A Python Tool for Multi-Gage Calibration of SWAT Models using the NSGA-II Algorithm

Mehmet B. Ercan \textsuperscript{a}, Jonathan L. Goodall \textsuperscript{b,a}

\textsuperscript{a} University of South Carolina, Department of Civil and Environmental Engineering, 300 Main St., Columbia, SC 29208 USA (ercanm@engr.sc.edu)

\textsuperscript{b} University of Virginia, Department of Civil and Environmental Engineering, 351 McCormick Road, P.O. Box 400742, Charlottesville, VA 22904-4742 (goodall@virginia.edu)

Abstract: Calibration of large watershed models requires multi-gage calibration, however there are limited tools available for performing these calibrations of SWAT models. Non-Dominated Sorting Genetic Algorithm II (NSGA-II) has been shown to be an effective and efficient multi-objective calibration algorithm in various disciplines. Although NSGA-II has been used with SWAT before, there is no publically available software tool for easily applying the calibration approach for SWAT models. Therefore, the objective of this study was to create an open source tool for multi-gage calibration of SWAT models using the Python programming language. This tool is demonstrated through an application for the Upper Neuse Watershed in North Carolina, USA. The objective functions used for the calibration were Nash-Sutcliffe (E) and Percent Bias (PB), and the objective sites were the Flat, Little, and Eno watershed outlets. Results from the chosen parameter set in the Pareto front were E values ranging between 0.65 and 0.75 and PB values ranging between 0.02 and 0.08 for the objective sites. Similar to previous studies, the results show that the use of multi-objective calibration algorithms for SWAT calibration improved model performance.

Keywords: Hydrology; Watershed; NSGA II; Calibration; SWAT

1 INTRODUCTION

The use of multi-objective calibration algorithms increases model performance for large watersheds [Andersen et al., 2001]. Multiple-objective rather than single-objective calibration better constrains the calibration process, resulting in a calibrated model that better matches the physical conditions within the watershed [Niraula et al., 2012]. Of the optimization approaches used for watershed model calibration, genetic algorithms (GA) have been a popular option in water resources because GA offers the ability to effectively solve highly non-linear optimization problems.

One of the most popular multi-objective GA that can be applied for calibration of watershed models is the Non-Dominated Sorting Genetic Algorithm II (NSGA-II). NSGA-II is a population-based, fast, and efficient optimization technique that can be parallelized. The algorithm can significantly reduce calibration time compared to other approaches for semi-distributed watershed models [Confesor and Whittaker, 2007]. As a result, the NSGA-II algorithm has been used in many studies for watershed model calibration using multi-objective functions [Confesor and Whittaker, 2007; Bekele and Nicklow, 2007].

The semi-distributed watershed model Soil and Water Assessment Tool (SWAT) is a widely used model that can benefit from multi-objective calibration, especially for large watersheds where streamflow observations are available at more than one location [White and Chaubey, 2005; Bekele and Nicklow, 2007]. However, there are limited tools available for multi-objective calibration of SWAT models. Furthermore, there is no known implementation of the NSGA-II for multi-objective calibration of SWAT models that is freely available to the community. One study did report on multi-objective calibration of SWAT using NSGA-II [Bekele and
Nicklow, 2007], however, based on personal communication with the authors, the source code for this implementation is no longer available. The primary calibration tool available for SWAT 2009 is SWAT-CUP [Abbaspour et al., 2007] which has five different calibration algorithms, but does not have an implementation of the NSGA-II algorithm [Abbaspour, 2013].

The goal of this work is therefore to create an open source and freely-available NSGA-II tool for SWAT model calibration. We chose to implement the tool using the Python programming language in part because of its popularity in the scientific computing community. The tool is designed to provide easy access to efficient and effective multi-gage functions for calibrating SWAT models. This paper provides a brief background on the NSGA-II algorithm and on the SWAT model, then describes the tool design, and finally presents a case study applying the tool for calibration of a SWAT model. The source code for this tool is available from https://github.com/mehmetbercan/NSGA-II_Python_for_SWAT_model.

2 BACKGROUND

Multi-Objective Genetic Algorithms (MOGA) have been popular in many disciplines including rainfall-runoff models. These algorithms are based on the process of natural selection using similar theory to natural evolution. The MOGA tool utilized in this study to calibrate SWAT model is the Non-dominated Sorting Genetic Algorithm (NSGA-II) [Deb et al., 2002], which is able to find optimal solutions (Pareto front) in an efficient way. In the coming subsections, we provide a brief background on the NSGA-II algorithm and the SWAT model.

2.1 NSGA-II method

NSGA-II is a fast and efficient Multi-Objective Evolutionary Algorithm (MOEA) that out performs other MOEAs including the Pareto-archieved evolution strategy and the strength-pareto evolutionary algorithm [Deb et al., 2002]. NSGA-II has improvements over NSGA including reducing the complexity of the non-dominated sorting procedure and replacing a sharing function with a crowded-comparison function. A standard NSGA-II process starts with a random parent population (P\(_t\)) that goes through selection, crossover, and mutation to generate a child population (Q\(_t\)) (Figure 1). Selection of the next generation (P\(_{t+1}\)) is based upon non-domination (or Pareto ranking) and crowding distance to P\(_t\) and Q\(_t\). The next generation is used to create Q\(_{t+1}\) and the process repeats until termination criteria are met.

2.2 Soil and Water Assessment Tool (SWAT)

SWAT is a semi-distributed conceptual watershed model that is capable of running on a daily and sub-daily time steps over long time periods. It is able to simulate large watersheds with different management scenarios where the impact on water supplies and non-point source pollution can be assessed [Arnold et al., 1998]. It has been widely used in the US and around the world for water quantity and quality simulations [Cools et al., 2011; Liu et al., 2013]. Although the SWAT model is capable of different discretization schemes, modelers typically use a semi-distributed discretization scheme in which the concept of a Hydrologic Response Unit (HRU) is used for representing variability within subbasins of a watershed. HRUs, which are representations of unique land cover, soil, and management characteristics, are not spatially contiguous and therefore are often composed of many disjointed parcels.

3 PYTHON TOOL DEVELOPMENT

The tool was developed following the general approach described by Bekele and Nicklow [2007]. Using Latin Hypercube Sampling (LHS) to generate the initial population is more efficient than using a random initial population when calibrating SWAT models with NSGA-II method [Bekele and Nicklow, 2007]. Therefore, we included the LHS method in our Python Tool. Also, we designed the input files for the tool to be similar to those required by the widely used SWAT calibration tool SWAT-CUP. We did this so that users would be familiar with the input file structure, and to ease eventual migration of this algorithm into SWAT-CUP if that
becomes an option in the future.

The algorithm that the tool follows is described in Figure 1. The tool first executes a SWAT model with LHS members and defines objective function values. Then, non-dominated sorting and crowding distance sorting are applied to reduced the size of LHS to the population size by separating the front. After this initial step, standard NSGA-II process are applied using binary tournament selection (crossover and mutation) until termination criteria are met.

**Figure 1.** Architecture of the NSGA-II Python tool for SWAT calibration

4 Case Study

A case study demonstrating the tool's application is conducted for a SWAT model of the Upper Neuse watershed in North Carolina. The new calibration tool is used to calibrate this model using streamflow records at three observation sites. In the following subsections we first discuss how we created a SWAT model for Upper Neuse watershed, second show how we used our NSGA II tool to calibrate the SWAT model, and third we present the results of the calibration.

4.1 SWAT Model Development

The Upper Neuse watershed (Figure 2) includes the Flat, Little, and Eno River watersheds defined by the United States Geological Survey (USGS) level-8 watershed codes 02085500, 0208521324 and 02085070, respectively. The study area has a mild climate and gently rolling topography. The soil type of the watershed is dominated by silty clay and loam, and the land cover of the watershed is dominated by forest and cultivated crops.

Terrain and land cover data were obtained from the United States Geological Survey (USGS) National Elevation Dataset (NED) and the 2006 version of the National Land Cover Database (NLCD). Soil data were obtained from the State Soil Geographic (STATSGO) dataset provided by the United States Department of Agriculture (USDA). Weather data including temperature, wind speed, and humidity were obtained from the National Climatic Data Center (NCDC). Precipitation data was obtained from National Weather Service (NWS) for Nexrad-derived rainfall estimates and from NCDC for gauge observed rainfall estimates. These two different precipitation estimates were combined using the approach described by Ercan and Goodall [2012] to create a composite rainfall dataset for the watershed area. Lastly, daily average streamflow data from the USGS National Water Information System (NWIS) were downloaded using the Consortium of Uni-
versities for the Advancement of Hydrologic Science, Inc. (CUAHSI) Hydrologic Information System (HIS) [Tarboton et al., 2009].

We divided the watershed into subbasins based on the USGS streamflow station locations and homogeneity of land characteristics. We used threshold values of 10% for soil, slope, and land cover to reduce variability within the subbasins. The result was a total of 837 Hydrologic Response Units (HRUs) for the 93 subbasins in the watershed, which is within the HRU/subbasin ratio range recommended in SWAT documentation. The commonly used settings were chosen to configure the model that includes the Natural Resources Conservation Service (NRCS) Curve Number (CN) surface runoff method, the Penman-Monteith potential evapotranspiration method, and the variable storage channel routing method. The ArcSWAT software program was used for much of the data preprocessing steps required to create the Upper Neuse watershed SWAT model.

4.2 Calibration Methodology

Flat, Little and Eno watershed outlets were set as objective-sites including the Nash-Sutcliffe (E) and Percent Bias (PB) as goodness of fit. Therefore, we had six objective functions (3 sites * 2 fitness) in the study area. We ran Generalized Likelihood Uncertainty Estimation (GLUE) [Beven and Binley, 1992] available in SWAT-CUP to find the sensitivity of the flow parameters on streamflow prediction. The six most sensitive parameters were chosen for model calibration (Table 1).

We then used the following settings for calibrating the Upper Neuse watershed model with the NSGA-II Python tool. The LHS size was set to 1000 and crossover probability was set to 0.5 using uniform crossover. Mutation probability and the seed for random number generation were 0.5. Population size and generation number were 80. Since our parameters do not have a wide range, we used 8 bits for binary crossover and mutations.

4.3 Calibration Results

The Pareto front is shown in Figure 3 where axes are fitness values and parameter sets are optimal when their values are zero. Because of difficulties of showing all six objective functions on a single graph, we averaged fitness values over the Flat, Little and Eno watershed outlets in Figure 3b. Significant tradeoffs are illustrated between the two fitnesses as was also shown by Bekele and Nicklow [2007]. This illustrates the utility of multi-objective calibration of SWAT models by attempting to balance multiple competing objectives when selecting optimal parameter sets. We selected a parameter set for illustration purposes that had good fitness for both PB and E (Figure 3b) and differences in parameter sets between sites were small (Figure 3a and Table 2).
Using the chosen parameter set (Table 1), the daily and monthly statistics showed good agreement between simulated and observed streamflows for each sites (Table 2). PB values are considered to be “very good” for both the calibration and validation periods except that the Flat River watershed for the validation period, which was considered to be only “good” [Moriasi et al., 2007]. Monthly E values, on the other hand, were considered to be “good” for the calibration period and “very good” for the validation period [Moriasi et al., 2007]. Lastly, daily statistics showed very good accuracy compare to previous SWAT studies [Gassman et al., 2007], indicating the strength of the calibration method.

**Table 1. Parameter Values and Their Limits**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Range</th>
<th>Operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha_Bf</td>
<td>0.99</td>
<td>0.01-1.00</td>
<td>Replaced</td>
</tr>
<tr>
<td>Crn2</td>
<td>0.07</td>
<td>±0.25</td>
<td>% Relative</td>
</tr>
<tr>
<td>Ch_K2</td>
<td>30.59</td>
<td>0.01-150.00</td>
<td>Replaced</td>
</tr>
<tr>
<td>Canmx</td>
<td>9.53</td>
<td>0.01-10.00</td>
<td>Replaced</td>
</tr>
<tr>
<td>Esco</td>
<td>0.94</td>
<td>0.01-1.00</td>
<td>Replaced</td>
</tr>
<tr>
<td>Sol_Aw c</td>
<td>-0.06</td>
<td>±0.25</td>
<td>% Relative</td>
</tr>
</tbody>
</table>

**Table 2. Simulation Output Statistics**

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Statistic</th>
<th>Flat</th>
<th>Little</th>
<th>Eno</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005-2008b</td>
<td>E</td>
<td>0.74</td>
<td>0.75</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>Ea</td>
<td>0.73</td>
<td>0.72</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.75</td>
<td>0.76</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>R²a</td>
<td>0.74</td>
<td>0.73</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>PB</td>
<td>0.04</td>
<td>0.08</td>
<td>0.02</td>
</tr>
</tbody>
</table>

| 2009-2012c  | E         | 0.62  | 0.61   | 0.59|
|             | Ea        | 0.80  | 0.80   | 0.77|
|             | R²        | 0.62  | 0.61   | 0.64|
|             | R²a       | 0.82  | 0.81   | 0.82|
|             | PB        | -0.13 | -0.09  | -0.11|

### 5 Conclusion

The limited availability of SWAT model calibration tools and lack of multi-site and -objective evolutionary algorithms restricts SWAT users. In fact, the only available calibration tool for SWAT model (after 2005 version) is
SWAT-CUP which does not accommodate genetic algorithms including the powerful Non-Dominated Sorting Genetic Algorithm II (NSGA-II) that has been shown to be an effective and efficient multi-objective calibration algorithm for rainfall-runoff models including SWAT. Therefore, we created an open source NSGA-II tool for multi-gage calibration of SWAT model using the Python programming language.

We demonstrated the tool through a case study using the Upper Neuse Watershed as a study area. There were six objective functions: Nash-Sutcliffe (E) and Percent Bias (PB) as the fitness coefficients for three streamflow stations located at the outlets of the Flat, Little, and Eno watersheds. Six parameters were identified using the GLUE sensitivity analysis procedure. The fitness coefficient values did not vary much between the sites but we did end up with large tradeoffs between fitness coefficients, similar to results shown in previous NSGA-II studies. The Eno watershed had worse E values compare to other two watersheds, and we suspect that this is due to urbanization within the Eno watershed as the two other watersheds had very little urbanization.

We chose a parameter set within the Pareto front as optimal. The results from running the SWAT model using this parameter set for calibration period were E values ranging between 0.65 and 0.75 and PB values ranging between 0.02 and 0.08. The results for the evaluation period were E values ranging between 0.59 and 0.62 and PB values ranging between -0.13 and -0.09. All results for calibration and evaluation periods were considered satisfactory performance compare to previous SWAT studies. The multi-gage, multi-objective calibration therefore resulted in a well calibrated SWAT model that increases our confidence in the model's
predictive capabilities.

References


