HYDROLOGICAL MODELING OF HARO RIVER WATERSHED, PAKISTAN

Mona A. Hagras 1 & Rasikh Habib 2
1 Faculty of Engineering, Ain Shams University, Cairo, Egypt
2 National University of Sciences and Technology (NUST), Islamabad, Pakistan

ABSTRACT
Geographical Information System based semi distributed model, Soil and Water Assessment Tool (SWAT) is used to model the hydrology of Haro river watershed 40 kilometer North West Islamabad, Pakistan. Thus, it aims to simulate the stream flow, establish the water balance and estimate the monthly volume inflow to Khanpur dam located at the basin outlet. SWAT model was calibrated for a ten years period from 1994 to 2003 and validated for a seven years period from 2004 to 2010. Statistical indicators; Pearson Correlation, Coefficient of Determination and Nash-Sutcliffe Efficiency were used to verify the simulation abilities of the model. Results illustrate a good performance for both calibration and validation periods and acceptable agreement between measured and simulated values of both annual and monthly discharge. The water balance components were estimated and presented here for the studied watershed. Finally, it is concluded that SWAT model can be used in semi-arid regions for the water resources management of the studied watershed.

Key words: Hydrological modeling, Soil and Water Assessment Tool, Haro River, Khanpur Dam.

1. INTRODUCTION
Pakistan is classified as one of the extremely high level water stress country in the world (Tianyi Luo et al., 2015). Agriculture in Pakistan uses well over 95% of the freshwater resources in addition to the high losses in the sprawling irrigation system. Rapid and unsustainable development too, has polluted and disturbed some major watersheds and river plains, (Shaimaa, 2015). The availability and the sustainable use of the water resources become the core of the local and national strategies and policies in many regions which face scarcity of freshwater or subject to pollution. For efficient water management issues, different elements of hydrologic processes should be analyzed and quantified on watershed basis to address the range of water resources, environmental and social problems.

Recently, the development in computer technology and accessibility of spatial datasets has enriched the competencies to precisely define the watershed features in approximation of runoff response from precipitation. The GIS and Remote Sensing technology has also played its role in enhancement of hydrological models. It has revolted the hydrological models to physically based distributed models. This has improved the abilities of models to effectively integrate the spatial heterogeneity in complex hydrological processes, (Mohan and Madhav, 2000). Various models had been established for watershed hydrology but the temporal and spatial datasets was the key limitation hindering the application of these models mainly in developing countries (Singh and Woolhiser, 2002). SWAT which is a river basin or watersheds scale model has been used globally for hydrological and water quality modeling. It was developed at the USDA-ARS Grassland, Soil and Water Research Laboratory. The model is capable of predicting and assessing the impact of management on water, sediment, and nutrients at watershed scale. SWAT (Arnold et al., 1998) is a continuous–time, quasi–physically based, distributed model designed to simulate water, sediment, and agricultural chemical transport on a river–basin scale. SWAT was developed from earlier hydrologic and nutrient assessment models, including SWRRB, CREAMS, GLEAMS, EPIC, and ROTO (Borah and Bera, 2003), for predicting the long–term impact of different management scenarios and/or climate changes on watershed–scale hydrology and nonpoint–source pollution. It represents the hydrological cycle by interception, evapotranspiration, surface runoff, infiltration, soil percolation, lateral flow, groundwater flow, and channel routing processes, (Grunwald, 2005). (A.D. Khan et al., 2014) used the SWAT model for the Upper Indus Basin. The calibration of the model was effectively carried out at Nowshera and Kalabagh. It was tested and used in many regions of Africa by Fadil et al. (2011), Ashagre (2009) and Schuol et al. (2008). It also applied to simulate St. Joseph River watershed in US by Kieser & Associated (2005). Swat model was used successfully to estimate the water balance components in South eastern Ethiopia by Shawul et al. (2013) and in Nigeria by Adeniyi et al. (2014).

The objective of this study is to model the hydrology of Haro river watershed and to calibrate the hydrological processes at Khanpur. This will improve knowledge about the hydrological cycles in the river basin. It is beneficial for development and management of water resources for irrigation and water supply for the portable purposes. It also provides the baseline for study of climate changes and variability on various parameters of hydrological cycles and to manage water balance, agriculture and environmental flow.
A spatially distributed modeling would enhance the understanding of core hydrologic patterns in the watershed. In this perspective, soil and water assessment tool (SWAT) was chosen. As such as soils, land use and land cover and topography impact hydrology, hence it is beneficial to manage hydrological patterns of the watersheds.

2. STUDY AREA
Haro River originates from the hilly areas of Ayubia, Murree and Margallah Hills as shown in Figure 1. Its catchment falls in Khyber Pakhtunkhwa (KPK) and Islamabad. The river flows through the parts of KPK and Punjab. It has a length of 43.65 km. The inflow of the river is derived from snow and rainfall upstream. The upper part of the watershed receives snow in winter season. The inflows depend on the seasonal variation. The climate change has an effect on inflows and peak discharge. It is fed by four major tributaries, which are The Lora Haro, The Stora Haro, The Neelan and The Kunhad.

Figure 1: Studying Area

Khanpur Dam located at 33° 48' 06" N and 72° 55' 50" E, has been built across the Haro River at Haripur at a distance of 40 Km North West Islamabad. It has a catchment area of 800 Km². It was conceived to a height of 167 feet to store 106,000 acre-feet of water. It has a gross command area of 124,000 acres from which 36,470 acres is cultivable command area. It provides 407 cusecs of irrigation water in the districts of Abbottabad, Attock and Rawalpindi. It is also used to provide 102.37 MGD municipal water supply for the twin city of Rawalpindi and Islamabad besides providing 28.5 MGD water for Industrial complexes at Taxila.

3. Materials and Methodology
ArcSWAT version 2.3.4 which was built for ArcMap 9.3 is used in this study to model the hydrology of Khanpure Dam watershed in Haro river basin. Hydrological model and the special dataset which used in the simulation are described in the following sections.

3.1 Data
SWAT requires three spatial datasets to describe the study area which includes topography, soil and land use/land cover. The hydrological processes like flow direction, runoff, infiltration and water quality are controlled by these spatial datasets. The SRTM DEM of 90 meter resolution (HTML: CGIAR-CSI), FAO soil data (HTML: FAO-AGL, 2003), and European Union (EU) global land use / land cover datasets (HTMAL: EU-GEM, 2000) have been used in this study. Delineation of the studied watershed was carried out as a first step in using SWAT model.

Land Use
Land use and the land cover in catchment was distributed into 5 classes as identified in Figure 2a based on the regional scaled standardized European Union land use data of 2000. Table 1 shows the percentage of each class.
Table 1: Distribution of land use/land cover classes used for ArcSWAT in Khanpur Dam Watershed

<table>
<thead>
<tr>
<th>Land use - Land cover class</th>
<th>SWAT classes</th>
<th>Watershed Area %</th>
</tr>
</thead>
<tbody>
<tr>
<td>PINE</td>
<td>PINE</td>
<td>26.47</td>
</tr>
<tr>
<td>OAK</td>
<td>OAK</td>
<td>18.65</td>
</tr>
<tr>
<td>Deciduous Forest</td>
<td>FRSD</td>
<td>9.95</td>
</tr>
<tr>
<td>Row Crops</td>
<td>AGRR</td>
<td>34.27</td>
</tr>
<tr>
<td>Urban High Density</td>
<td>URHD</td>
<td>10.66</td>
</tr>
</tbody>
</table>

Soil Classes

The hydrological behavior of soil is characterized by its physical properties. To input the soil properties in the model, the FAO regional scale soil vector maps are used where each cartographic unit was associated with one or two delineations corresponding to sub soil group of USDA. Due to soil data limitations, the USA soils are compared with the watershed area to use their properties to define HRUs. Two soils delineated in the catchment; CM (Cambisol) and LP (Leptosol) have their corresponding USA series of Brewster and Merino series respectively as shown in Figure 2(b). The Khanpur catchment covers 82.47% by Brewster (CM) and 17.53% by Marino (LP).

Figure 2: Catchment Spatial Input Datasets. (a) Land use Classes and (b) Soil Classes

The Merino series: consists of very shallow and shallow, well drained soils formed in residuum and colluvium from monzonite and other granitic rocks, gneiss, tuff, and breccia. Merino soils are on undulating plateaus, ridgetops, and side slopes of intermontane basins and on mountainsides and mountain ridges. Slope ranges from 5% to 65%. The mean annual precipitation is about 22 in., and the mean annual temperature is about 38 F. The Brewster series: consists of very shallow or shallow, well drained, moderately permeable soils that formed in loamy materials weathered from igneous bedrock. These soils are on rolling to very steep hills and mountains. Slopes range from 5% to 60%, Khan et al. (2014). The soil units were then extracted and completed by additional information from the soil properties listed in Table 2.
Table 2: Derived Soil Properties delineated in the catchment

<table>
<thead>
<tr>
<th>Soil name</th>
<th>Merino</th>
<th>Brewster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil hydrologic group</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Maximum rooting depth (mm)</td>
<td>2000</td>
<td>2000</td>
</tr>
<tr>
<td>Porosity fraction from which anions are excluded</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Crack volume potential of soil</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Texture 1</td>
<td>Grv_Sl</td>
<td>Grv_CL</td>
</tr>
<tr>
<td>Depth (mm)</td>
<td>330</td>
<td>300</td>
</tr>
<tr>
<td>Bulk density moist (g/cc)</td>
<td>1.38</td>
<td>1.61</td>
</tr>
<tr>
<td>Ave. AW Incl. Rock Frag</td>
<td>0.13</td>
<td>0.10</td>
</tr>
<tr>
<td>K_{sat} (est.) (mm/h)</td>
<td>883</td>
<td>672</td>
</tr>
<tr>
<td>Organic carbon (weight %)</td>
<td>0.5</td>
<td>1.25</td>
</tr>
<tr>
<td>Clay (weight %)</td>
<td>16</td>
<td>27</td>
</tr>
<tr>
<td>Silt (weight %)</td>
<td>40</td>
<td>38</td>
</tr>
<tr>
<td>Sand (weight %)</td>
<td>44</td>
<td>35</td>
</tr>
<tr>
<td>Rock fragments (vol. %)</td>
<td>27</td>
<td>47</td>
</tr>
<tr>
<td>Soil albedo (moist)</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Erosion K</td>
<td>0.18</td>
<td>0.13</td>
</tr>
<tr>
<td>Salinity (EC, Form 5)</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Meteorological Data

The long term meteorological datasets of Precipitation, Temperature, Wind Speed, Solar Radiation and Relative Humidity of the stations falling in the study area are required for the hydrological modeling. The records of precipitation and temperature are the minimum mandatory inputs of the SWAT Model. Other parameters are optional. If the datasets of these optional parameters are not available then the model has the capability to generate the data against these parameters.

The meteorological data including the precipitation and temperature was taken from Pakistan Meteorological Department (PMD). The stations available adjacent to the catchment area were Khanpur dam, Islamabad and Murree. The stations record maximum and minimum temperature, relative humidity, solar radiation and snow melt equivalence of water on daily basis. Figure 3 shows sample of meteorological data at Khanpur station.

![Figure 3: Sample of Precipitation and Temperature data at Khanpur station](image-url)
Stream Flow Data
The flow of the Haro River at upstream Khanpur is collected from Water and Power Development Authority (WAPDA) from 1994 – 2010 for calibration and validation.

3.2. Methodology
SWAT allows simulating the major watershed processes as hydrology, sedimentation, nutrients transfer, crop growth, environment and climate change. The aim is to depict the physical functioning of these different components and their interactions as simply and realistically as possible through conceptual equations and using available input data so that it can be useful in routine planning and decision making of large catchments management (Fadil et. al., 2011). SWAT model simulate the hydrologic cycle based on the water balance equation

\[ SW_t = SW_o + \sum_{i=1}^{t} (R_{day} - Q_{surf} - E - w_{seep} - Q_{gw}) \]  

Where \( t \) is the time in days, \( SW_t \) is the final soil water content (mm), \( SW_o \) is the initial soil water content (mm), \( R_{day} \) is amount of precipitation on day \( i \) (mm), \( Q_{surf} \) is the amount of surface runoff on day \( i \) (mm), \( E \) 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), \( Q_{gw} \) is the amount of return flow on day \( i \) (mm)

Surface runoff can be estimated by the model using the Soil Conservation Service (SCS) curve number method, Arnold et al. (1998) This method is a widely used for the prediction of approximate amount of runoff from a given rainfall event. It is mainly based on the soil properties, land use and hydrologic conditions. The SCS curve number equation is

\[ Q_{surf} = \frac{(R_{day} - 0.2 \cdot S)^2}{(R_{day} - 0.8 \cdot S)} \]  

where \( Q_{surf} \) is the daily surface runoff (mm), \( R_{day} \) is the rainfall depth for the day (mm), and \( S \) is the retention parameter (mm). The retention parameter \( S \) can be estimated using Equation (3).

\[ S = 25.4 \left( \frac{1000}{CN} - 10 \right) \]  

where \( S \)= drainable volume of soil water per unit area of saturated thickness (mm/day); \( CN \)=curve number. SCS defines three antecedent moisture conditions: I – dry (wilting point), II – average moisture and III – wet (field capacity). The moisture condition I curve number is the lowest value the daily curve number can assume in dry conditions. The curve numbers for moisture conditions I and III are calculated with the Equations. (4) and (5), respectively. For Lateral flow it can be predicted by Equation (6)

\[ CN1 = CN2 - \frac{20 \cdot (100 - CN2)}{(100 - CN2 + e^{[2.533 - 0.0636(100-CN2)]})} \]  

\[ CN3 = CN2 \cdot e^{[0.00673(100-CN2)]} \]  

\[ q_{lat} = 0.024 \frac{(2SSC \cdot \sin \alpha)}{\theta_d L} \]  

where \( q_{lat} \)=lateral flow (mm/day); \( S \)=drainable volume of soil water per unit area of saturated thickness (mm/day); \( SC \)= saturated hydraulic conductivity (mm/h); \( L \)= flow length (m), \( \alpha \)=slope of the land, \( \theta_d \)=drainable porosity.
The SWAT model is generally divided in two parts: Inputs and outputs. The first section includes the input from land system and atmospheric system in the form of Digital Elevation Model (DEM), land use / land cover, soil, and precipitation and temperature datasets respectively. These are the minimum essential inputs which are required by the SWAT model. Figure 4(a) shows the DEM of the study area catchment.

In data processing, first of all, all the datasets are projected using UTM Zone-43 to make them under same the projection system so that they could be properly overlaid as well as calculations could be made easily. The SRTM DEM of 90 meter resolution is processed for the extraction of flow direction, flow accumulation, slope, stream network generation and finally delineation of catchment area. These topographic parameters play an important role in the hydrological modeling and are the foremost requirement of the any hydrological model.

The land cover / land use, soil and meteorological datasets will be required for the Hydrological Response Unit (HRU) definition / estimation and weather generation. After delineation the total catchment area upstream Khanpur is about 800 km$^2$. The catchment area is segmented in a total number of 84 sub-basins topographically which are divided into 187 HRUs. The HRU is the basic hydrological unit at which the model simulates the watershed hydrology. Next, the climatic datasets collected from PMD were processed in the model input format. For this purpose, MATLAB was used and a code written for each of precipitation and temperature file for its conversion according to model required format. The MATLAB files are then transferred into database files which are actually required for model input. After climatic data input and weather generation, the model is ready to run for the simulation of watershed hydrology at HRU and sub-basin level. The calibration was made for the verification of the model’s results. At the end, the model is verified to check the reliability of the model. The detailed flow chart methodology is explained in Figure 5.

![Digital Elevation Model (DEM) and Delineation](image)

*Figure 4: Catchment (a) Digital Elevation Model (DEM). (b) Delineation*
4. **MODEL RUN**

The simulation is done for a period of 10 years (1994 - 2003) on yearly and monthly basis for flow. A sensitivity analysis is also commenced prior to calibration for the identification of the most responsive hydrological parameters. The sensitivity analysis was made using a built-in SWAT sensitivity analysis tool that uses the Latin Hypercube simulation combined with One-factor-At-a-Time sampling (LH-OAT), Van Griensven,(2003).

The model has ranked the Curve Number (CN2), Soil Evaporation Compensation (ESCO) and Available water capacity of the soil layer (Sol_AWC) as most sensitive parameters respectively for this study area. This result supports those found by many similar studies confirming that these three parameters are the crucial sensitive parameters for water balance White and Chaubey (2005).

The model is calibrated at Khanpur, upstream Khanpur reservoir on the Haro River which is the main tributary of the reservoir. The model simulated results of flow are compared with the observed data collected by WAPDA at their gauge on upstream Khanpur for the model baseline period of 10 years.

5. **RESULTS AND DISCUSSION**

5.1. **Model Calibration**

The model is calibrated for flow at upstream Khanpur on Haro River which is the main tributary of the Khanpur Dam. The yearly and monthly basis results of flow produced by the model are accepted as shown in Figure 6 and Figure 7. The value of correlation is 0.81 whereas standard deviation of data is 5.32 which indicate that the results produced by the model are reliable. The value of coefficient of determination ($R^2$) is 0.656 which recognizes the accuracy of the results shown.

The peak flow is observed in the year 1997 whereas the lowest flow in the year 2002. The fact of the low flow from 1999 - 2001 is the drought period which was observe in the country. The simulation results show excellent match from 1999 - 2003 whereas during the initial years from 1994 - 1998 the satisfactory match is observed due to the missing of data of few stations and other data constraints. For the missing data, the value of -1999 is used so that the model can generate data itself for this time (days) by adopting the built-in equations / methods. The model was unable to simulate in 1997 as there was high monsoon rainfall. This results in low simulated flow. This shows the model limitation that a model is unable to respond a high rainfall.

Figure 6 clearly explains the same trend between the observed and simulated flow other than 1997 which complements the accuracy of the research as well as model accuracy. There is excellent match in 1999 and 2001.
Figure 6: Annual Flow Calibration for Haro at Khanpur

Figure 7: Monthly Calibration of Haro at Khanpur
5.2. Model Validation

After the calibration, the model was validated for a period of seven years (2004-2010). Figure 8 shows that the yearly basis results of flow produced by the model are quite accepted and satisfactory. The validation results are better than the calibration as the period does not have any extreme event as that of calibration in 1997.

The results of validation show a good correlation of observed and model simulated. The value of correlation is 0.88 whereas standard deviation of data is 1.44 which indicates that the results produced by the model are reliable. The peak flow is observed in the year 2007 whereas the lowest flow in the year 2004.

![Figure 8: Comparison of annual observed and simulated flow](image)

5.3. Model Efficiency

The commonly statistical parameters; Co-relation, Co-efficient of Determination ($R^2$) and Nash-Sutcliffe Efficiency Index (NSE) are used in this study to assess the observed versus simulated watershed data. The model has given the Correlation ($r$) values of 0.81 for calibration and 0.88 for validation flow. This shows the model results are good and equally accepted.

The value of Coefficient of Determination ($R^2$) comes out to be 0.656 for calibration and 0.775 for validation. It indicates that model results produced for are very good. Generally, values larger than 0.50 are considered adequate (Van Liew et al., 2003). Similarly the higher values represent the less error between the simulated and observed values (Santhi et al., 2001). Nash-Sutcliffe Efficiency Index (NSE) is another statistical method used for the prediction of relative amount of noise compared with information (Nash and Sutcliffe, 1970) and is calculated from the Equation (7).

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Y_{obs}^i - Y_{sim}^i)^2}{\sum_{i=1}^{n}(Y_{obs}^i - Y_{mean}^i)^2}$$  (7)
Where $Y_{i}^{obs}$ is the $i$th observation (stream flow), $Y_{i}^{sim}$ is the $i$ th simulated value, $Y_{mean}$ is the mean of observed data and $n$ is the total number of observations.

NSE ranges from $-\infty$ to 1.0 (1 inclusive), whereas NSE=1 being the ideal value. The values of NSE between 0.0 and 1.0 are usually regarded as suitable levels of performance. Generally, the model simulation is judged as very good if NSE > 0.65 (Moriasi et al., 2007). According to NSE method, the model results 0.79 for calibration and 0.89 for validation which are quite acceptable. After applying all three model efficiency test mentioned above, it is found that the model is quite efficient and results are reliable.

5. WATER BALANCE COMPONENTS

It is important to analyze and quantify the components of hydrological process taking place within the study area for efficient water management. The most important water balance components of water balance of a catchment are precipitation, surface runoff, lateral flow, base flow and evapotranspiration. Among these, all the variables, except precipitation, need prediction for quantifying as their measurement is not easy.

Figure 9 shows the distribution of average annual values for different water balance components of Haro river watershed through all sub-catchments; and their average annual values are presented in Table 3.

<table>
<thead>
<tr>
<th>Water balance component</th>
<th>Average annual value (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>757</td>
</tr>
<tr>
<td>Water yield</td>
<td>351.14</td>
</tr>
<tr>
<td>Surface runoff</td>
<td>22.27</td>
</tr>
<tr>
<td>Evapo-transpiration</td>
<td>383.6</td>
</tr>
<tr>
<td>Total aquifer recharge</td>
<td>142.84</td>
</tr>
<tr>
<td>Contribution of groundwater to steam flow</td>
<td>121.61</td>
</tr>
</tbody>
</table>
Figure 9: catchment Water balance Component
6. CONCLUSIONS
This study had revealed the capability of ArcSWAT interface implemented in the ArcGIS software to model the hydrology of Khanpure Dam catchment. It also had verified that the SWAT model works well for semi-arid regions and in hilly areas. The model was first calibrated using observed stream flow data during period of 1994-2003 and a seven year period of 2004-2010 for model validation. The calibration and validation of the model showed good simulation results. The efficiency of the model has been tested by Nash Sutcliffe Efficiency (NSE) in addition to Pearson’s correlation (r) and coefficient of determination (R²). For calibration phase, the Correlation, Coefficient of Determination, and Nash Sutcliffe Efficiency (NSE) were 0.81, 0.656 and .79 respectively and 0.88, 0.775 and 0.89 for validation phase, which indicate predictive ability of the model. Therefore, the calibrated model can be used successfully to predict the volume inflow to the Khanpur Dam and facilitate the storage and release water management. Water balance components such as Precipitation, Water yield, Surface runoff, Evapo-transpiration, Total aquifer recharge and Contribution of groundwater to stream flow have been estimated.

7. RECOMMENDATIONS
The calibrated model can be used for the assessment and prediction in other small watersheds like the effect of climate change on Haro river flow patterns. The performances can be improved using some other global climate data as well as data sets with good distribution of meteorological stations. The calibration of water quality parameters should be carried out to predict the effect of climate change and change in land use pattern on water quality.

8. REFERENCES


