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \overline{Y})^2} \qquad (3)$$

Where Y_i denotes the value of the i^{th} dependent variable, $\overline{Y}$ is the mean of the dependent variable and $\hat{Y}_i$ is the i^{th} fitted value.

RESULTS AND DISCUSSION

Sensitivity analysis

Sensitivity analysis has been carried out for each available observed streamflow data of each LULC SWAT project. Sequential Uncertainty Fitting (SUFI2) for the calibration of uncertainty in procedure was used for this analysis. Six parameters were found to be sensitive, with relative sensitive values in the range of 0.031 to 0.034. The most sensitive parameters are threshold depths of water in the shallow aquifer for “revap” to occur (REVAPMN.gw), base flow alpha factor (Alpha_Bf), groundwater "revap" coefficient (Gw_Revap), soil evaporation compensation factor (ESCO), initial SCS Curve Number II value (CN2) respectively. These sensitive parameters were considered for model calibration. The remaining parameters had no significant effect on streamflow simulations. Changes in their values do not cause significant changes in the model output.

Calibration and validation to estimate streamflow: simulation result at the observation station (Figure 4) was discussed.

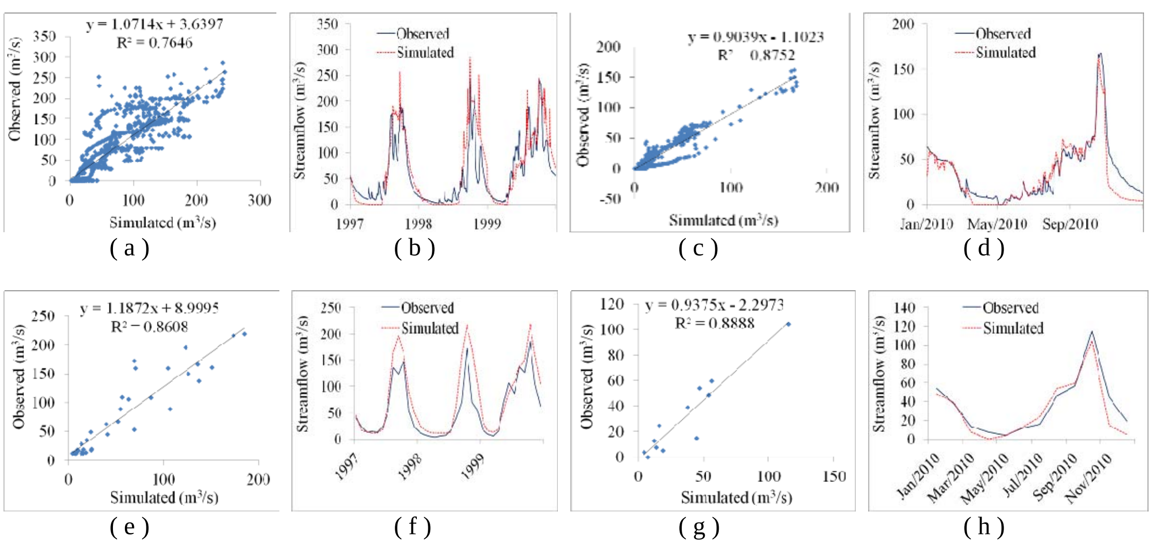

Figure 4 Simulation results: (a) and (b) Daily calibration result; (c) and (d) Monthly calibration result;

(e) and (f) Daily validation result; (g) and (h) Monthly validation result.

The monthly calibration results have shown better agreement between monthly observed and simulated flows in both calibration and validation processes. The overall result of the E_{NS} and R^2 are as high as 0.87 and 0.89 respectively (Table 2). Based on the statistical analysis of model evaluation results, conclusion of whether MODIS can be effectively applied as an input to SWAT interface for hydrological streamflow modeling is noticeable.

Table 2 Calibration and validation results

	Calibration		**Validation**	
	E_{NS}	R^2	E_{NS}	R^2
Daily	0.75	0.76	0.85	0.88
Monthly	0.76	0.86	0.87	0.89

According to Benaman *et al.*, (2005), model simulation can be judged as satisfactory if R^2 is greater than 0.6 and E_{NS} is greater than 0.5. Hence, the results of the study are consistent with these accuracy simulations of LULC parameters. Any inaccuracies in the model are caused by gaps in the data and lack of accurate and efficient input data such as rainfall, temperature and evapotranspiration. Therefore, increases in the efficiency of the model depend on such data inputs as well as on suitable distribution of the measuring stations over the watershed.

CONCLUSIONS

This study has demonstrated that moderate resolution of non-commercial and freely-available satellite imagery like MODIS holds considerable potential for application in hydrological modeling.

However, the use of other hydrological models would be more beneficial for the hydrological modeler in order to enhance our understanding of alternative MODIS-based LULC as an input parameter for hydrological modeling.

In addition to the modeling tool, the assessment of LULC data input capability would be more beneficial if simulation is tested by a number of hydrological parameters other than streamflow, such as surface run-off, water quality, etc.

ACKNOWLEDGEMENTS

The author and co-authors would like to express their appreciation to colleagues from the Department of Water Resources in Cambodia and the Mekong River Commission Secretariat for facilitating and providing data for this research. Special thanks to Dr. Lia Genovese for editing this manuscript.

REFERENCES

Arnold, J. G., P. Srinivasan and R. S. Muttiah. 1998. Large area hydrologic modeling and assessment. Part I. Model development. *The Journal of the American Water Resources Association* (JAWRA) 34, pp. 73–89

Benaman, J., A. S. Christine and A. H. Douglas. 2005. Calibration and validation of soil and water assessment tool on an agricultural watershed in upstate New York. *Journal of Hydrologic Engineering*, 10(5), pp. 363-374.

Burian, S., M. Brown and T. McPherson. 2002. Evaluation of land use and land cover datasets for urban watershed modeling. *Water Science and Technology*. 45, 269–276.

Chen P., M. D. Luzio and J. G. Arnold. 2005. "Impact of Two Land-cover Data Sets on Streamflow and Total Nitrogen Simulations Using A Spatially Distributed Hydrologic Model". *Pecora 16 "Global Priorities in Land Remote Sensing"*, October 23-27, 2005. Sioux Falls, South Dakota.

Chaplot, V., A. Saleh and D. B. Jaynes. 2004. Predicting water, sediment and NO3–N loads under scenarios of land-use and management practices in a flat watershed. *Water, Air, & Soil Pollution*. 154, pp. 271–293.

Eisenhauer, J. G. 2003. *Regression through the origin*. Teaching Statistics 2003. 25, pp. 76–80.

FAO, 1988. FAO/UNESCO Soil Map of the World, Revised Legend, with corrections and updates. *World Soil Resources Report 60, FAO, Rome*. Reprinted with updates as Technical Paper 20, ISRIC, Wageningen, 1997.

Gassman P. W., J. Arnold, R. Srinivasan and M. Reyes. 2010. The Worldwide Use of the SWAT Model, Technological Drivers, Networking Impacts, and Simulation *Trends. 21st Century Watershed Technology, Improving Water Quality and Environment Conference Proceedings*, 21-24 February 2010, Universidad EARTH, Costa Rica 701P0210cd. (doi, 10.13031/2013.29418) .

Huang, J., P. Zhou, Z. Zhou and Y. Huang. 2013. Assessing the Influence of Land use and Land Cover Datasets with Different Points in Time and Levels of Detail on Watershed Modeling in the North River Watershed, China. *International Journal of Environmental Research and Public Health* 10, pp. 144-157.

Johnston R., P. Rowcroft, K. G. Hortle and C. McAlister. 2003. Hydrological models of the Lower Mekong Basin at Mekong River Commission, Resources Allocation and Optimization Model. Environmental Inventory and Valuation, Environmental Flow, Wetlands Inventory and Valuation, Valuation of Mekong Fisheries. Paper presented at *Workshop on Integrating Environmental Impacts into Water Allocation Models of the Mekong River Basin*, University of Economics, Ho Chi Minh City, 15 December 2003.

Mainuddin, M., C. T. Hoanh, K. Jirayoot, A. S. Halls, M. Kirby, G. Lacombe and V. Srinetr. 2010. Adaptation options to reduce the vulnerability of Mekong water resources, food security and the environment to impacts of development and climate change. *CSIRO, Water for a Healthy Country National Research Flagship*. 152.

Mekong River Commission. 2003. *State of the Basin Report*. 2003, Mekong River Commission (MRC), Phnom Penh, Cambodia. 316. ISSN, 1728-3248.

Nash, J. E. and J. Sutcliffe. 1970. River flow forecasting through conceptual models. Part I - A discussion of principles. *Journal of Hydrology* 10, pp. 282–290.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry and J. R. Williams. 2001. *Soil and Water Assessment Tool* - Version 2000-Theoritical Documentation, Texas, USA.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, J. R. Wiliams and K.W. King. 2002. Soil and Water Assessment Tool Theoretical Documentation Version 2000. *GSWRL Report 02-01*, *BRC Report 02-05, TR-191*. College Station, Texas. Texas Water Resources Institute.

Pai, N. and D. Saraswat. 2011. Swat2009_Luc: A Tool To Activate The Land Use Change Module In Swat 2009. *American Society of Agricultural and Biological Engineers*. 54(5), pp. 1649-1658.

Rossi, C. G., R. Srinivasan, K. Jirayoot, T. Le Duc, P. Souvannabouth, N. Bin and P. W. Gassman. 2009. Hydrological Evaluation of the Lower Mekong River Basin with the Soil and Water Assessment Tool Model. *International Agricultural Engineering Journal*. 18(1-2), pp. 1-13

WRP, 1994. The Corps of Wetlands Research Program (WRP), 2004. Remote Sensed Data, Information for Monitoring Dynamic Wetland Systems. *Remote Sensing/GIS Support Center*, U.S. Army Cold Regions Research and Engineering Center, May, 2004. Available source. http://el.erdc.usace.army.mil/elpubs/pdf/wgsw2-1.pdf. November 2013.