The Use of Soil Taxonomy as a Soil Type Identifier for the Shasta Lake Watershed Using SWAT

D. L. Ficklin, Y. Luo, M. Zhang, S. E. Gatzke

Published in *Transactions of the ASABE* 57(3): 717-728 (doi: 10.13031/trans.57.9557).

Submitted for review in December 2011 as manuscript number SW 9557; approved for publication by the Soil & Water Division of ASABE in May 2014.

The authors are **Darren L. Ficklin,** Assistant Professor, Department of Geography, Indiana University, Bloomington, Indiana; **Yuzhou Luo,** Research Scientist, and **Minghua Zhang,** Professor, Department of Land, Air, and Water Resources, University of California, Davis, California; **Sarah E. Gatzke,** Water Resource Control Engineer, State Water Resources Control Board, Sacramento, California. **Corresponding author:** Minghua Zhang, 131 Veihmeyer Hall, University of California, Davis, CA 95616; phone: 530-752-4953; e-mail: mhzhang@ucdavis.edu.

Abstract. *We tested a methodology for aggregating soil properties across multiple soil survey areas according to soil taxonomic information available within the Soil Survey Geographic (SSURGO) dataset. Most hydrologic modeling studies using SSURGO assume that soil property data are adequately grouped into a "soil type" that is represented by a map unit key within SSURGO. The map unit key in SSURGO, however, is not intended for this purpose because the map unit key is not guaranteed to be the same between adjacent surveys. As a result, similar soil types are assigned a different map unit key across soil survey boundaries resulting in an artificial increase of soil types. The Soil and Water Assessment Tool (SWAT) was used to simulate hydrology using data from the low-resolution State Soil Geographic Database (STATSGO) dataset, the SSURGO dataset using map unit key as a soil type identifier (MUKEY), and the SSURGO dataset using soil taxonomy as a soil type identifier (TAXSUB; TAXa by SUBbasin) in the Shasta Lake watershed in northern California and southern Oregon. Results indicate that, of the soil properties tested, only TAXSUB maximum soil depth was significantly different from the STATSGO maximum soil depth, while both TAXSUB maximum soil depth and saturated hydraulic conductivity were significantly different from the MUKEY dataset. Based on SWAT streamflow output, the TAXSUB soil dataset generated streamflow results closer to observed streamflow data as compared to the STATSGO or MUKEY inputs, which generated streamflow values much higher than observed data. The TAXSUB soil dataset had a greater maximum soil depth, resulting in more soil water infiltration. During large precipitation events, the soil column for TAXSUB may still be able to accommodate more water, while the STATSGO and MUKEY soil columns are completely saturated with water, leading to surface runoff. The soil taxonomy grouping method within SWAT produced more accurate streamflow results than using MUKEY and STATSGO but should still be further tested in other environmental settings.*

Keywords. Hydrology, Soils, SSURGO, STATSGO, Streamflow, SWAT, Watershed modeling.

Large-scale hydrological models require a great deal of spatial input data, such as elevation, land cover, and soils coverages. Ideally, field work and laboratory tests should be performed to determine these parameter values for a specific study area; however, this type of work is not always feasible when working with large watersheds or regions due to cost and time limitations. Therefore, hydrologic modelers frequently use freely available governmental or state databases for initial parameter values and then subsequently adjust the sensitive values to match simulated data with observed hydrologic or water quality data. Commonly used state and governmental databases in hydrologic modeling include: digital elevation model (DEM) data from the U.S. Geological Survey (USGS), National Land Cover Database (NLCD) land use data from the Multi-Resolution Land Characteristics Consortium (MLRC), and soil survey data from the USDA National Resources Conservation Service (NRCS).

The NRCS soil survey maps are the most comprehensive freely available spatial environmental datasets and are used as the primary source of soil physical and chemical properties for many surface and subsurface hydrologic models. Two spatially different NRCS soil survey databases are available for public use: the State Soil Geographic Database (STATSGO) and the Soil Survey Geographic Database 2.0 (SSURGO). STATSGO is a state-level soils database (1:250,000), while SSURGO is a detailed soil survey database, ranging in scale from 1:12,000 to 1:24,000 depending on the survey location (USDA-NRCS, 2007).

Both STATSGO and SSURGO can be used to provide the needed soil input for hydrologic modeling, depending on the modeling objective and resolution. Several studies in the literature have examined the differences in hydrologic output using both STATSGO and SSURGO (Di Luzio et al., 2004; Gowda and Mulla, 2005; Romanowicz et al., 2005; Anderson et al., 2006; Peschel et al., 2006; Wang and Melesse, 2006; Geza and McCray, 2008; Mednick, 2010; Moriasi and Starks, 2010; Gatzke et al., 2011; Sheshukov et al., 2011), with no general consensus on which soil dataset is better for hydrologic simulations (Mednick et al., 2008). Mednick et al. (2008) reviewed 18 studies comparing SSURGO and STATSGO inputs and found that, while hydrologic simulations with SSURGO inputs were more often closer to observed data, STATSGO inputs led to accurate hydrologic simulations in several cases. Due to the varying environmental settings and hydrological model setups between these studies, no general conclusion can be made about STATSGO and SSURGO inputs.

Hydrologic models query SSURGO and STATSGO differently, with STATSGO using a map unit identification (MUID) and SSURGO using a map unit key (MUKEY) for soil type identification. MUID and MUKEY are linked to a soil property database and, spatially, hydrologic models use these identification codes to extract soil properties for each spatial polygon. Because of the low STATSGO spatial resolution, MUID transcends governmental and county boundaries, while MUKEYs are unique within each soil survey area (SSA), which is often at the county level. SSURGO polygons crossing SSA boundaries should have approximately the same soil properties because they are dissected by an artificial boundary line. However, mapping style differences between project leaders for each county or state, as well as changing views on how soil resources are utilized, have led to some irreconcilable irregularities between survey areas (Gatzke et al., 2011). For example, in the San Joaquin Valley of California, Gatzke et al. (2011) examined SSURGO data across county boundaries and found that of the 1,460 soil polygons, 762 of the soil types were different across the corresponding boundary. These differences can artificially increase the amount of soil polygons used for modeling, thus increasing model complexity and inaccuracies.

Many hydrologic models, such as the Precipitation Runoff Modeling System (PRSM) (Leavesley et al., 1983), Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998), U.S. Army Corps of Engineers Hydrologic Engineering Center Hydrologic Modeling System (USACE-HEC, 1998), and Hydrologic Simulation Program - Fortran (HSPF) (Donigian et al., 1995), use the concept of hydrologic response unit (HRU) to balance input resolution and model complexity (Gatzke et al., 2011). An HRU is an area of "homogeneous" hydrologic characteristics determined by the spatial overlay of datasets such as elevation, land use, and soil type. The use of HRUs in a hydrologic model conserves some of the spatial variation in input parameters, while reducing model complexity by lumping areas with similar hydrologic characteristics into a single modeling unit (Neitsch et al., 2005).

For this study, we used the SWAT model to test the difference in soil aggregation strategies. SWAT performs hydrologic balance calculations at the HRU scale and then sums the individual hydrologic components to the subbasin scale. ArcSWAT (a preprocessor for spatial data for SWAT) was developed for using soil input in the STATSGO (or MUID) format, and therefore the SSURGO dataset cannot be used unless converted to a SWAT-specific format. Therefore, several tools are available to preprocess SSURGO data into SWAT format (Di Luzio et al., 2004; Peschel et al., 2006; Sheshukov et al., 2011; Luo et al., 2012). Most of these SSURGO preprocessing tools use the dominant soil component in a particular soil polygon to represent all soil properties. The problem with this method is that the dominant soil component may not accurately represent the other minor components. Further, the dominant soil component may not be the dominant soil component at all, covering only a small fraction of the soil polygon (Luo et al., 2012). The tool developed by Luo et al. (2012), however, uses the area-weighted average of all components within a particular soil polygon so that all soil components within a polygon are represented. We used the method developed by Luo et al. (2012) for this study.

In terms of SWAT, the previously mentioned differences of soil data between soil survey boundaries will artificially inflate the number of HRUs, thereby increasing the number of "different" soil types, resulting in increased model complexity. We follow the same procedure as a previous study (Gatzke et al., 2011) that used Great Group soil taxonomy as a defining feature that relates SSURGO polygons rather than MUKEY in the highly managed San Joaquin River watershed. However, this work assesses the method used by Gatzke et al. (2011) in a mountainous region rather than an agricultural region. Great Group soil taxonomy provides a robust and mature framework for grouping soils based on physical and chemical properties at several levels of generalization (USDA-NRCS, 1999). Results from Gatzke et al. (2011) indicate that using Great Group soil taxonomy as a defining feature should be further tested in an unmanaged watershed, as the reservoir outflows in the San Joaquin River watershed may have masked hydrologic differences in the different soil data aggregation methods. We also compare the use of Great Group soil taxonomy against the MUID identifier in STATSGO. The objective of this study, then, is to further test the use of Great Group soil taxonomy as a definition of soil type rather than the database construct of MUID or MUKEY on the Shasta Lake watershed in northern California.

Materials and Methods

SWAT Hydrologic Model

SWAT is a river basin scale model designed to simulate watershed and water quality processes (Arnold et al., 1998) that has been successfully applied throughout the world (Gassman et al., 2007). SWAT simulates the entire hydrologic cycle, including surface runoff, lateral soil flow, evapotranspiration, infiltration, deep percolation, and groundwater return flows. For this study, surface runoff was estimated using the USDA-SCS curve number (CN) method (USDA-SCS, 1984), and evapotranspiration was estimated using the Penman-Monteith method (Penman, 1956; Monteith, 1965). Soil water can be removed by evapotranspiration, deep percolation for aquifer recharge, or move laterally in the soil column for streamflow contribution. Groundwater return flow is estimated based on the groundwater balance, where shallow and deep aquifers can contribute to streamflow. SWAT uses a temperature index-based approach to estimate snow accumulation and snowmelt processes. A full description of SWAT can be found in Neitsch et al. (2005).

Spatial parameterization within SWAT is performed by dividing the watershed into subbasins based on topography. The subbasins are further divided into HRUs based on unique combinations of land use and soil characteristics. Surface runoff, infiltration, evaporation, plant water uptake, lateral flow, and percolation to the shallow and deep aquifer are modeled for each HRU, and the components are summed to the subbasin level. The components are then routed through the channel network to the watershed outlet.

Model Input Data

SWAT input parameter values from topography, land cover, and soils data were compiled using databases from governmental agencies. A 30 m digital elevation model (DEM) was obtained from the USGS for watershed and stream delineation and estimation of stream slopes. The 2001 National Land Cover Database from the USGS was used to define land cover. STATSGO and SSURGO data were extracted from the NRCS Soil Data Mart (USDA-NRCS, 2012). Natural flow data at the Shasta Lake watershed outlet from 1950 to 2005 for streamflow comparisons were gathered from the California Data Exchange Center (CDEC) developed by the California Department of Water Resources. The CDEC natural flow data are based on measured flow that was corrected for upstream diversions, impoundments, and other manmade alterations and therefore is a suitable dataset for comparison of SWAT simulations. Daily climate data of a 1/8 degree (~12 km) spatial resolution from 1950 to 2005, including precipitation, maximum and minimum temperature, and wind speed, were obtained from gridded observed meteorological data (Maurer et al., 2002).

SSURGO, MUKEY, and Taxonomic Profile Aggregation Algorithm

A large amount of effort has been invested in standardizing MUKEY legends across SSURGO survey boundaries (USDA-NRCS, 1999). Nevertheless, irregularities between survey areas still remain due to evolving soil taxonomy definitions, mapping style differences among soil surveyors, and changing views on how soil resources are utilized (Gatzke et al., 2011). This may have implications when using SSURGO data for large-scale modeling applications, as the soil landscapes are artificially dissected and may differ along political boundaries despite being SSURGO-certified (Drohan et al., 2003; USDA-NRCS, 2009).

Distributed hydrologic models, such as SWAT, rely on spatial maps of horizon-level soil property information linked by an identification code such that any given soil type within a profile will correspond to a single profile of soil properties. In SSURGO, soil profile information is associated with soil components, which are grouped together into map units (fig. 1). A soil component is defined as a group of soils with similar properties, interpretations, and productivity. In the SSURGO database, map units are spatially defined, but the soil components within the map units have no spatial reference. Therefore, when map units are used to define soil type, the user must either select the largest soil component to represent the entire map unit or create a soil horizon template that combines soil property data from each component into a single set of horizon-level information for every map unit. The second method is most appropriate when using a collection of map units (and their components) that is defined by some higher-level criterion, such as Great Group soil taxonomic membership (Gatzke et al., 2011). In this study, SSURGO profile aggregation by MUKEY (hereafter MUKEY) and Great Group taxonomy (TAXa by SUBbasin; hereafter TAXSUB) were weighted according to component area so that larger components contribute most to the final representative profile without losing information from smaller components. The aggregation algorithm is fully documented and publicly available in the “aqp” package for R(Beaudette and O’Geen, 2010). It is important to note that STATSGO was not aggregated for this study. Further, STATSGO is seamless across state and county boundaries, and thus there are no edge-matching problems. Other SSURGO studies have used the dominant soil component rather than the weighted average of the components (Di Luzio et al., 2004; Peschel et al., 2006; Sheshukov et al., 2011).

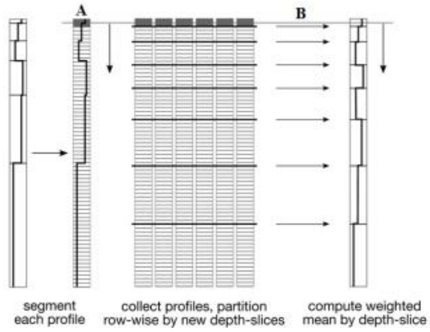


Figure 2. Soil profile algorithm showing (A) the soil property depth-function algorithm and (B) soil profile aggregation into a representative soil profile.

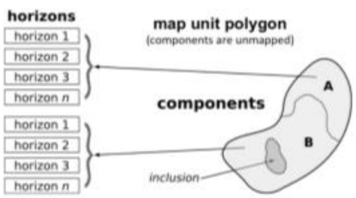


Figure 1. SSURGO data architecture (from Gatzke et al., 2011).

The aggregation algorithm is briefly explained here and explained in detail by Gatzke et al. (2011). First, the soil data profiles for MUKEY and TAXSUB must be converted into a standardized soil profile structure for each soil property in every soil profile (fig. 2). For this study, the soil horizon structure used is as follows: 0-5 cm, 5-10 cm, 10-15 cm, 15-30 cm, 30-60 cm, 60-90 cm, 90-150 cm, and 150-250 cm. If the soil horizon structure depth is greater than 250 cm, the soil layer structures are set to a 100 cm increment until the maximum soil depth is reached. The soil properties necessary for SWAT include: sand, silt, and clay percentages, bulk density, cation exchange capacity, saturated hydraulic conductivity, soil depth, water storage, organic carbon, and soil erodibility factor.

Second, a step is required to reduce the collection of soil profiles to a single representative soil profile with a set of depth functions for each soil property (fig. 2). This second aggregation step varies slightly depending on the grouping strategy (MUKEY or TAXSUB). For the MUKEY dataset, representative soil profiles are derived from the collection of components within each map unit or polygon. For the TAXSUB soil input, representative soil profiles are derived from the collection of components associated with polygons derived from merging adjacent map unit polygons of the same Great Group taxa within each subbasin (and thus, TAXSUB).The contribution of a single component's data to a representative depth function is weighted by the area of that component within an individual polygon for the MUKEY soil input or by the associated subbasin TAXSUB input (fig. 2). We assume that the soil properties for a taxonomy are approximately equal within an individual subbasin but may differ between subbasins. Therefore, soil profiles in TAXSUB are tagged with the subbasin number at the end of the taxonomic ID, such as "Haploxeralfs-3" for the taxonomic group Haploxeralfs found in subbasin 3.

The entire aggregation algorithm can be broken down into three steps (Gatzke et al., 2011):

1. Each soil profile from the SSURGO database is extracted and parsed into 1 cm slices and then aggregated into the previously mentioned soil horizon structure for the MUKEY and TAXSUB datasets. STATSGO is not aggregated.

2. The collection of segmented soil profiles defined by a grouping variable such as map unit key or combination of subbasin and Great Group level taxonomic units are combined, column-wise, into a matrix. If needed, columns in this matrix are padded with "not available" (or NA) values to the depth of the deepest profile within the group. The resulting matrix is partitioned, row-wise, at the standardized horizon boundaries previously described.

3. Weighted-mean soil property values are computed within each partition and assigned to each standardized horizon.

This process was repeated for each soil property needed by the SWAT model. See Gatzke et al. (2011) and Beaudette and O'Geen (2010) for a complete description of how the SSURGO and TAXSUB datasets were processed for SWAT. See the ArcSWAT manual (Winchell et al., 2007) for a complete description of how STATSGO was processed for this study.

HRU Generation

Following soil data aggregation, HRUs are generated for SWAT. The ArcSWAT version of the SWAT model provides users with the graphic user interface tool to create HRUs. The user imports land cover, soil, and slope coverages, and the automated ArcSWAT tool calculates HRUs by overlaying land use, soil, and elevation/slope coverages. Slope was evenly partitioned into four classes: 0% to 8%, 8.1% to 16%, 16.1% to 150%, and 150.1% to 300%, where the watershed slope median was 16%. Slope percentages were calculated as the rise of elevation divided by the length of the elevation and then multiplied by 100 to obtain a percent value. A slope of 300% represents a 71.5° slope, and a slope of 100% represents a 45° slope. The HRU calculation depends on a user threshold for land cover, soil area thresholds, and slope area coverages. For example, a specific HRU land unit may contain sandy loam, evergreen forest, and a 10% slope. Only land cover, soil types, and slope areas occurring at or greater than the user-defined threshold percentage within a subbasin are represented by HRUs. The land use, soil, and slope areas that are not above the user-defined threshold are not modeled. The land area occupied by the land use, soil type, and slope area combinations below the user-defined threshold are redistributed between the combinations above the user-defined threshold so that 100% of the watershed is represented. For this study, the land use, soil area, and slope area threshold percentages were the same for each soil input dataset at 20%, 5%, and 20%, respectively. The median HRU area for the STATSGO, SSURGO, and TAXSUB datasets were 5,028, 4,475, and 5,431 acres, respectively, which is approximately 18 to 20 km^2. These HRU threshold percentages are similar to our previous work in the Sacramento and San Joaquin River watersheds in California (Luo et al., 2008; Ficklin et al., 2013). Similar land uses, soil types, and slope areas are lumped together within the subwatershed for model simplicity and therefore do not have a spatial location within the subbasin. In this study, the elevation and land use data remained constant for each soil input so that the terrain characteristics remained the same. Therefore, the output of the hydrologic model depends only on differences in the soil inputs.

We recognize that the large spatial resolution of HRUs for this study could lead to biased results. With an HRU threshold percentage of 20% for land cover/use, 5% for soils, and 20% for slopes, a number of land covers/uses can be potentially "lumped" in with other land covers/uses, assuming that other major land covers/uses are near, but below, the 20% land cover/use cutoff. However, the Shasta Lake watershed is dominated by only two land cover/use types: evergreen forest (FRSE) and brush rangeland (RNGB). The evergreen forest land cover encompasses 53% of the watershed, and the brush rangeland encompasses 23% of the watershed, which add up to approximately 76% of the entire watershed. The only other land cover that has a considerable land cover percentage is range grassland (RNGE) with 10%. No other land cover/use is above 3% coverage. Therefore, three land covers encompass 86% of the entire Shasta Lake watershed, and using a threshold of 20% would likely only include the major land covers/uses. Additionally, the small "potential" HRUs (which would be defined with <20% land use for the watershed and >5% soil map unit for each land use) were not excluded in the simulation but aggregated in the major HRUs.

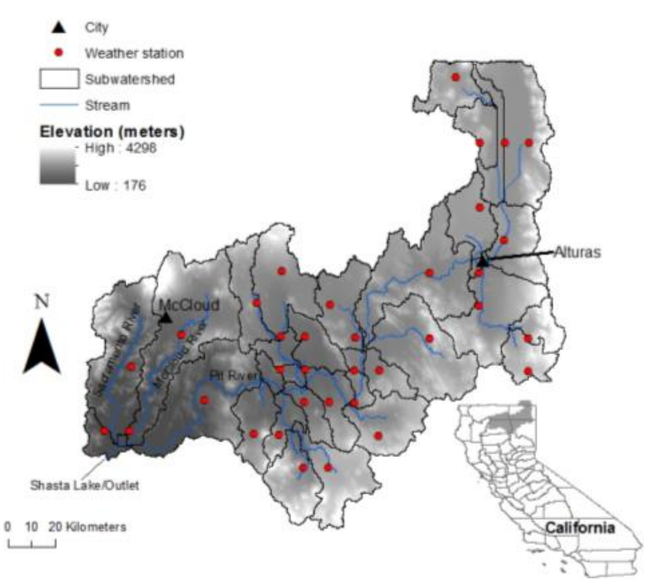


Figure 3. Map of the Shasta Lake watershed.

Inherently, the soil aggregation method presented in this article decreases HRU resolution by lumping soils by Great Group taxonomy, leading to a decrease in model complexity and inaccuracies. While the large spatial resolution of HRUs could lead to biased results, there is no general consensus on this bias within SWAT. Several studies have found that increasing HRU resolution did not lead to better streamflow simulation (Fitzhugh and Mackay, 2000; Jha et al., 2004; Tripathi et al., 2006; Muleta et al., 2007; Kumar and Merwade, 2009; Cho et al., 2010). One study (Mamillapalli et al., 1996) found that HRU subdivision led to better streamflow simulation, but only up to 215 km^2. Therefore, for this study, it is difficult to conclude that increased HRU resolution will lead to better streamflow results without including simulations with differing HRU resolutions.

Study Area

The Shasta Lake watershed (SLW) is located in northeastern California, with a small portion in southeast Oregon (fig. 3). The total study watershed area for this study is 18,839 km^2. Three major rivers flow into Shasta Lake: Upper Sacramento River (average monthly flow of 35 $m^3\ s^{-1}$), McCloud River (49 $m^3\ s^{-1}$), and Pit River (122 $m^3\ s^{-1}$), with Pit River being the dominant watershed that feeds Shasta Lake (82% of the SLW watershed area). The total number of watersheds delineated based on elevation for this study was 35 (fig. 3). The SLW is composed of three major land uses: evergreen forest (53% of watershed area), brush range (24%), and grass range (10%). Soils in the watershed differ depending on their location within the SLW, with the valley soils dominated by clayey soils, the plateau and foothill soils dominated by gravelly and sandy soils, and the mountain soils dominated by gravelly soils (table 1).

Table 1. Description of soil series in Shasta Lake watershed (adapted from Vestra, 2004).

Soil Series		Description
Valley soils		
	Modoc-Oxendine-Bieber	Gravelly sandy clay
	Pittville-Dudgen-Esperanza	Clay and sand
	Aikman-Cardon	Clay, silty clay, and clay loam
	Deven-Bieber-Pass-Canyon	Clay and cobbly clay loam
Plateau and foothill soils		
	Jellycamp-Jellico-Adinot	Gravelly sandy clay
	Alcot-Sadie-Germany	Loam, clay loam, and cobbly clay loam
	Loveness-Hunsinger-Lava Flows	Gravelly clayey clay
	Jimmerson-Gasper-Scarface	Equal parts sand, silt, and clay
Mountain soils		
	Gosch-Witcher-Trojan	Gravelly sandy clay
	Divers-Lapine-Kinzel	Gravelly coarse sand and gravelly sandy loam

	Rivalier-Tionesta-Blankout	Equal parts sand, silt, and clay
	Anatone-Bearskin-Merlin	Gravelly clay and cobbly silty clay loam
	Cheadle-Superior-Behanin	Gravelly sandy loam and cobbly clay loam
	Canyoncreek-Hermit	Clayey silty sand

The SLW has a montane Mediterranean climate with hot, dry summers and cool, wet winters. For Alturas (elevation 1,332 m; location shown in fig. 3), average monthly precipitation ranges from 0.81 cm in July to 3.98 cm in January with an average annual precipitation of 31 cm. The highest maximum temperatures occur in July at approximately 31°C, while the lowest minimum temperatures occur in January at approximately -9°C. For McCloud (elevation 991 m; location shown in fig. 3), average monthly precipitation ranges from 0.5 cm in July to 21 cm in January with an average annual precipitation of 120 cm. The highest maximum temperatures occur in July at approximately 31°C, while the lowest minimum temperatures occur in January at approximately -4.5°C.

Soil Dataset Assessment Methods

The differences of three soil datasets (STATSGO, MUKEY, and TAXSUB) were assessed by comparing the differences in uncalibrated streamflow at the Shasta Lake inlet. All model scenarios were left uncalibrated, as calibration may disguise the differences that result from applying different soil inputs. We chose to leave the SWAT models uncalibrated so that the manipulation of surface runoff (curve number, surface runoff travel time, etc.) and physical soil parameters (bulk density, sand percentage, etc.) would not have an effect on the model results. It is likely that changing pertinent parameters would result in an adequate matching of streamflow for all soil datasets; however, the point of the study is to compare the raw (or un-manipulated) data within the soil datasets. Additionally, the uncalibrated model results show how well each dataset predicts before calibration, which is an indication of the amount of effort needed to calibrate the model (Geza and McCray, 2008). This is also the same method as our previous study (Gatzke et al., 2011). Daily simulations were performed and aggregated to monthly and annual time steps. Listed streamflow descriptive statistics were used for soil dataset comparison: average, standard deviation (SD), maximum, minimum, kurtosis, and skewness.

For comparison with observed unimpaired streamflow data, five efficiency statistics were calculated: Nash-Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970), coefficient of determination (R^2), percent bias (PBIAS; Gupta et al., 1999), and the ratio of root mean square error (RMSE) to the standard deviation of the observations, also known as RSR. NSE indicates how well the plot of observed versus simulated values fits the 1:1 line. NSE ranges from negative infinity to 1, where 1 indicates a perfect simulation. R^2 is a determination coefficient that measures the degree of variation of the simulations explained by the observed data. R^2 ranges from 0 to 1, with higher values indicating higher collinearity. PBIAS measures the tendency of simulated data to be larger or smaller than observed data. An optimal PBIAS value is 0. Positive PBIAS values indicate model underestimation, while negative PBIAS values indicate overestimation. RSR standardizes RMSE using the standard deviation of the observations, combing both an error index and a normalization statistic that can be compared against different simulations. The Mann-Whitney rank sum test ($p < 0.05$, two-sided) was used to compare the soil datasets since the hydrologic outputs exhibited non-normal distributions (Mann and Whitney, 1947). The Mann-Whitney rank sum test is based on the concept that if two groups come from the same distributions, the sum of the ranks of the values should be somewhat equally distributed between the two.

The soil properties of available water capacity (AWC), saturated hydraulic conductivity (K), and maximum soil depth were chosen for analysis, as they are properties most relevant for hydrology. AWC is defined as the amount of water that an increment of soil depth can store that is available to plants. A higher AWC indicates a higher water storage capacity. K is a numerical description of how fast water can move through the soil. The maximum soil depth can be an indicator of the volume of water that a soil column can hold, where greater soil depths can hold more water. The curve number (CN) was also chosen for analysis. CN values are empirical parameters generated from soil and land use values; a lower CN indicates low runoff potential and a higher CN indicates high runoff potential. CN values are calculated based on soil hydrologic groups A, B, C, and D, where A has the highest infiltration rate and D has the lowest. Each land use unit (e.g., agricultural) has a different set of CN values based on its soil hydrologic groups. For CN values, the descriptive groupings are aggregated based on numerical conversion, with group A = 1, B = 2, C = 3, and D = 4. These values are aggregated by area weight and then converted back to letters using the same conversion. Numerical values ranging from 0 to 1.5 result in an A hydrologic group, 1.5 to 2.5 in a B hydrologic group, 2.5 to 3.5 in a C hydrologic group, and 3.5 to 4 in a D hydrologic group. Since the same land use dataset is used for every soil dataset, the CN differences are entirely related to differences in soil properties.

Results and Discussion

Soil Aggregation Results

The total number of different soil types within the SLW was 63 for the STATSGO dataset, 158 for the MUKEY dataset, and 172 for the TAXSUB dataset (fig. 4). Area-weighted means of the soil properties relevant to SWAT showed mostly minor differences between soil datasets (table 2). The main difference between soil datasets was average maximum soil depth, where the TAXSUB soil dataset was approximately 600 mm deeper than the STATSGO or MUKEY soil datasets (table 2). Based on the Mann-Whitney rank sum tests of several important soil characteristics, the TAXSUB dataset showed a significantly different ($p < 0.05$) value of saturated hydraulic conductivity compared to MUKEY and a larger ($p < 0.05$) value of maximum soil depth than both STATSGO and MUKEY. There were no other differences in the area-weighted means of the soil properties identified with statistical significance between TAXSUB and STATSGO and between TAXSUB and SSURGO (table 2).

Spatial patterns of AWC, *K*, CN, and maximum soil depth are shown in figure 5. AWC was commonly between 0 and 0.18 cm cm^{-1} for all soil datasets, with approximately 88%, 77%, and 80% of the STATSGO, MUKEY, and TAXSUB subbasins, respectively, within this range. The MUKEY and TAXSUB soil datasets contained an equal percentage of high AWC (0.24 to 0.44 cm cm^{-1}) subbasins, while the STATSGO dataset did not contain any subbasins in this range. The weighed-average *K* subbasin values varied spatially between soil datasets, with *K* generally being highest in the southwestern and northeastern portions of the SLW. There was little variation of *K* between the soils of the subbasins in the TAXSUB soil dataset, with approximately 68% of the subbasins within the 33.6 to 94.3 mm h^{-1} range and no subbasins containing a weighted-average *K* greater than 144 mm h^{-1}. The MUKEY soil dataset had a *K* distribution similar to that of the TAXSUB dataset. Weighted-average *K* for the STATSGO soil dataset was spread somewhat evenly throughout the SLW, with 9% of the subbasins within the 144.1 to 225.5 mm h^{-1} range and 31% within the 5.5 to 33.5 mm h^{-1} range. For CN, the weighted-average values were generally evenly distributed throughout the SLW, resulting in no clear spatial pattern. CN values for the TAXSUB soil dataset showed the least variation, with over 45% of the subbasins between a CN of 57.9 and 67.4. The STATSGO and MUKEY spatial patterns for the maximum soil depths are similar, while the TAXSUB maximum soil depth is larger throughout the SLW.

Table 2. Soil properties of first soil layer.

Soil Property	STATSGO	MUKEY	TAXSUB
Sand (%)	49.4	52.0	51.7
Silt (%)	32.1	30.5	30.3
Clay (%)	14.7	16.8	17.1
Bulk density (g cm^{-3})	1.4	1.3	1.4
K (mm h^{-1})	109.3	59.3	70.6[a]
AWC (cm cm^{-1})	0.10	0.20	0.10
Max. soil depth (mm)[b]	1,130	1,127	1,749[c],[a]
Curve number[d]	66.4	66.4	66.3

[a] Statistically different from MUKEY at $p = 0.05$.

[b] Depth of the entire soil column.

[c] Statistically different from STATSGO at $p = 0.05$.

[d] Based on land use and soil data.

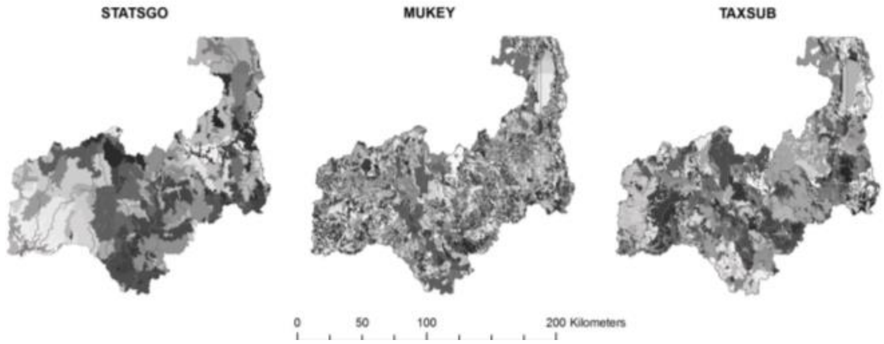


Figure 4. Maps of soil types for each soil dataset within the Shasta Lake watershed. The STATSGO dataset contained 63 soil types, the MUKEY dataset contained 158 soil types, and the TAXSUB dataset contained 172 soil types.

HRU Generation

The HRU soil threshold percentage of 5% led to the generation of 462, 385, and 406 HRUs for the STATSGO, MUKEY, and TAXSUB datasets, respectively. The MUKEY and TAXSUB soil datasets have a higher spatial resolution than the STATSGO soil dataset, and therefore a soil threshold percentage of 5% excluded more soil types than the STATSGO soil dataset, resulting in a lower number of HRUs. Due to STATSGO's lower resolution, there is a greater chance that a STATSGO soil polygon has an area greater than 5% of the subbasin area as compared to the MUKEY and TAXSUB soil datasets.

Dataset Effects on Simulated Streamflow

Based on the monthly streamflow descriptive statistics (table 3 and fig. 5), the TAXSUB soil input generated streamflow results more similar to the observed data as compared to the STATSGO or MUKEY inputs for all statistical categories except R^2 (table 3). For all soil datasets, the standard deviation and maximum streamflow results were inaccurately predicted, suggesting that further calibration is needed to capture the natural streamflow variability of the SLW. Minimum streamflow values generated from the STATSGO and MUKEY soil datasets simulated periods of no flow, which, according to the observed data, does not occur in an unimpaired setting. However, the TAXSUB monthly streamflow output did not simulate any period of no flow, which is a result of the greater soil depth that is able to provide soil water to the streams during the summer when precipitation is negligible.

The model efficiency statistics also indicate that using the TAXSUB soil dataset led to better predictions of monthly streamflow. Based on the work by Moriasi et al. (2007), if the model efficiency statistics are NSE > 0.5, RSR = 0.7, and PBIAS = ±25%, then the model calibration can be deemed satisfactory. The model efficiency statistics of the uncalibrated simulation using the TAXSUB soil dataset are closer to these satisfactory requirements (table 3), indicating that completing the model calibration may take much less effort as compared to the STATSGO and MUKEY soil datasets.

A plot of average monthly streamflow results also indicated that the simulations with the TAXSUB input are closer to the observed data compared to STATSGO and MUKEY (figs. 6 and 7). The largest differences between

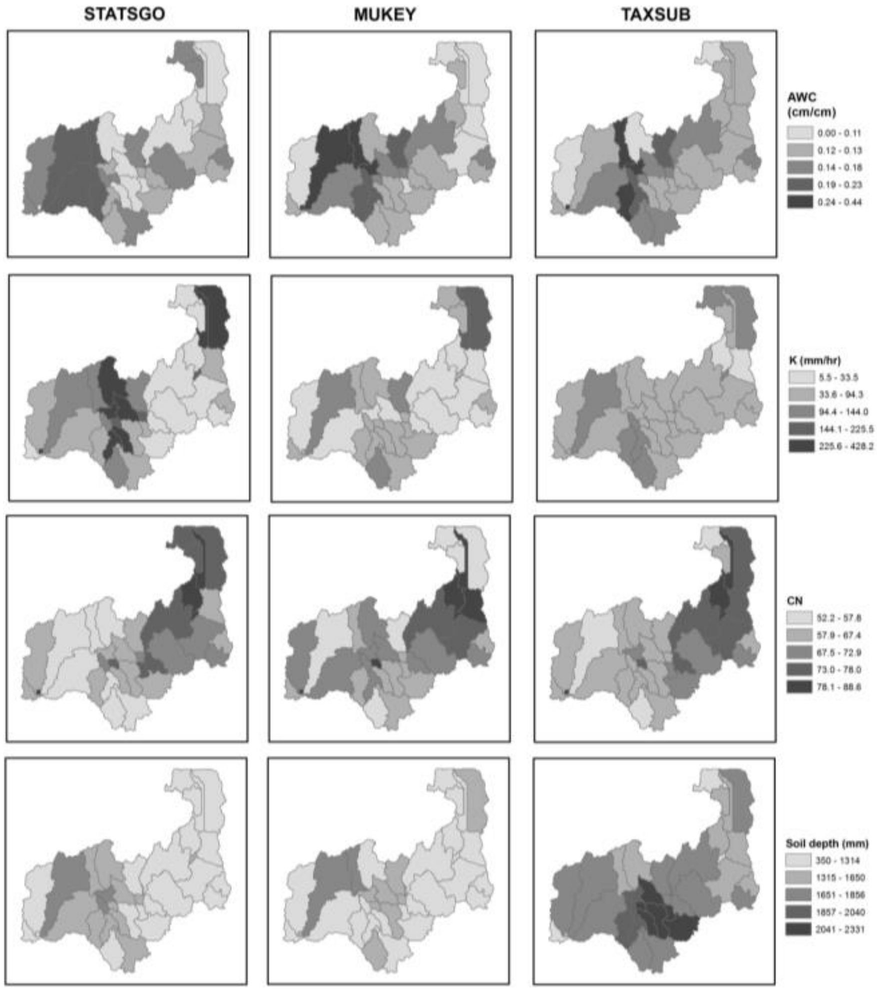


Figure 5. Area-weighted soil properties for each subbasin.

all soil dataset simulations occur during the winter months (December through March) when most of the precipitation and snowmelt occurs. On average, the winter streamflow simulations by STATSGO, MUKEY, and TAXSUB are 292, 255, and 118 $m^3 s^{-1}$ greater than the observed data, respectively (fig. 6). As previously mentioned, the summer streamflow simulations with the STATSGO and MUKEY soil datasets predict periods of low flow (August and September), while the TAXSUB soil dataset simulations predict nearly double the amount of streamflow than predicted by the other soil datasets for this time period.

Streamflow percent exceedances are shown in figure 7. For the monthly percent exceedance streamflow values, the STATSGO and MUKEY streamflow simulations generally follow the same percent exceedance curve, which is due to the fact that they share similar soil property values (table 2). Using the STATSGO and MUKEY soil datasets, monthly streamflow was overpredicted compared to the observed streamflow approximately 57% of the time, while using the TAXSUB soil dataset overpredicted streamflow approximately only 35% of the time. It is important to note that using the STATSGO and MUKEY soil datasets predicted periods of no flow approximately 10% of the time.

Table 3. Descriptive statistics of monthly streamflow.

Statistic	Observed	STATSGO	MUKEY	TAXSUB

Mean (m^3 s^{-1})	239.9	334.4	327.1	255.6
SD (m^3 s^{-1})	194.0	385.5	360.9	306.1
Max. (m^3 s^{-1})	1347.0	2111.0	1892.0	1856.0
Min. (m^3 s^{-1})	73.8	0.0	0.0	9.1
Kurtosis	6.4	2.9	2.5	6.1
Skewness	2.3	1.7	1.6	2.3
NSE	-	-0.57	-0.18	0.43
R^2	-	0.87	0.89	0.86
PBIAS	-	-40.7	-37.6	-6.5
RSR	-	1.3	1.1	0.8
RMSE (m^3 s^{-1})	-	243.2	211.0	146.9

As with the monthly simulations, using the TAXSUB soil dataset led to more accurate annual streamflow predictions than using the STATSGO and MUKEY soil datasets (table 4). However, using the STATSGO and MUKEY soil datasets led to more accurate predictions of the annual streamflow minimum and extreme distribution values (kurtosis and skewness), which suggests that using the STATSGO and MUKEY soil datasets may lead to better predictions of extreme (wet or dry) years compared to the TAXSUB soil datasets at the annual time scale. Based on the model efficiency statistics in table 4, similar results were found as with the monthly simulations, with the TAXSUB simulations producing near satisfactory calibration requirements. These findings are further illustrated in figure 6.

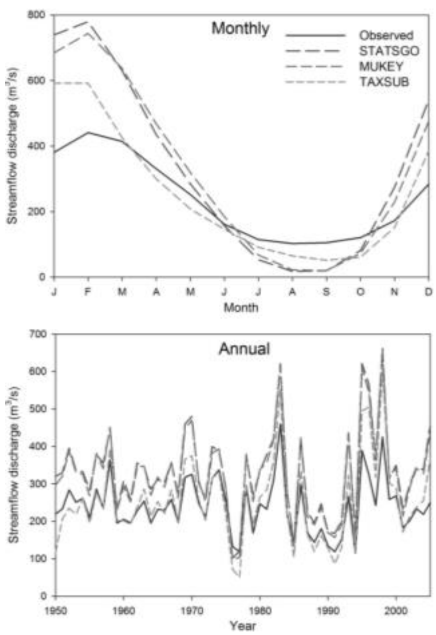


Figure 6. Average monthly streamflow simulated from 1950 to 2005 at the Shasta Lake watershed outlet for each soil dataset.

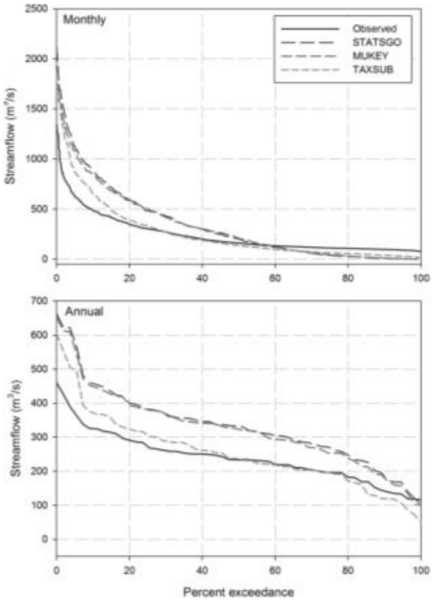


Figure 7. Streamflow percent exceedance at the Shasta Lake watershed outlet for each soil dataset.

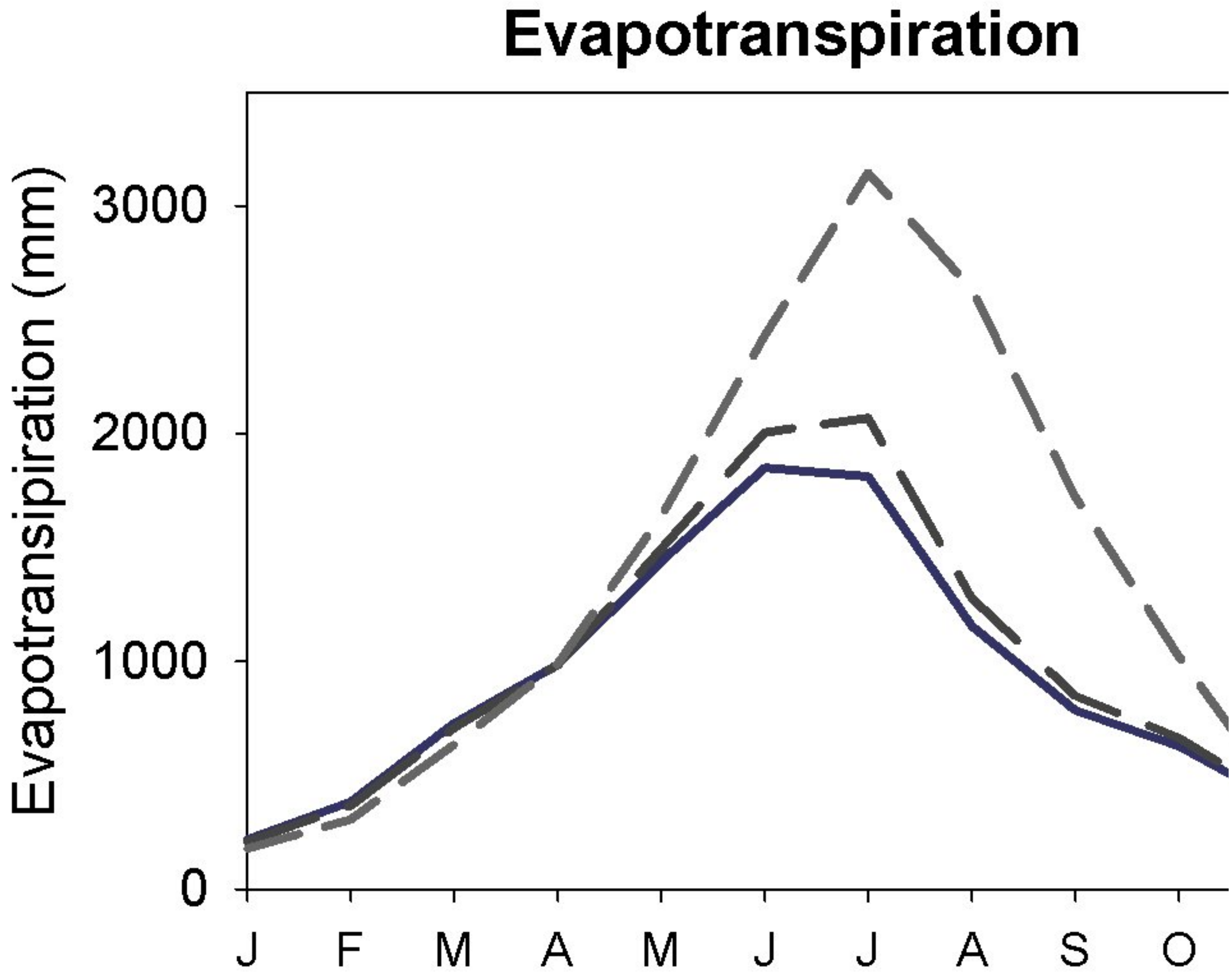
Table 4. Descriptive statistics of annual streamflow.

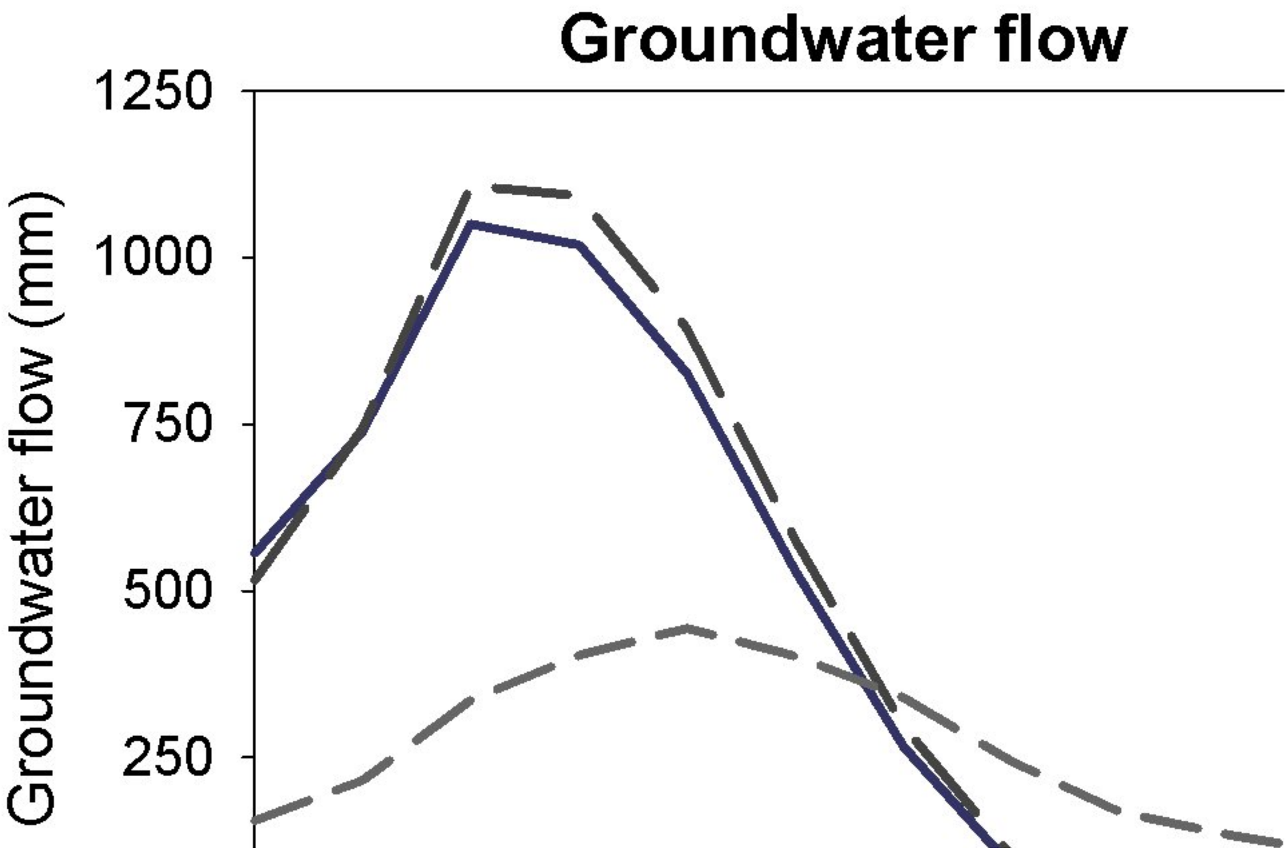
Statistic	Observed	STATSGO	MUKEY	TAXSUB
Mean (m^3 s^{-1})	239.9	332.3	325.1	254.2
SD (m^3 s^{-1})	73.8	118.0	117.9	113.1
Max. (m^3 s^{-1})	460.0	662.8	653.2	603.4
Min. (m^3 s^{-1})	115.9	102.1	100.2	49.8
Kurtosis	0.8	1.0	0.9	1.4
Skewness	0.7	0.7	0.6	0.9
NSE	-	-1.14	-0.86	0.48
R^2	-	0.89	0.91	0.88
PBIAS	-	-38.6	-35.5	-5.9
RSR	-	1.45	1.35	0.71
RMSE (m^3 s^{-1})	-	107.0	99.9	52.5

Using STATSGO and MUKEY led to an overprediction of annual streamflow approximately 96% of the time, compared to 67% of the time when using the TAXSUB soil dataset (fig. 7). As previously stated, the extreme annual streamflow events were better captured when using STATSGO and MUKEY, but based on the efficiency statistics using the TAXSUB soil dataset led to an overall better annual streamflow prediction. This can be seen in figure 7, where using the TAXSUB soil dataset led to moderately accurate annual streamflow predictions 20% to 90% of the time and accurate annual streamflow predictions 40% to 85% of the time.

Reasons for Differences in Streamflow Predictions

Based on the spatial and tabular summaries of the soil property differences (table 2 and fig. 5), the main reasons for differences in streamflow are the differences in *K* and maximum soil depth. It is expected that *K* should be different between soil datasets, given that the distribution of *K* is lognormal and largely dependent on extreme values. Maximum soil depth is larger for TAXSUB due to the method of aggregation. For the TAXSUB aggregation, all soils sharing a Great Group taxonomy within a particular subbasin are aggregated. Similar to the MUKEY method, the maximum soil depth of soils with the same taxonomy is set to the maximum value of the taxonomic set. For example, a TAXSUB soil type may contain three soil components with soil depths of 500, 1000, and 1500 cm, respectively. The final soil depth for this TAXSUB will then be 1500 cm. Therefore, soils within TAXSUB are usually larger because the maximum soil depth is based on the greatest maximum soil depth within a subbasin, as compared to the STATSGO and MUKEY datasets where the largest maximum soil depth is chosen within a soil polygon.


Evapotranspiration
Evapotranspiration (mm)
3000
2000
1000
0
J
F
M
A
M
J
J
A
S
O


Groundwater flow
Groundwater flow (mm)
1250
1000
750
500
250

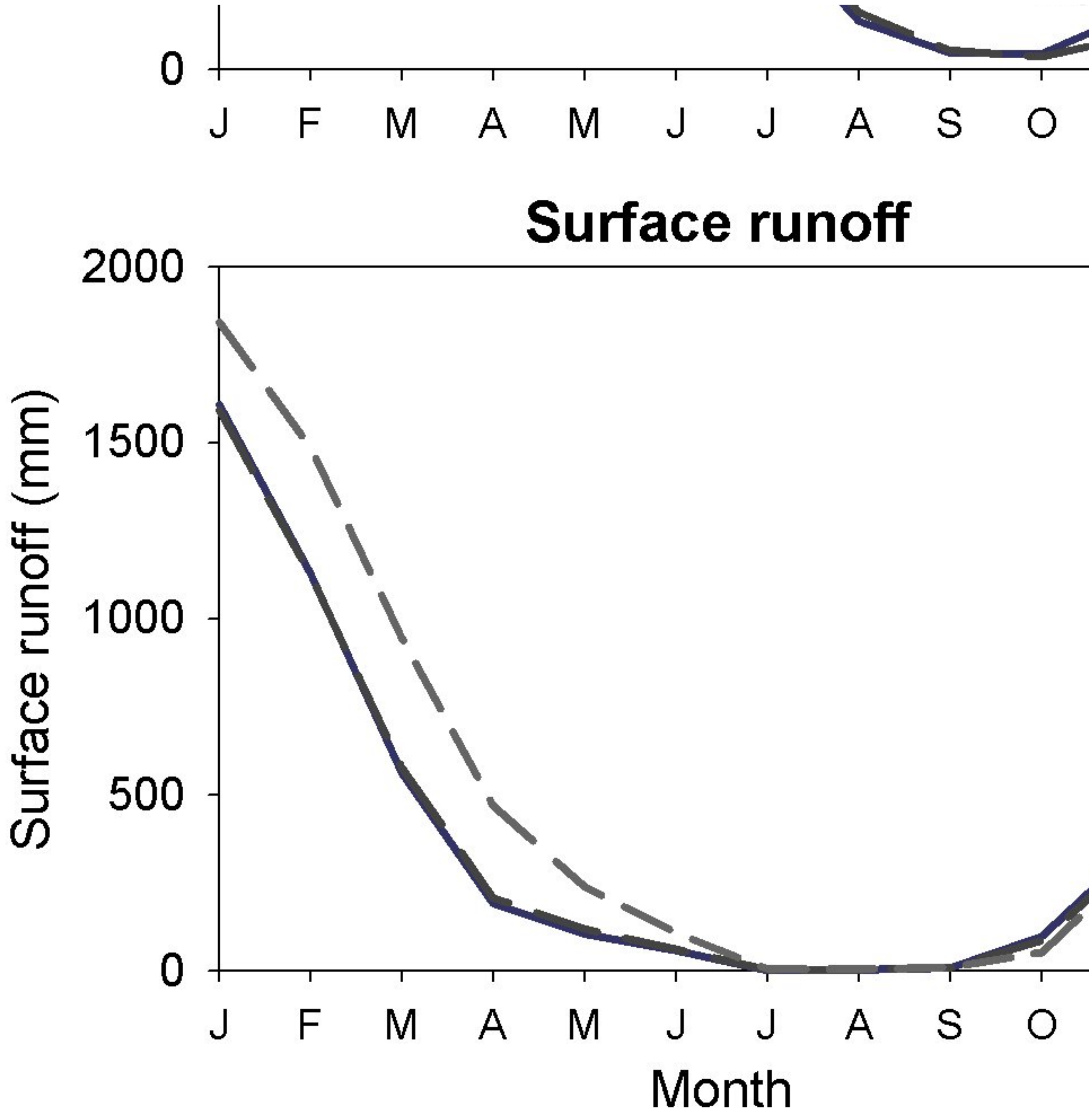

0
J F M A M J J A S O
Surface runoff
2000
1500
1000
500
0
Surface runoff (mm)
J F M A M J J A S O
Month

Figure 8. Individual hydrologic components of the Shasta Lake watershed for each soil dataset.

These differences in soil properties resulted in changes in streamflow and hydrologic components. The most obvious difference in figure 8 is the difference of soil water storage. Because TAXSUB had a greater maximum soil depth compared to STATSGO and MUKEY, the TAXSUB soils are able to store more soil water, which results in more soil infiltration. Further, with a greater soil depth, the soil is able to accommodate more precipitation. During large precipitation events, the soil column for TAXSUB may still be able to accommodate more water, while the STATSGO and MUKEY soil columns are completely saturated with water. If the STATSGO and MUKEY soil columns are completely saturated with water, then the excess water will be shifted to four hydrologic components: surface runoff, lateral soil water flow, groundwater flow, and aquifer percolation (fig. 8). The deeper TAXSUB soil column allows for more soil water storage, thus reducing (1) aquifer percolation, which results in less groundwater flow into the stream, and (2) lateral soil flow because the soil water is being held within the soil and not contributing to runoff. Further, a shallower soil depth allows the hydrologic components of groundwater, lateral soil flow, and surface runoff to be more “flashy,” which is a result of faster saturation of the soil column, therefore moving water to other hydrologic components faster. Compared to TAXSUB, the STATSGO and SSURGO simulations more likely overestimate streamflow during wet seasons and underestimate streamflow during dry seasons. For the viewpoint of model calibration, this pattern of differences between predictions and observations suggests too little base flow and too much surface runoff (Neitsch et al., 2005). Corresponding parameter adjustments are to increase AWC and decrease CN in SWAT simulations. This was consistent with the simulation results for the three different soil datasets. The modeling application parameterized by TAXSUB, with higher water storage, generated better results for streamflow compared to the other two soil data sources, indicating that the actual soil properties in the study area are better captured by the taxonomy-based aggregations.

Conclusions

In this study, we built upon the previous work by Gatzke et al. (2011), which used Great Group soil taxonomy as a defining feature that relates SSURGO polygons rather than MUKEY. Results of the soil aggregation indicate that, of the soil properties tested, TAXSUB maximum soil depth was significantly different from the STATSGO dataset, while both maximum soil depth and saturated hydraulic conductivity were significantly different from the MUKEY dataset.

Based on streamflow output, the TAXSUB soil dataset generated streamflow results closer to the observed streamflow data compared to the STATSGO or MUKEY inputs, which generated streamflow values much higher than the observed data. Model efficiency statistics also indicate that using the TAXSUB soil dataset led to better predictions of monthly streamflow. The largest differences between all soil dataset simulations occurred during the winter months (December through March) when most of the precipitation occurs. On average, the winter streamflow simulations with STATSGO, MUKEY, and TAXSUB were 292, 255, and 118 $m^3 s^{-1}$ greater than the observed data, respectively. Throughout the Shasta Lake watershed, the TAXSUB soil dataset had a much greater maximum soil depth, resulting in more surface water infiltration. During large precipitation events, the TAXSUB soil column may still be able to accommodate more water, while the STATSGO and MUKEY soil columns are completely saturated with water, leading to surface runoff. Using the soil taxonomy grouping method better captured the soil properties in the study area, and using the TAXSUB soil dataset more accurately simulated streamflow than using MUKEY and STATSGO. This is in agreement with our previous study (Gatzke et al., 2011). While this study and the study by Gatzke et al. (2011) resulted in more accurate streamflow simulations using the TAXSUB soil dataset, this new soil aggregation technique should also be tested in other study areas.

Acknowledgements

The authors would like to thank the University of California Kearney Foundation for the financial support for this study (Grant

No. 2008.002).

References

Anderson, R. M., Koren, V. I., & Reed, S. M. (2006). Using SSURGO data to improve Sacramento Model *a priori* parameter estimates. *J. Hydrol., 320*(1-2), 103-116. http://dx.doi.org/10.1016/j.jhydrol.2005.07.020.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., & Williams, J. R. (1998). Large-area hydrologic modeling and assessment: Part I. Model development. *JAWRA, 34*(1), 73-89. http://dx.doi.org/10.1111/j.1752-1688.1998.tb05961.x.

Beaudette, D. E., & O'Geen, A. T. (2010). Algorithms for quantitative pedology: A toolkit for soil scientists. Retrieved from aqp.r-forge.r-project.org.

Cho, J., Lowrance, R. R., Bosch, D. D., Strickland, T. C., Her, Y., & Vellidis, G. (2010). Effect of watershed subdivision and filter width on SWAT simulation of a coastal plain watershed. *JAWRA, 46*(3), 586-602.

Di Luzio, M., Arnold, J. G., & Srinivasan, R. (2004). Integration of SSURGO maps and soil parameters within a geographical information system and nonpoint-source pollution model system. *J. Soil Water Cons., 59*(4), 123-133.

Donigian, A. S., Bicknell, B. R., & Imhoff, J. C. (1995). Chapter 12: Hydrological Simulation Program - Fortran (HSPF). In V. P. Singh (Ed.), *Computer Models of Watershed Hydrology* (pp. 395-442). Littleton, Colo.: Water Resources Publications.

Drohan, P. J., Ciolkosz, E. J., & Petersen, G. W. (2003). Soil survey mapping unit accuracy in forested field plots in northern Pennsylvania. *SSSA J., 67*(1), 208-214.

Ficklin, D. L., Luo, Y., & Zhang, M. (2013). Watershed modeling of hydrology and water quality in the Sacramento River watershed, California. *Hydrol. Proc, 27*(2), 236-250. http://dx.doi.org/10.1002/hyp.9222.

Fitzhugh, T. W., & Mackay, D. S. (2000). Impacts of input parameter spatial aggregation on an agricultural nonpoint-source pollution model. *J. Hydrol., 236*(1-2), 35-53. http://dx.doi.org/10.1016/S0022-1694(00)00276-6.

Gassman, P., Reyes, M., Green, C., & Arnold, J. A. (2007). The Soil and Water Assessment Tool: Historical development, applications, and future research directions. *Trans. ASABE, 50*(4), 1211-1250. http://dx.doi.org/10.13031/2013.23637.

Gatzke, S. E., Beaudette, D. E., Ficklin, D. L., Luo, Y., & Zhang, M. H. (2011). Aggregation strategies for SSURGO data: Effects of SWAT soil inputs and hydrologic outputs. *SSSA J., 75*(5), 1908-1921. http://dx.doi.org/10.2136/sssaj2010.0418.

Geza, M., & McCray, J. E. (2008). Effects of soil data resolution on SWAT model streamflow and water quality predictions. *J. Environ. Mgmt., 88*(3), 393-406. http://dx.doi.org/10.1016/j.jenvman.2007.03.016.

Gowda, P. H., & Mulla, D. J. (2005). Scale effects of STATSGO vs. SSURGO soil databases on water quality predictions. In *Proc. 3rd Conf.: Watershed Management to Meet Water Quality Standards and Emerging TMDLs.* St. Joseph, Mich.: ASAE.

Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1999). Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *J. Hydrol. Eng., 4*(2), 135-143. http://dx.doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135).

Jha, M., Gassman, P. W., Secchi, S., Gu, R., & Arnold, J. (2004). Effect of watershed subdivision on SWAT flow, sediment, and nutrient predictions. *JAWRA, 40*(3), 811-825. http://dx.doi.org/10.1111/j.1752-1688.2004.tb04460.x.

Kumar, S., & Merwade, V. (2009). Impact of watershed subdivision and soil data resolution on SWAT model calibration and parameter uncertainty. *JAWRA, 45*(5), 1179-1196. http://dx.doi.org/10.1111/j.1752-1688.2009.00353.x.

Leavesley, G. H., Lichty, R. W., Troutman, B. M., & Saindon, L. G. (1983). *Precipitation-Runoff Modeling System: User's Manual.* USGS Water Resources Investigations Report 83-4238. Denver, Colo.: U.S. Geological Survey, Water Resources Division.

Luo, Y., Zhang, X., Liu, X., Ficklin, D. L., & Zhang, M. (2008). Dynamic modeling of organophosphate pesticide load in surface water in the northern San Joaquin Valley watershed in California. *Environ. Pollution, 156*(3), 1171-1181. http://dx.doi.org/10.1016/j.envpol.2008.04.005.

Luo, Y., Ficklin, D. L., & Zhang, M. (2012). Approaches of soil data aggregation for hydrologic simulations. *J. Hydrol., 464-465*, 467-476. http://dx.doi.org/10.1016/j.jhydrol.2012.07.036.

Mamillapalli, S., Srinivasan, R., Arnold, J. G., & Engel, B. A. (1996). Effect of spatial variability of basin-scale modeling. In

Proc. 3rd Intl. NCGIA Conf. and Workshop on Integrating GIS and Environmental Modeling. Santa Barbara, Cal.: National Center for Geographic Information and Analysis. Retrieved from www.ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/sf_papers/mamillapalli_sudhakar/my_paper.html.

Mann, H., & Whitney, D. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat., 18*(1), 50-60. http://dx.doi.org/10.1214/aoms/1177730491.

Maurer, E. P., Wood, A. W., Adam, J. C., Lettenmaier, D. P., & Nijssen, B. (2002). A long-term hydrologically based data set of land surface fluxes and states for the conterminous United States. *J. Climate, 15*(22), 3237-3251. http://dx.doi.org/10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2.

Mednick, A. C. (2010). Does soil data resolution matter? State Soil Geographic Database versus Soil Survey Geographic Database in rainfall-runoff modeling across Wisconsin. *J. Soil Water Cons., 65*(3), 190-199. http://dx.doi.org/10.2489/jswc.65.3.190.

Mednick, A. C., Sullivan, J., & Watermolen, D. J. (2008). Comparing the use of STATSGO and SSURGO soils data in water quality modeling: A literature review. *Research/Management Findings 60* (Dec. 2008). Madison, Wisc.: Wisconsin Department of Natural Resources, Bureau of Science Services.

Monteith, J. I. L. (1965). *Evaporation and Environment.* New York, N.Y.: Academic Press.

Moriasi, D. N., & Starks, P. J. (2010). Effects of resolution of soil dataset and precipitation dataset on SWAT2005 streamflow calibration parameters and simulation accuracy. *J. Soil Water Cons., 65*(2), 63-78. http://dx.doi.org/10.2489/jswc.65.2.63.

Moriasi, D. N., Arnold, J. G., Liew, M. W. V., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE, 50*(3), 885-900. http://dx.doi.org/10.13031/2013.23153.

Muleta, M. K., Nicklow, J. W., & Bekele, E. G. (2007). Sensitivity of a distributed watershed simulation model to spatial scale. *J. Hydrol. Eng., 12*(2), 163-172. http://dx.doi.org/10.1061/(ASCE)1084-0699(2007)12:2(163).

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models: Part I. A discussion of principles. *J. Hydrol., 10*(3), 282-290. http://dx.doi.org/10.1016/0022-1694(70)90255-6.

Neitsch, S. L., Arnold, J. G., Kiniry, J. E., Williams, J. R., & King, K. W. (2005). *Soil and Water Assessment Tool Theoretical Documentation. Version 2005.* College Station, Tex.: Texas Water Resource Institute.

Penman, H. L. (1956). Evaporation: An introductory survey. *Netherlands J. Agric. Sci., 4*, 9-29.

Peschel, J. M., Haan, P. K., & Lacey, R. E. (2006). Influences of soil dataset resolution on hydrologic modeling. *JAWRA, 42*(5), 1371-1389. http://dx.doi.org/10.1111/j.1752-1688.2006.tb05619.x.

Romanowicz, A. A., Vanclooster, M., Rounsevell, M., & Junesse, I. L. (2005). Sensitivity of the SWAT model to the soil and land use data parametrisation: A case study in the Thyle catchment, Belgium. *Ecol. Modelling, 187*(1), 27-39. http://dx.doi.org/10.1016/j.ecolmodel.2005.01.025.

Sheshukov, A. Y., Daggupati, P., & Douglas-Mankin, K. R. (2011). High spatial resolution soil data for watershed modeling: 2. Assessing impacts on watershed hydrologic response. *J. Natural and Environ. Sci., 2*(2), 32-41.

USACE-HEC. (1998). *HEC-HMS Hydrologic Modeling System User's Manual.* Davis, Cal.: U.S. Army Corps of Engineers, Hydrologic Engineering Center.

USDA-NRCS. (1999). *Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys.* Washington, D.C.: USDA Natural Resources Conservation Service.

USDA-NRCS. (2007). Soil Survey Geographic Database. Washington, D. C.: USDA Natural Resources Conservation Service.

USDA-NRCS. (2009). *National Soil Survey Handbook*. Title 430-VI. Washington, D.C.: USDA Natural Resources Conservation Service. Retrieved from www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs142p2_054242.

USDA-NRCS. (2012). Soils Data Mart. Washington, D.C.: USDA Natural Resources Conservation Service. Retrieved from http://datagateway.nrcs.usda.gov.

USDA-SCS. (1984). *SCS National Engineering Handbook.* Washington, D.C.: USDA Soil Conservation Service.

Tripathi, M. P., Raghuwanshi, N. S., & Rao, G. P. (2006). Effect of watershed subdivision on simulation of water balance components. *Hydrol. Proc., 20*(5), 1137-1156. http://dx.doi.org/10.1002/hyp.5927.

Vestra. (2004). Upper Pit River watershed assessment. Redding, Cal.: Vestra Resources, Inc. Retrieved from www.pitriveralliance.net/pitrcd/library/pdfs/UPPER_PIT_RIVER_WATERSHED_ASSESSMENT_TextOnly.pdf.

Wang, X., & Melesse, A. M. (2006). Effects of STATSGO and SSURGO as inputs on SWAT model's snowmelt simulation. *JAWRA, 42*(5), 1217-1236. http://dx.doi.org/10.1111/j.1752-1688.2006.tb05608.x.

Winchell, M., Srinivasan, R., Luzio, M. D., & Arnold, J. (2007). *ArcSWAT Interface for SWAT2005: User's Guide.* Temple, Tex.: Blackland Research Center, Texas Agricultural Experiment Station.