

**CHAPTER 7. A COMPARISON OF SWAT MODEL CALIBRATION
TECHNIQUES FOR HYDROLOGICAL MODELING IN THE GANGA RIVER
WATERSHED**

7.1 Introduction

In terms of the land and climate characteristics of the southern part of Uttar Pradesh Basin, tendencies toward sudden floods and soil erosion present some of the most extreme issues. Because of soil erosion, the soil of this watershed is continuously degrading and losing nutrients, thus affecting agriculture in the state. The severity of the issue is intensified in arid and semi-arid lands, where short-duration intense rainfall and unsustainable land use have quickened soil losses by erosion. Detailed and accurate information regarding soil erosion and surface water discharge is helpful for watershed managers to manage better and conserve natural resources such as soil and water, and to promote sustainable development. At present, various procedures are used for hydrological modeling, utilizing precipitation, land use, and soil characteristics information. Among these strategies, the SWAT is a process-based hydrological model developed by the United States Department of Agriculture (USDA) Agricultural Research Service (ARS) (Kumar et al., 2015). New algorithms have been developed to best estimate the parameters used in hydrological modeling.

The SWAT is widely used for hydrological and sediment yield modeling. It is a time-consistent and spatially appropriate test model that was developed to assist watershed managers in anticipating the effects of land use management activities on runoff, soil erosion, and agricultural chemical yields (Khalid et al., 2016). Specialists have effectively used the SWAT for runoff assessments (Salimi et al., 2016, Noori et al., 2016), water

quality modeling (Qiu and Wang 2013, Pinaras et al., 2010), hydrological and river basin modeling (Omani et al., 2000, Fukunaga et al., 2015, Shi et al., 2013), sediment yield modeling (Briak et al., 2016, Vigiak et al., 2017), and the management of erosion-prone areas (Sardar et al., 2014). (Psomas et al., 2016) utilized the SWAT and WEAP to develop water efficiency measures in Greece.

To modify the elements that influence SWAT yields, calibration using observed values and evaluated estimations of runoff, evapotranspiration, and other SWAT outputs are necessary. Validation is a way that compares SWAT results with the observed data, without modifying the values of the influencing factors. Calibration of various parameters is necessary for proper hydrological modeling. There are 26 parameters distinguished for runoff, more than 30 for soil disintegration, and 41 for water quality, all of which can be utilized for calibration. The calibration of such a circulated parameterized watershed model entails some significant issues that require careful attention on the part of investigators, particularly in regard to uncertainty. Qiu and Wang (2014) performed validation and calibration for the periods 1997–2002 and 2003–2008, respectively, for the hydrological and water quality evaluation of central New Jersey. Vigiak et al., 2015 performed calibration and validation for hillslope erosion modeling of the Danube Basin.

Earlier researchers applied deterministic approaches such as the trial and error method for uncertainty analysis, calibration, and validation. When using such methods, it is necessary to continue to alter the parameters until a sensible match is achieved between the simulation and observations. However, these approaches are now outdated; scientists have developed many stochastic algorithms for calibration, validation, and uncertainty analysis. Some of the widely used algorithms for calibration are sequential uncertainty fitting version 2 (SUF2), parallel solution (ParaSol), particle swarm optimization (PSO), generalized

likelihood uncertainty estimation (GLUE), artificial neural network (ANN), and Markov chain Monte Carlo (MCMC). (Vilaysane et al., 2015) used the SUFI-2 technique to calibrate the values for hydrological stream flow modeling of the Xedone River Basin. Researchers utilized SUFI-2 and SWAT-CUP for sediment yield modeling in Ethiopia; they reported that SUFI-2 utilized a computerized advancement system that was equipped to perform affectability, alignment, approval, and vulnerability examinations with an enhanced model run time proficiency. Ercan et al., 2014 used cloud techniques for SWAT model calibration. Talebizadeh et al., 2010 used ANN and the SWAT for sediment load modeling and uncertainty analysis. Calibration and validation have also been done using genetic algorithms and Bayesian model averaging (Zhang et al., 2009). (Tuo et al., 2016) applied the SUFI-2 algorithm for uncertainty analysis when evaluating precipitation input for the SWAT model. Noori and Kalin (2016) coupled ANN with the SWAT and SWAT-CUP for daily streamflow prediction. Similarly, runoff and sediment fluxes modeling (Vigiak et al., 2017) have been done using the SUFI-2 algorithm. An approximation of the SWAT model can also be constructed using ANN and a support vector machine (SVM) (Zhang et al., 2009). Out of all these algorithms, it is difficult to determine the best algorithm for the calibration of SWAT outputs. Therefore, in the present study, the three calibration-uncertainty analysis algorithms of GLUE, SUFI-2, and Parasol are considered.

In the present chapter, For calibration, the three widely used algorithms GLUE, SUFI-2, and ParaSol were applied and compared in order to determine the best calibration technique. In this investigation, the results of ArcSWAT Interface from chapter 6 were taken for calibration and validation. These results were associated with the Ganga River Basin for depicting and organizing sub-watersheds in order to gauge and model runoff, sediment yield, and evapotranspiration. The SWAT model was calibrated for the observed

data from the years 2004–2009 at the sub-watershed and watershed level using the SUFI-2, GLUE, and ParaSol algorithms. By analyzing the results, assumptions, and limitations of these calibration techniques, the best technique among the three was determined based on five categories, as described in Section 7.3. The model calibrated using the best technique was then validated for the period of 2009–2015. Finally, the estimated and corrected results of the hydrological parameters were calculated, and modeling was conducted.

7.2 Calibration and Validation

As noted earlier, calibration is the modification of the elements influencing the SWAT yields, and is assisted by observed values and evaluated estimations of runoff, evapotranspiration, and other SWAT outputs. Validation compares the SWAT results with the observed data without modifying the values of the influencing factors. Calibration is necessary for proper hydrological modeling. In this chapter Sufi-2, GLUE and Parasol methods were used and compared for calibrating the SWAT results, and the best method is then applied for final calibration. There are 34 parameters distinguished for runoff and soil erosion that can be utilized for calibration. Table 7.1 describes the parameters used in the present study for calibration, along with their minimum and maximum absolute SWAT values. In this step, SWAT-CUP was used as an interface between the SWAT and calibration algorithms to perform the uncertainty analysis, calibration, and validation of the hydrological modeling outputs. The TxtInOut folder of the SWAT model was imported into the SWAT-CUP software for input. The observed values of surface water discharge of the Varanasi sub-watershed were calculated using the acoustic Doppler current profiler (ACDP) for the years 1996–2015. These values were used for comparison and calibration. The input files of the SWAT-CUP model were updated as necessary. Figure 7.1 shows the

detailed procedure as a flowchart.

Table 7.1 Parameters used for calibration, with their absolute ranges.

Parameters	Description	Minimum	Maximum
GW_DELAY	Ground water delay (d)	0	500
OV_N	Manning's n value for the overland flow	0.01	30
SOL_K	Saturated hydraulic conductivity	0	2000
ALPHA_BF	Base flow alpha factor	0	1
CH_N2	Manning's n value for the main channel	-0.01	0.3
CH_K2	Effective hydraulic conductivity in main channel alluvium	-0.01	500
CN2	SCS runoff curve number f	35	98
USLE_P	USLE equation support parameter	0	1
SOL_AWC	Available water capacity of the soil layer	0	1
SURLAG	Surface runoff lag time	0.05	24
USLE_K	USLE equation soil erodibility (K) factor	0	0.65

Note : SCS: soil conservation service; USLE: universal soil loss equation.

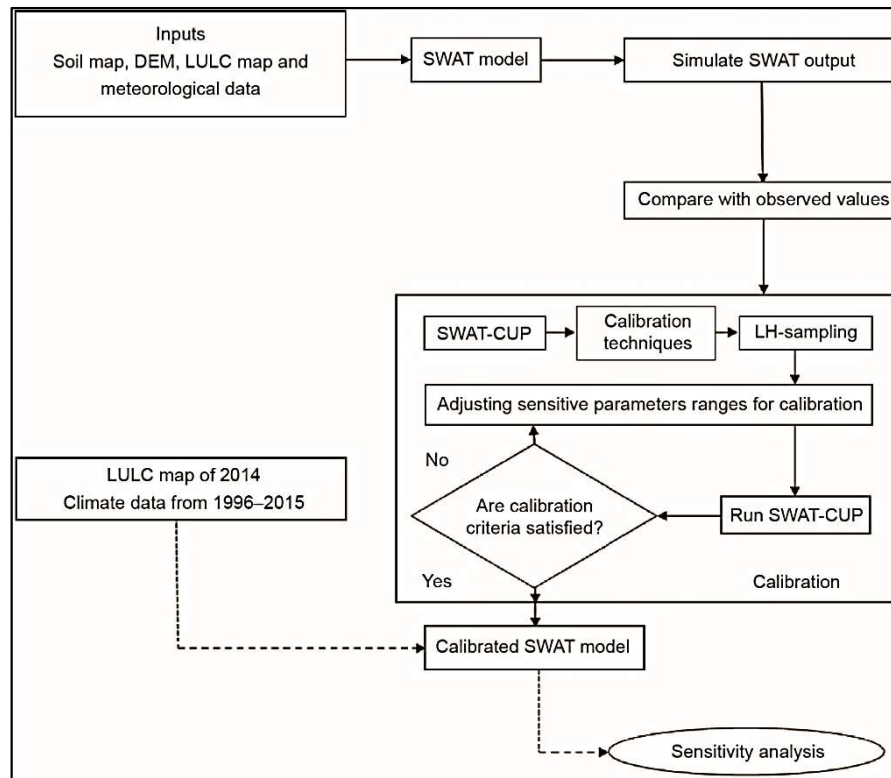


Figure 7.1 Calibration flowchart.

7.2.1 Sequential Uncertainty Fitting Version 2 (Sufi-2)

The SUFI-2 technique is a stochastic algorithmic approach most frequently used by scientists for evaluating uncertainty. In this algorithm, the extent to which all uncertainties are represented is assessed by a parameter referred to as the P-factor. This parameter is the percentage of estimated information sectioned by the 95% prediction uncertainty, also called 95PPU. Another parameter evaluating the quality of the calibration is the R-factor, which is the standard thickness of the 95PPU band separated by the standard deviation of the predetermined information. Hypothetically, the estimation of the P-factor ranges from 0 to 100%, while that of the R-factor ranges from 0 without limit. A P-factor of 1 and R-factor of 0 signify a recreation that precisely compares to the predetermined information. In this study, SWAT-CUP was the software used to consolidate the SWAT 2012 simulated

values and observed values. Using these values and the SUFI-2 algorithm, the uncertainty analysis and calibration were done.

A short well-ordered depiction of the SUFI-2 algorithm is as follows:

1. In the first step, the objective function is characterized. Then, the minimum and maximum absolute ranges of the physically important parameters being optimized are identified.
2. In this step, a sensitivity analysis for each of the parameters is completed; afterward, the initial uncertainty ranges are relegated to the parameters for the first round of Latin hypercube testing.
3. Latin hypercube sampling is performed, and the corresponding objective functions are assessed. The sensitivity matrix J and the parameter covariance grid C are ascertained by

$$J_{ij} = \frac{\Delta g_i}{\Delta \theta_j}, \quad i = 1, \dots, C_2^m, \quad j = 1, \dots, n, \quad (7.1)$$

$$C = S_g^2 (J^T J)^{-1} \quad (7.2)$$

where S_g^2 is the variance of the objective function values resulting from the model runs.

4. The 95PPU is then calculated, followed by the P-factor and R-factor.

7.2.2 Generalized Likelihood Uncertainty Estimation (GLUE)

GLUE was applied to consider the non-uniqueness of parameter sets during the estimation

of model parameters in over-parameterized models. The method is simple, and GLUE accepts, that for vastly over-parameterized models, there is no unusual arrangement of parameters that optimize the decency-of-fit criteria. In GLUE, all sources of parameter uncertainties are considered in the parameter uncertainty. The probability esteem, which is related to a parameter set, mirrors all sources of error and the impacts of covariation of parameter esteem on the model execution. A GLUE examination comprises the following three stages:

1. After defining the “generalized likelihood measure,” $L(h)$, many parameter sets are randomly sampled from the earlier dissemination; each parameter set is assessed as either “behavioral” or “non-behavioral” through a comparison of the “likelihood measure” with a selected threshold value.
2. Each behavioral parameter set is given a “likelihood weight” according to

$$W_i = \frac{L(\phi)}{\sum_{k=1}^N L(\phi)} \quad (7.3)$$

where n is the number of behavioral parameter sets.

3. Finally, the prediction uncertainty is depicted as a forecast quintile from the aggregate dissemination acknowledged from the weighted behavioral parameter sets. The most frequently utilized likelihood measure for GLUE is the Nash–Sutcliffe coefficient (NSE).

7.2.3 Parallel Solution

ParaSol depends on an adjustment to the worldwide optimization calculation SCE-UA. The motivation is to utilize the simulation performed during optimization to derive prediction uncertainty because the simulations accumulated by SCE-UA are extremely important as

the calculation tests over the whole parameter space with an emphasis on arrangements close the ideal/Optima.

The process of ParaSol is as per the following:

1. After streamlining the application of the changed SCE-UA (the arbitrariness of the SCE-UA calculation is expanded to enhance the scope of the parameter space), the reproductions performed are partitioned into “great” recreations and “not great” recreations using a limit estimation of the target work, as in GLUE. This yields the “great” parameter sets and “not great” parameter set.
2. A prediction of uncertainty is built by similarly weighing every “great” recreation. The target work utilized as a part of ParaSol is the sum of the squares of the residuals (SSQ):

$$SSQ = \sum_{tj=1}^n (y_{ti}^M(\varnothing) - y_{ti})^2 \quad (7.4)$$

Brief descriptions of SWAT-CUP, SUFI-2, GLUE, and ParaSol model components are found in the SWAT-CUP user manual by Abbaspour (2014) and the SWAT-model use, calibration and validation document by Arnold et al., 2012.

7.3 Methodology and Criteria for Comparison

Different challenges appear when comparing calibration techniques used for soil and water modeling. The most important concerns that we addressed while comparing these techniques are as follows:

- (1) Most algorithms are diverse in their theories. Personal decisions must be made in their formulation concerning prior parameter distributions and objective functions. In this study, this issue was addressed by choosing objective functions for each algorithm as they would

be utilized in hydrological applications. Fundamentally, this leads to various objective functions for various algorithms. In discussing the outcomes, we indicate whether an issue is caused by the general definition of a specific algorithm or by choice of the function.

(2) Every algorithm has its own fundamental concept and objective functions, which impairs comparison. To solve this problem, the value of each function behind every algorithm was calculated to permit fair comparison. Similarly, the measures of computational proficiency and an appraisal of the conceptual basis criteria are utilized for better comparison.

(3) It is fairly apparent that each algorithm yields different results. To address this problem, the results of all the algorithms were considered for all possible criteria; it is then outlined them all so that the reader can draw his or her own particular conclusions.

(4) The outcomes of the comparison innately depend upon the applications. To solve this issue, the specific results of each application of the algorithm are separated from the generic ones.

Following five categories were used for the comparison of the three (SUFI-2, GLUE, and ParaSol) calibration algorithms:

1. There are several parameters used by the calibration techniques for calibration. For each technique, the best estimate and uncertainty range of these parameters differ, so the first comparison is based on the best estimate and the minimum and maximum uncertainty ranges of each algorithm and the parameter correlation.
2. There are various objective functions used by the various algorithms, so the second comparison is based on the values of NS, R2, and other objective functions.

3. The third comparison is based on the values of the R-factor, which is the average width of the band divided by the standard deviation of the corresponding measured variable, and the P-factor, which is the percentage of data bracketed by the 95PPU band of each algorithm.
4. The fourth comparison was made based on theoretical concepts, testability, and fulfillment of statistical assumptions.
5. The final comparison was made based on the difficulty of implementation.

7.4 Results

7.4.1 Comparison Output

Table 7.2 summarizes the comparison of the calibration algorithms. The comparison was performed over the five categories described earlier, with the following results:

Criteria 1: According to this criteria, GLUE is better than SUFI-2 and ParaSol, as GLUE has the widest range of uncertainties. Thus, most of the intervals considered by SUFI-2 and ParaSol are covered by GLUE. In this criteria, SUFI-2 is the second best.

Criteria 2: According to this Criteria, ParaSol is the best for its objective function, that is, NSE, as it is based on a global optimization algorithm. As shown by the table values, it can be concluded that SUFI-2 and GLUE have similar values, so they hold the same position.

Criteria 3: According to this criterion, ParaSol is worse than the other two algorithms because of its narrow prediction uncertainty bands. From the other values, it can be concluded that SUFI-2 works better than GLUE.

Criteria 4: According to this criterion, all sources of uncertainties are considered by SUFI-2 and GLUE, whereas ParaSol only considers parameter uncertainty and ignores all other

uncertainties. On a conceptual basis, SUFI-2 is better than GLUE and ParaSol because it gives the most effective results.

Criteria 5: According to this criterion, GLUE is the best, as it is the easiest of the three algorithms.

Table 7.2 Comparison of calibration algorithms, according to the criteria given in Section 7.4.1

Criteria	Criteria	GLUE	Parasol	SUFI-2
1 ^a	GW_DELAY	160.1(9.52,79.29)	108.70(92.23,114.2)	191.1(100.04, 300)
	SOL_K	0.16 (0.36, 0.78)	0.37(0.41, 0.34)	0.10 (0.58, 0.34)
	ALPHA_BF	0.12 (0.06, 0.97)	0.12(0.08, 0.13)	0.51 (0.23, 0.74)
	OV_N	0.05 (0.00, 0.20)	0.11(0.07, 0.10)	0.06 (0.00, 0.11)
	CH_N2	0.30(-0.01, 0.30)	0.20 (0, 0.30)	0.20 (0, 0.30)
	CH_K2	78.19(6.01,144.82)	35.7(27.72,37.67)	83.95(69.42,150.0)
	CN2	17.68 (28.85, 8.90)	21.17 (20.93, 20.08)	27.00 (29.00, 7.23)
	USLE_P	0.40 (0, 1)	0.50 (0, 1)	0.30 (0, 1)
	SOL_AWC	0.11 (0.01, 0.15)	0.07 (0.08, 0.08)	0.07 (0.05, 0.15)
	SURLAG	2.00 (0.05, 24)	3.20 (2.00, 20)	2.00 (0.05, 23)
	USLE_K	0.17(0.01, 0.64)	0.19 (0.07, 0.60)	0.17 (0.01, 0.63)
	Parameter correlations	Yes	Yes	No
2 ^b	NSE	0.76	0.77	0.76
	R^2	0.77	0.66	0.78
	$b R^2$	0.63	0.62	0.62

3	<i>P</i> -factor	0.76	0.18	0.79
	<i>R</i> -factor	0.65	0.08	1.51
4 ^c	Source of parameter uncertainty	All sources	Parameter only	All sources
	The conceptual basis of parameter uncertainty	i. Normalization of generalized likelihood measure ii. Primitive random sampling strategy	i. Least squares (probability theory) ii. SCE-UA based sampling strategy	i Generalized objective function ii. Latin hypercube sampling; restriction of sampling intervals
	Testability of statistical assumption	No	Yes	Yes
5	Difficulty of implementation	Very easy	Easy	Easy

^a Best estimate values of parameters and their minimum and maximum range.

^b Values of objective functions.

^c Uncertainty described by parameter uncertainty.

From the comparison and description, it is concluded that the performances of SUFI-2 and GLUE are very similar. However, by looking at the 95PPU plot (Figures. 7.2 and 7.3), it is concluded that SUFI-2 performs better than GLUE, as the results are more efficient and the algorithm considers uncertainties better. SUFI-2 can be run with the smallest number of parameters. GLUE works with a large number of simulations, but cannot provide better results than SUFI-2. The main drawback of GLUE is the length of the computational techniques because of its random sampling strategy. Thus, SUFI-2 is considered to be the best for this study, and further analysis of sediment yield calibration and validation was

done using only the SUFI-2 algorithm.

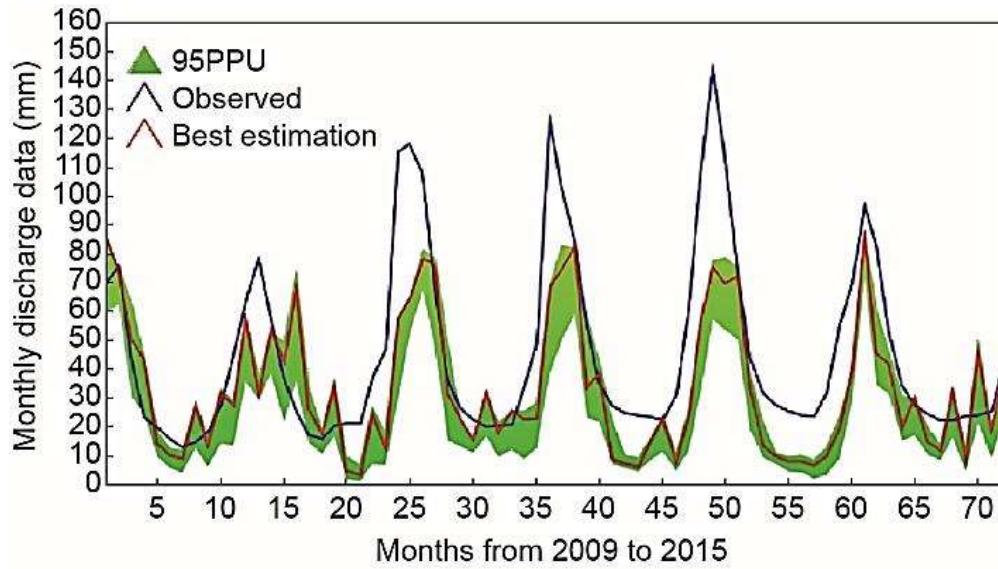


Figure 7.2 The plot of estimated and observed values after calibration using GLUE.

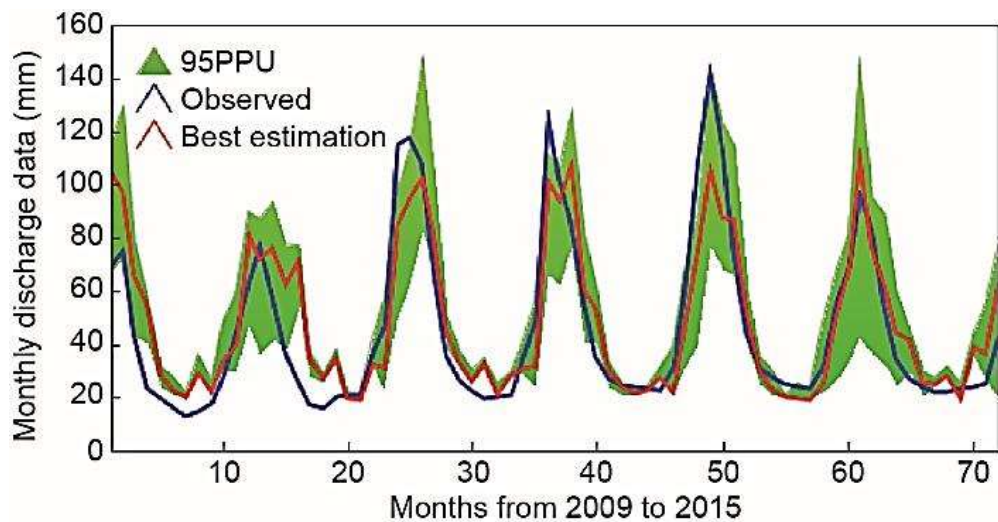


Figure 7.3 The plot of estimated and observed values after calibration using SUFI-2.

7.4.2 SUFI-2 Output

In the present analysis, the model was calibrated using the sensitive parameters given in Table 7.1. As indicated by the execution evaluation criteria of the model, which is suggested for a month-to-month time cycle, five parameters were determined to be the most

sensitive parameters for modeling the present watershed. These parameters are the effective channel hydraulic conductivity (CH_K2), universal soil-loss equation (USLE), support parameter (USLE_P), Manning’s n value for the main channel (CH_N2), surface runoff lag time (SURLAG), and available water capacity of the soil layer (SOL_AWC). The behavioral threshold for calibration was set to 0.500000. The period from 2000 to 2004 was considered to be a warm-up period. Data from the period 2004–2009 were used for calibration, and the data period 2009–2015 was used for validation. The results of the calibration and validation performed by the SUFI-2 algorithm are shown in Figure 7.3 and 7.4. Figure 7.3 shows a plot of the estimated and observed values of discharge and Figure 7.4 shows a plot of the estimated and observed values of sediment data after calibration. Estimated outputs were compared at the same outlet point in sub-basin #27.

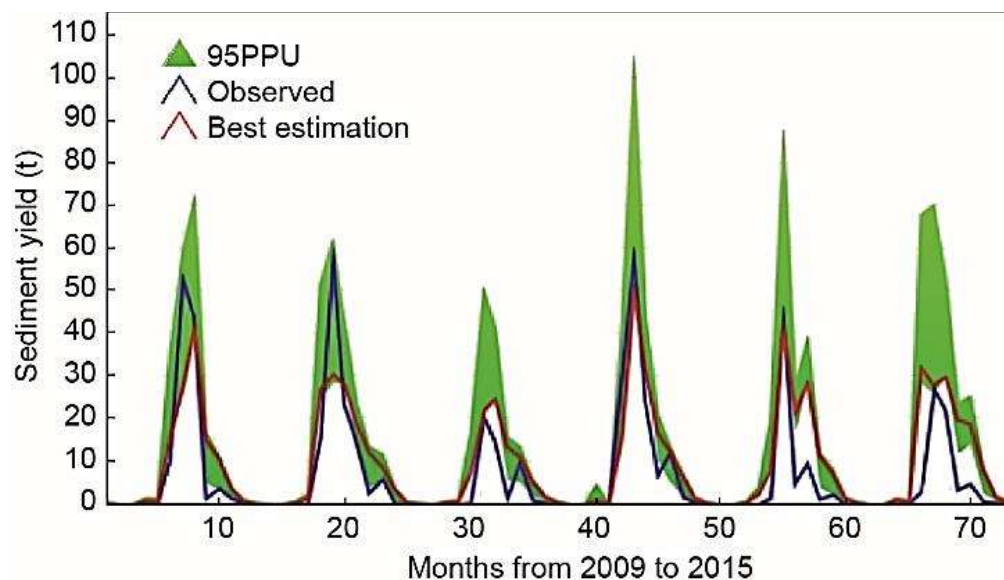


Figure 7.4. The plot of estimated and observed values of sediment yield (mm).

Table 7.3 lists the evaluation coefficients of the simulated monthly sediment yield of different objective functions for the study watershed. The assessment coefficients of the simulated month-to-month sediment yield of various target functions for the study area are

recorded in Table 7.2. The P-factor, or the percentage of observations bracketed by the 95PPU, was 0.69. The R-factor was 0.763. The outcomes demonstrated that SUFI-2 sectioned a lower percentage of the observed sediment yield. Similarly, $R^2 = 0.78$, $NSE = 0.76$, percent bias (PBIAS) = 2.4×10^2 , and the observation standard deviation ratio (RSR) = 0.49 are not strong indicators of the goodness of fit; however, they are within adequate evaluation ratings. The outcomes demonstrate that the SWAT can simulate the hydrological attributes of the Indian watershed extremely well. Hence, this model can be used for further hydrological studies in the basin.

Table 7.3 Objective function values.

Method	P-factor	R-factor	R^2	NSE	bR^2	PBIAS	KGE	RSR
SUFI-2	0.69	0.763	0.78	0.76	0.62	2.4×10^2	0.82	0.49

KGE: Kling-Gupta efficiency

7.5 Conclusion

In this chapter, the sediment and hydrological modeling results obtained using SWAT in chapter 6 were calibrated and validated. A significant outcome was obtained using 20 years of daily meteorological data from 1996 to 2015. The period from 2000 to 2004 was used as a warm-up period, and the results of the period from 2004 to 2015 were divided into two halves. The first half, from 2004 to 2009, was used for calibration, and the second half, from 2009 to 2015, was used for validation. The calibration was done using the three algorithms of GLUE, SUFI-2, and ParaSol; the results were used to compare these algorithms. The comparison showed that of the three calibration algorithms, SUFI-2 performed the best, as it accounted better for uncertainties and required the smallest number of computational parameters for calibration. The main limitations of SUFI-2 are as follows:

(1) Without proper knowledge of the parameters' effects on the stream flow and soil erosion, running SUFI-2 may become difficult for the user.

(2) SUFI-2 does not consider parameter correlations, which decreases its ability.

The P-factor and R-factor calculated with SUFI-2 yielded excellent outcomes, by bracketing value higher than 75% of the observed data. SUFI-2 calculation was therefore considered to be a viable strategy, although it requires extra emphasis and an additional requirement for changing parameter ranges. Despite information uncertainty, the SWAT model creates excellent reproduction consequences for monthly time steps, which are helpful for water resource management in this watershed.