Evaluating the impact of land use uncertainty on the simulated streamflow and sediment yield of the Seyhan River basin using the SWAT model

Ashraf EL-SADEK[1,*], Ahmet İRVEM[2]

[1]Desert Research Center, Ecology and Dryland Agriculture Division, Cairo, Egypt
[2]Biosystem Engineering Department, Faculty of Agriculture, Mustafa Kemal University, Antakya, Turkey

Abstract: As a result of the increased availability of spatial information in watershed modeling, several easy to use and widely accessible spatial datasets have been developed. Yet, it is not easy to decide which source of data is better and how data from different sources affect model outcomes. In this study, the results of simulating the stream flow and sediment yield from the Seyhan River basin in Turkey using 3 different types of land cover datasets through the soil and water assessment tool (SWAT) model are discussed and compared to the observed data. The 3 land cover datasets used include the coordination of information on the environment dataset (CORINE; CLC2006), the global land cover characterization (GLCC) dataset, and the GlobCover dataset. Streamflow and sediment calibration was done at monthly intervals for the period of 2001–2007 at gauge number 1818 (30 km upstream of the Çatalan dam). The model simulation of monthly streamflow resulted in good Nash–Sutcliffe efficiency (NSE) values of 0.73, 0.71, and 0.68 for the GLCC, GlobCover, and CORINE datasets, respectively, for the calibration period. Furthermore, the model simulated the monthly sediment yield with satisfactory NSE values of 0.48, 0.51, and 0.46 for the GLCC, GlobCover, and CORINE land cover datasets, respectively. The results suggest that the sensitivity of the SWAT model to the land cover datasets with different spatial resolutions and from different time periods was very low in the monthly streamflow and sediment simulations from the Seyhan River basin. The study concluded that these datasets can be used successfully in the prediction of streamflow and sediment yield.

Key words: CORINE, GLCC, GlobCover, sediment, Seyhan River, SWAT

1. Introduction

Geographic information system (GIS) based distributed hydrologic models simulate the hydrologic processes using spatial parameters derived from geospatial data. These data mainly have information about relief, soil and land cover types, and intensity. Land cover has a great impact on the water quantity and quality in a river basin. Better estimation of land cover parameters improves the performance of the hydrologic model used. Appropriate spatial and temporal resolution of the used land cover improves the prediction of the hydrologic model (Huang et al., 2013). Several studies have been conducted to study the impact of land cover change on hydrology and water quality by (1) using readily available data (Cai et al., 2012; Yan et al., 2013), (2) using artificial land cover scenarios including farming practices (Chaplot et al., 2004; De Girolamo and Lo Porto, 2012; Mbonimpa et al., 2012), and (3) generating land use change scenarios using the land use change models (one such land use change model is the conversion of land use and its effects model (CLUE-s, Verburg et al.,

2002, 2004), which was used by Lin et al. (2007), Zhang et al. (2011), Zhou et al. (2013), and Park et al. (2011).

The oldest and most widely used global land cover dataset, the global land cover characterization (GLCC), was produced by the US Geological Survey, the University of Nebraska–Lincoln, and the European Commission's Joint Research Centre (JRC). The 1-km resolution results are based on unsupervised classification of the 1-km advanced very high-resolution radiometer (AVHRR; Eidenshink and Faundeen, 1994) 10-day normalized difference vegetation index composites spanning the period from April 1992 through March 1993. The first version (version 1.2) of the GLCC database was released to the public in November 1997. The updated version (version 2.0; Loveland et al., 2000) represents the same time period (April 1992 through March 1993) and is the one used for our study. The data can be obtained from http://edc2.usgs.gov/glcc/glcc.php. Numerous hydrological studies using the soil and water assessment tool (SWAT) model have relied on the GLCC dataset as a source of land cover information. For example,

* Correspondence: elsadek@nmsu.edu

Vu et al. (2011) used the GLCC dataset along with other internet-based datasets to derive the SWAT model for the streamflow simulation of the Da River across the transboundary regions of China and Vietnam, Chen et al. (2005) compared the SWAT simulation results from the GLCC land cover dataset and the national land cover data over the continental US, and Schoul et al. (2008) applied SWAT to model the blue and green water from 1496 subbasins over the entire continent of Africa.

In 2005, the European Space Agency, in collaboration with the European Environment Agency (EEA), Food and Agriculture Organization (FAO), Global Observation of Forest and Land Cover Dynamics, International Geosphere–Biosphere Program, European Commission's Joint Research Center (JRC), and the United Nations Environment Program, developed the GlobCover project (Arino et al., 2007), whose purpose was to create a land cover map based on Envisat's medium resolution imaging spectrometer (MERIS). The time period covered by this dataset is from January 2005 to June 2006. It was released in 2008 as GlobCover 2005. The data have a spatial resolution of 300 m. A second updated version of the data was delivered in 2010. Its land cover map is derived from an automatic and regionally tuned classification of a time series of global MERIS full resolution mosaics for the year 2009 (Arino et al., 2010). The global land cover map provides 22 land cover classes defined with the United Nations land cover classification system. The data can be downloaded free of charge from http://due.esrin.esa.int/globcover/. The GlobCover dataset has been used to provide land cover information for SWAT modeling of streamflow and water quality in the Kaiping reservoir, China (Nielsen et al., 2013).

The highest resolution of land cover that exists for Europe is the third version of the CORINE (CLC 2006; coordination of information on the environment) dataset of land cover data for 2006 (earlier versions were for 1990 and 2000) produced by the EEA. The data covers most of the EU countries and 13 partner countries in eastern and central Europe including Turkey. CLC 2006 was produced using a combination of images from 2 satellites (French Spot 4&5 and Indian IRS P6) in a multitemporal satellite image coverage. That data covers 38 European countries with a spatial resolution of 100 m (http://www.eea. europa.eu/data-and-maps/data/corine-land-cover-2006-raster-2). Many research studies have used the CORINE dataset as input for SWAT to model the hydrology of European catchments; examples of such studies include Nasr et al. (2007) in Ireland, Candela et al. (2012) in Italy, and Koch et al. (2013) in Germany.

The relative impacts of different land cover datasets on streamflow and sediment yield have not yet been described sufficiently. Many studies have shown that streamflow is not highly affected by either the origin of the chosen dataset or the spatial resolution of the land cover dataset; however, sediment, nitrogen and total phosphorus loads are highly affected by which land cover dataset is used as well as the spatial resolution of that data. For example, Huang et al., (2013) compared 3 types of land cover datasets derived from Landsat thematic mapper satellite imagery from 2007 and 2010 and an ETM+ image of 2002. They used 3 different categories of land cover detail (10, 5, and 3) in the SWAT model. The model was applied to the Jiulong River Basin in China. Their results showed that there is relatively little impact on the daily and monthly streamflow by using the 3 land cover datasets at different levels of detail. However, ammonia–nitrogen (NH4+–N) and total phosphorus (TP) loads were highly affected by choice of land cover dataset. The relative differences in predicted monthly NH4+–N using the 2007 and 2010 LULC datasets were −11.0% and −7.8% as compared to the 2002 LULC dataset, respectively. However, for the predicted monthly TP loads, they were −4.8% to −9.0%, respectively when using the 2 LULC datasets from 2007 and 2010 compared with that from 2002. Cotter et al. (2003) also showed that stream flow is not significantly affected by the land cover data resolution; however, sediment, NO_3–N, and TP were greatly affected by the land cover resolution with relative errors of 19%, 11%, and 41%, respectively, when the coarsest land cover data were used (1000 m) as compared to the 30-m resolution data.

Known as the longest river in the country that flows into the Mediterranean Sea, the Seyhan River is located in the Eastern Mediterranean region of Turkey, and it runs from the Taurus Mountains in the north to the Mediterranean Sea in the south. The basin has 4 dams along the Seyhan River and its tributaries: Seyhan dam, Çatalan dam, and Nergizlik dam located at the south of the basin and the Bahçelik dam at the north of the basin. Çatalan Dam Lake has a storage capacity of 2.1 km³ with a reservoir surface area of 82 km². The mean annual discharge at the Çatalan dam is 163 m³ s⁻¹ (Acar and Dincer 2005). The measured averaged sediment yield for station number 1818 (30 km upstream from the Çatalan dam), which has a contributing area of 13,846 km², is 1.51 t ha⁻¹ year⁻¹ and the total amount of sediment is 2,090,746 t year⁻¹ (Irvem et al., 2007).

The watershed model used in this study is SWAT. The land cover data in SWAT is used in interaction with soil data and terrain to create the hydrologic response units (HRUs) that are used to assess water, sediment, and chemical yields. In this study, we describe how land cover data from different sources can affect the SWAT model outputs for the Seyhan River basin. The objective of this study was to determine the impact of using 3 different land cover datasets, i.e. GLCC, GlobCover, and CORINE, that are different in terms of the time period being represented and spatial details on the stream flow and sediment yield from the Seyhan River basin based on SWAT model simulation.

2. Materials and methods

2.1. Study area

The Seyhan River basin is located between 34.25–37.0°E and 36.5–39.25°N, and the basin covers an area of 20,164 km². The study area includes the drainage area of station number 1818 (13,910 km²), which is located 30 km upstream of the Çatalan dam (Figure 1). The dominant soil type is loam (49.69%). The watershed receives a mean annual precipitation of 708.5 mm with annual average maximum and minimum temperatures of 19.78 and 7.74 °C, respectively, as determined by data from the period of 2000 to 2012.

2.2. Model description

SWAT is a semidistributed, physically based, time continuous model designed to predict the impact of land management practices on water, sediment, and agricultural chemical yield in large river basins (Arnold et al., 2012). Major model components describe processes associated with weather, hydrology, soil properties, water movement, plant growth, nutrients, pesticides, bacteria and pathogens, and land management. The model uses a digital elevation model (DEM) to delineate the watershed boundary and to divide the watershed into multiple subwatersheds that are then further subdivided into hydrologic response
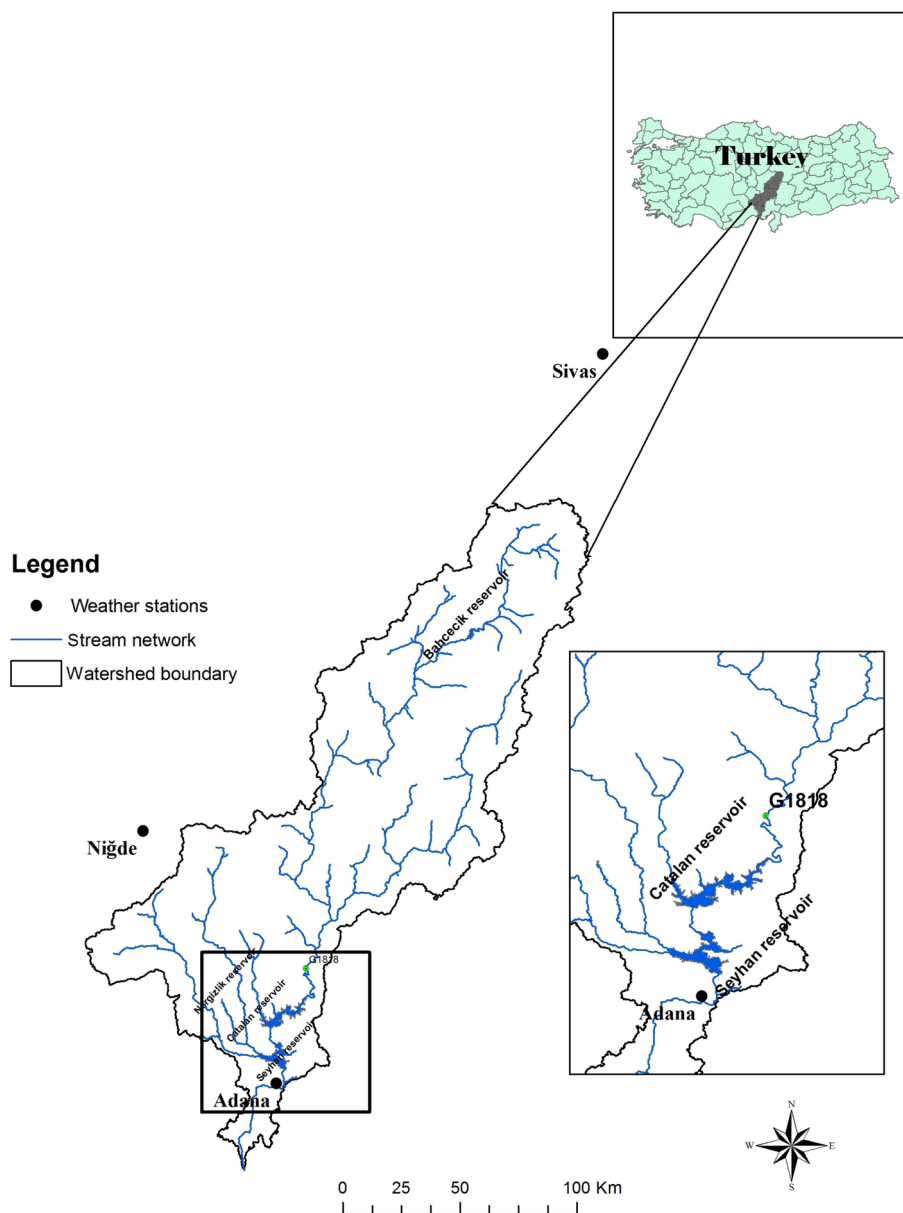

Figure 1. Location of the study area, gauging station (G1818), and weather stations.

units (HRUs) that consist of similar land use, slope, and soil characteristics (Arnold et al., 2012). Model outputs include surface runoff, evapotranspiration, groundwater, lateral flow, sediment, nutrient, and pesticide yields.

The surface runoff can be simulated by using the USDA Natural Resources Conservation Service curve number method (USDA-NRCS, 2004) or the Green and Ampt infiltration model (Green and Ampt, 1911). The evapotranspiration can be estimated by using the Hargreaves, Priestly–Taylor, and/or the Penman–Monteith method. Ground water, including shallow and deep aquifer recharge, is routed using empirical and analytical techniques in SWAT (Neitsch et al., 2005). The hydrological cycle in SWAT is based on the water balance equation (Eq. (1)):

$$SW_t = SW_0 + \Sigma_{i=1}^{t}(R_{day} - Q_{surf} - E_a - W_{seep} - Q_{gw}), \quad (1)$$

where SW_t is the final soil water content (mm), SW_0 is the initial soil and water content (mm), t is time (days), R_{day} is the amount of precipitation on day i (mm), Q_{surf} is the amount of surface runoff on day i (mm), E_a is the amount of evapotranspiration on day i (mm), W_{seep} is the amount of water entering the vadose zone from the soil profile on day i (mm), and Q_{gw} is the amount of return flow on day i (mm).

Erosion is estimated using the modified universal soil loss equation (MUSLE) (Williams, 1975; Neitsch et al., 2005) as given by Eq. (2):

$$Sed = 11.8(Q_{surf} \times q_{peak} \times area_{hru})^{0.56} \\ \times K_{USLE} \times C_{USLE} \times P_{USLE} \times LS_{USLE} \times CFRG, \quad (2)$$

where Sed is the sediment yield on a given day [t], Q_{surf} is the surface runoff volume [mm ha^{-1}], q_{peak} is the peak runoff rate [m^3 s^{-1}] (Eq. (3)), $area_{hru}$ is the area of an HRU [ha], K_{USLE} is the USLE soil erodibility factor, C_{USLE} is the USLE soil cover factor, P_{USLE} is the USLE support practice factor, L_{SUSLE} is the terrain shape factor (slope and length of a slope), and $CFRG$ is the coarse fragment factor.

The peak runoff rate in the SWAT model is given by

$$q_{peak} = \frac{a \times q \times A}{360 \times tc}, \quad (3)$$

where q_{peak} is the peak runoff rate (m^3 s^{-1}), q is runoff (mm), A is the HRU area (ha), tc is the time to concentration (h), and α is a dimensionless parameter that expresses the proportion of total rainfall that occurs during tc.

2.3. SWAT input
A GIS interface to SWAT, ArcSWAT (Winchell et al., 2009), was used to automate the development of model input parameters. A 30-m digital elevation model from ASTER-GDEM (advanced spaceborne thermal emission

and reflection radiometer - global digital elevation model) was obtained for the study area from http://asterweb.jpl. nasa.gov/gdem.asp, and the DEM layer is presented in Figure 2. The dataset was used to derive information about the topographic characteristics of the watershed: elevation, watershed boundary, flow path, subbasin area, slope, and river channel elevation. Eight years of daily weather data (i.e. daily minimum and maximum temperature, solar radiation, relative humidity, and average wind speed) were obtained from the US National Climatic Data Center, global summary of the day for the period from 2000 to 2007. The data are available online at ftp://ftp. ncdc.noaa.gov/pub/data/gsod/. However, the tropical rainfall measurement mission (product 3B42) described in Huffman et al. (2007) was used as the source of the daily precipitation data. These data have a spatial resolution of 0.25° × 0.25°, so that 17 grid points (one at each of the 17 pixel locations) covering the study area have been used. All soil data were obtained from the FAO/UNESCO 2003 Soil Map of the World CD-ROM; the soil map for our study area is shown in Figure 2.

2.4. Land cover datasets
Three sources of land cover datasets were used, i.e. GLCC, GlobCover, and CORINE. Tables 1–3 and Figure 3 show the distribution of the different land cover classes from each dataset over the study area and the corresponding SWAT land cover codes that were used in the simulation. Because each of the land cover datasets has a different number of land cover classes, a different number of HRUs were defined using the threshold values of 10% for land use, 10% for soil, and 20% for slope for the dominant land use, soil, and slope of individual subbasin areas. For the GLCC data, 357 HRUs were defined representing 11 land cover classes (Table 1), which reveals that the most dominant land cover type is grassland (42%). For the GlobCover dataset, the mosaic cropland and vegetation class are the dominant land use types (52.72%). This produced 519 HRUs representing 13 land use classes (Table 2). It was obvious that the CORINE land cover dataset, which has the highest number of land use classes (28), would produce the highest number of HRUs (595). The dominant land cover types are sparsely vegetated area (18.44%), natural grassland (13.23%), and arable land (11.77%). We included our land cover classes and soil types in the SWAT model's database by modifying the user soil and land cover files.

To more easily compare the impact that the different land cover datasets have on model output, the land cover categories in each dataset were classified according to the well-established Anderson land use/land cover classification system (Anderson et al., 1976) as presented in Tables 1–3. The system consists of 9 land cover categories (urban or built-up, agricultural, range, forest, water, wetland, barren, tundra, and perennial snow and ice), and
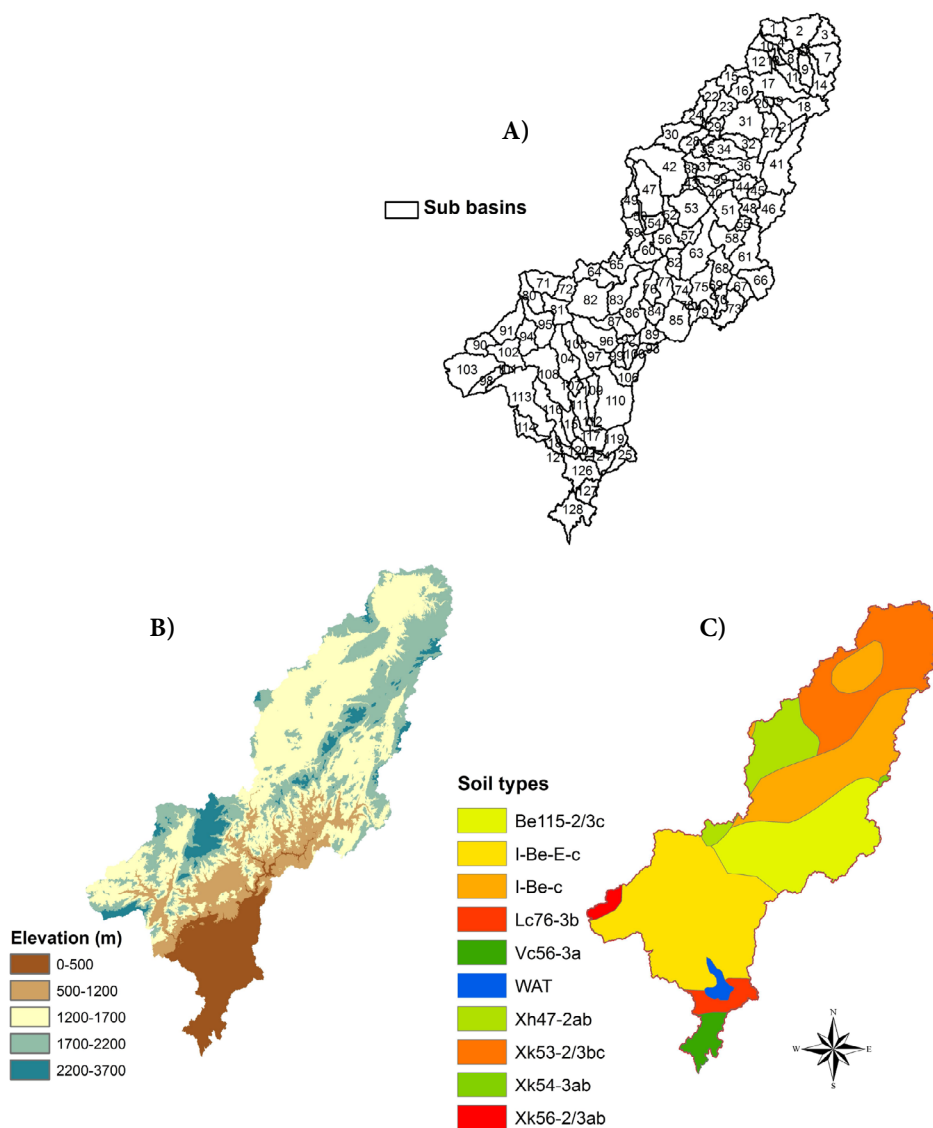
Figure 2. Map showing A) the subbasins, B) the digital elevation model, and C) soil types.

37 subcategories (for example, varieties of agricultural land class like cropland, pasture, and vineyard).

2.5. Observed data

The mean monthly flow data and the total monthly sediment yield for station 1818 (37°22′50″N, 35°28′05″E) were obtained from the Electrical Power Resources Survey and Development Administration of Turkey (EIE). The mean daily streamflow measurement made by the EIE is calculated based on the rating curve of the given hydrometric station. However, the suspended sediment sampling (expressed as ppm) is collected by the depth-integrated method using USDH-48 and USD-49 sampling equipment on a monthly basis. The suspended sediment samples are then used to develop a relationship between sediment load (t day^{-1}) and water discharge at the time of

sampling, which is known as the sediment rating curve for each gauging station. The monthly data from the period 2001 to 2007 were used for model calibration with a 1-year-long warm-up period.

2.6. Model calibration

The autocalibration was done using the SWAT-calibration uncertainty programs (SWAT-CUP; Abbaspour et al., 2007a) package using the sequential uncertainty fitting (SUFI-2) algorithm (Abbaspour et al., 2004, 2007b). SUFI-2 considers all sources of uncertainties; for example, model structure, observation data error, and model input. A detailed description of the procedure is given in Abbaspour et al. (2007b). The model was simultaneously calibrated for streamflow and sediment yield. For calibration, 14 flow parameters and 9 sediment yield parameters were

Table 1. Distribution of land cover for the study area based on GLCC classes and the modified land cover classes of Anderson level I.

Code	Modified land cover classes of Anderson level I	Type	SWAT code	% of the watershed
100	Urban or built-up land	Urban and built-up land	URBN	0.12
211		Dryland and cropland pasture	CRDY	10.61
212	Agricultural land	Irrigated cropland and pasture	CRIR	3.79
280		Mosaic cropland/grassland	CRGR	21.72
290		Mosaic cropland/woodland	CRWO	9.56
311		Grassland	GRAS	41.99
321	Rangeland	Shrubland	SHRB	1.85
330		Mixed shrubland/grassland	MIGS	0.37
332		Savanna	SAVA	8.31
411		Deciduous broad-leaf forest	FODB	0.01
422	Forest	Evergreen needle-leaf forest	FOEN	0.47
430		Mixed forest	FOMI	0.69
500	Water	Water bodies	WATB	0.34
810	Tundra	Wooded tundra	TUWO	0.15
850		Mixed tundra	TUMI	0.02

Table 2. Distribution of land cover for the study area based on GlobCover classes and the modified land cover classes of Anderson level I.

Code	Modified land cover classes of Anderson level I	Type	SWAT code	% of the watershed
190	Urban or built-up land	Artificial surfaces and associated areas (urban areas >50%)	URBN	0.06
14		Rainfed crops	CRDY	4.22
20	Agricultural land	Mosaic cropland (50%–70%) / vegetation (grassland/shrubland/forest) (20%–50%)	AGRR	16.41
30		Mosaic vegetation (grassland/shrubland/forest) (50%–70%) / cropland (20%–50%)	CRGR	36.31
110		Mosaic forest or shrubland (50%–70%) / grassland (20%–50%)	MISG	5.35
120	Rangeland	Mosaic grassland (50%–70%) / forest or shrubland (20%–50%)	MIGS	1.89
130		Closed to open (>15%) (broad-leaved or needle-leaved, evergreen or deciduous) shrubland (<5 m)	SHRB	14.5
50		Closed (>40%) broad-leaved deciduous forest (>5 m)	FRSD	0.30
70	Forest	Closed (>40%) needle-leaved evergreen forest (>5 m)	FRSE	9.58
100		Closed to open (>15%) mixed broadleaved and needle-leaved forest (>5 m)	FRST	0.64
150	Barren land	Sparse vegetation (<15%)	BSVG	8.10
200		Bare areas	BARE	2.08
210	Water	Water bodies	WATB	0.55

Table 3. Distribution of land cover for the study area based on CORINE land cover classes and the modified land cover classes of Anderson level I.

CLC Code	Modified land cover classes of Anderson level I	Type (level 3)	SWAT code	% of the watershed
111		Continuous urban fabric	URHD	0.05
112		Discontinuous urban fabric	URML	0.62
121		Industrial or commercial units	UCOM	0.1
124	Urban or built-up land	Airports	UTRN	0.03
131		Mineral extraction sites	UIDU	0.08
133		Construction sites	URLD	0.04
141		Green urban areas	GRUR	0
142		Sport and leisure facilities	FESC	0.05
211		Nonirrigated arable land	CRDY	11.77
212		Permanently irrigated land	AGRC	6.04
221		Vineyards	GRAP	0.07
222	Agricultural land	Fruit trees and berry plantations	ORCD	0.43
231		Pastures	PAST	0.91
242		Complex cultivation patterns	AGRL	5.24
243		Mixed agriculture and natural vegetation	CRGR	11.5
321	Rangeland	Natural grasslands	RNGE	13.23
324		Transitional woodland-shrub	SHRB	11.55
311		Broad-leaved forest (deciduous)	FRSD	0.7
312	Forest	Coniferous forest (evergreen)	FRSE	8.21
313		Mixed forest	FRST	4.93
331		Beaches, dunes, sands	BESU	0.06
332	Barren land	Bare rocks	BARE	5.09
333		Sparsely vegetated areas	BSVG	18.44
411		Inland marshes	WEHB	0.04
511		Water courses	WATC	0.06
512	Water	Water bodies	WATB	0.75
521		Coastal lagoons	WTCL	0
522		Estuaries	WATM	0

optimized to obtain the best fit between simulated and measured monthly streamflow and sediment yield as shown in Table 4, which shows the selected parameters with their calibration ranges and final values.

2.7. Model evaluation

To evaluate the model performance and compare the simulated versus the observed results, 4 statistical measurements were used: the coefficient of determination (R^2), Nash–Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970), percent bias (PBIAS; Gupta et al., 1999), and the root mean square error (RMSE) observation's standard deviation ratio (SR), collectively called RSR (Eqs. (4)–(7)).

R^2 describes the degree of collinearity between simulated and measured data and describes the proportion of the variance in measured data explained by the model. R^2 ranges from 0 to 1, with higher values indicating less error variance.

$$R_2 = \frac{\left[\sum_{i=1}^{n} (O_i - \overline{O})(P_i - \overline{P})^2\right]}{\left[\sum_{i=1}^{n} (O_i - \overline{O})^2\right]\left[\sum_{i=1}^{n} (P_i - \overline{P})^2\right]}, \tag{4}$$

where P_i are the predicted values, O_i are the observed values, n is the total number of observations, $\overline{O}$ is the mean of the observed data, and $\overline{P}$ is the mean of the predicted data.
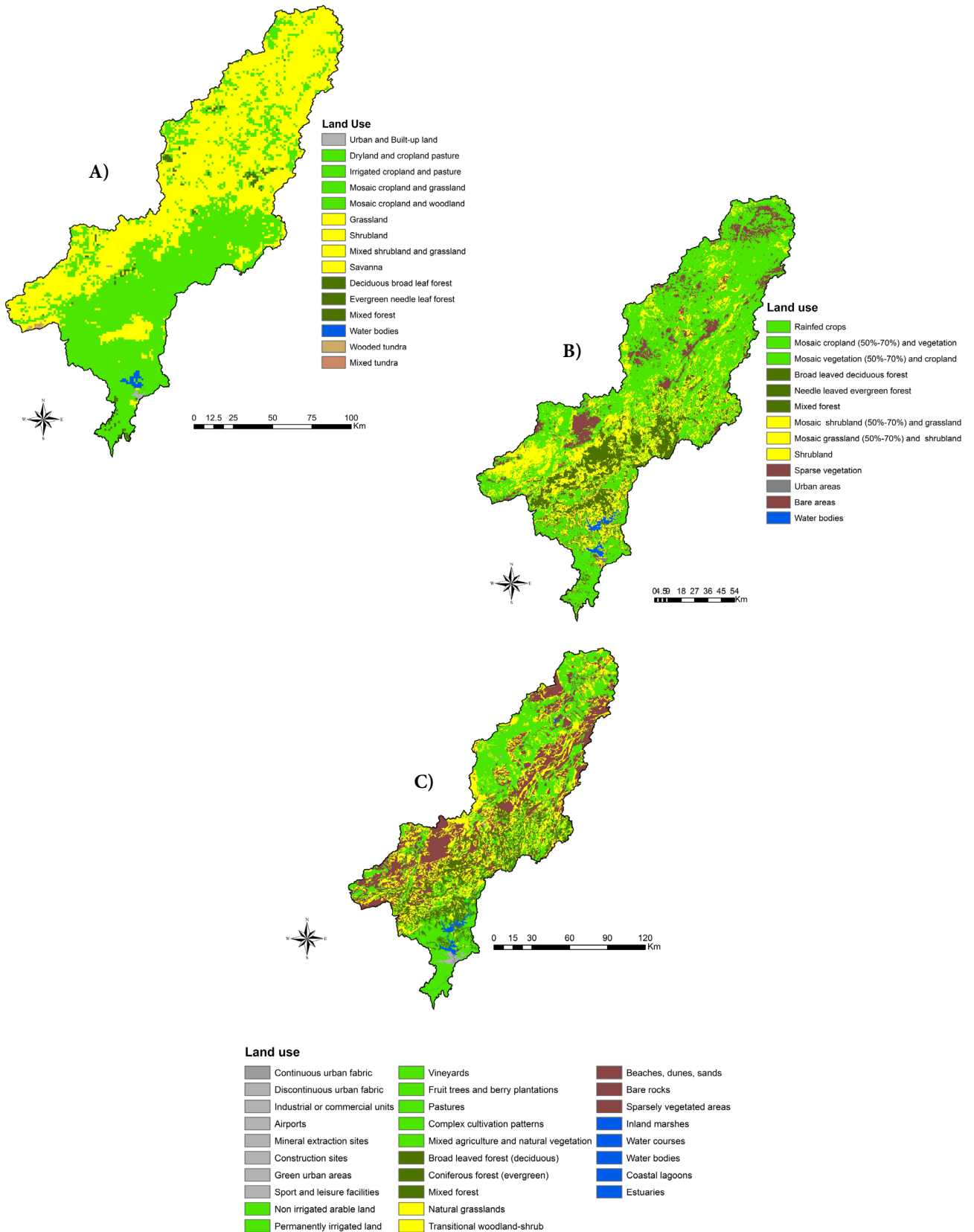
Figure 3. Land use maps from the A) GLCC, B) GlobCover, and C) CORINE datasets.

Table 4. Listing of SWAT model parameters, their descriptions, calibration ranges, and best fit parameter values used for streamflow and sediment yield calibration.

Parameters	Definition	Calibration range		Fitted values		
		Min	Min	GLCC	GlobCover	CORINE
Streamflow						
ALPHA_BF	Base flow alpha factor (days)	0	1	0.12	0.11	0.11
Ch_K2	Effective channel hydraulic conductivity (mm h⁻¹)	0	150	131.8	101.79	103.9
Ch_N2	Manning coefficient for main channel	0.01	0.3	0.11	0.15	0.17
CN2*	SCS curve number for moisture condition II	−50	50	14.51	10.05	17.63
ESCO	Soil evaporation compensation factor	0	1	0.97	0.99	0.59
GW_DELAY	Ground water delay (days)	0	500	477.4	312.63	207.1
GW_REVAP	Groundwater revap coefficient	0.02	0.2	0.2	0.14	0.12
GWQMN	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)	0	5000	1204	97.64	1949
REVAPMN	Threshold depth of water in the shallow aquifer required for revap to occur (mm)	0	500	232	265	312
SOL_AWC*	Available water capacity of the soil layer (mm mm⁻¹)	−50	50	50	−21.48	−14.62
SOL_K*	Soil saturated hydraulic conductivity (mm h⁻¹)	−50	50	37.77	49.57	11.3
SOL_Z*	Depth from soil surface to the bottom of layer (mm)	−50	50	−50	−49.79	−20.77
OV_N*	Overland Manning roughness	0	0.8	0.17	0.39	0.28
HRU_SLP*	Average slope steepness (m m⁻¹)	−20	20	6.72	12.48	3.42
Sediment						
USLE_K	Soil erodability factor in USLE	0	0.65	0.22	0.18	0.16
SPCON	Linear parameter for calculating the channel sediment routing	0.0001	0.01	0.009	0.003	0.009
SPEXP	Exponent parameter for calculating the channel sediment routing	1	1.5	1.08	1.37	1.17
CH_EROD	Channel erodibility factor	0	1	0.24	0	0.34
CH_COV	Channel cover factor	0	1	0.88	0.88	0.05
USLE_P	USLE equation support parameter	0	1	0.48	0.008	0.006
PRF	Peak rate adjustment factor for sediment routing	0	2.0	0.11	0.58	0.96
USLE_C	Min value of USLE C factor applicable to the land cover/plant	0.001	0.5	0.12	0.29	0.26
ADJ_PKR	Peak rate adjustment factor for sediment routing in the subbasin (tributary channels)	0.5	2	1.61	1.24	0.93

Type of change: * = relative change, all other parameters have absolute change.

NSE, ranging between $-\infty$ and 1, measures the predictive power of the hydrological model and how well the plot of observed versus simulated value fits the 1:1 line. The value of NSE = 1 corresponds to a perfect match between predicted and observed data, whereas values ≤ 0 indicate that the mean observed value is a better predictor than the simulated value (Moriasi et al., 2007).

$$NSE = \frac{\sum_{i=1}^{n}(O_i - \overline{O})^2 - \sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2}, \qquad (5)$$

PBIAS measures the percent deviation between simulated and observed data. A negative PBIAS indicates that simulated values are higher than observed (model bias overestimation), and a positive PBIAS indicates that simulated values are lower than observed (bias underestimation) (Gupta et al., 1999).

$$PBIAS = \left[\frac{\sum_{i=1}^{n}(O_i - P_i) * 100}{\sum_{i=1}^{n}(O_i)}\right], \qquad (6)$$

RSR is a commonly used error index that is calculated as a ratio of the RMSE and standard deviation of the measured data. The RSR value varies from the optimal value of 0, which indicates 0 RMSE or residual variation and a perfect model simulation, to a large positive value. The lower the RSR, the better the model performance.

$$RSR = \frac{RMSE}{STDEV_{obs}} = \frac{\left[\sqrt{\sum_{i=1}^{n}(O_i - P_i)^2}\right]}{\left[\sqrt{\sum_{i=1}^{n}(O_i - \overline{O})^2}\right]}, \qquad (7)$$

where $STDEV_{obs}$ is the standard deviation of the observed values.

Moriasi et al. (2007) suggest that NSE, PBIAS, and RSR are better quantitative statistics to evaluate model performance. They also reviewed the range of values for these statistics and derived corresponding qualitative performance ratings. They concluded that model simulation for monthly streamflow is satisfactory if $0.5 < NSE \leq 0.65$, $0.60 < RSR \leq 0.70$, and $\pm15 \leq PBIAS < \pm25$. However, it is satisfactory for the sediment yield if $\pm30 \leq PBIAS < \pm55$ with same values for NSE and RSR as reported for streamflow. Other studies have suggested that an R^2 value greater that 0.5 is considered acceptable (Santhi et al., 2001; Van Liew et al., 2003)

3. Results

3.1. Comparison of the land cover datasets
The distribution of the 3 land cover datasets within the 7 Anderson classes was analyzed to determine the agreement. Our results show that the watershed was dominated by agricultural land for the GlobCover and CORINE datasets (56.97% and 35.89%, respectively); however, the dominant land cover according to the GLCC is range land (52.52%) and agricultural land represents 45.68%, as shown in Table 5. A higher percentage of rangeland for the GLCC dataset is due to a large area being identified as grassland (41.99%), as shown in Table 1. This suggests that the barren land in the GlobCover and CORINE datasets, which represent 10.18% and 23.54%, respectively, was classified as grassland area in the GLCC dataset.

Another difference among the land cover datasets is the amount of forest land, which occupies 10.52% and 13.84% of the total area for GlobCover and CORINE, respectively; however, it only occupies 1.17% according to GLCC. The land cover datasets generally agreed on the size of the water body class, which occupied 0.34%, 0.55%, and 0.85% for the GLCC, GlobCover, and CORINE datasets, respectively. There are 2 land cover classes that are missing

Table 5. Representation of Anderson level I land use/land cover classes in each of the land cover datasets.

Modified land cover classes of Anderson level I	GLCC	GlobCover	CORINE
Urban or built-up land (1)	0.12	0.06	0.97
Agricultural land (2)	45.68	56.94	35.89
Rangeland (3)	52.52	21.74	24.78
Forest land (4)	1.17	10.52	13.84
Water (5)	0.34	0.55	0.85
Barren land (7)	–	10.18	23.59
Tundra (8)	0.17	–	–

The numbers in parentheses represent land-cover codes.

from some of the datasets. The tundra class in the GLCC dataset represents 0.17% of the watershed area; however, this class was not included in the other datasets. Sparsely vegetated and barren land covered an area of 10.18% and 23.59% of the watershed area according to GlobCover and CORINE, respectively, but was not represented in the watershed according to the GLCC.

3.2. Model calibration

3.2.1. Streamflow output

Table 6 lists the values of the statistical measures for the model performance. Figure 4 presents the comparison between the simulated versus the measured monthly streamflow for the calibration period from 2001 to 2007.

Simulated and observed monthly streamflow matched well in the calibration period with good NSE values of 0.73, 0.71, and 0.68 for the GLCC, GlobCover, and CORINE datasets, respectively. The R^2 values for monthly streamflow simulation in calibration were 0.76, 0.73, and 0.69 for the GLCC, GlobCover, and CORINE datasets, respectively. The NSE and R^2 values indicate that the model performed slightly better when the GLCC was used as a source of land cover information than when the GlobCover and CORINE datasets were used, indicating that the sensitivity of SWAT modeling of LULC datasets with different spatial and temporal data resolution is very low for the streamflow simulation.

According to Moriasi et al. (2007), when considering the PBIAS, the average magnitude of simulated monthly streamflow values was within the very good range (PBIAS < ±10) with values of 1.07%, 2.52%, and –0.495% for the GLCC, GlobCover, and CORINE datasets, respectively. For RSR, there was no difference (0.523, 0.537, and 0.562) among the GLCC, GlobCover, and CORINE datasets, respectively, and they are rated as good according to Moriasi et al. (2007). Visual inspection of the comparison between monthly simulated and observed streamflow

(Figure 4) shows that the model adequately simulated the streamflow for the entire simulation period and captured all the peaks, except for April 2002, in which the simulated flow was underestimated to be almost half of the observed flow (405.13 m^3 s^{-1}) for all of the used land cover datasets.

3.2.2. Sediment simulation

The calibration effort significantly improved the accuracy of the model's sediment yield prediction, especially when the GlobCover dataset was used. The model was calibrated for monthly sediment yield using the parameters listed in Table 4. The statistics of the calibrated SWAT model for monthly sediment yield are given in Table 6. Based on the NSE values, the model simulated the sediment yield satisfactorily with a slightly better accuracy using the GlobCover dataset compared to the GLCC and CORINE datasets. The NSE monthly value for sediment yield was 0.51 when the GlobCover dataset was used, while the NSE values of the GLCC and CORINE land cover datasets were 0.48 and 0.45, respectively.

According to the qualitative assessments suggested by Moriasi et al. (2007), the 3 land cover datasets performed similarly in regard to the other statistical measures; for example, the 3 datasets produced PBIAS values within the range of good (15± ≥ PBIAS < ±30), satisfactory RSR within the range 0.60 < RSR ≤ 0.7, and R^2 values > 0.5 (Table 6). The negative values of the PBIAS indicate that the sediment yield was overestimated. Figure 5 shows the comparison between the simulated and observed monthly sediment yield. Streamflow and sediment yield are highly correlated; i.e. high monthly streamflow peaks are associated with high sediment load. The model could simulate the sediment peak for most of the simulation period, except for April 2002 and March 2004, with no differences due to using different land cover datasets, which can partially be explained by the underprediction of streamflow for these 2 months.

Table 6. Statistics of the comparison between measured and simulated streamflow.

Land cover dataset	Evaluation statistics							
	R^2	NSE	PBIAS	RSR	R^2	NSE	PBIAS	RSR
	Streamflow				Sediment			
GLCC	0.76	0.73	1.07	0.523	0.52	0.48	−21.69	0.626
GLOBCOVER	0.73	0.71	2.52	0.537	0.52	0.51	−12.21	0.608
CORINE	0.69	0.68	−0.495	0.562	0.50	0.46	−18.15	0.637

R^2 = the coefficient of determination, NSE = Nash–Sutcliffe efficiency, PBIAS = percent bias, RSR = RMSE – SR.
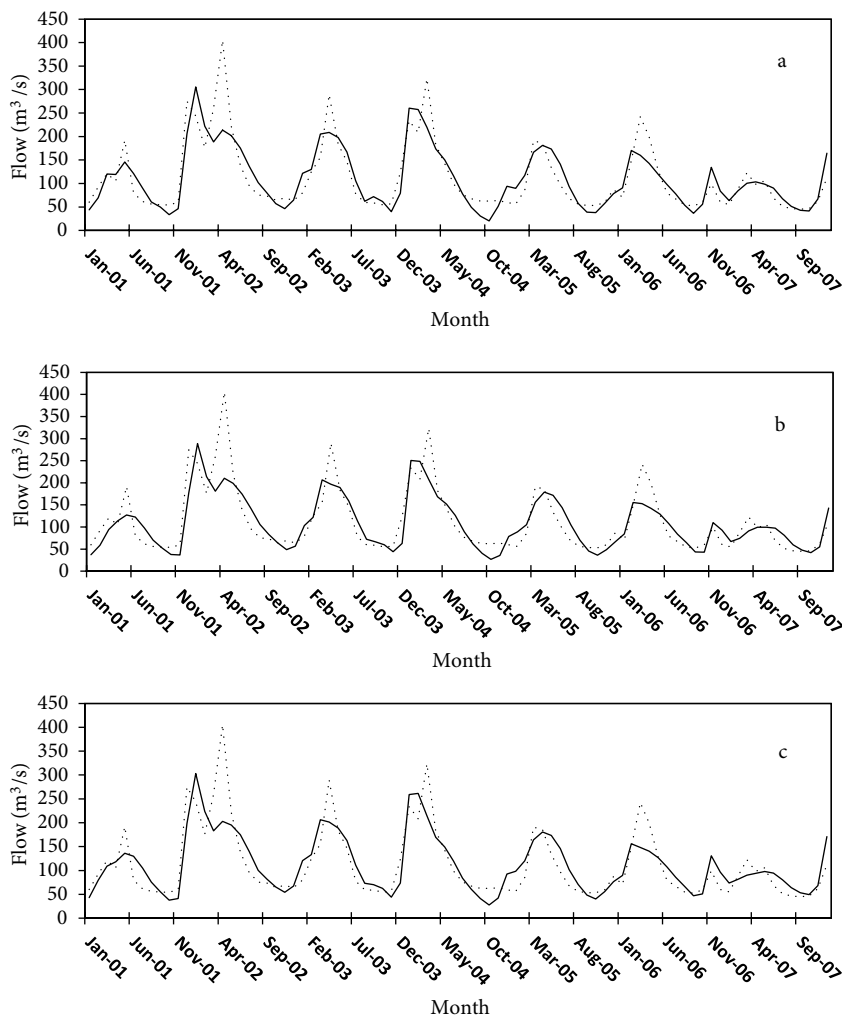
Figure 4. Comparison between the simulated (solid line) and observed (dashed line) monthly streamflow using the A) GLCC, B) GlobCover, and C) CORINE datasets.

4. Discussion

The results show that the land cover distribution in the GlobCover and CORINE datasets were mostly represented by 4 major land cover classes: agricultural land, range land, forest land, and barren land. In the GLCC land cover dataset, it is represented by only 2 major land cover classes (agricultural land and range land) due to the coarse resolution (1 km), which suggests that some of the land cover classes were combined with other land cover classes, i.e., orchards/vineyards, sparsely vegetated, and barren land. The CORINE land cover dataset contained the largest number of land cover classes corresponding to the Anderson land cover classes. For example, the CORINE land cover has more classes that were categorized as urban land (8 classes), agricultural land (9 classes), and water (5 classes) that are included in the Anderson land cover level I as compared to the GLCC and GlobCover datasets.

The results suggest that the GLCC simulated the monthly flow more accurately according to the NSE and R^2 values; however, the PBIAS statistic indicated the CORINE land cover dataset ranks first with a slightly better PBIAS over the other 2 datasets. The model could not match all the peaks with the observed flow, i.e. in April 2002 and March 2004, which is mainly caused by differences in measured precipitation. High actual evapotranspiration (ET) during the months of April 2002 and March 2004 associated with insufficient monthly precipitation of 115 mm and 38.75 mm, respectively, could be the reason for inaccurate simulation of the streamflow. This resulted in a reduction in the NSE values to be in the good range. However, the model succeeded in capturing the low flow peaks and the base flow. In general, it can be concluded that model accuracy was not greatly affected by using different land cover data sources for simulating the streamflow from the Seyhan River basin.
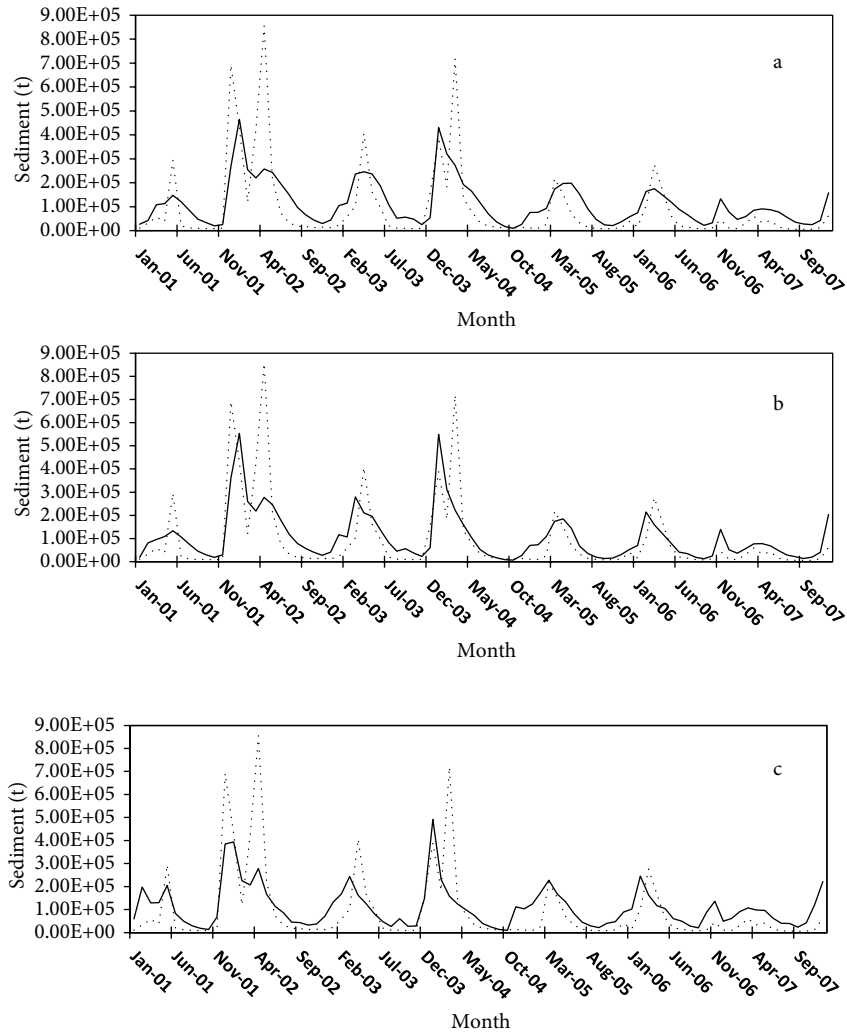
Figure 5. Comparison between the simulated (solid line) and observed (dashed line) monthly sediment yields using the A) GLCC, B) GlobCover, and C) CORINE datasets.

The uncalibrated results revealed that the monthly sediment yield was severely overpredicted in the winter months during the rainy season; however, it was underestimated during the summer months when there is no precipitation; i.e. all NSEs for the 3 models were negative. The observed sediment yield obtained from the EIE was estimated by using the sediment rating curve approach, which basically is controlled by a suspended sediment discharge/streamflow relationship. However, the SWAT model simulates not only the suspended sediment load (smaller bed material particles) but also the bedload fraction (larger particles transported by rolling sliding or saltation), which together are known as bed material load (Ndomba and van Griensven, 2011).

The highest uncalibrated sediment loading was associated with using the GLCC land cover followed by GlobCover and finally the CORINE dataset. Two

reasons may give an explanation for this. The first reason is explained by Muleta et al. (2007) and FitzHugh and Mackay (2000), who reported that increasing the number of HRUs decreases the average HRU area and subsequently decreases the runoff term in the MUSLE equation. This decreases the generated sediment yield. From Eq. (3), the peak runoff change is mostly controlled by the change of the HRU area, indicating the slow change of the time of concentration as a result of the HRU area change. The second reason is that the majority of the land type for the GLCC is occupied by rangeland (52.52%) and agricultural land (45.68%) with a very low portion of forest; therefore, this land cover pattern results in a very high sediment yield. The GlobCover and CORINE datasets have a significant portion of forest, which caused a slight decrease in sediment yield. The influence of land cover type changes on streamflow and sediment is presented by Yan et al.

(2013). They found that the streamflow of the upper Du catchment in China is highly affected by the changes in the farmland, forest, and urban areas between 1978 and 2007 with regression coefficients of 0.232, –0.147, and 1.256, respectively, and variable influence on projection (VIP) values greater than 1. However, sediment yield is mostly influenced by changes in farmland (with a VIP and regression coefficient of 1.762 and 14.343, respectively) and forest (with a VIP and regression coefficient of 1.517 and –7.746, respectively)

The weak model performance for sediment load for the different land cover datasets may be related to using the simple MUSLE equation for sediment estimation in the SWAT model. The model assumes that all the soil eroded by the runoff will be delivered to the channel, ignoring the sediment deposition process in the surface catchment area (Oeurng et al., 2011). Moreover, we have not been provided sufficient information about the sediment and erosion control structures, and that could be a reason for the low observed sediment load at the gauging station compared to the model output.

In conclusion, the simulation results do not show significant differences using land cover datasets with different spatial and temporal details for the streamflow and sediment simulations. This leads to the idea that a very detailed dataset, i.e. CORINE, may not necessarily provide the best results; moreover, important information could be lost due to the aggregation of major land cover classes (Romanowicz et al., 2005). The different land cover datasets had little impact because there was little change in the land cover conditions over the study period. To further contribute to the understanding of the land cover change in the basin, land cover change detection has been done using the CORINE land cover change maps for the years 1990–2000 and 2000–2006 (Figure 6). These data show that the proportion of the land area that experienced land cover change over the period of 1990–2000 was 2.63% of the total area of the basin, and it was 1.03% over the period of 2000–2006.

Improvements to sediment loading predictions may be possible with better information regarding agricultural and management practices such as grazing, tillage, and irrigation in concert with a longer record of sediment sampling. This will enhance the SWAT model to provide
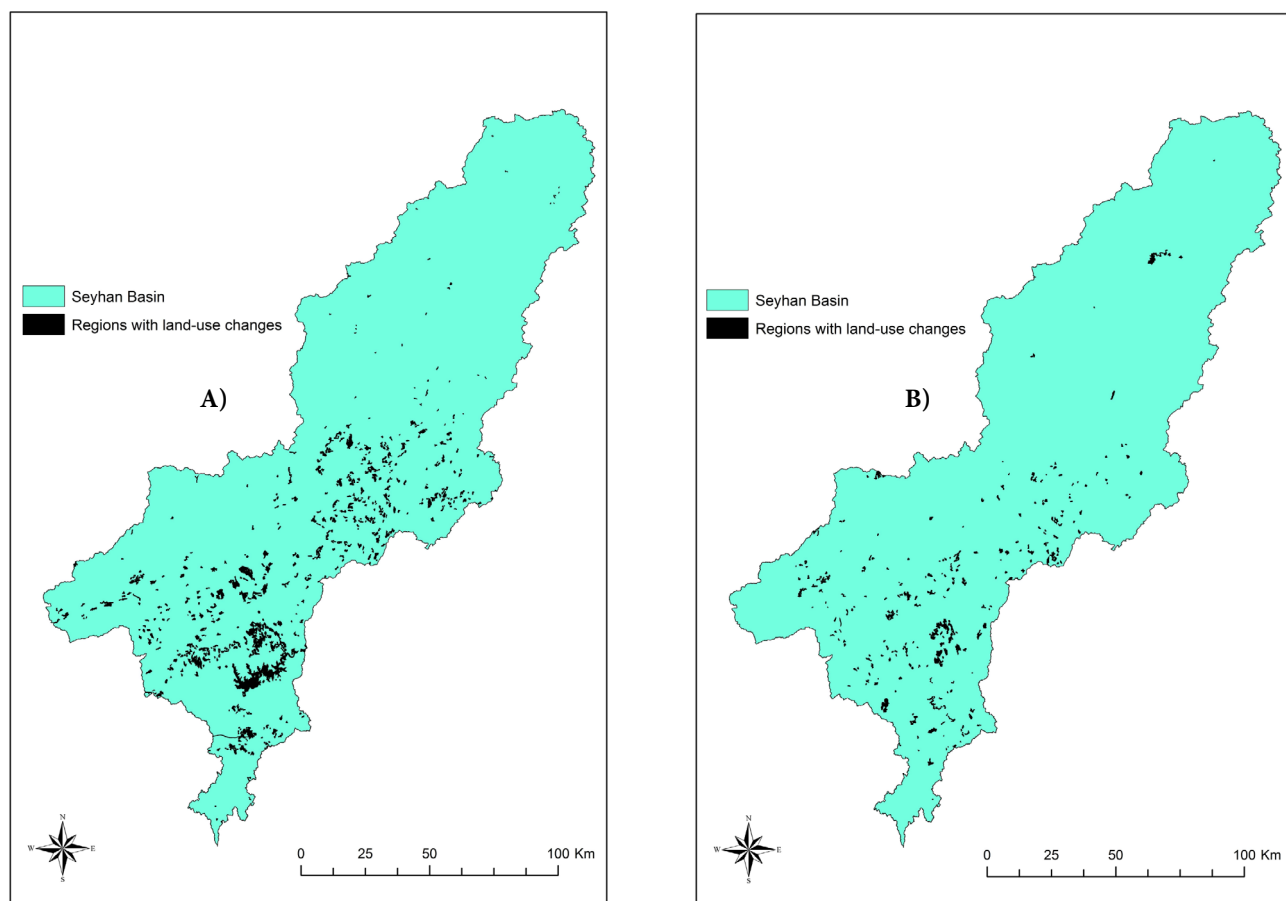

Figure 6. Changes in land cover for the years A) 1990–2000 and B) 2000–2006 in the Seyhan River basin. (Figures are produced from the CORINE land cover change 1990–2000 and 2000–2006 maps.)

greater insight into potential sediment sources and sinks. We suggest more studies can be conducted to examine the effect of implementing the structural and nonstructural best management practices, i.e. contour terracing, vegetated buffer strips along water courses, ponds, and grade stabilization structures on the streamflow and the sediment load in the river. In addition, further investigation is needed to evaluate the different land cover datasets under conditions of significant land cover change and to study the effect of these datasets on the estimation of other water quality parameters such as total nitrogen and total phosphorus.

References

Abbaspour KC, Johnson CA, van Genuchten MT (2004). Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. Vadose Zone J 3: 1340–1352.

Abbaspour KC, Vejdani M, Haghighat S (2007a). SWATCUP calibration and uncertainty programs for SWAT. In: Oxley L, Kulasiri D, editors. MODSIM 2007 Proceedings of the International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand, December 2007; Melbourne, Australia, pp. 1603–1609.

Abbaspour KC, Yang J, Maximov I, Siber R, Bogner K, Mieleitner J, Zobrist J, Srinivasan R (2007b). Modelling hydrology and water quality in the Pre-Alpine / Alpine Thur watershed using SWAT. J Hydrol 333: 413–430.

Acar A, Dincer I (2005). Left upstream slope design for the Catalan Dam, Adana, Turkey and its behavior under actual earthquake loading. Eng Geol 82: 1–11.

Anderson JR, Hardy EE, Roach JT, Witmer RE (1976). A land use and land cover classification system for use with remote sensor data. Reston, VA, USA: USGS Professional Paper 964.

Arino O, Gross D, Ranera F, Leroy M, Bicheron P, Brockmann C, Defourny P, Bourg L, Latham J, Di Gregorio A, et al. (2007). GlobCover: ESA service for global land cover from MERIS. In: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 2412–2415.

Arino O, Ramos, J, Kalogirou, V, Defourny, P, Achard, F (2010). GlobCover 2009, Proceedings of the Living Planet Symposium SP-686, June 2010.

Arnold JG, Moriasi DN, Gassman PW, Abbaspour KC, White MJ, Srinivasan R, Santhi C, Harmel RD, van Griensven A, Van Liew MW et al. (2012). SWAT: Model use, calibration and validation, T ASABE 55: 1491–1508.

Cai T, Li Q, Yu M, Lu G, Cheng L, Wei X (2012). Investigation into the impacts of land-use change on sediment yield characteristics in the upper Huaihe River basin, China. Phys Chem Earth 53–54: 1–9.

Candela C, Freni G, Mannina G, Viviani G (2012). Receiving water body quality assessment: an integrated mathematical approach applied to an Italian case study. J Hydroinform, 14: 30–47.

Chaplot V, Saleh A, Jaynes D, Arnold J (2004). Predicting water, sediment and NO_3-N loads under scenarios of land-use and management practices in a flat watershed. Water Air Soil Poll 154: 271–293.

Chen PY, Di Luzio M, Arnold JG (2005). Spatial assessment of two widely used land-cover datasets over the continental US. IEEE T Geosci Remote 43: 2396–2404.

Cotter AS, Chaubey I, Costello TA, Soerens TS, Nelson MA (2003). Water quality model output uncertainity as affected by spatial resolution of input data. J Am Water Resour Assoc 39: 977–986.

De Girolamo AM, Lo Porto A (2012). Land use scenario development as a tool for watershed management within the Rio Mannu Basin. Land Use Policy 29: 691–701.

Eidenshink JC, Faundeen JL (1994). The 1 km AVHRR global land dataset: first stages in implementation. Int J Remote Sens 15: 3443–3462.

FitzHugh TW, Mackay DS (2000). Impacts of input parameter spatial aggregation on an agricultural nonpoint source pollution model. J Hydrol 236: 35–53.

Green WH, Ampt G (1911). Studies on soil physics 1: the flow of air and water through soils. J Agr Sci 4: 1–24.

Gupta HV, Sorooshian S, Yapo PO (1999). Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. J Hydrol Eng 4: 135–143.

Huang J, Zhou P, Zhou Z, Huang Y (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. Int J Environ Res Public Health 10: 144–157.

Huffman GJ, Bolvin DT, Nelkin EJ, Wolff DB (2007). The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J Hydrometeorol 8: 38–55.

Irvem A, Topaloğlu, F, Uygur, V (2007). Estimating spatial distribution of soil loss over Seyhan River basin in Turkey. J Hydrol 336: 30–37.

Koch S, Bauwe A, Lennartz B (2013). Application of the SWAT model for a tile-drained lowland catchment in North-Eastern Germany on subbasin scale. Water Resour Manage 27: 791–805.

Lin YP, Hong NM, Wu PJ, Wu CF, Verburg PH (2007). Impacts of land use change scenarios on hydrology and land use patterns in the Wu-Tu watershed in Northern Taiwan. Landscape Urban Plan 80: 111–126.

Loveland TR, Reed BC, Brown JF, Ohlen DO, Zhu J, Yang L, Merchant JW (2000). Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data. Int J Remote Sens 21: 1303–1330.

Mbonimpa EG, Yuan Y, Mehaffey MH, Jackson MA (2012). SWAT model application to assess the impact of intensive corn-farming on runoff, sediments and phosphorous loss from an agricultural watershed in Wisconsin. J Water Res Prot 4: 423–431.

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

Muleta M, Nicklow J, Bekele E (2007). Sensitivity of a distributed watershed simulation model to spatial scale. J Hydrol Eng 12: 163–172.

Nash JE, Sutcliffe JV (1970). River flow forecasting through conceptual models. Part 1: discussion of principles. J Hydrol 10: 282–290.

Nasr A, Bruen M, Jordan P, Moles R, Kiely G, Byrne P (2007). A comparison of SWAT, HSPF and SHETRAN/GOPC for modelling phosphorus export from three catchments in Ireland. Water Res 41: 1065–1073.

Ndomba PM, van Griensven A (2011). Suitability of SWAT model for sediment yields modelling in the Eastern Africa In: Chen D, editor. Advances in Data, Methods, Models and Their Applications in Geoscience. Rijeka, Croatia: InTec, pp. 261–284.

Neilsen A, Trolle D, Me W, Luo L, Han B-P, Liu Z, Olesen JE, Jeppesen (2013). Assessing ways to combat eutrophication in a Chinese drinking water reservoir using SWAT. Mar Freshwater Res: 475–492.

Neitsch SL, Arnold, JG, Kiniry JR, Williams JR (2005). Soil and Water Assessment Tool Theoretical Documentation, version 2005. Temple, TX, USA: Agricultural Research Service and Texas Agricultural Experiment Station.

Oeurng C, Sauvage S, Sanchez-Pérez J (2011). Assessment of hydrology, sediment and particulate organic carbon yield in a large agricultural catchment using the SWAT model. J Hydrol 401: 145–153.

Park JY, Park MJ, Joh HK, Shin HJ, Kwon HJ, Srinivasan R, Kim SJ (2011). Assessment of MIROC3.2 hires climate change and CLUE-s land cover change impacts on watershed hydrology using SWAT. T ASABE 54: 1713–1724.

Romanowicz AA, Vanclooster M, Rounsevell M, La Junesse I (2005). Sensitivity of the SWAT model to the soil and land use data parametrisation: a case study in the Thyle catchment, Belgium. Ecol Model 187: 27–39.

Santhi C, Arnold JG, Williams JR, Dugas WA, Srinivasan R, Hauck LM (2001). Validation of the SWAT model on a large river basin with point and nonpoint sources. J Am Water Resour Assoc 37: 1169–1188.

Schuol J, Abbaspour KC, Yang H, Srinivasan R, Zhender AJB (2008). Modeling blue and green water availability in Africa. Water Resour Res 44: 1–18.

USDA-NRCS (2004). NRCS National Engineering Handbook Part 630: Hydrology. Chapter 10: Estimation of Direct Runoff from Storm Rainfall: Hydraulics and Hydrology: Technical references. Washington, DC, USA: USDA National Resources Conservation Service.

Van Liew MW, Arnold JG, Garbrecht JD (2003). Hydrologic simulation on agricultural watersheds: Choosing between two models. T ASAE 46: 1539–1551.

Verburg PH, Soepboer W, Veldkamp A, Limpiada R, Espaldon V (2002). Modeling the spatial dynamics of regional land use: the CLUE-s model. Environ Manage 30: 391–405.

Verburg PH, van Eck JR, de Hijs TC, Dijst MJ, Schot P (2004). Determination of land use change patterns in the Netherlands. Environ Plan B: Plan Design 31: 125–150.

Vu MT, Raghavan SV, Liong SY (2011). Simulating stream flow over data sparse areas – an application of internet based data. Hydrol Earth Syst Sci Discuss 8: 11015–11037.

Williams JR (1975). Sediment-yield prediction with universal equation using runoff energy factor. In: Present and Prospective Technology for Predicting Sediment Yield and Sources. Proceedings of a Sediment-yield Workshop, Nov. 28–30, 1972; Oxford, MS, USA. USDA Sedimentation Laboratory, ARS-S-40.

Winchell M, Srinivasan R, Di Luzio M, Arnold J G (2009). ArcSWAT 2.3.4 Interface for SWAT2005. Temple, TX, USA: Grassland, Soil and Research Service.

Yan B, Fang NF, Zhang PC, Shi ZH (2013). Impacts of land use change on watershed streamflow and sediment yield: an assessment using hydrologic modelling and partial least squares regression. J Hydrol 484: 26–37.

Zhang P, Liu Y, Pan Y, Yu Z (2011). Land use pattern optimization based on CLUE-S and SWAT model for agricultural non-point source pollution control. Math Comput Model 58: 588–595.

Zhou F, Xu Y, Chen Y, Xu CY, Gao Y, Du J (2013). Hydrological response to urbanization at different spatio-temporal scales simulated by coupling of CLUE-S and the SWAT model in the Yangtze River Delta region. J Hydrol 485: 113–125.