Effects of soil data input on catchment streamflow and soil moisture prediction

A. Binesh¹, I.Y. Yeo¹ and G. Hancock²

 School of Engineering, Faculty of Engineering and Built Environment, University of Newcastle
 School of Environmental and Life Sciences, Faculty of Earth Sciences, University of Newcastle Email: <u>ali.binesh@uon.edu.au</u>

This study attempts to investigate the impacts of soil data input on hydrologic model performance Abstract: in simulating streamflow and soil moisture. Two different soil datasets available in Australia were considered: the Digital Atlas of Australian Soil (AoAS) and Soil and Landscape Grid of Australia (SLGA). We quantified the impacts of these two soil databases on hydrologic simulations using Soil Water Assessment tool (SWAT) model. Two separate calibration schemes were set up with two soil databases while keeping other inputs the same. For both cases, SWAT was calibrated to the daily streamflow at the catchment outlet over 2006-2012 (including wet and dry periods) and validated against the dataset over 2013-2015 (wet period), after 3 years warm up period (2003-2005). The soil moisture estimation from calibrated SWAT was then compared with the two radiometric satellite soil moisture products, the Soil Moisture Active Passive (SMAP)-Enhanced 9 km (L3SMP-E) and Soil Moisture and Ocean Salinity (SMOS) 25 km gridded (SMOS CATDS L3 SM 3-DAY) obtained during 2015. This study was conducted in Merriwa catchment, located in the upper part of the Goulburn River basin in Upper Hunter Region of NSW. The simulation results showed very little difference in streamflow prediction due to two different soil inputs. Both models showed very similar streamflow patterns (with similar NSE value of ~ 0.61 for calibration and ~ 0.45 for validation), but different soil moisture estimates. When catchment average near surface soil moisture estimates were compared with the satellite soil moisture products, SWAT calibrated with SLGA showed improved results (with R² value of 0.52 and 0.66 against SMAP-9 km and SMOS-25 km and RMSE of ~10 %) than that with AoAS (with R² value of 0.35 and 0.49 against SMAP-9 km and SMOS-25 km and RMSE of 18-22 %). The large differences in simulated soil moisture indicate importance of improved soil data input to capture soil moisture change patterns and significantly different water and energy partitioning for the catchment.

Keywords: SWAT, streamflow, soil moisture, Soil and Landscape Grid of Australia, Atlas of Australian Soil

Binesh et al., Effects of soil data input on catchment streamflow and soil moisture prediction

1. INTRODUCTION

Prediction of hydrologic variables heavily relies on the accuracy of input dataset and its spatial representation. Several studies investigated the impacts of the spatial resolution of input soil data and hydrologic parameters on hydrologic prediction (Anderson et al., 2006; Daggupati et al., 2018; Di Luzio et al., 2009; Geza & McCray, 2008; Levick et al., 2004; Li et al., 2012; Moriasi & Starks, 2010; Robinson et al., 1995; Singh et al., 2011). Some studies reported improved prediction with soil database with higher spatial resolution (Anderson et al., 2006; Di Luzio et al., 2009), while other argued that high resolution soil data has limited or no effect on streamflow prediction at the catchment outlet with carefully calibrated simulation models (Cotter et al., 2003; Moriasi & Starks, 2010). For example, Moriasi and Starks (2010) showed insignificant difference in daily

streamflow simulation (either at the sub-catchment or catchment scale) based on two different soil databases available in the US (i.e., Soil Survey Geographic database-SSURGO with spatial scale of 1:250,000 and State Soil Geographic-STATSGO with spatial scale range between 1:15,850 to 1:31,680). Similarly, Wang and Melesse (2006) showed detailed information on soil characteristics may not improve streamflow prediction at the catchment outlet, though it may improve estimates of soil moisture in upland areas. It is because the streamflow observed at the catchment outlet is more likely influenced by surface runoff from its immediate upstream areas and channel routing processes.

Recently, the Soil and Landscape Grid of Australia (SLGA ~90 \times 90 m) was released to provide improved information on physical and chemical characteristics of soils at the continental scale in Australia. This study aims to quantify the impacts of improved soil data input on streamflow and soil moisture prediction. This research was undertaken in the Merriwa catchment, located in the northern portion of Goulburn River Basin in the Upper Hunter Region of NSW, Australia. Soil and Water Assessment Tool (SWAT) was used to simulate the catchment hydrological processes across different spatial scales (William et al., 2012).



Figure 1. The Study area, the Merriwa catchment located in the Goulburn River basin.

2. MATERIALS AND METHODS

2.1. Study area

The Merriwa catchment with areal size of $\sim 651 \text{ km}^2$ is located in the northern part of the Goulburn River basin (Figure 1). The study site, as a semi-arid to temperate catchment, has experienced extreme climate variability over the last two decades (Rüdiger et al., 2007). The site is covered by a range of soil textures, from sandy soils in the downstream to clay soils in the upstream (Rüdiger et al., 2007). Elevations within the basin range from nearly 200 m above the sea level at the catchment outlet (southern part of the region) to 1100 m across the northern ridge. The spatial precipitation patterns follow the topographic gradient. The low lying south region receives relatively low precipitation (from 241 mm in 2008 to 760 mm in 2007 with annual average of 515 mm during 2006-2015), while the northern ridge received the highest precipitation amount (from 464 mm in 2006 to 1231 mm in 2010 with annual average of 812 mm during 2006-2015). The land was mostly used for grazing and pasture with a very small part used for cropping.

2.2. Catchment Model

SWAT is a semi-distributed, physically-based, continuous model that simulates the hydrologic processes at hydrologic response unit (HRU) level. Hydrological response unit (HRU) is specified by the unique combination of soil, land use type, and slope. Daily soil water balance is simulated at HRU level, as follow:

$$SW_t = SW_0 + \sum_{t=1}^{n} (PRCP_t - SURQ_t - AET_t - W_t - PERC_t)$$
 Eq. 1

Binesh et al., Effects of soil data input on catchment streamflow and soil moisture prediction

where SW_t and SW_0 are the amount of soil water at time t and the initial time step (mm/d), PRCP_t is precipitation (mm/d), SURQ_t is surface runoff (mm/d), AET_t is actual evapotranspiration (mm/d), GW_t is the groundwater flow (mm/d), and PERC_t is the amount of water percolation from the soil profile (mm/d) to deep soils.

The SURQt and storage terms (S) are estimated based on the curve number (CN) method (William et al., 2012):

$$SURQ_{t} = \frac{[PRCP_{t} - 0.2S_{t}]^{2}}{[PRCP_{t} + 0.8S_{t}]}, PRCP_{t} > 0.2S_{t} \text{ and } SURQ_{t} = 0, PRCP_{t} < 0.2S_{t}$$
Eq. 2

$$S_t = 254 \times [\frac{100}{CN_t} - 1]$$
 Eq. 3

where S_t is the retention parameter (mm/d).

The surface runoffs simulated at HRU level are aggregated at the sub-catchment scale, and then routed to the catchment outlet following the stream network. SWAT estimates the amount of water entering the soil profile as difference between rainfall and surface runoff. The vertical water flow throughout the soil profile is estimated based on the storage routing technique in each soil layer (up to 10 layers), considering the difference between available water content and field capacity (FC) at each soil layer. Water percolates downward to the next layer if it is unsaturated and then redistribution occurs until the water content throughout the entire profile reaches equilibrium. The variable storage method is used to simulate streamflow for each of the reach segment based on the continuity equation (Williams et al., 2012).

SWAT Model Setup

The ArcSWAT-GIS (the SWAT model extension, http://swat.tamu.edu/software/arcswa) was used for preparing input dataset listed in table (1). SWAT requires detailed information on topography, soil, land use distribution, and climate data. Daily stream flow and soil moisture from radiometric satellite data products were used to assess the model performance on stream flow and soil moisture prediction.

Dataset	Source	Temporal Resolution
Digital Elevation Model (DEM)	SRTM-derived 1 Second Digital Elevation Models, Geoscience of Australia	
Soil	 (a) Atlas of Australian Soils from Australian Soil Resource Information System (ASRIS) (b) Soil Grids from Soil and Landscape Grid of Australia 	
Land Use	Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES) for 2010-2011.	
Climate	Australian Government, Bureau of Meteorology and Scaling and Assimilation of Soil Moisture And Streamflow project (SASMAS), including climate data such as precipitation, temperature, wind speed and humidity for 2005-2015.	Daily
Streamflow	Department of Primary Industries-Office of Water	Daily
Soil Moisture	European Space Agency led Soil Moisture and Ocean Salinity (SMOS) mission version 5.51 (RE02) data set provided by the Centre Aval de Traitement des Données SMOS (CATDS, Level 3 product-lunched in 2009) & Soil Moisture Active Passive (SMAP) 9 km enhanced radiometric soil moisture products (L3SMP-E)	Daily global composite (every three days)

 Table 1. Data sources for the SWAT model

Soil Database: Two soil datasets available in Australia were considered. AoAS was compiled by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) in the 1960's to provide consistent national description of Australia's soils (McKenzie and Hook, 1992). The scale of the published maps is 1:2,000,000, but the original compilation was at scales from 1:250,000 to 1:500,000. AoAS represents the soil information using mapped units (shown with large polygons) (Figure 2a). AoAS includes information on hydrologic parameters compiled from McKenzie and Hook (1992), such as soil hydrologic properties (i.e., soil bulk density, soil saturation hydraulic conductivity, soil water capacity, etc.) and the soil texture (clay, silt and sand) using the factual key soil classification system across Australia (Northcote, 1979). The produced digitized map has been widely used for water resource management, sediment yield and water quality studies.

SLGA is a recently released soil dataset as a part of the Australian Soil Resource Information System (ASRIS) to provide consistent, comprehensive, nation-wide data in an easily accessible format. SLGA is a part of Global Soil Map project which aims to provide soil chemical (e.g., EC, pH, etc) and textural (clay, sand, silt) information at a high spatial resolution. Special attention was paid to generate spatially consistent, and detailed soil map to inform hydrologic (i.e., soil water holding capacity, soil hydraulic conductivity and soil bulk





Table 2. Descriptive soil characteristics of the first soil layer of AoAs (15cm) and SLGA (5cm) at catchment scale.

Soil information	SLGA	AoAS
Clay (%)	40.15	51.74
Silt (%)	8.96	11.78
Sand (%)	56.08	36.48
Saturated Hydraulic Conductivity (mm/hr)	9.3	20.48
Soil water capacity (fraction)	0.15	0.14
Bulk density (Mg/m ³)	1.43	1.23

density) soil properties across Australia (Grundy et al., 2012). This dataset contains soil attribute information at 6-depth (i.e., 0-50mm, 50-150mm, 150-300mm, 300-600mm, 600-1000mm and 1000-2000mm) throughout the profile up to 2 m for lower (5th percentile) and median (estimated value) upper (95th percentile) percentile. In this research, median values are used to describe soil properties (Figure 2b).

Figure 2 and Table 2 summarise similarities and differences between two soil databases. Figure 2 shows that the northern part

of the study site is predominantly covered by tertiary basalt soils with thick clay layers, while southern part generally by sandy soil type. The catchment average clay content was shown to be higher in AoAS (51.74%) than in SLGA (40.15%). Due to the inherent nature of the soil representation (i.e., in vector format), AoAS shows more uniform soil texture across upper part of the catchment, while SLGA (in raster format) shows more heterogeneous soil distribution for the same region (Fig 2c and 2d). This difference was most noticeable in the upper portion of the study site. However, soil texture pattern close to the catchment outlet is quite similar

between two soil datasets. Both show higher sandy soils. Given similar soil texture pattern in downstream areas, it is most likely to obtain similar stream flow responses at the catchment outlet with two soil datasets. However, difference in spatial soil characteristics in upstream areas as shown by saturation hydraulic conductivity from the two soil databases may lead to different estimations of hydrologic components in upstream areas. Difference in saturated hydraulic conductivity was due to different soil texture and pedotransfer function applied in each dataset (Grundy et al., 2015).

2.3. Model Setup and Evaluating Its Performance

The Merriwa catchment is divided into 31 sub-catchments and further divided into 464 HRUs under SLGA and 339 HRUs using AoAS. The number of HRUs cannot be imposed to the same as the delineation of HRUs is inherently determined by



Figure 3. Remotely sensed soil moisture products across the study catchment.

unique combination of land use, soil, and slope maps (Moriasi and Starks, 2010). Therefore, the differences in soil dataset influenced the HRU delineations despite special efforts were made to generate similar HRU delineation for both set-ups.

As there was limited data availability for different water balance components, soil moisture estimation from calibrated SWAT was compared against the satellite products: Soil Moisture Active Passive (SMAP)-Enhanced 9 km (L3SMP-E) and Soil Moisture and Ocean Salinity (SMOS) 25 km gridded (SMOS CATDS L3 SM 3-DAY) radiometric products obtained during 2015 (Figure 3). These two satellite products aim to provide global estimate of surficial soil moisture (~5cm) with frequently ~3-day revisit period at an expected accuracy of 0.04 v/v (Entekhabi et al., 2010a; Karthikeyan et al., 2017b; Senanyake et al., 2019). Both SMAP and SMOS products were used for the catchment level analysis while SMAP 9-km product was applied at sub-catchment level at the pixel X (Figure 3). As the satellite products only show soil moisture information at a shallow soil depth (~5 cm), SWAT was revised to extract soil water stored at the corresponding soil depth at the HRU level.

Binesh et al., Effects of soil data input on catchment streamflow and soil moisture prediction

The extracted soil moisture at the HRU level was then aggregated to sub-catchment level and to catchment based on weighted area average approach. Note that the default SWAT model provides estimation for profile soil water content only.

Two SWAT setups (with different soil data inputs) were calibrated separately against daily stream flow obtained from 2006-2012 and validated for 2013-2015 after 3-years warm up (2003-2005) period. The objective function for model calibration was based on the agreement between simulated and observed variables as measured by Nash-Sutcliff efficiency (NSE). In addition, root mean square error (RMSE), correlation of determination (R^2), bias were calculated to evaluate the goodness of fit between the observation and simulation.

NSE=1-
$$\frac{\sum_{i=1}^{n} (O_i - S_i)^2}{\sum_{i=1}^{n} (O_i - \overline{O})^2}$$
 Eq. 4 RMSE= $\frac{\sqrt{\sum_{i=1}^{n} (O_i - S_i)^2}}{n}$ Eq. 5

$$R^{2} = \left[\frac{\left(\sum_{i=1}^{n} (S_{i} - \hat{S}) \times (0_{i} - \hat{O}) \right)^{2}}{\sum_{i=1}^{n} (S_{i} - \hat{S})^{2} \times \sum_{i=1}^{n} (0_{i} - \hat{O})^{2}} \right]$$
 Eq. 6 Bias = $\left[\frac{\sum_{i=1}^{n} (S_{i} - 0_{i})}{\sum_{i=1}^{n} (0_{i})} \right]$ Eq. 7

where S and O represent simulated and observed variables respectively. $\hat{0}$ and \hat{S} are also the mean of observed and simulated variables respectively. i is the daily time step (i=1..n) and n is the total number of observations. In total, 21 parameters (related to the rainfall-runoff processes, sub-surface water storage, and channel routing) have been calibrated using the Sequential Uncertainty Fitting ver.2 (SUFI-2) from SWAT-CUP package (Abbaspour, 2015).

3. RESULTS AND DISCUSSIONS

3.1. Model calibration and validation

The streamflow simulation with two soil datasets showed good and satisfactory performance for both calibration periods (Table 3). However, the set up with AoAS showed slightly better performance for both calibration (NSE = 0.66) and validation (NSE = 0.46) compared to that with SG (NSE value of 0.61 and 0.43, respectively for calibration and validation). SWAT with SLGA predicted streamflow with lower bias (~0) than AoAS (0.14) during calibration. However, during the validation, the simulation with AoAS showed lower bias (-0.29) than that with SLGA (-0.53). Overall, the results showed that introducing high spatial resolution soil dataset does not necessarily guarantee better streamflow prediction (Moriasi and Starks, 2010). As the streamflow generation is most likely influenced by downstream runoff processes and channel hydraulic properties in SWAT, spatial difference in soil physical characteristics between the datasets tend to show insignificant impacts on streamflow prediction. Note that downstream soil characteristics in the study site were shown very similar between the two datasets; hence, it was most likely to have similar downstream responses leading to the streamflow generation.

Table 3. Descriptive statistics of streamflow in calibration (unshaded) and validation (shaded) considering different soil datasets.

Soil Dataset	F	R ²	BI	AS	NS	SE
SLGA	0.61	0.68	~0	-0.53	0.61	0.43
AoAS	0.67	0.69	0.14	-0.29	0.66	0.46

3.2. Surficial soil moisture estimation

Table 4 showed the statistical comparison between the satellite soil moisture products and estimates of surficial soil moisture (top 5 cm) from April to December 2015. The SWAT output using SLGA provided very good correlation and performance, with R² value of 0.52 and 0.66 against SMAP-9 km and SMOS-25 km respectively. However, the SWAT model with AoAS overestimated surficial soil moisture and did not capture its drying down process very well. This set up showed the RMSE value of 17.7% and 22.3% against SMAP-9 km and SMOS-25 km respectively. The positive bias resulted from AoAS setup was attributed from the input soil characteristics in upland areas (i.e., clay soils with high soil water storage capacity). Considering simulated streamflow results (in comparison to the observed flow), SWAT with AoAS would result in higher soil water storage and leading to generate higher streamflow. This could explain overestimation of low flows (i.e., large BIAS) in comparison to the SWAT with SLGA (Table 3).





Figure 4. Streamflow (SF) simulation with SLGA and AoAS during the study period (2006-2015) at daily time scale.

Analysis of the soil moisture prediction at corresponding sub-catchment level against the pixel (X) showed very similar results, i.e. underestimation with SLGA but overestimation with AoAS (Figure 5). The SWAT with SLGA showed a good correlation ($R^2=0.62$), while AoAS prediction was less powerful ($R^2=0.44$) with higher RMSE at sub-catchment level. Further analysis with consideration of soil input uncertainty (e.g., using 90 % CI of soil attributes) as shown by SLGA may help to resolve this model bias.



Figure 5. Time series of the remotely sensed soil moisture datasets, near surface soil moisture estimates wth AoAS and SLGA at (a) catchment level and (b) the pixel (X).

Table 4. Descriptive statistics of the surficial soil moisture and the model results using Soil Grids and Atlas of Australian Soil at catchment scale and in pixel (X).

Soil Dataset	R ²		RMSE (%)	
	SMAP_9km	SMOS_25km	SMAP_9km	SMOS_25km
SLGA-Catchment wise	0.52	0.66	10.6	8.9
AoAS-Catchment wise	0.35	0.49	17.7	22.3
SG-Pixel (X)	0.62		12.3	
AoAS-Pixel (X)	0.44		15.8	

4. CONCLUSION

This study investigated the impacts of soil data input on prediction of hydrologic variables. Two soil datasets currently available in Australia were considered to simulate hydrological processes in a semi-arid catchment across spatial scales, using SWAT model. Due to the lack of observation for water balance components, the

daily streamflow at the catchment outlet and the satellite measured surface soil moisture from SMAP and SMOS were used to evaluate the impacts of soil dataset on the streamflow and near surface soil moisture prediction. Similar to previous studies, our results indicated the soil dataset with high resolution and improved information content may not lead to better streamflow prediction at the catchment outlet. However, higher resolution SLGA, compared to AoAS, enhanced the surficial soil moisture prediction at catchment as well as upland areas simultaneously. This can have important implication for ecosystem and land surface feedback to the atmosphere. Further investigation will consider the uncertainty associated with soil attributes of SLGA on soil moisture estimation and streamflow prediction.

REFERENCES

- Abbaspour, K. C. (2015). SWAT-CUP 2012: SWAT Calibration and Uncertainty Programs: A User Manual. Department of Systems Analysis. Integrated Assessment and Modelling (SIAM), Eawag, Swiss Federal Institute of Aquatic Science and Technology, Duebendorf, Switzerland.
- Anderson, R. M., Koren, V. I., & Reed, S. M. (2006). Using SSURGO data to improve Sacramento Model a priori parameter estimates. Journal of Hydrology, 320(1-2), 103-116.
- Cotter, A. S., Chaubey, I., Costello, T. A., Soerens, T. S., & Nelson, M. R. (2003). Water quality model output uncertainty as affected by spatial resolution of input data. Journal of American Water Resources Association, 39(4), 977-986.
- Daggupati, P., Shukla, R., Mekonnen, B., Rudra, R., Biswas, A., Goel, P., . . . Yang, W. (2018). Hydrological Responses to Various Land Use, Soil and Weather Inputs in Northern Lake Erie Basin in Canada. Journal of Water, 10(2).
- Di Luzio, M., Arnold, J. G., & Srinivasan, R. (2009). Integration of SSURGO maps and soil parameters within a geographic information system and non point source pollution model system. Journal of Soil and Water Convservation, 59(4), 123-133.
- Entekhabi, D, Njoku, E.G., O'Neil, P.E., Kellogg, K,H., Crow, W.T., Edelstein, W.N., Entine, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, K.C., Mcdonald, K.C., M. Moghaddam, S. Moran, R. Reichle, J.C. Shi, M.W. Spencer, S.W. Thurman, L. Tsang, J. Van Zyl. (2010) The soil moisture active passive (SMAP) mission Proc. IEEE, 98 (5) (2010), pp. 704-716
- Geza, M., & McCray, J. E. (2008). Effects of soil data resolution on SWAT model stream flow and water quality predictions. Journal of Environmental Management, 88(3), 393-406.
- Grundy, M.J., Viscarra Rossel, R.A., Searle, R.D., Wilson, P.L., Chen, C., Gregory, L.J. (2015). Soil and landscape grid of Australia. Journal of Soil Research, 53. 835-844.
- Karthikeyan, L. Pan, M., Wanders, N., Kumar, D.N., Wood, E.F. (2017) Four Decades of microwave satellite soil moisture observations: Part 2. Product validation and inter-satellite comparisons.
- Levick, L. R., Semmens, D. J., Guertin, D. P., Burns, I. S., Scott, S. N., Unkrich, C. L., & Goodrich, D. C. (2004). Adding Global Soils Data to the Automated Geospatial Watershed Assessment Tool (AGWA). Proceeding 2nd SAHRA, (Sustainability of Semi-Arid Hydrologic and Riparian Areas), University of Arizona, International Symposium of Transboundary Water Management, Tucson, Az.
- Li, R., Zhu, A. X., Song, X., Li, B., Pei, T., & Qin, C. (2012). Effects of spatial aggregation of soil spatial information on watershed hydrological modelling. Hydrological Processes, 26(9), 1390-1404.
- Moriasi, D. N., & Starks, P. J. (2010). Effects of the resolution of soil dataset and precipitation dataset on SWAT2005 streamflow calibration parameters and simulation accuracy. Journal of Soil and Water Conservation, 65(2), 63-78.
- Northcote, K.H. (1979). A factual key for the recognition of Australian soils. 4th Edition. (Rellim Tech. Publ.: Glenside, S.A.).
- Robinson, J. S., Sivapalan, M., & Snell, J. D. (1995). On the relative roles of hillslope processes, channel routing, and network geomorphology in the hydrologic response of natural catchments. Journal of Water Resources Research, 31(12), 3089-3101.
- Rüdiger, C., Hancock, G., Hemakumara, H. M., Jacobs, B., Kalma, J. D., Martinez, C., . . . Willgoose, G. R. (2007). Goulburn River experimental catchment dataset. Journal of Water Resources Research, 43(10).
- Senanayake, I. P., Yeo, I. Y., Tangdamrongsub, N., Willgoose, G. R., Hancock, G. R., Wells, T., . . . Walker, J. P. (2019). An in-situ data based model to downscale radiometric satellite soil moisture products in the Upper Hunter Region of NSW, Australia. Journal of Hydrology, 572, 820-838.
- Singh, H. V., Kalin, L., & Srivastava, P. (2011). Effect of Soil Data Resolution on Identification of Critical Source Areas of Sediment. Journal of Hydrologic Engineering, 3, 253-262.
- Wang, X., Melesse, A.M. (2006). Effects of STATSGO and SSURGO as inputs on SWAT model's snowmelt simulation. Journal of American Water Resources Association. 42(5), 1217-1236.
- Williams, J. R., Kannan, N., Wang, X., Santhi, C., & Arnold, J. G. (2012). Evolution of the SCS Runoff Curve Number Method and Its Application to Continuous Runoff Simulation. Journal of Hydrologic Engineering, 17(11), 1221-1229.