Optimal Calibration and Uncertainty Analysis of SWAT for an Arid Climate

Milad Jajarmizadeh¹, Lariyah Mohd Sidek¹, Sobri Harun² and Mohsen Salarpour²

¹College of Engineering, Universiti Tenaga Nasional, Kuala Lumpur, Malaysia. ²Department of Hydraulic and Hydrology, Universiti Teknologi Malaysia, Skudai, Malaysia.

ABSTRACT: One of the major issues for semidistributed models is calibration of sensitive parameters. This study compared 3 scenarios for Soil and Water Assessment Tool (SWAT) model for calibration and uncertainty. Roodan watershed has been selected for simulation of daily flow in southern part of Iran with an area of 10570 km². After preparation of required data and implementation of the SWAT model, sensitivity analysis has been performed by Latin Hypercube One-factor-At-a-Time method on those parameters which are effective for flow simulation. Then, SWAT Calibration and Uncertainty Program (SWAT-CUP) has been used for calibration and uncertainty analysis. Three schemes for calibration were followed for the Roodan watershed modeling in calibration analysis as evolution. These include the following: the global method (scheme 1), this is a method that takes in all globally adjusted sensitive parameters for the whole watershed; the discretization method (scheme 2), this method considered the dominant features in calibration such as land use and soil type; the optimum parameters method (scheme 3), this method only adjusted those sensitive parameters by considering the effectiveness of their features. The results show that scheme 3 has better performance criteria for calibration and uncertainty analysis. Nash-Sutcliffe (NS) coefficient has been obtained 0.75 for scheme 3. However, schemes 1 and 2 resulted in NS 0.71 and 0.74, respectively, between predicted and observed daily flows. Moreover, percentage bias (P-bias) obtained was 6.7, 5.2, and 1.5 for schemes 1, 2, and 3, respectively. The result also shows that condition of parameters (parameter set) during calibration in SWAT-CUP program model has an important role to increase the performance of the model.

KEYWORDS: Calibration, discharge, Iran, SWAT, SWAT-CUP, uncertainty


DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Milad Jajarmizadeh, College of Engineering, Universiti Tenaga Nasional, Jalan Ikrarn-Uniten, Kajang, Selangor 43000, Malaysia. Email: milad_jajarm@yahoo.com

Introduction

Calibration is the procedure of adjusting parameter value to optimize model performance according to a set of predefined criteria. Each model uses a set of one or more parameters that are used to determine the basic behavior of a modeled system. Many numbers of calibration procedures to find such a solution in an iterative style has been tested in 4 decades. Generally, they are changeable in details but include 4 elements, namely, objective function, calibration data, adjustment strategy, and termination criterion. The model can be calibrated for having determined appropriate measures of performance. This may be performed automatically, semiautomatically, or manually. Manual calibration needs an experienced hydrologist who understands the behavior of the model. It includes running the model, looking the output, and adjusting the parameters until a satisfactory level of model behavior is obtained. The process of manual model calibration can be very time-consuming. A substitute to the manual procedure is automatic optimization. In this case, the model is applied to undertake multiple simulations, each searching for an optimal set of parameter values. The effectiveness of this depends on the sophistication of the search procedure and the utility of the measure of goodness of fit. Finally, the other option is semiautomatic calibration which combines both methods, with the user control the automatic parameterization method, to ensure that unrealistic parameter combinations are excluded from consideration.

Uncertainty can be a supplement part of any hydrological modeling undertaken. The uncertainty in the modeling process includes 4 major sources. (1) Data uncertainty such as errors attributed to the measurement itself, by the temporal and spatial discretization by data preprocessing or measurements. (2) Model structural uncertainty such as simplifications or inaccuracies in the explanation of actual world phenomena. The unavoidable deficiencies in the model building often result in the problem that different parameter sets fit one mode of system response at the expense of other response modes that are reproduced less precisely. (3) Model specification uncertainty such as the inability to converge to single best model using the information provided by available data. This uncertainty results mainly from data and model structure uncertainties. (4) Uncertainty due to unknown initial circumstances such as the states of the model is usually unknown at the beginning of any simulation period. However, this uncertainty can be reduced either by adjusting the initial conditions or via using a warming up period that allows the internal states to adjust.

Regarding application of Soil and Water Assessment Tool (SWAT) in this study, Arnold et al. reported that several calibration techniques have been developed for SWAT, including
Manual calibration procedures and automated procedures using the shuffled complex evolution method. Recently, SWAT Calibration and Uncertainty Program (SWAT-CUP) has been developed which provides a decision-making framework that incorporates a semiautomated approach Sequential Uncertainty Fitting–2 (SUFI-2) using both manual and automated calibration and incorporating sensitivity and uncertainty analysis. Abbaspour et al. applied SWAT model including a general calibration for the large area based on data availability by European Water Framework Directive. Recently, calibration and uncertainty analysis are implemented in an open-source cloud platform and generalized likelihood uncertainty estimation method by Zhang et al. The result shows that it can be an optional method to solve by speeding up the procedure of processing the model depending on complexity and flexibility of model. Therefore, calibration of SWAT model is still one of the progressing issues for researchers.

The SUFI-2 algorithm has been used for several studies in large and small scales in Iran. The main reason of using SUFI-2 algorithm for calibration of SWAT model is the availability of many parameters regarding water balance modeling. Indeed, calibration of the semidistributed hydrological model is a challenging task due to the inclusion of spatial and temporal features. Rooodan watershed has been modeled via SWAT in a southern part of Iran with objectives those elaborate calibration and validation of parameters as a benchmark in one scenario. The main objective of this study is a comparison of 3 scenarios for calibration of SWAT using SUFI–2 algorithm in an arid climate. The importance of this study can be reviewing different methods for exploring the optimum parameter set, which is needed for calibration of SWAT model in a large arid area. It can be helpful for similar studies in future that have the same climate. This study investigates the accuracy of daily flow prediction including uncertainty analysis in an evolution calibration.

Case Study
The study area known as Rooodan watershed is located in the south of Iran between Hormozgan and Kerman provinces. The area of the catchment is 10570 km². The average annual precipitation of the study area is 215 mm. Esteghlal Dam that has an important role in collecting the surface waters for the development of downstream areas is located at the outlet of Rooodan watershed. Field data can be related to the observation of land use and recording some soil samples data. Soil type of watershed is usually modeled by a range of equations; no hydrological features. Rooodan watershed has been modeled via SWAT in a southern part of Iran with objectives those elaborate calibration and validation of parameters as a benchmark in one scenario. The main objective of this study is a comparison of 3 scenarios for calibration of SWAT using SUFI–2 algorithm in an arid climate. The importance of this study can be reviewing different methods for exploring the optimum parameter set, which is needed for calibration of SWAT model in a large arid area. It can be helpful for similar studies in future that have the same climate. This study investigates the accuracy of daily flow prediction including uncertainty analysis in an evolution calibration.

Tools
Soil and Water Assessment Tool
Soil and Water Assessment Tool is a semidistributed and continuous calculation model. As a short overview, Groundwater Loading Effects of Agricultural Management Systems (GLEAMS) model, the Chemicals, Runoff, and Erosion from Agricultural Management Systems (CREAMS) model, and the Environmental Policy Integrated Climate model are the basic part of the Simulator for Water Resources in Rural Basins (SWRRB) model. Gradually, the first version of SWAT was built based on Routing Outputs to Outlet model by interfacing in SWRRB. Then, pollution transport capabilities have been embed in SWAT model such as a reservoir, pond, wetland, point source, and sediment routing. In addition, improved representations of conservation and management practices have been included in SWAT such as temporal accounting of management practice, evaluation of land use changes, plant growth, and irrigation plans. The goal of development for SWAT model was for the prediction of land use impact and management on water, sediment, and agricultural chemical yields in ungauged watersheds.

Soil and Water Assessment Tool model has 2 phases for simulation of watershed hydrology, namely, land phase and routing phase. A complete description of theoretical and input/output data for SWAT version 2009 can be found in the works by Arnold and colleagues.

For those who are interested in SWAT model application and development, sufficient information is available at http://swatmodel.tamu.edu. Moreover, Gassman et al. reported a review on climatic inputs and pollutant losses and flow routing across the globe. Gassman et al. presented that innovative application and adaptations for SWAT code and simulation capabilities. Daniel et al. focused on popular context for application of watershed modeling and related new technologies involved with SWAT model. Regarding current development and presentation of performance statistics for SWAT model, 20 types of research have been summarized by Douglas-Mankin et al.

Sequential Uncertainty Fitting–2
Clearly, in a direct model, a single valued parameter results in a single model signal. However, in an inverse model (IM), an observed signal can produce many sets of a different parameter. This nonuniqueness is a natural characteristic of the IM. In recent years, IM has become an acceptable and motivating procedure for calibration. Inverse model solves the problem of figuring out the physical systems from measuring the output variables of the model. Inverse modeling is popular due to its straightforward and direct measurement of parameters which opposes the physical system that is usually described to be time-consuming, costly, and boring. Often, measured outputs have limitations for application, and almost all measurements are subjected to some uncertainties. Generally, the derivations are statistical. Furthermore, the other reason is that only a limited number of (noisy) data can be measured and the physical systems are usually modeled by a range of equations; no hydrological...
An example of IM is the SUFI-2, which is developed for calibration and uncertainty analysis of SWAT model. In SUFI-2, parameter uncertainty calculates for the attributed sources of uncertainties in a semidistributed hydrological model such as the uncertainties in driving variables, the concept of model,
parameters, and measured data. The 2 important factors (P-factor and R-factor) in this calculation become the index of the evaluation of the results. Therefore, the degree, which all uncertainties are calculated for uncertainties, is quantified by a measurement referred to as the P-factor. P-factor is the percentage of measured data bracketed by the 95% prediction uncertainty (95PPU). In SUFI-2, the 95PPU is calculated at 2.5% and 97.5% degree of the cumulative distribution of an output variable obtained through Latin hypercube sampling, by forbidding 5% of the very bad simulations. The R-factor, however, is related to the strength of the calibration and uncertainty analysis. It is the average thickness of the 95PPU band divided by the standard deviation of the measured data. Sequential Uncertainty Fitting-2 tries to bracket most of the measured data with the smallest possible uncertainty band. In this research, SWAT-CUP has been used as a computer program for calibration of SWAT model. The SWAT-CUP is a public domain program and as such may be used and copied freely. The program links SUFI-2 procedures to SWAT. It enables sensitivity analysis, calibration, and uncertainty analysis of a SWAT model.

**Model Performance**

For evaluation of SWAT model calibration, graphical and statistical regarding streamflow can be sufficient. Nash-Sutcliffe (NS) coefficient is used for statistic evaluations mostly. The other criteria can be percentage bias (P-bias). Krause et al. and Moriasi et al. interpreted comprehensively equations and related features of mentioned statistical model performance that have been used in this research (ie, NS and P-bias). Nash-Sutcliffe efficiency values can range between \(-\infty\) and 1 and provide a measure how well the simulated output matches the observed data along a 1:1 line. A perfect fit between the simulated and observed data is indicated by an NS value of 1. Nash-Sutcliffe efficiency values ≤0 indicate that the observed data mean is a more accurate predictor than the simulated output. Moriasi et al. stated that absolute value of P-bias ranging from 15 to 25 shows that the SWAT model is rated as satisfactory, rated good when from 10 to 15, and very good when less than 10. Also, it has been used mean squared error which is a measure of the quality of an estimator. It is nonnegative and should be close to 0.

**Performing SWAT Model in Roodan**

Soil and Water Assessment Tool has been used for delineation of Roodan watershed. In summary, the watershed simulation involves major ordinal steps: (1) digital elevation map set-up, (2) stream burning, (3) outlet and inlet localized, (4) basin outlets selection, and (5) definition and calculation of subbasin parameters. For overlaying the digital streams, the threshold-based stream burning was used by considering the minimum size of the subbasin. Next, land use, soil, and slope data sets were imported as well overlaid and linked to the SWAT database. According to with topographic map (digital elevation model), different slope classifications are possible due to the spatial distribution of watershed and better identification of spatial location hydrological response units (HRUs). This is particularly important if subbasins have a range of slopes occurring within HRUs. Therefore, 3 slope classifications have been defined for Roodan from 0% to 5%, 5% to 20%, and more than 20%. After importing the land use, soil data, and slope classifications, the distribution of HRUs within the watershed must be determined. Hydrological response units are the smallest elements to contribute to increasing the accurate calculation of streamflow and other hydrological conditions with various land uses, soils, and slopes. Therefore, short amount threshold in accordance with the percentage (5%) was identified for land use, soil, and slope distribution as suggested by Raneesh et al. This scheme leads us to avoid generalization as result of the dominant land use, soil, and slope class. The Roodan watershed was divided to 513 HRUs for whole catchment and 45 subbasins. Weather data were set for the land phase of the hydrological cycle. Hargreaves method has been chosen for calculation potential evapotranspiration as well as SCS Runoff Curve Number (SCS-CN) for calculation of runoff volume. The surface runoff is estimated using a modification of the SCS-CN method with daily rainfall amounts. The curve number values are based on soil type, land use/land cover, and land management conditions and are adjusted according to soil moisture conditions. Percolation is estimated applying the combination of a storage routing technique and a crack-flow model. The lateral flow is calculated at the same time with percolation with the application of a kinematic storage model. The baseflow is accounted based on the hydraulic conductivity of shallow aquifer, distance from subbasin to the main channel, and water table height.

In this study, reach evaporation coefficient (EVRCH.bsn) is adjusted for Roodan watershed based on the measured value of daily streamflow and considering the contribution of baseflow for main channels. Reach evaporation adjustment factor in original equation tends to overestimate evaporation from riches in arid areas. Therefore, it is recommended that adjusting for this factor due to the revival of neglected runoff can be helpful in channels for the arid area before modeling streamflow. Finally, yet importantly, prepared model was run for 1988 to 2002 by considering warm-up period for daily streamflow. For simulation of 5 years or less, equilibration and warm-up by SWAT are recommended; meanwhile, for long simulation, it can be optional. Therefore, in this study, 1988 as a starting date was used for warm-up, the SWAT due to having a complete hydrological cycling. The length of the burn-in period depends on the availability of longitudinal data and objective of the study. Then, SWAT was run for the daily time scale for simulation of Roodan watershed.

**Calibration Schemes for Roodan**

Primary in this research, the sensitivity analysis was performed with Latin Hypercube One-factor-At-a-Time (LH-OAT) method, which is embedded in SWAT (version 2009) package.
model. Usually, before the calibration, it has been required sensitivity analysis due to recognizing the sensitive parameters, model components, with respect to the model's performance. For Roodan watershed, sensitivity analysis has been done with a number of 26 parameters that have high sensitivity on flow using SWAT model as suggested by Winchell et al. For better evaluation due to the final selection of sensitive parameters in Roodan watershed, it has been made a decision on sensitive parameters regarding LH-OAT analysis, cognition from our case study, and literature in Iran. Then, sensitive parameters were applied for calibration and a more in depth sensitivity analysis by SUFI-2 algorithm. This procedure helps to boost dominant characteristic (eg, land use, soil texture) and relative change due to maintaining the spatial variation of parameters. In our study, only the sensitive parameters were adjusted to avoid over-parameterization in calibration scheme, as well as to involve spatial variation in the sensitive parameters, which lead us to a reasonable calibration and uncertainty analysis. Due to approach a confident calibration, 3 schemes for calibration have been followed for Roodan watershed modeling by SUFI-2. These include the following: the global method (scheme 1), which is adjusted for sensitive parameters globally for the whole watershed, which means that the calibrated parameters are considered heterogeneous for the whole watershed; the discretization method (scheme 2), which is considered for dominant features (eg, land use and soil type) in calibration, which means that dominant parameters are calibrated separately for watershed and watershed considered with dominant feature; the optimum parameters method (scheme 3), which is adjusted for only those sensitive parameters by considering effectiveness their features according to SUFI-2 algorithm, which means that only those parameters which are sensitive are considered for watershed based on dominant feature or globally feature. Indeed, every scheme leads us to achieve to an optimum parameter set calibration and finding the degree of sensitivity parameters by their effectiveness features on modeling. In this study, the model calibration period is 1989 to 2002, and location of data for calibration is the outlet of Roodan watershed.

Results and Discussion

In total, 13 parameters have been identified as sensitive parameters for Roodan watershed. They are baseflow alpha factor (ALPHA_BF.gw), effective hydraulic conductivity of main channel (CH_K2.rte), available water capacity of the soil layer (SOL_AWC.sol), SCS-CN for antecedent moisture condition type II for whole catchment (CN2.mgt), moist soil bulk density(SOL_BD.sol), maximum canopy index (CANMX.hru), Manning coefficient for channel (CH_N2.rte), soil evaporation compensation factor (ESCO.hru), surface runoff lag coefficient (SURLAG.bsn), soil conductivity (SOL_K.sol), plant uptake compensation factor (EPCO.hru), groundwater recharge to deep aquifer (RCHRG_DP.gw), and threshold depth of water in the shallow aquifer required for return flow to occur (GWQMN.gw). Then, calibration and uncertainty calculation involved with 3 schemes for recognition of optimum calibration. These are categorized as follows: (1) global scheme (considering the catchment's similar features), (2) discretization scheme (considering dominant features, for example, land use types), and (3) optimum scheme (final parameter set based on dominant or global scheme), respectively. For calibration and uncertainty evaluation, 500 simulations have been performed in each iteration in every scheme (global, discretization, and optimum). Figures 2 to 4 show values of NS (vertical axis) versus corresponding values of each parameter (horizontal axis). The dotty plots show the distribution of the number of simulations in parameter sensitivity analysis after comparing the parameter values with the objective function (NS) for the daily calibrations, respectively, via SUFI-2.

Figures 2 to 4 are the outputs of the SUFI-2 algorithm, which show values of each parameter’s range in every simulation (500 sampling). These were obtained after 500 runs in the last iteration of the SUFI-2 algorithm. Figures 2 to 4 (global, discretization, and optimum schemes) jointly showed that the ALPHA_BF graph has a desirable parameter range which is between 0.8 and 1 (horizontal axis); this gives a reasonable value for NS (vertical axis). The CH_K2 parameter for all the schemes (global, discretization, and optimum schemes) has also given better values for the objective function with an overall small range of values. The graphs of CH_K2 show that by increasing the values of CH_K2, the NS will decrease. This is explained as such when all the graphs are being interpreted concurrently. A comparison between Figures 2 to 4 shows that the optimum scheme (Figure 3) has a better consistency for NS (vertical axis) against parameter range (horizontal axis). This is because the dispersal (distribution) of selected parameter values (green points) have decreased and gradually shifted to a better NS from global scheme to optimum scheme. Hence, all parameters have narrow dispersal regarding objective function (vertical axis in Figure 3). This evolution shows that a selected parameter set has more stability for uncertainty which derived via trial and errors with a combination of scheme 1 (global) and scheme 2 (discretization). Indeed, it keeps the meaning of semiautomatic calibration.

Scheme 3 (optimum parameter set) was chosen based on its calibration results because it had better objective functions and uncertainty than the other 2 schemes. Table 1 shows the performance of SWAT after calibration between performed schemes for Roodan watershed. Scheme 3 had the highest value for NS concurrently marked at 75% for calibration. Its P-bias was 1.5%. The uncertainty results reported by the SUFI-2 algorithm calculations, specifically for the SWAT model, show an acceptable range for a large basin modeling. As it can be seen from Table 1, the P-factor (measured data bracketed by 95% prediction uncertainty) and R-factor (strength of calibration) were 50% and 0.18, respectively. Generally, it is desirable to seek a larger P-factor and a smaller R-factor in SUFI-2; these 2 values are ideally assumed as 100% and 0,
Figure 2. (Continued)
Figure 2. Sequential Uncertainty Fitting-2 results for the global scheme (vertical axis: value of Nash-Sutcliffe; horizontal axis: value of parameter).
Figure 3. (Continued)
Figure 3. Sequential Uncertainty Fitting-2 results for discretization scheme (vertical axis: value of Nash-Sutcliffe; horizontal axis: value of parameter).
Figure 4. Sequential Uncertainty Fitting-2 results for the optimum scheme (vertical axis: value of Nash-Sutcliffe; horizontal axis: value of parameter).
Table 1. Comparison of different schemes for simulation of streamflow.

<table>
<thead>
<tr>
<th>INDEX</th>
<th>SCHEME 1</th>
<th>SCHEME 2</th>
<th>SCHEME 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-factor %</td>
<td>38</td>
<td>43</td>
<td>50</td>
</tr>
<tr>
<td>R-factor</td>
<td>0.24</td>
<td>0.16</td>
<td>0.18</td>
</tr>
<tr>
<td>NS coefficient %</td>
<td>71</td>
<td>74</td>
<td>75</td>
</tr>
<tr>
<td>P-bias %</td>
<td>6.7</td>
<td>5.2</td>
<td>1.5</td>
</tr>
<tr>
<td>MSE</td>
<td>2942</td>
<td>2664</td>
<td>2631</td>
</tr>
</tbody>
</table>

Abbreviations: MSE, mean squared error; NS, Nash-Sutcliffe; P-bias, percentage bias.

Figures 5 to 7 show trend analysis for streamflow modeling during the calibration period. All the graphs have logical trend for peak flows that shows SWAT has a fairly cognition of peak flows as a physically based model. An evaluation has been performed on peak flows more than 1000 m$^3$/s for all schemes. Relative errors include overestimation and underestimation. Figure 8 reveals that optimum scheme has shorter relative errors for flows more than 1000 m$^3$/s. Then, global scheme calibration has shorter relative errors in comparison with discretization only. Figure 8 reveals that for largest flow during 1989 to 2002, which is 4209.5 m$^3$/s, the global scheme has shortest relative error value. Both optimum and discretization methods have a similar error. Generally, all schemes have the same trend overestimation and underestimation for flow prediction for high flows more than 1000 m$^3$/s. These 3 schemes show that SWAT has an acceptable prediction for flow in the arid region, but the selection and discretization of parameters for calibration task have an important role to increase the accuracy the results of high flows. Singh et al$^{32}$ have performed sensitivity and uncertainty analyses with SUFI-2 in a large catchment (14 429.36 km$^2$) in India. They reported that sensitive parameters include curve number, baseflow alpha factor, threshold depth of water in shallow aquifer required for return flow, soil evaporation compensation factor, effective hydraulic conductivity in main conductivity, and soil available water capacity. It seems that Roodan study has a fair agreement with previously published research as large scale. Singh et al$^{32}$ reported that P-factor has been obtained 42% for
daily flow simulation; meanwhile, P-factor has been obtained 38%, 43%, and 50% for global, discretization, and optimum schemes, respectively, in Roodan study. This comparison shows that type of calibration with considering the type of parameters result in shorter uncertainty and better performance. Indeed, uncertainties increased with large variations in topography and rainfall in the form of land use and soil types. Therefore, step-by-step calibration can help the modeler to find an optimum result for a semidistributed model such as SWAT via SUFI-2 algorithm when increasing the accuracy is under evaluation (Table 2).

**Conclusions**

This research has been performed in southern part of Iran to daily streamflow modeling with SWAT model. Then, calibration and uncertainty analysis are involved with SUFI-2 algorithm. Three scenarios as evolution have been performed for calibration and uncertainty analysis. (1) The global method, which is adjusted for sensitive parameters globally for whole watershed; (2) discretization method, which is considered for dominant features (eg, land use and soil type) in calibration; (3) the optimum parameters method, which is adjusted for only those sensitive parameters by considering effectiveness of their features according to SUFI-2 algorithm. According to NS coefficient, all scenarios (1, 2, and 3) are logical and satisfactory and they have a fair tendency with observed data. However, optimum scenario outperformed regarding calibration and uncertainty indexes (NS, P-factor, and R-factor). Sensitivity and uncertainty analyses reveal that last scenario (optimum) has more stability in corresponding with NS coefficient. An assessment of scenarios 1, 2, and 3 shows that the dispersion of parameter values is shorter for scenario 3 (optimum). It can be concluded that the impact of parameter types’ adjustment (ie, lumped, semidistributed, or fully distributed) has considerable significance on strength of model performance.

**Acknowledgements**

The authors appreciate the cooperation and help given by the Department of Hydraulic and Hydrology and Centre of Information and Communication Technology (CICT) of Universiti Teknologi Malaysia. They are thankful for all members of consultant engineers of Ab Rah Saz Shargh Corporation in Iran and the Regional Water Organization, Agricultural Organization, and Natural Resources Organization of the Hormozgan Province, Iran.

**Author Contributions**

MJ is the main author, LMS is the advisor for this paper, SH is the supervisor, and MS is the editor.

**REFERENCES**


