Multi-Objective Automatic Calibration of a Semi-Distributed Watershed Model using Pareto Ordering Optimization and Genetic Algorithm

Remegio B. Confesor Jr., Research Associate (confesor@onid.orst.edu)
Oregon State University, 3450 SW Campus Way, Corvallis, OR 97331

Gerald Whittaker, Research Hydrologist
NFSPRC-USDA-ARS, 3450 SW Campus Way, Corvallis, OR 97331

Written for presentation at the
2006 ASABE Annual International Meeting
Sponsored by ASABE
Portland Convention Center
Portland, Oregon
9 - 12 July 2006

Abstract. This study explored the application of a multi-objective evolutionary algorithm (MOEA) and Pareto ordering in the multiple-objective automatic calibration of the Soil and Water Assessment Tool (SWAT). SWAT was calibrated in the Calapooia watershed, Oregon, USA, with two different pairs of objective functions in a cluster of 24 parallel computers. The non-dominated sorting genetic algorithm (NSGA-II), a fast MOEA, and SWAT were called from a parallel genetic algorithm library (PGAPACK) to determine the Pareto optimal set. One hundred fifty-five parameters were explicitly calibrated (9 for each 17 hydrologic response units [HRUs] and 2 for the whole watershed). With the root mean square error (RMSE) and mean absolute error (MAE) of the daily flows as objective functions, the Pareto front converged to a narrow range of solution set. A wider Pareto optimal front was formed when the RMSE of high and low flows were used as objective functions. The calibrated SWAT model simulated well the daily streamflow of the Calapooia River for a 3-year period. The daily Nash-Sutcliffe efficiency was 0.85 at calibration and 0.80 at validation. Automatic multi-objective calibration of a complex process-based watershed model such as SWAT was successfully implemented using Pareto ordering optimization and an MOEA. Simultaneous automatic-calibration of flows and water quality parameters for the whole watershed and for different sub-basins, dynamic link with economic models, and integration of uncertainty and sensitivity methods are now explored.

Keywords. SWAT, NSGA-II, multi-objective, automatic calibration, Pareto, genetic algorithm.
Introduction

Calibration is the estimation of model parameters to achieve a system that closely resembles the actual system that the model represents (Yapo et al., 1998). Model calibration is necessary since most of the model's parameters cannot be directly measured or are not exactly known because of spatial variability and measurement error (Gupta et al., 1998; Eckhardt and Arnold, 2001). Manual trial and error calibration is time consuming and depends on the modeler's experience, skill, and knowledge of the model's processes and dynamics. While overcoming the shortcomings of manual trial and error calibration, automatic calibration of complex hydrologic models requires tremendous amount of computation. Multiple objective optimizations improve model calibration but further increase the computational complexity. Classical multi-objective optimization involves transforming multiple objectives into a single function with user-defined parameters. Van Griensven and Bauwens (2003) auto-calibrated an extended SWAT model (ESWAT) with multiple objectives by normalizing and putting weights in their objective functions to create a global optimization criterion (GOC).

To avoid the disadvantages of converting multi-objective functions into a single optimization problem, Pareto optimization has gained use and popularity in hydrologic modeling. In problems with at least two conflicting objectives a set of optimal solutions exists because of the trade-offs between these objectives. A Pareto optimal solution set is achieved when no solution can be improved without degrading another solution. The optimal Pareto front for most if not all real-world problems is not known. One of the major challenges in multi-objective optimization then is finding the true or near true Pareto optimal front. Previous multi-objective automatic calibration of hydrologic models (Yapo et al., 1998; Boyle et al., 2001; Vrugt et al., 2003; Khu and Madsen, 2005) used dominance and Pareto ranking so that the solutions evolved toward the Pareto optimal set in the search space. These studies generated the final Pareto solution set of the optimization process but did not illustrate the evolution and movement of the solutions. The cardinality of the Pareto optimal set is one if the objective functions are not conflicting to each other (Deb, 2001). This means that the optimum solution corresponding to any objective is the same and the Pareto front in the search space will converge to a single solution. Otherwise, a front of different solutions should be searched for conflicting objectives. This optimal front can be determined after a number of iterations and function evaluations that require tremendous amount of computation.

Gupta et al. (1998) discussed the use of different error statistics as objective functions in the multi-objective calibration of a model to simulate one aspect (i.e., daily flow) of a watershed system. These error statistics are correlated and in most cases do not conflict to each other. They used a general-purpose global multi-objective optimization algorithm (MOCOM-UA) and successfully estimated the Pareto solution space using the Pareto ranking suggested by Goldberg (1989). They further cited that the objectives should be "unrelated" functions that measure different aspects of the observed and simulated data related to the parameters of the system. Boyle et al., (2001) suggested a different approach wherein they partitioned the streamflow hydrograph into three components. The hydrograph were classified as "driven," "nondriven quick," and "nondriven slow" assuming that the behavior of the watershed is different during the periods "driven" by rainfall and periods without rain (Boyle et al., 2001). The periods immediately after rainfall (non-driven quick) should be dominated by interflow and the latter periods (non-driven slow) by baseflow. In SWAT, these flow components are governed by different model parameters (Eckhardt and Arnold, 2001; Van Liew et al., 2005). The multi-objective calibration of SWAT can then be formulated such that the objective functions are related to different model parameters.
Commodity hardware components with tremendous computational power have become cheaper and affordable. Beowulf clusters, offering high performance parallel computation, can be constructed using these components and run with free open-source software such as Linux (Sterling et al., 1999) and Open Source Cluster Application Resources (OSCAR). Whittaker (2004) built a Beowulf cluster consisting of a server (P4 3.2 GHz dual processor) and 11 computation nodes (11 P4-2.4 GHz) at the National Forage Seed Production Research Center (NFSPRC), a unit of the United States Department of Agriculture - Agricultural Research Service (USDA-ARS) in Corvallis, Oregon. The cluster was just upgraded with additional 12 nodes of P4-2.8 GHz. With the NFSPRC Beowulf cluster providing the computational power required in Pareto ranking and multi-objective evolutionary algorithm, this study explored the applications of Pareto ordering optimization using a fast multi-objective evolutionary algorithm in the automatic calibration of a complex process-based, and semi-distributed hydrologic models such as the Soil and Water Assessment Tool (SWAT). This study also demonstrated the behavior of the Pareto optimal solutions generated from different pairs of objective functions in calibrating SWAT.

**Genetic Algorithms**

Genetic algorithms are adaptive heuristic search algorithms based on Darwin’s theory of natural selection and genetics (inheritance, mutation, recombination, etc.) formally introduced by Holland (1975). A population (chromosome) of possible solutions (genes or individuals) is randomly generated and the individuals are represented by binary, integer, real, or character strings. The fitness of each individual in this "parent" population is evaluated using an objective or fitness function. Individuals are stochastically selected from the parent population based on their fitness. The primary objective of the selection operator is to duplicate good solutions and eliminate the bad ones from the population. The most common method is the tournament selection where tournaments based on fitness are played between two solutions and the better solution is placed in the new population (Deb, 2001). Two other solutions are compared again until the new population is filled. Each solution can be made to participate in two tournaments and the best solution will win both times while the worst solution will lose twice. Any solution in the original population will then have zero, one, or two copies in the new population.

The selection operator only copies good solutions into the new population and does not create any new solution. New solutions are created through crossover and mutation. In single-point bitwise crossover, two strings (or solutions) are randomly chosen from the new population and all their bits on the right side of the crossing site are exchanged (Deb, 2001). A crossover probability sets how many of the strings (or solutions) in the population are used in the crossover. After crossover, the solutions in the new population can undergo mutation determined by a mutation probability. In a bitwise mutation, 1 is changed to 0 or vice versa, and a simple mutation clock operator is used to determine its location in the string. The resulting population after selection, crossover, and mutation is the new "child" population. This "child" population becomes the "parent" population in the next iteration of the algorithm. The iteration is terminated by conditions set by the user (e.g., acceptable solution set or a maximum number of iterations).

**NSGA-II**

Deb et al., (2002) proposed the nondominated sorting genetic algorithm II (NSGA-II), a fast and efficient MOEA characterized by a nondominated sorting algorithm, an elitist selection method, and the elimination of a sharing parameter. Unlike single objective GAs, NSGA-II assigns fitness by Pareto ranking (nondomination) and crowding distance to the combined parent and child
populations. A solution (or individual) is nondominated if it performs better in at least one objective functions and as well in all the other objective functions. The individual is then ranked according to the number of solutions that dominates it. Crowding distance is the average distance between an individual and its nearest neighbors in the search space (Deb et al., 2002). With the objective functions as minimization problems, individuals that are dominated by fewer solutions (i.e., has a lower rank) are given a better fitness than the dominated ones. In cases where the solutions have the same nondomination rank, the individual with larger crowding distance is preferred, thus ensuring diverse and well-spread population. The new parent population is chosen from the combined parent and child population based on the individuals’ fitness or rank, thus the elitist selection. NSGA-II had been tested and yielded solutions that converged to the true Pareto front of problems with convex, non-convex, non-convex disconnected, convex disconnected, and non-convex non-uniformly distributed solutions (Deb et al., 2002).

NSGA-II’s nondominated sorting algorithm has a computational complexity $O(MN^2)$, where $M$ is the number of objectives and $N$ is the size of a population $P$. For each solution two parameters are calculated in the algorithm: 1) the domination count $n_i$, which is the number of solutions that dominate the solution $i$ and 2) a set of solutions $S_i$ that the solution $i$ dominates (Deb, 2001; Deb et al., 2002). The solutions of the first non-dominated front are identified in Steps 1 to 3 of the algorithm and the solutions in higher fronts are searched in Steps 4 to 6. The algorithm is as follows:

Step 1: For each $i \in P$, set $n_i = 0$ and $S_i = \emptyset$.

Step 2: For all $j \neq i$ and $j \in P$, if $i$ dominates $j$,

Add $j$ to the set of solutions dominated by $i$: $S_i = S_i \cup \{j\}$.

Otherwise, increment the domination count of $i$: $n_i = n_i + 1$.

Step 3: If $n_i = 0$, keep $i$ in the first non-dominated front $P_1$ and set the front counter $k = 1$.

Step 4: While $P_k \neq \emptyset$, initialize $Q = \emptyset$ for storing the next non-dominated solutions.

Step 5: For each $i \in P_k$ and for each $j \in S_i$, update $n_j = n_j - 1$.

If $n_j = 0$, $j$ belongs to the next front and update $Q = Q \cup \{j\}$.

Step 6: Set $k = k + 1$ and $P_k = Q$, goto Step 4.

**PGAPack Library**

PGAPack is a general purpose, data-structure-neutral, parallel genetic algorithm library developed at the Argonne National Laboratory (Levine, 1996). Its key features include: (1) Parallel portability across uni-processors, multiprocessors, multi-computers, and workstation networks, (2) Callable in C and Fortran languages, (3) Binary-, integer-, real-, and character-valued native data types, (4) Simple interface for novice and expert application users, (5) Large set of example problems, and (6) Parameterized population replacement. The structure and usage of PGAPack is discussed in detail in Levine, 1996. PGAPack can be downloaded by anonymous ftp from ftp.mcs.anl.gov in pub/pgapack as pgapack.tar.Z.
METHODS

Model Description

The Soil and Water Assessment Tool (SWAT; Arnold et al., 1998) was developed by the United States Department of Agriculture - Agricultural Research Service (USDA-ARS) "to predict the impact of land management practices on water, sediment and agricultural chemical yields in large complex watersheds with varying soils, land use and management conditions over a long period of time." SWAT is physically based, uses readily available inputs, is computationally efficient, and is a continuous model that operates on a daily time step. SWAT is not designed to simulate single-event storms. The buildup of pollutants and their impact on water bodies can be studied with SWAT simulation runs spanning several decades. In SWAT, the entire watershed can be divided into several subbasins and each subbasin is further divided into unique combinations of land use and soil properties called the hydrologic response unit (HRU). However, the location of each HRU is not specified in the subbasin. A graphical user Geographic Information Service (GIS) interface (AVSWAT2000) can be used to input and designate land use, soil, weather, groundwater, water use, management, pond and stream water quality data, and the simulation period (Di Luzio et al., 2001). GIS input files include digital elevation model (DEM), land use and soil properties layers, and weather database.

Watershed Description

The Calapooia river watershed (US Geological Survey, USGS, 10 digit HUC 1709000303) is a tributary of the Willamette river basin west of the Cascades mountain range in Oregon (Figure 1). It has drainage area of 963 km² as delineated from a USGS streamflow gaging station (44°37'15" N, 123°07'40" W) in Albany, Linn County, Oregon. Its elevation ranged from 56 m to 1576 m and its land use is mainly agriculture (43%), forest (41.8%), and hay/pasture/range areas (11.2%). The remaining areas were composed of wetlands, urban areas, and water bodies. The watershed was delineated with a threshold size of 150 hectares using the Arcview interface to SWAT2000 (Di Luzio et al., 2001). The 10-m DEM used in delineating the watersheds was taken from the Regional Ecosystem Office http://www.reo.gov/reo/data/DEM_Files/indexes/orequadindex.asp). The observed daily streamflow data used in calibrating SWAT were obtained from the USGS National Water Information System (NWIS) website (http://nwis.waterdata.usgs.gov/nwis/discharge). The state soil geographic (STATSGO) database for Oregon was from the US Department of Agriculture - National Resources Conservation Service, USDA-NCRS (http://www.ncgc.nrcs.usda.gov/products/datasets/statsgo). Land use for the Willamette basin was acquired from the USGS National Water-Quality Assessment (NAWQA) Program (http://or.water.usgs.gov/projs_dir/ pn366/landuse.html). Climate data were taken from the Oregon Climatic Service (http://www.ocs.oregonstate.edu/).

Automatic Calibration

The SWAT model was initially set up using the Arcview interface (AVSWAT2000) to SWAT (Di Luzio et al., 2001). HRU distribution was defined using a 10% / 10% (Land use over subbasin area / Soil class over land use area) threshold and resulted in four dominant land uses (mixed forest; evergreen forest; perennial grass; and hay, pasture, and rangelands) and nine major soil groups in the watershed. After overlaying the land use and soil properties layers there were a total of 17 HRUs. Based from the SWAT user’s manual (Neitsch et al., 1999) and previous SWAT sensitivity analysis studies (Eckhardt and Arnold, 2001; Van Liew et al., 2005), eleven
variables (9 for each HRU and 2 for the whole watershed) were used in the calibration, resulting in 155 parameters to be explicitly optimized. The limits of the variables for calibration were fixed to ensure realistic and acceptable values representative of the watershed characteristics (Table 1). The simulation period was set for three water years (October 1973 to September 1976) due to the availability of streamflow data.

PGAPACK was called in FORTRAN to randomly generate the initial parent population of 120 solutions. The parameters in each solution were assigned in 10 bits of gray-coded binary strings. A child population (size = 120) was then generated through selection, crossover, and mutation of the parent population with PGAPACK. In the first iteration (or generation) the child and parent populations were evaluated for the two objective functions (Figure 2). In the evaluation step, SWAT was called as a subroutine and was executed for each solution. SWAT’s source code was modified so that the values of the parameters of each solution were read instead of the values from the input files previously generated by AVSWAT2000. The evaluation steps were implemented in a Beowulf cluster of 24 parallel computers. NSGA-II’s nondominated sorting algorithm was then called to rank the individuals of the combined population and assign their fitness. The best 120 individuals were selected to form the new parent population for the next iteration. This new parent population was used to form the new child population through crossover and mutation. After the first iteration, only the new child population is evaluated for the two objective functions since the parent population was previously evaluated and its fitness was already known. Previous simulation runs by the authors showed that there was minimal change in the Pareto front and the objective function values at 1000th generation, thus the maximum number of iterations was set to 1000.
Table 1. Range of Values of Variables Selected for the Calibration of the SWAT model.\(^a\)

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>DESCRIPTION</th>
<th>MIN</th>
<th>MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>GWDELAY</td>
<td>Groundwater delay time (days).</td>
<td>0.001</td>
<td>62.000</td>
</tr>
<tr>
<td>ALPHABF</td>
<td>Baseflow alpha factor (days).</td>
<td>0.040</td>
<td>1.000</td>
</tr>
<tr>
<td>GWQMN</td>
<td>Threshold depth of water in the shallow aquifer required for return flow to occur (mm).</td>
<td>0.000</td>
<td>100.000</td>
</tr>
<tr>
<td>GWREVP</td>
<td>Groundwater 'revap' coefficient.</td>
<td>0.020</td>
<td>0.250</td>
</tr>
<tr>
<td>REVAPMNP</td>
<td>Threshold depth of water in the shallow aquifer for 'revap' or percolation to the deep aquifer to occur (mm).</td>
<td>0.001</td>
<td>200.000</td>
</tr>
<tr>
<td>RCHRGP</td>
<td>Deep aquifer percolation factor.</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>AWHC(^b)</td>
<td>Available water capacity of the upper-most soil layer (mm H2O/mm-soil).</td>
<td>0.110</td>
<td>0.300</td>
</tr>
<tr>
<td>ESCO(^b)</td>
<td>Soil evaporation compensation factor.</td>
<td>0.700</td>
<td>1.000</td>
</tr>
<tr>
<td>CN2(^b)</td>
<td>SCS runoff curve number.</td>
<td>55.000</td>
<td>89.000</td>
</tr>
<tr>
<td>CHK2(^c)</td>
<td>Effective hydraulic conductivity in the main channel alluvium (mm/hr).</td>
<td>0.000</td>
<td>125.000</td>
</tr>
<tr>
<td>SURLAG(^c)</td>
<td>Surface runoff lag coefficient.</td>
<td>1.000</td>
<td>21.000</td>
</tr>
</tbody>
</table>

\(^a\) Unless specified, the range of values is the same for all the hydrologic response units (HRUs).
\(^b\) For brevity, only the range of values across all HRUs is shown. Each HRU has a different range depending on soil type or land use.
\(^c\) Only one value of this variable is calculated for the whole watershed.

Objective functions

Two pairs of objectives were used in this present study. The first pair of objective functions were to minimize the RMSE of the observed with simulated daily flows and to minimize the mean absolute error (MAE) of the observed with simulated daily flows of the main watershed outlet. These objective functions are error statistics that are correlated and do not necessarily conflict with each other. The RMSE and MAE were defined as:

\[
RMSE = \left[ \frac{1}{n} \sum_{i=1}^{n} (Q_{obs,i} - Q_{sim,i})^2 \right]^{0.5}
\]

(1)

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} \text{abs}(Q_{obs,i} - Q_{sim,i})
\]

(2)

where, \(n\) is the total number of daily flow events, \(Q_{obs,i}\) is the observed streamflow at flow event \(i\), and \(Q_{sim,i}\) is the simulated streamflow at flow event \(i\). The second pair of objective functions were to minimize the average root mean square error (RMSE) of the observed vs. simulated high (driven) flows and to minimize the average RMSE of the observed vs. simulated low (nondriven) flows. These objective functions may not strictly conflict but are independent of each other. A baseflow filter (Arnold et al., 1995; Arnold and Allen, 1999) was used to estimate the baseflow component of the observed streamflow. The streamflow was designated as "driven" when the 1st pass baseflow was less than 80% of the observed streamflow; otherwise the streamflow was classified as "nondriven."

Evaluation criteria

Two statistical parameters were used to evaluate the simulation output of SWAT at calibration and validation. The first parameter was the percent bias (PBIAS) defined as the average tendency of simulated flow to underestimate or overestimate the observed values (van Liew et al., 2005). In equation form (all the variables are the same as previously defined):
The second parameter is the Nash-Sutcliffe model efficiency (NSE; Nash and Sutcliffe, 1970) defined as:

\[
PBIAS = \left[ \sum_{i=1}^{n} \left( Q_{\text{obs},i} - Q_{\text{sim},i} \right)(100) \right] / \left[ \sum_{i=1}^{n} Q_{\text{obs},i} \right]
\]

(3)

\[
NSE = 1 - \frac{1}{n} \sum_{i=1}^{n} \left( Q_{\text{obs},i} - Q_{\text{sim},i} \right)^2 / \left( \frac{1}{n} \sum_{i=1}^{n} Q_{\text{obs},i} - \bar{Q}_{\text{obs}} \right)^2
\]

(4)

---

Figure 2. Genetic Algorithm Parallel Computational Scheme in PGAPACK Linking NSGA-II and SWAT with 2 Objective Functions.
where, $Q_{obs}$ is average of the observed daily flows and all the other variables are as previously defined. The Nash-Sutcliffe efficiency ranges from negative infinity to 1, with 1 indicating a perfect fit. Plots of the daily flow hydrographs (observed and simulated) were compared. The fit of the observed and simulated mean monthly flows in a 1:1 line were also analyzed.

**Results and Discussion**

The optimization run of 1000 iterations took 13.2 hours or an average of 47.5 seconds per iteration for each pair of objective functions. Except for the first iteration, there were 240 function evaluations (2 objective values x 120 solutions) and SWAT was executed in each evaluation. Thus, during the entire calibration period there were 240,240 SWAT runs. Khu and Madsen (2005) observed that a "small" Pareto front was generated for highly correlated objective functions. However, they stopped their optimization run at 20 generations and may not have searched the global optimum Pareto front. In this present study it was apparent that at the 100th iteration the Pareto front of the first pair objective functions (RMSE vs. MAE of the observed and simulated daily flows) converges to a narrow range of closely spaced solutions (Figure 3). This was not surprising since these two functions were correlated and do not conflict with each other. Theoretically, the front should even converge to a single solution. Furthermore, NSGA-II's non-crowding distance algorithm forced and stretched the front into these closely spaced solutions. These results imply that to calibrate a model based in single aspect of the watershed such as daily flow, a pair of error statistics (that are correlated and not conflicting to each other) between the observed and simulated values would be enough as objective functions. However, focus should be directed to searching the true or near true Pareto front that corresponds to the global optimum. Minimal changes in the objective function values were observed after 250 iterations and the Pareto front stabilized with negligible changes in the objective function values at the 1000th generation. In this final generation, the NSE was 0.85 and the PBIAS ranged from 3.65 to 4.38 for all the solutions. Van Liew et al. (2005) categorized PBIAS values of $0 \pm 20\%$ as good solutions.

A wider and distinct Pareto front was formed for the 2nd pair of objective functions (RMSE of the observed vs. simulated high flows and RMSE of the observed vs. simulated low flows) implying

![Figure 3. Evolution of the Pareto front with the RMSE and MAE of observed vs. simulated flows as objective functions.](image-url)
that Pareto ordering took into account the trade-offs between these objectives (Figure 4). The front also stabilized after 1000 iterations indicated by negligible changes in objective function values. NSE values ranged from 0.83 to 0.85 and PBIAS ranged from 0.61 to 9.20. The PBIAS values indicate that the model tend to underestimate the observed values. Although the NSE and PBIAS values for both pair of objectives were not considerably different, using the 2nd pair of objective functions has some physical meaning because of the relationships of the functions with the calibrated parameters. The curve number (CN2), soil evaporation compensation factor (ESCO), and available soil water holding capacity (AWHC) govern the surface water response in SWAT (Eckhardt and Arnold, 2001; van Liew et al., 2005). The optimized values of these parameters favor the contribution of direct runoff to the streamflow. Thus, optimizing the RMSE of peak flows would in turn optimize these parameters. In the same manner, optimizing the RMSE of low flows would optimize the parameters that govern the subsurface water response in SWAT. These parameters were the groundwater “revap” coefficient (GWRVAP), groundwater delay (GWDELAY), threshold depth of water in the shallow aquifer for return flow (GWQMN), threshold depth of water in the shallow aquifer for “revap” or percolation to the deep aquifer to occur (REVAPMN), the baseflow alpha factor (ALPHABF), and the deep aquifer percolation fraction (RCHR_GDP). The other parameters govern SWAT’s basin response and further adjust the shape of the hydrographs. The surface runoff lag time (SURLAG) governs the release of the surface runoff to the main channel. The main channel effective hydraulic conductivity (CHKE) shows the relationship of the stream with the groundwater and directs the water movement from streambed to the subsurface (or vice-versa) depending on the stream type.

The calibrated SWAT model simulated well the daily streamflow of the Calapooia River watershed for a 3-year period (October 1973 to September 1976). The selected solution resulted in a daily Nash-Sutcliffe efficiency of 0.85, which was a large improvement from 0.28 calculated from the simulated daily streamflow using the initial model setup (uncalibrated) with AVSWAT2000. Figure 5 shows the observed daily streamflow plotted with the simulated daily streamflow. To highlight the differences in observed and simulated values, only the rainfall period in 1973 to 1974 (simulation day 28 to 196) was shown. A daily NSE of 0.97 was calculated from 36 monthly means of the same calibration period (Figure 6). Streamflow validation in the following 3 water years (October 1976 - September 1979) resulted

Figure 4. Evolution of the Pareto front with the RMSE (observed vs. simulated) of the non-driven and driven flows as objective functions.
in an overall model efficiencies of 0.80 (daily) and 0.94 (monthly). The PBIAS at validation was -0.81, which means that the simulated values tend to slightly overestimate the observed values. Nonetheless, these good evaluation parameter values further verified the calibration results.

The solution (Solution 3) used in the calibration and validation was taken from the Pareto set at 1000th generation of the 2nd pair of objectives (Figure 4). The selection of the solution was arbitrary since all the solutions were Pareto-optimal and the choice between solutions depend on the interest of the user. This method can be extended to include the error statistics of different measurements in the watershed (i.e., flow, sediment, nitrogen, and phosphorus concentrations) as objective functions.

**Figure 5.** Calibration results: observed and simulated daily flows for the Calapooia watershed.

**Figure 6.** Calibration results: observed and simulated monthly flows for the Calapooia river fitted in a 1:1 Line.
Summary and Conclusions

This study explored the application of multi-objective evolutionary algorithm (MOEA) and Pareto ordering optimization in the automatic calibration of the Soil and Water Assessment Tool (SWAT), a complex, process-based, semi-distributed, and continuous hydrologic model. SWAT was automatically calibrated for the Calapooia watershed in Oregon, USA, with two different pairs of objective functions. The non-dominated sorting genetic algorithm (NSGA-II), a fast and modern MOEA, and SWAT was called in FORTRAN from a parallel genetic algorithm library (PGAPACK) to determine the Pareto optimal set. For non-conflicting objective functions such as error statistics of one hydrologic measurement, the Pareto optimal front converges to a narrow range of closely spaced solution set. To calibrate a model based in a single aspect of the watershed such as daily flow, a pair of different error statistics would be enough as objective functions. In conflicting and unrelated objective functions, Pareto ordering optimization gives the modeler a set of optimal solutions that accounts the trade-offs between the objectives. The selection of a solution depends on the modeler's preference and interest since all the solutions are Pareto optimal. Focus should be directed to searching the true or near true Pareto front that corresponds to the global optimum. The calibrated SWAT model simulated well the daily streamflow of the Calapooia watershed for a 3-year period. The daily Nash-Sutcliffe coefficients were 0.85 at calibration and 0.80 at validation. Future studies include: 1) simultaneous automatic calibration of water quality and quantity parameters for the whole watershed and for different sub-basins, 2) link with economic models, and 3) integration of uncertainty and sensitivity methods. The application of Pareto optimization in decision and policy-making problems related to conflicting objectives of economics and environmental quality is now being explored.

Acknowledgements

This study was funded by the United States Department of Agriculture - Agricultural Research Service (USDA-ARS) through the Conservation Effects Assessment Project (CEAP).

References


