Water Resour Manage (2010) 24:4469–4489
DOI 10.1007/s11269-010-9668-y

Selecting Model Parameter Sets from a Trade-off Surface Generated from the Non-Dominated Sorting Genetic Algorithm-II

**Gift Dumedah · Aaron A. Berg ·
Mark Wineberg · Robert Collier**

Received: 9 November 2009 / Accepted: 11 May 2010 /
Published online: 25 May 2010

Abstract There is increasing trend in the use of multi-objective genetic algorithms (GAs) to estimate parameter sets in the calibration of hydrological models. Multi-objective GAs facilitate the evaluation of several model evaluation objectives, and the examination of massive combinations of parameter sets. Typically, the outcome is a set of several equally-accurate parameter sets which make-up a trade-off surface between the objective functions, usually referred to as Pareto set. The Pareto set is a set of incomparable parameter sets as each solution has unique parameter values in parameter space with competing accuracy in the objective function space. As would be required for decision making purposes, a single parameter set is usually chosen to represent the model calibration procedure. An automated framework for choosing a single solution from such a trade-off surface has not been thoroughly investigated in the model calibration literature. As a result, this study has outlined an automated framework using the distribution of solutions in objective space and parameter space to select solutions with unique properties from an incomparable set of solutions. Our Pareto set was generated from the application of Non-dominated Sorting Genetic Algorithm-II (NSGA-II) to calibrate the Soil and Water Assessment Tool (SWAT) for simulations of streamflow in the Fairchild Creek watershed in

G. Dumedah · A. A. Berg (✉)
Department of Geography, University of Guelph, Guelph, ON, Canada N1G 2W1
e-mail: aberg@uoguelph.ca

G. Dumedah
e-mail: gdumedah@uoguelph.ca

M. Wineberg · R. Collier
Department of Computing, and Information Science, University of Guelph,
Guelph, ON, Canada N1G 2W1

M. Wineberg
e-mail: wineberg@cis.uoguelph.ca

R. Collier
e-mail: collierr@uoguelph.ca

southern Ontario. Using cluster analysis to evaluate the distribution of solutions in both objective space and parameter space, we developed four auto-selection methods for choosing parameter sets from the trade-off surface to support decision making. Our method generates solutions with unique properties including a representative pathway in parameter space, a basin of attraction (or the center of mass) in objective space, a proximity to the origin in objective space, and a balanced compromise between objective space and parameter space (denoted BCOP). The BCOP method is appealing as it is an equally-weighted compromise for the distribution of solutions in objective space and parameter space. That is, the BCOP solution emphasizes stability in model parameter values and in objective function values—in a way that similarity in parameter space implies similarity in objective space.

Keywords Hydrological calibration model • Multi-objective genetic algorithms • NSGA-II • Pareto set • Robust

1 Introduction

As mathematical simplifications of watershed processes, majority of hydrological models are calibrated through parameter estimation. Estimation of model parameter values improves the performance of the model to match observed watershed responses. Calibration may be manual or automated depending on whether the matching between simulated and observed values are evaluated manually, based on a physical adjustment of parameter sets, or computer algorithms are used to search for best model parameter values. Manual adjustment of parameter sets is time intensive and the assessment of the goodness of fit can be subjective. Manual parameter adjustment can be complicated by the effects of model parameter intercorrelations which usually result in compensating errors (Leavesley 1994). As a result of limitations in manual calibration, extensive research and a growing body of literature (Duan et al. 1993; Gupta et al. 1998; Yapo et al. 1998; Beven 2001) have focused on the development of automated parameter adjustment methods. Automated (or computer based) methods employ the speed and power of a computer to search for optimal parameter sets. However, a recurring theme in the model calibration literature has been finding a robust approach to search for an optimum model parameter set during model calibration, and a subsequent evaluation of the distribution of solutions in objective space and parameter space.

In the past, studies (Liang et al. 1994; Young and Beven 1994; Tabrizi et al. 1998) have employed different function optimization techniques such as multiple input transfer functions (Ahsan and O'Connor 1994) to automate parameter estimation for hydrological models. These traditional optimization methods use deterministic transition rules to find the optimum solution for an optimization problem. But the assumption of deterministic transition rules may be difficult to satisfy for most practical model calibrations procedures, which often include several model parameters that exhibit nonlinearities and complex relationships. In addition, traditional optimization methods use different initialization methods to guide the search process which usually result in different optimal solutions. Some methods use direct objective functions and constraint variables while others employ the first and second derivatives of the objective functions and constraints to initiate the search procedure.

Several of the automated approaches such as hill-climbers often get stuck on a local optimum or a sub-optimal solution, particularly in multimodal problems where there are several peaks on the search surface.

Recently, the use of genetic algorithms (GAs) for automated model calibration (Confesor and Whittaker 2007; Gill et al. 2006; Khu and Madsen 2005; Madsen 2003; Wöhling et al. 2008; Tang et al. 2006) has emerged as an important optimization technique in hydrology. GAs constitute a particular class of a large body of computational techniques called evolutionary algorithms (EAs). EAs mimic the concept of natural evolution to address problems that are based on stochastic trial-and-error or generate-and-test problems (Eiben and Smith 2003). GAs consider problem solving as an evolution of a population of individuals in an environment with limited resources. Given that there are limited resources, individuals compete among themselves and are selected to reproduce for the next generation (or the new population). As a result, individuals are selected and undergo reproduction to generate a new population which evolves as a whole. In other words, the competition based selection and variation among members of the population result into a fitter population (or a better solution to a problem).

Recent studies (Khu and Madsen 2005; Gill et al. 2006; Nazemi et al. 2006; Tang et al. 2006; Bekele and Nicklow 2007; Shafii and Smedt 2009) have employed multi-objective GAs to calibrate hydrological models for generating trade-off surface between several objectives. While these studies have progressed from evaluating single objective to multiple objectives the outcomes usually comprise several parameter sets that fit the hydrograph in different ways (Gupta et al. 1998) to form the trade-off surface. As a set of incomparable solutions, the trade-off surface does not have a single definitive solution for decision making but rather a suite of solutions for different trade-offs between the objectives. Selecting a single solution for decision making is a natural follow-up step after generating a trade-off surface for hydrological model calibration. But a definitive or an automated method of selecting solutions from such a trade-off surface is limited in the model calibration literature. It is desirable, for decision making purposes, to evaluate the resulting parameter sets on the trade-off surface for their distributions in objective space and parameter space and thereby identify solutions with unique properties. Additionally, the choice of solution from the trade-off surface can be made on the basis of the level of uncertainty associated with each solution.

Note that each solution on the trade-off surface is valid in itself and different solutions could be preferable to different users or even the same user at different times. If user dependent preferences are not a dominant factor or if these preferences are focused on properties of the Pareto frontier then an auto-selection method can be applied to choose solutions based on properties of the trade-off surface, natural clustering in solution space of solutions in the Pareto set, or a combination of the two.

The decision making process in multi-objective genetic algorithms are generally categorized into three subgroups based on when the decision is made. These categories are decisions made before, after and during the search (Coello Coello et al. 2002; Marler and Arora 2004). Decision made before the search allows the user to rank the objectives in an order of preference/importance (Weinert et al. 2009; Xuebin 2009; Vamvakeridou-Lyroudia et al. 2006; Miyamoto et al. 2006). Decisions made during the search are referred to as interactive searches where the user is allowed input to guide the search (Hee-Su and Sung-Bae 2000; Tappeta et al. 2000). Decision making after the search allows the generation of non-dominated solutions from which the user can select one solution. This paper focuses on this category of decision making after search by providing four methodologies to select solutions from the non-dominated set.

Various studies (Crispim and de Sousa 2009; Taboada and Coit 2006) have applied cluster analysis to select compromise solutions from Pareto frontier. The study by Khu and Madsen (2005) have applied preference ordering and compensation between model objectives to screen through Pareto-optimal solutions and select fewer solutions. Other researches have randomly selected one non-dominated solution which does not degrade the objectives (Tang et al. 2006; Bekele and Nicklow 2007; Ferreira et al. 2007; Grierson 2008). However, a formal methodology to selecting solutions using properties of the Pareto frontier is limited in the literature.

The objective of this paper is to provide an automated framework for selecting solutions from a trade-off surface based on the distribution of parameter sets in objective space and parameter space. To achieve this objective, we use the Non-dominated Sorting Genetic Algorithm II (NSGA-II) (Deb et al. 2002) to calibrate the Soil and Water Assessment Tool (SWAT) for streamflow simulations based on two conflicting objectives: root mean square error (RMSE) and model bias. The modeling approach determines several parameter sets that have competing accuracies of matching simulation to observed streamflows. The resulting solutions constitute a trade-off surface between RMSE and bias, while the parameter values for the solutions represent points in parameter space.

As a follow-up step, the decision making procedure selects one solution from the trade-off surface with a unique property using patterns observed in both objective space and parameter space. The method generates solutions with unique properties including a representative pathway in parameter space, a basin of attraction (or the center of mass) in objective space, a proximity to the origin in objective space, and a balanced compromise between objective space and parameter space. In this analysis, we used cluster analysis (CA) to facilitate this decision making process. The CA evaluates the distribution of solutions on the trade-off surface to find relationships in both objective space and parameter space.

The rest of the paper is organized as follows. Following is the materials and methods section which outlines an automated framework for selecting solutions with unique properties from a trade-off surface. This section also introduces the study area and describes datasets used in the study and the SWAT model. The application of the NSGA-II framework to calibrating SWAT for simulations of streamflow is reported in the results and discussion section. The results section implements the automated framework for selecting solutions from the trade-off surface. The paper concludes with a discussion about the utility of selecting unique solutions through the evaluation of parameter sets in both objective space and parameter space.

2 Materials and Methods

This section describes the study region, datasets and their sources, the SWAT model and its set-up, and the NSGA-II framework to calibrating the model for simulations of streamflow.

2.1 Automated Framework for Selecting Unique Solutions from a Trade-off Surface

As would be required for decision making purposes a single solution is usually chosen from the trade-off surface (also called Pareto frontier). The choice of a solution on the Pareto frontier can be a daunting task because the solutions are incomparable and have competing accuracies as no one solution is better than the other. The Pareto frontier is, therefore, a decision front where a solution can be selected on the basis of a property that is considered important for the problem at hand. Usually, a user may specify a problem-specific property emphasizing the relative importance between the objectives. However, if user preference is focused on the properties of the trade-off surface, a natural clustering of solutions in parameter space and objective space, or a combination of the two then an automated method for choosing solutions can be applied.

In this section, we outline four auto-selection methods to evaluate the distribution of solutions on the Pareto frontier. The distribution of solutions encompasses the clustering of parameter values in parameter space and the spread of solutions in objective space. Since the solutions on the Pareto frontier are incomparable, the selection methods do not in any way provide indication if one solution is better than the other, rather the methods aim to find solutions which have unique properties on the Pareto frontier.

We measure the distribution of solutions in parameter space by clustering the model parameter values of the solutions found on the Pareto frontier. An application of a specific distance function is used to compare parameter vectors, which can be weighted to emphasize one parameter over another. As we have no preference towards one parameter over another, we normalize each parameter and take the Euclidean distance between the normalized vectors. We call the clusters thus formed "parameter clusters" (PC), and index them so that the ith parameter cluster is denoted as PC_i.

Similarly we cluster the distribution of solutions in objective space using a distance measure which combines the model evaluation functions (bias and RMSE in this study, that is, solutions discovered on the Pareto frontier). Specifically, Euclidean distance was used to combine bias and RMSE, although any distance measure tailored to the system under investigation could be used. These clusters are called "fitness clusters" (FC) and the jth fitness cluster is denoted as FC_j.

Other classifiers different from cluster analysis could be applied to analyze the distribution of solutions, so clustering should not be thought of as the only approach. One suggestion is that cluster analysis should be used to analyze the distribution of solutions if and only if a pattern of clustering is observed following the test of clustering outlined in Thorndike (1953). The procedure shown in Thorndike (1953) also illustrates the 'knee' approach to determine the appropriate number of clusters to be used in clustering a specific dataset based on the relative size of reduction of variance introduced by increasing the number of clusters.

2.1.1 Parameter Based Clustering: CAP Method

The first auto-selection method focuses on the distribution of solutions in parameter space. The method applies a clustering analysis based on parameter values (denoted

CAP) to identify persistent sub-categories for parameter clusters by evaluating the trade-off between parameter values. The PC_i with the largest membership (denoted PC_k) is chosen as it represents an area of maximum concentration of solutions (i.e. a representative pathway in parameter space) or a unique region on the Pareto frontier with respect to the parameter space. The cluster center for PC_k is defined as a parameter set that has the minimum averaged Euclidean distance to each of the other members in the cluster (computed after normalizing the parameter values) for the set of solutions in the cluster PC_k. This solution at the cluster center of PC_k is chosen as the CAP solution, that is, a representative pathway in parameter space. Note that the parameter values are weighted equally when computing the Euclidean distance, as a result, no weighting factors are assigned to the parameter values. Since CAP solution is chosen from PC_k, a level of uncertainty around the CAP solution is the interval between minimum and maximum values for objectives (RMSE and bias in this study) for all solutions in PC_k.

2.1.2 Objective Based Clustering: BAPF and IPO Methods

The second auto-selection method focuses on the distribution of solutions in objective space. The method is based on the concept of basin of attraction (or center of mass) on the Pareto frontier (denoted BAPF) to evaluate the trade-off between the objectives (e.g. RMSE and bias). The purpose of the BAPF method is to select a compromise solution whose RMSE and bias values represent the center of mass on the Pareto frontier where mass is a function of the objective function values. The FC_i with the largest membership (denoted FC_k) is chosen as it captures the dominant variability in objective space (i.e. a dominant cluster in objective space) or a unique region on the Pareto frontier with respect to objective space. The cluster center for FC_k is defined as a parameter set that has the minimum averaged Euclidean distance to each of the other member in the cluster for the set of solutions in cluster FC_k. The solution at the cluster center of FC_k is chosen as the BAPF solution. Like the CAP solution, the level of uncertainty around the BAPF solution is the interval between minimum and maximum values for RMSE and bias for all solutions in FC_k.

The third auto-selection method also focuses on the distribution of solutions in objective space. The method evaluates all solutions on the Pareto frontier based on their index of proximity to the origin of the objective function values. We referred to this technique as index of proximity to the origin (IPO) method. The minimum of all the maximum values for solutions on the Pareto frontier represents the parameter set with the least distance to the origin. This method of computing IPO uses a distance based on L_∞ norm which takes the maximum value of a vector just as the Euclidean distance is based on the L_2 norm, which sums the squared values of a vector. It is important to note that this method of finding the IPO parameter set is similar to computing the Euclidean distance between the origin and each ith solution $\left(\text{i.e. } \sqrt{\text{RMSE}_i^2 + \text{Bias}_i^2}\right)$, where RMSE and Bias are the objectives applied in this study. The solution with the minimum Euclidean distance is chosen, and will frequently return the same point. Note that the IPO method has been applied in past studies (Bekele and Nicklow 2007; Tang et al. 2006), and is formalized in this study with the aim to compare its properties to the other selection methods.

2.1.3 Linking Parameter Space and Objective Space: BCOP Method

The fourth auto-selection method, denoted BCOP, determines a balanced compromise between objective space and parameter space. The above three methods of selecting solutions from the trade-off surface have independently considered the distribution of solutions in either objective space or parameter space. Although a dominant pathway in parameter space, the CAP method has no linkage to objective space. Similarly, the BAPF method focuses on the dominant variability in objective function values with no reference to parameter space. The CAP and BAPF solutions are compromise solutions in either parameter space or objective space but these solutions do not represent a compromise for the distribution of solutions in both spaces. The linkage between the two spaces—that is, a dominant pathway in parameter space and the center of mass in objective space—describes the level of robustness for parameter sets. Hence we determine BCOP solutions by evaluating the linkages between the patterns observed in objective space and those in parameter space.

To achieve this, we find a cluster with the largest membership in both objective space (FC_k) and parameter space (PC_k). A solution has the BCOP property if it belongs a representative pathway in parameter space and also in a dominate cluster in objective space. A BCOP solution is chosen by finding the parameter set at the center of the overlapped region between PC_k and FC_k using model parameter values. The uncertainty around the BCOP solution is the interval between minimum and maximum values for RMSE and bias for all solutions in the overlapped region between PC_k and FC_k. Note that if no solution exists in the intersection between PC_k and FC_k then PC_k is overlapped with clusters in objective space until solution(s) are found in the overlapped region. The process starts with clusters with high memberships to clusters with the lowest membership. The chosen BCOP solution represent a compromise between PC_i and FC_i in a way that similarity in parameter space implies similarity objective space. Note that selections methods presented in this paper can apply to a trade-off surface (Pareto front) that is generated using any algorithm. As a result, the selection methods are independent of the algorithm used.

2.2 Study Area

Fairchild Creek (FCr) watershed is located in the central portion of the Grand River Basin in southern Ontario. The FCr watershed shown in Fig. 1 runs from north to south and drains an area of about 400 km^2. Physiographically, the FCr watershed is located on the south-eastern slope of the Galt Moraine, a major moraine system which forms part of the southern Ontario Moraines physiographic region. The FCr watershed is a dendritic drainage system, having elevation range between 184 to 325 m and a slope interval between zero and 18.9°. The FCr watershed, as a dendrite stream, has a typical pathway pattern of decreasing slope downstream. The headwater section is mostly a plain physiographic region with shallow soils, gentle-sloping bedrocks and several wetlands. The upstream section also has a drumlin field comprising small and isolated drumlins. The middle and lower sections have massive deposits of fine-grained sediments such as silt and clay and are eroded by surface runoff. The hummocky terrain accompanied with high erosion soils in this area presents a persistent soil conservation problem. The FCr watershed is extensively

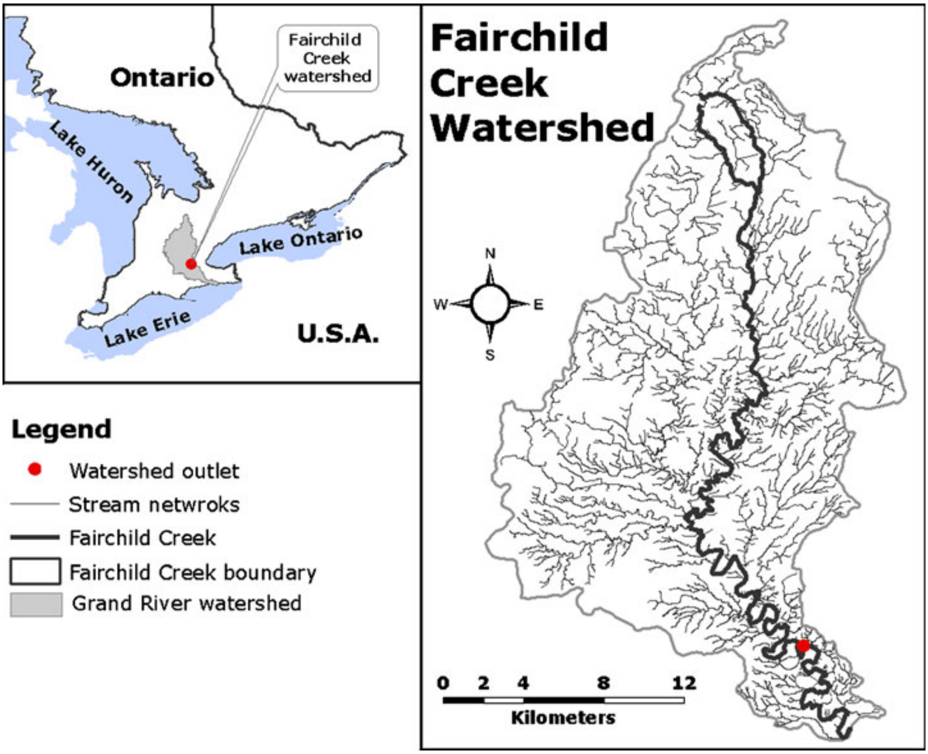


Fig. 1 Study area – Fairchild Creek watershed

agricultural with scattered wetlands and small rural communities. Major land uses and their respective proportions in the watershed include agriculture (64%), forest (21%), pasture (9%), urban (5%) and open water and wetlands (1%).

2.3 Data Description and Sources

Datasets that are used in this study are shown Table 1. A description of each of the individual datasets follows. A 30 m grid land use/cover data which was generated from August 21, 2003 Landsat Thematic Mapper (TM) image was provided by (Bonnycastle 2006). The dataset has specific classifications for corn, soybean and cereal which are predominant crops in the FCr watershed. The soil dataset was provided by Ontario Ministry of Agriculture and Food, and Agriculture and Agri-Food Canada. The dataset was extracted for the counties of Brant, Waterloo, Wellington, and Wentworth as these counties cover different parts of the FCr watershed. The soil data is based on the Ontario Soil Survey data model, which uses the NSDB (National Soil DataBase) data model for CanSIS (Canadian Soil Information System) Detailed Soil Surveys. The FCr watershed has a high proportion of fine textured soils (e.g., clay and silt), which makes the basin very sensitive to surface runoff, particularly, during snowmelt periods.

Meteorological information such as precipitation, temperature, wind, humidity, and solar radiation values are used in SWAT for a number of calculations including

Table 1 Fairchild Creek watershed data for SWAT model

Data	Data uses	Scale/Date	Data sources
Digital elevation model (DEM)	For computing hydrologic parameters	1:10 000; 10 m resolution	Grand River Conservation Authority
Stream network	For watershed delineation	1:10 000	Grand River Conservation Authority
Land use/cover	For computing HRUs and related parameters	1:1 000 000; 30 m resolution	Bonnycastle (2006)
Soil data	For computing HRUs and related parameters	Brant: 1:15840/1989; Waterloo: 1:20000/1971; Wellington: 1:63360/1963; Wentworth: 1:63 360/1965	Grand River Conservation Authority
Weather data	Include daily precipitation, solar radiation, wind speed, temperature, and relative humidity	Brantford MOE, Hamilton A, Millgrove, Roseville, Waterloo and Wellington A	Environment Canada
Stream flow data	For calibrating simulated watershed response	Temporal duration: 1990–2003 time step for discharge: daily	Water Survey of Canada: Ontario Ministry of Environment

crop growth, snowmelt, and evapotranspiration. Specifically, SWAT requires daily precipitation, minimum/maximum air temperature, solar radiation, wind speed and relative humidity. Daily precipitation and temperature data from 1990 to 2003 were provided by Meteorological Service of Canada (Environment-Canada 2006). The streamflow data for the FCr watershed were provided by the Water Survey of Canada (WSC) of the Ontario Ministry of Environment (MOE). The FCr watershed outlet for the MOE station is located near the city of Brantford and has a gross drainage area of about 360 km^2. The discharge data were collected using a recording gauge based on a continuous water level records from which discharge records are computed. The discharge data is available from 1964 to 2005 at a daily interval but only data from 1990 to 2003 were used in this study as temporal duration for climate data are limited.

2.4 The Soil and Water Assessment Tool

Soil and Water Assessment Tool (SWAT) is a conceptual, continuous time step model that was developed by United States Department of Agriculture in the early 1990s for assessing the impact of management and climate on water supplies in watersheds and large river basins (Arnold and Fohrer 2005). SWAT is a physically-based and spatially distributed water quality model that requires extensive input data such as soil, land cover, and terrain information and can operate at a daily and sub-daily time steps. During simulation, SWAT divides the watershed into sub-basins which are grouped based upon input data categories: climate, hydrologic response units (HRU), ponds, ground water, and main channels (Borah and Bera 2003a; Tolson and Shoemaker 2007). SWAT's conceptual basis is the HRU which are lumped land areas within the sub-basin comprising unique land use/cover, soil, and management practices (Borah and Bera 2003a, b). The partitioning of sub-basins facilitates spatial evaluation of site characteristics such as soils or management practices for different locations in the watershed. For example, evapotranspiration and runoff can be computed to reflect various crops and soils for different sub-basins. Runoff volume is calculated using either runoff curve number approach or Green and Ampt infiltration equation (Green and Ampt 1911). SWAT has been applied alongside with Riparian Ecosystem Management Model (REMM) in the FCr watershed to simulate streamflow (Liu and Yang 2007). A detailed description of the SWAT model is presented in various sources (Arnold et al. 1998; Neitsch et al. 2001; Arnold and Fohrer 2005).

The SWAT model was calibrated for streamflow using 24 parameters shown in Table 2. The table shows the description of each parameter and their lower and upper bounds that define their corresponding initial uncertainties. It is worth noting that the lower and upper bounds for the parameters are predetermined by the SWAT model, and that parameters which are changed using a relative percentage value are restricted within their predetermined intervals. The overall parameters changed is number of parameters changed at HRU multiplied by the number of HRUs in one hand, plus number of parameters changed at sub-basin level multiplied by the number of sub-basins. The model is calibrated for a six year period from 1992 to and 1997 and validated for another six year period from 1998 to 2003.

2.5 NSGA-II Framework for Calibrating the SWAT Model

Non-dominated sorting genetic algorithm II (NSGA-II) is a multi-objective genetic algorithm that was developed by Kalyanmoy Deb and his colleagues in Deb et al. (2000, 2002). A key concept applied in NSGA-II is non-domination which states that a candidate solution A dominates B if and only if:

- the performance of A in every dimension (i.e., objective function) is at least as good as the performance of B, and
- the performance of A in at least one dimension is better than the performance of B.

However, if A and B are assigned the same non-dominance level then A does not dominate B and B does not dominate A. In the NSGA-II, the algorithm begins by generating a random population of solutions that is usually referred to as the parent population, P_o of size N which is sorted into different non-domination levels such that each solution is assigned a fitness based on its non-domination level. The result is different non-dominated fronts starting from the best non-dominated front to the second non-dominated front and so on. A child population, Q_o also of size N is created by applying binary tournament selection, crossover and mutation operators to P_o. The generation of Q_o allows a check for global non-domination for both parent and offspring solutions (Deb 2001; Deb and Goel 2001; Deb et al. 2000, 2002).

The NSGA-II has been applied widely in several different studies (Rakesh and Chandan 2005; Agrawal et al. 2006; Atiquzzaman et al. 2006; Jeong and Abraham 2006; Sarkar and Modak 2006; Shafii and Smedt 2009). Specifically in hydrological model calibration studies (Bekele and Nicklow 2007; Wöhling et al. 2008; Confesor and Whittaker 2007; Gill et al. 2006; Tang et al. 2006; Khu and Madsen 2005; Madsen 2003) have used NSGA-II to generate a trade-off surface for different objectives.

Table 2 SWAT streamflow parameter intervals and their descriptions

Model parameter	Change type	Lower bound	Upper bound	Description
TIMP	Absolute	0.01	1.00	Snow pack temperature lag factor
SURLAG	Absolute	0.00	24.00	Surface runoff lag time (days)
SFTMP	Absolute	−5.00	5.00	Snowfall temperature (°C)
SMTMP	Absolute	−5.00	5.00	Snowmelt base temperature (°C)
SMFMX	Absolute	0.00	10.00	Melt factor for snow on Jun. 21 (mm/°C-day)
SMFMN	Absolute	0.00	10.00	Melt factor for snow on Dec. 21 (mm/°C-day)
MSK−CO1	Absolute	0.00	10.00	Calibration coefficient used to control impact of the storage time constant for normal flow (km)
MSK−CO2	Absolute	0.00	10.00	Calibration coefficient used to control impact of the storage time constant for low flow (km)
MSK−X	Absolute	0.00	0.30	Weighting factor controlling relative importance of inflow rate and outflow rate in determining water storage in reach segment
CH−K1	Absolute	0.00	150.00	Effective hydraulic conductivity in tributary channel alluvium (mm/h)
CN2	Relative	−0.35	0.35	Initial SCS runoff curve number for moisture condition II
CH−N2	Relative	−0.50	0.50	Manning's n value for main channel
CH−K2	Absolute	0.01	150.00	Effective hydraulic conductivity in main channel alluvium (mm/h)
ALPHA−BF	Absolute	0.00	1.00	Baseflow alpha factor (days)
GWQMN	Absolute	0.00	5,000.00	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)
GW−REVAP	Absolute	0.02	0.20	Groundwater 'revap' (transfer of groundwater to upper soil layers) coefficient
GW−DELAY	Absolute	0.00	500.00	Groundwater delay (days)
RCHRG−DP	Absolute	0.00	1.00	Deep aquifer percolation fraction
REVAPMN	Absolute	0.00	500.00	Threshold depth of water in the shallow aquifer for 'revap' to occur (mm)
CANMX	Absolute	0.00	100.00	Maximum canopy storage (mm H2O)
ESCO	Absolute	0.00	1.00	Soil evaporation compensation factor
EPCO	Absolute	0.00	1.00	Plant uptake compensation factor
SOL−AWC	Relative	−0.50	0.50	Available water capacity of the soil layer (mm H2O/mm soil)
SOL−K	Relative	−0.80	0.80	Saturated hydraulic conductivity (mm/h)

The calibration process of most of these studies require a large number of model parameters. Large number of parameters means high dimensionality and likelihood of sub-optimal solutions, for the purpose of this study, however, NSGA-II has been found to be suitable for handling problems of this nature (Deb 2001; Deb and Goel 2001; Deb et al. 2000, 2002).

In the NSGA-II/SWAT framework, the program begins by using the upper and lower bounds for SWAT streamflow parameters to generate a solution space for each model parameter. The model parameter values were encoded such that each digit (including the whole number part and the decimal part) is represented separately as a gene. The calibration process is designed to minimize RMSE and model bias. Based on the initial population that is generated, each parameter set is supplied to SWAT where model parameter values are changed to update the model output. The simulated streamflow for the specified parameter set is used to compute its fitness (i.e., RMSE and bias). This procedure is repeated for all parameter sets in that population and non-dominated sorting is performed to sort parameter sets into different fronts. Selection and reproduction operations are performed on the population followed by recombination and mutation to create a new population of parameter sets for evaluation. The generation of the new population incorporates crowding operators by distributing solutions to less crowded areas in order to provide diversity among selected solutions. The process is repeated until the difference between the fitness values for the previous and current populations are minimal.

3 Results and Discussion

This section presents and discusses the NSGA-II/SWAT outputs that comprise several competing model parameter sets. First, we analyze model parameter clusters on the Pareto frontier which show varying trade-off information between model bias and RMSE to calibrating SWAT for streamflow. The massive parameter sets on the Pareto frontier have competing accuracies in predicting the observed streamflow given the model evaluation criteria. However, for decision making purposes it is desirable to select a single parameter set from the Pareto frontier based on the trade-off among the parameter sets. Second, using cluster analysis we select a mix of parameter values for decision making by evaluating the distribution of solutions in parameter space and the distribution of parameter sets in objective space. The resulting parameter sets emphasize the utility of the selection methods to extracting a mix of parameter values that improve streamflow prediction and reflects changing conditions in the watershed.

3.1 Parameter Sets from the NSGA-II Output

The NSGA-II framework was run for 400 generations with a population of 200 parameter sets which are evaluated using RMSE and model bias. A crossover probability of 0.8 and a mutation probability of $1/l$ (where l is the number of variables) as defined in the default configuration in Deb et al. (2002) were used. RMSE and model bias were chosen for this study because these objective functions have conflicting relationships (Gupta et al. 1998) and they measure different aspects of the calibration procedure. RMSE evaluates the proximity of the simulation to observations whereas model bias assesses the over-prediction or under-prediction of observed streamflow. The simultaneous evaluation of both objective functions generated a set of parameter sets that represent a trade-off between RMSE and model bias. The NSGA-II output comprises 115 parameter sets which constitute the Pareto frontier are shown in Fig. 2. Some parameter sets have similar values for RMSE and bias and their points

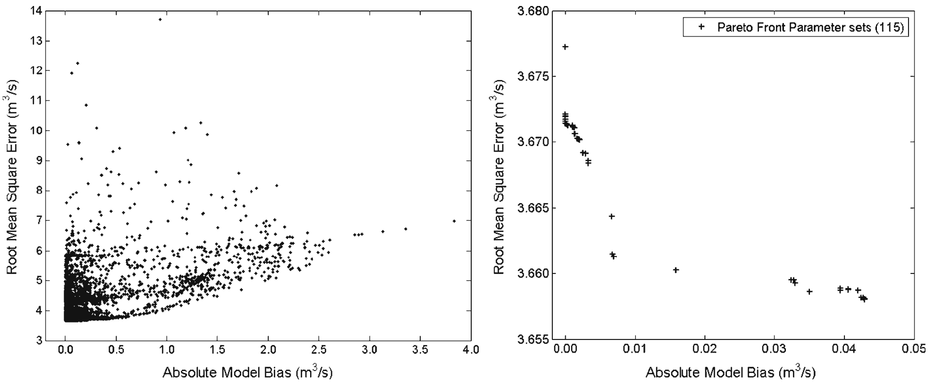


Fig. 2 A plot of solutions (*left-hand side* graph) explored to generate the Pareto frontier (*right-hand side* graph) for bias and RMSE comprising 115 parameter sets

almost overlap (at this scale for display purposes). Based on the 115 parameter sets on the Pareto frontier (Fig. 2), the absolute model bias ranges from 0.00003 to 0.01116 mm/day, and RMSE varies from 0.8160 to 0.8203 mm/day. The overall Pareto frontier shows a positive model bias indicating that several parameter sets have generally over-predicted the observed streamflow.

To examine the cluster of model parameters on the Pareto frontier, the parameter limits were re-scaled onto a unit scale between zero and one. This re-scaling does not change the distribution of the parameter values as the original parameter values can be determined by applying the inverse of the re-scaling factor. The re-scaling allows a graphical comparison between model parameters and shows the discontinuities as well as clusters within the parameter limits. Figure 3 shows the range of parameter values for the 115 scenarios on the Pareto frontier; each model parameter has 115 plotted values representing the Pareto frontier but some plotted values have almost converged (at this scale for display purposes) onto a point. On top of these 115 parameter sets are three plotted parameter sets representing the minimum bias, the minimum RMSE, and the compromise between bias and RMSE. The figure demonstrates that some model parameters have small spread between clusters while others have wide spread between their clusters, and that model parameter spaces are not continuous but exhibit distinct clusters.

An apparent justification for the varying parameter clusters is that some model parameters (e.g., TIMP, CN2, and ESCO) have almost converged onto a narrow cluster despite the trade-off between bias and RMSE. The narrow clusters for these model parameters can facilitate decision making when selecting optimal parameter values. Additionally, small changes in the values for these model parameters result into significant changes in watershed response. In contrast, model parameters with wide clusters suggest that they are resistant to converge to a narrow cluster and will usually require further evaluation. Additionally, the large distribution for some parameters with wide clusters could partly be due to their spatial variation within the watershed and the fact that SWAT modifies these parameters at watershed-wide. Selecting parameter values for model parameters with wide cluster spaces may be problematic but known conditions in the watershed can be used to guide the selection. Note that parameter-by-parameter clusters maybe useful but as will be

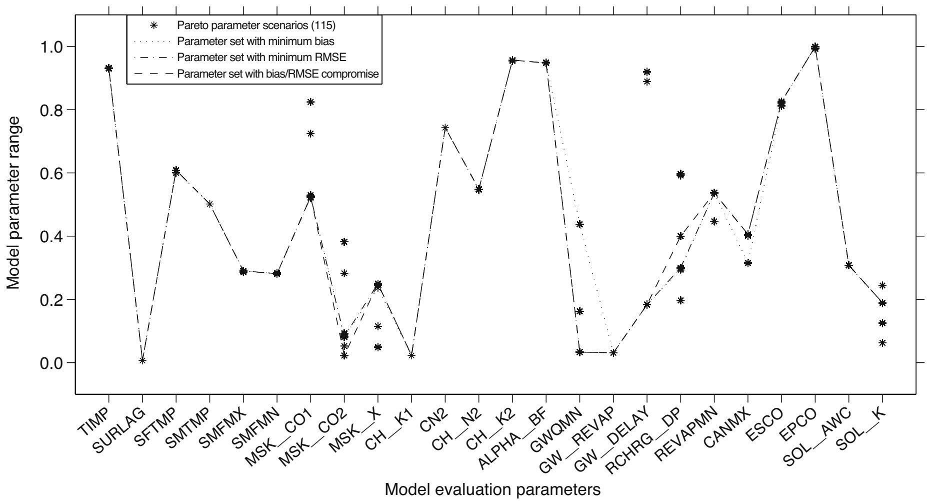


Fig. 3 Model parameters and their clusters on the Pareto frontier

shown in Section 3.2 evaluation of relationships between parameters using cluster analysis have a crucial impact on response at the watershed outlet.

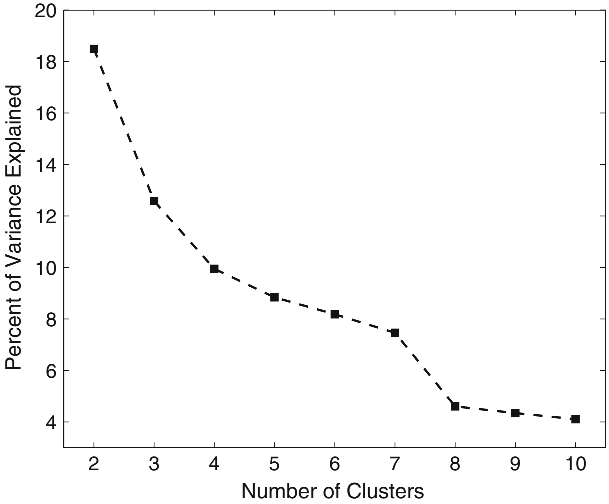
The resulting solutions on the Pareto frontier represent a problem-front in a way that a solution should be chosen from the front only when a problem with a known trade-off between the objectives has been specified. In our study, a specification of a water-resource decision relating RMSE and bias should precede the choice of solutions on the Pareto frontier. The choice of solutions can be a daunting task as the trade-off between the objectives is often unknown. In the discussion below, we examine the trade-offs between the parameter sets in order to choose a single mix of parameter values as would be required for decision making purposes. It is important to emphasize that the selection methods are purpose-dependent and should be applied when specific conditions such as clustering in parameter space or when a basin of attraction on the Pareto frontier are deemed to provide critical information about the watershed.

3.2 Selecting Model Parameter Sets on the Pareto Frontier

The following is an implementation of the automated framework described in Section 2.1 to select solutions with unique properties using the distribution of solutions in objective space and parameter space.

Using the CAP method, we determine the solution with the representative pathway in parameter space. The model parameter values were re-scaled (i.e. normalized) so that their values range between zero and unity. A k-means clustering method is applied to the 24 model parameters for 115 parameter sets on the Pareto frontier using Euclidean distance measure. The decision to use eight clusters was based on the clustering condition outlined in Thorndike (1953) where the percent of variance explained by different number of clusters k, and k = 8 provided the most suitable number of clusters as it falls at the knee (in Fig. 4). The eight clusters and their respective sizes are: PC_1 (has 14 solutions), PC_2 (has 33 solutions), PC_3 (has 20 solutions),

Fig. 4 Eight clusters identified in parameter space. Using the 'knee' technique to determining an appropriate number of clusters for parameter sets in parameter space



PC_4 (has 14 solutions), PC_5 (has 5 solutions), PC_6 (has 10 solutions), PC_7 (has 12 solutions) and PC_8 (has 7 solutions). Cluster PC_2 has the highest membership of 33 solutions and the CAP solution is chosen from this cluster of solutions. The parameter set at the cluster center for PC_2 is determined and its averaged streamflow is 3.8608 m^3/s, and the RMSE and model bias values are 0.8161 mm/day (uncertainty $= \pm 0.0014$) and 0.0091 mm/day (uncertainty $= \pm 0.00941$) respectively.

To determine the BAPF solution, Eq. 1 was applied to normalize the objective functions (RMSE and bias). Using clustering analysis on the Pareto frontier, six unique clusters are readily identified in the objective space (in Fig. 5). Again, k-means clustering was applied using the Euclidean distance measure. The six clusters and their cluster sizes are $\|FC_1\| = 60$, $\|FC_2\| = 7$, $\|FC_3\| = 18$, $\|FC_4\| = 13$, $\|FC_5\| = 10$ and $\|FC_6\| = 7$. The solution at the cluster center for FC_1 is determined and it's averaged streamflow is 3.8634 m^3/s, and the RMSE and model bias values are 0.8190 mm/day (uncertainty $= \pm 0.00083$) and 0.00004 mm/day (uncertainty $= \pm 0.00086$) respectively.

$$\mathrm{RMSE}_j = \frac{\mathrm{RMSE}_i - \mathrm{RMSE}_{\min}}{\mathrm{RMSE}_{\max} - \mathrm{RMSE}_{\min}}; \mathrm{Bias}_j = \frac{\mathrm{Bias}_i - \mathrm{Bias}_{\min}}{\mathrm{Bias}_{\max} - \mathrm{Bias}_{\min}} \tag{1}$$

The IPO solution is the parameter set closest to the origin of zero RMSE and zero bias. The IPO solution is a balanced compromise between the objectives as it equally reduces the objectives as aggressively as possible without damaging the response of the other objective(s). The chosen IPO solution has an averaged streamflow of 3.8376 m^3/s, and the RMSE and model bias values are 0.8167 mm/day and 0.0018 mm/day respectively.

A BCOP solution is determined by finding the parameter set at the center of the set of solutions in the overlapped region between PC_2 and FC_1. The overlapped region between PC_2 and FC_1 has four solutions. The solution at the center of these four solutions has an averaged streamflow of 3.8666 m^3/s, and the RMSE and model bias values with uncertainty are 0.8187 mm/day (uncertainty $= \pm 0.0022$ and 0.0004 mm/day (uncertainty $= \pm 0.0078$) respectively.

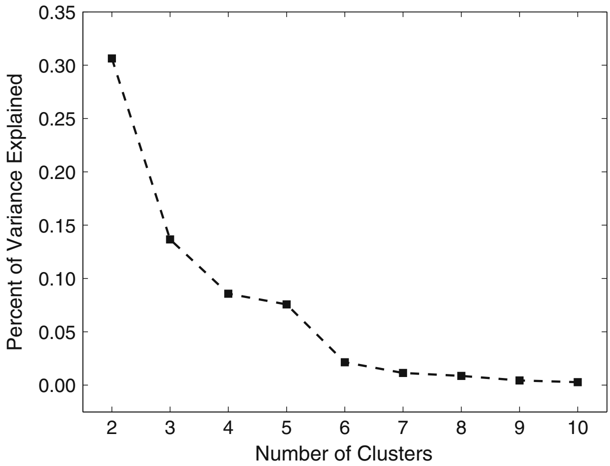


Fig. 5 Six solution clusters identified in objective space. Using the 'knee' technique to determining a suitable number of clusters in objective space

3.3 Comparison Between Selection Methods for Choosing Solutions on the Pareto Frontier

The CAP, BAPF, IPO and BCOP parameter sets have similar values for averaged streamflow, model parameter values and their associated RMSE and bias values when using the FCr watershed data set. Nevertheless, the four approaches are different and they have their own advantages on how they arrive at selecting a single mix of model parameter values to facilitate decision making. The CAP, BAPF, IPO and BCOP methods evaluate four different types of trade-offs between parameter sets on the Pareto frontier.

A comparison between model evaluation results for the chosen solutions are shown in Table 3 for model calibration and validation results. For these solutions, model bias has increased to about one order of magnitude and is consistently positively biased during the validation period while RMSE values have remained in the same order of magnitude as in the calibration period. There are more extreme streamflow events in the calibration period than there are in validation period.

Table 3 Model calibration and validation results for daily streamflow for CAP, BAPF, IPO and BCOP auto-selection methods

Evaluation parameter	Auto-selection methods				$RMSE_{best}$	$BIAS_{best}$
	CAP	BAPF	IPO	BCOP		
Calibration period: 1992–1997						
Model bias (mm/day)	0.00910	0.00004	0.00180	0.00040	0.01120	0.00003
RMSE (mm/day)	0.8161	0.8190	0.8167	0.8187	0.8160	0.8191
Nash–Sutcliffe efficiency	0.6059	0.6088	0.6062	0.6082	0.6089	0.6060
Validation period: 1998–2003						
Model bias (mm/day)	0.0816	0.0826	0.0780	0.0823	0.0836	0.0820
RMSE (mm/day)	0.6864	0.6794	0.6865	0.6795	0.6792	0.6866
Nash–Sutcliffe efficiency	0.5180	0.5080	0.5286	0.5092	0.5186	0.5078

Validation results are determined by using parameter sets obtained from the calibration period: 1992–1997

This condition influences model evaluation measures, for example, the validation period shows an improved RMSE values compared to the calibration period. For illustration purposes, Fig. 6 is used to illustrate differences/similarities between generated streamflows for CAP, BAPF, IPO and BCOP methods.

The CAP method is distinct because it selects a parameter set which describes the persistent/dominant cluster in the parameter space. The parameter values in the CAP parameter set represent the maximum likelihood values given the distribution of the parameter values on the Pareto frontier. The CAP method is considered a 'back end' approach because model parameter values are explored to select a persistent parameter cluster that gives a competitive streamflow prediction. Since the CAP selection is based on persistent model parameter cluster, model parameter values can be easily verified on the basis of how closely they reflect the watershed condition under investigation. The computed model parameter spaces are useful to specifying parameter values during future calibration operations as well as to further investigate individual model parameter values for accurate watershed description and characterization, beyond the evaluation of watershed responses at the outlet.

Similarly, RMSE and bias values in the BAPF parameter set represent maximum likelihood values given the distribution of RMSE and bias values on the Pareto frontier. The BAPF parameter set is equivalent to the center of mass representing the most commonly found compromise solution with respect to the concentration of solutions, and minimum and maximum values of RMSE and bias. The IPO parameter set is also distinct because it represents the most balanced compromise solution with the same magnitude of RMSE and bias, and is closest to the origin.

The BAPF and IPO methods are useful, particularly in streamflow prediction scenarios that are usually focused on watershed responses at the outlet. That is, the BAPF and IPO methods narrow-mindedly focus on predictability of watershed response but ignores the distribution of parameter values. The BAPF and IPO methods are considered a 'front end' approach as the parameter set selections are based on a compromise between RMSE and bias for predicted streamflows of

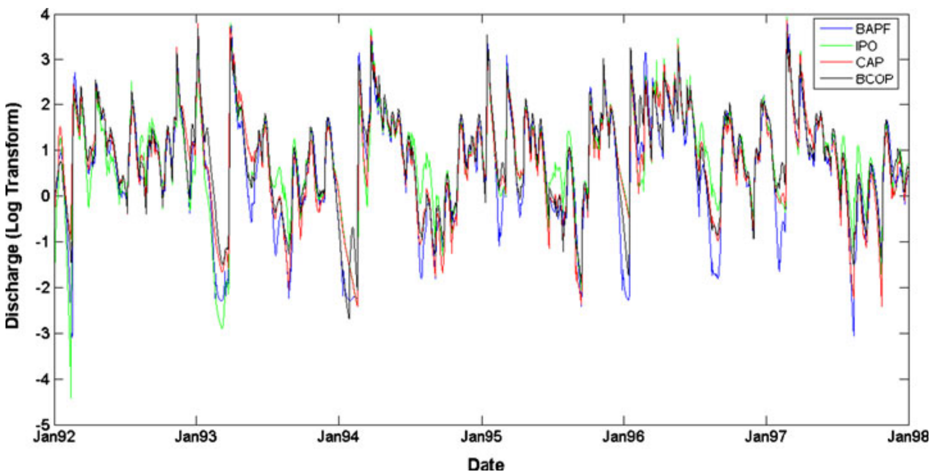


Fig. 6 Log transform of discharge for generated streamflows for CAP, BAPF, IPO and BCOP methods

several parameter sets that do not necessarily have similar parameter pathways. Since the BAPF and IPO methods evaluate the compromise between RMSE and bias to select a solution, the methods are useful to adjusting model estimates if the model persistently has predictions very far from observations. For example, in data assimilation operations the predicted streamflow from BAPF and IPO methods can be compared to the observed streamflow in order to penalize parameter sets to partially align the model to the observed streamflow.

The CAP, BAPF and IPO solutions are compromise solutions in either parameter space or objective space but these solutions do not capture a compromise for the distribution of solutions for both spaces. The BCOP solution is an equally-weighted compromise for the distribution of solutions in both objective space and parameter space. The BCOP solution captures the representative pathway in parameter space and the dominant variability in objective space; hence the desired solution to be determined.

4 Summary and Conclusion

This study has demonstrated the selection of solutions from an incomparable set of parameter sets that are generated from calibrating SWAT for simulations of streamflow using NSGA-II. The NSGA-II/SWAT integration results include trade-off information between model bias and RMSE and their corresponding competing solutions which make up the Pareto frontier. Analysis of solutions in parameter space demonstrates the stability of parameter clusters for some model parameters in achieving competing accuracies for parameter sets. Additionally, the trade-off information and the numerous parameter sets on the Pareto frontier emphasize the relevance of evaluating the Pareto frontier to select a parameter set for decision making. Evaluation of solutions on the Pareto frontier is important because the Pareto frontier is a trade-off surface where no one solution is better than the other unless a specific problem is able to quantify the relevance of one objective in relation to the other. However, solutions with unique properties on the Pareto frontier can be derived based on the distribution of solutions in objective space and parameter space.

Using cluster analysis to evaluate the distribution of solutions in objective space and parameter space, we have demonstrated four auto-selection methods: CAP, BAPF, IPO and BCOP—to choose solutions with unique properties on the Pareto frontier. The BAPF and IPO methods focus on the distribution of solutions in objective space. These two methods of evaluating the trade-off between bias and RMSE uses a ‘front end’ approach by finding a streamflow with a moderate RMSE and bias from several competing streamflow predictions. The BAPF and IPO approaches are useful in data assimilation contexts where streamflow predictions from the model can be used to align the model to better predict the observed streamflow for the future.

The CAP method examines the distribution of solutions in parameter space. In contrast to BAPF and IPO, the CAP method is tagged a ‘back end’ approach as it evaluates the parameter values themselves other than the predicted streamflows based on RMSE and bias to select a single parameter set. The CAP method uses several competing parameter sets and computes for each model parameter a persistent parameter cluster. An advantage of the CAP method is that it could be most suited for watershed characterization as model parameter values are thoroughly evaluated on their basis of representing known conditions in the watershed. Furthermore, the

CAP solutions describes a representative pathway in parameter space. The computed model parameter cluster for the CAP method can provide information for specifying model parameter values for future model parameter estimation operations.

The fourth auto-selection method, BCOP, determines a robust-type solution by finding a compromise for the distribution of solutions in objective space and parameter space. That is, the BCOP is a compromise between the representative pathway in parameter space (CAP) and the dominant variability in objective space (BAPF). The auto-selection method for BCOP solution emphasizes stability of model parameter values and objective function values in a way that similarity in parameter space implies similarity in objective space.

In sum, the key finding shown in the generation and analysis of the trade-off surface is the automated technique of selecting solutions with unique properties from a Pareto-optimal front. Following the generation of the trade-off surface between RMSE and bias we applied cluster analysis to develop four auto-selection methods to find parameter sets with unique properties on the trade-off surface. These properties on the trade-off surface include a representative pathway in parameter space (using CAP), a basin of attraction or the center of mass in objective space (using BAPF) and a proximity to the origin in objective space (using IPO). Finally, a BCOP solution which captures the properties of both CAP and BAPF is determined as it represents a compromise for the distribution of solution in objective space and parameter space. It is worth emphasizing that the solutions chosen by the auto-selection methods are in no way better than the remaining solutions on the Pareto frontier; rather the chosen solutions only possess unique properties in objective space and parameter space.

Acknowledgements This work is supported by NSERC (Natural Sciences and Engineering Research Council of Canada). We are thankful to Drs Wanhong Yang and Yongbo Liu, and Adam Bonnycastle for providing data and suggestions on model setup. We thank the anonymous reviewers for their comments and efforts.

References

Agrawal N, Rangaiah GP, Ray AK, Gupta SK (2006) Multi-objective optimization of the operation of an industrial low-density polyethylene tubular reactor using genetic algorithm and its jumping gene adaptations. J Ind Eng Chem Res 45(9):3128–3199

Ahsan M, O'Connor KM (1994) A simple non-linear rainfall-runoff model with a variable gain factor. J Hydrol 155:151–183

Arnold JG, Fohrer N (2005) Swat2000: current capabilities and research opportunities in applied watershed modelling. Hydrol Process 19:563–572

Arnold JG, Srinivasan R, Muttiah RS, Williams JR (1998) Large area hydrologic modeling and assessment part 1: model development. J Am Water Resour Assoc 34:73–89

Atiquzzaman M, Liong SY, Yu W (2006) Alternative decision making in water distribution network with nsga-ii. J Water Resour Plan Manage 132(2):122–126

Bekele EG, Nicklow JW (2007) Multi-objective automatic calibration of swat using nsga-ii. J Hydrol 341:165–176

Beven K (2001) Rainfall runoff modelling the primer. Wiley, Chichester, West Sussex, England

Bonnycastle A (2006) Environmental and economic implications of land management changes in agricultural watersheds. Thesis, Geography, University of Guelph, 50 Stone Road East, Guelph, Ontario, N1G 2W1 Canada

Borah DK, Bera M (2003a) Swat model background and application reviews. In: 2003 ASAE annual international meeting, ASAE: the society for engineering in agricultural, food, and biological systems. Riviera Hotel and Convention Center, Las Vegas, USA

Borah DK, Bera M (2003b) Watershed-scale hydrologic and nonpoint-source pollution models: review of mathematical bases. Trans ASAE 46(6):1553–1566

Coello Coello CA, Van Veldhuizen DA, Lamont GB (2002) Evolutionary algorithms for solving multi-objective problems. Kluwer/Plenum Publishers, New York, NY
Confesor RB, Whittaker GW (2007) Automatic calibration of hydrologic models with multi-objective evolutionary algorithm and pareto optimization. J Am Water Resour Assoc 43(4): 981–989
Crispim JA, de Sousa JP (2009) Partner selection in virtual enterprises: a multi-criteria decision support approach. Int J Prod Res 47(17):4791–4812
Deb K (2001) Multi-objective optimization using evolutionary algorithms. Wiley, Chichester, New York
Deb K, Goel T (2001) Controlled elitist non-dominated sorting genetic algorithms for better convergence. Swiss Federal Institute of Technology, Zurich Switzerland. Evolutionary multi-criterion optimization. Lecture notes in computer science, vol 1993/2001. Springer, pp 67–81
Deb K, Agrawal S, Pratap A, Meyarivan T (2000) A fast elitist non dominated sorting genetic algorithms for multi-objective optimization: Nsga-ii. Paris, France. Parallel Problem Solving from Nature VI (PPSN-VI). Lecture notes in computer science, vol 1917. Springer, pp 849–858
Deb K, Pratap A, Agrawal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: Nsga-ii. IEEE Trans Evol Comput 6(2):182–197
Duan QY, Gupta HV, Sorooshian S (1993) Shuffled complex evolution approach for effective and efficient global minimization. J Optim Theory Appl 76:501–521
Eiben AE, Smith JE (2003) Introduction to evolutionary computing. Springer, Berlin Heidelberg New York
Environment-Canada (2006) National climate data and information archive. Technical report, Environment Canada
Ferreira JC, Fonseca CM, Gaspar-Cunha A (2007) Methodology to select solutions from the pareto-optimal set: a comparative study. In: Genetic and evolutionary computation conference. Proceedings of the 9th annual conference on genetic and evolutionary computation. London, England, United Kingdom, GECCO, pp 789–796
Gill MK, Kaheil YH, Khalil A, McKee M, Bastidas L (2006) Multiobjective particle swarm optimization for parameter estimation in hydrology. Water Resour Res 42(7):W07417
Green WH, Ampt GA (1911) Studies on soil physics, 1: the flow of air and water through soils. J Agric Sci 4:1–24
Grierson DE (2008) Pareto multi-criteria decision making. Adv Eng Inf 22(3):371–384
Gupta H, Sorooshian S, Yapo PO (1998) Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. Water Resour Res 34:751–763
Hee-Su K, Sung-Bae C (2000) Application of interactive genetic algorithm to fashion design. Eng Appl Artif Intell 13(6):635–644
Jeong HS, Abraham DM (2006) Operational response model for physically attacked water networks using nsga-ii. J Comput Civ Eng 20(5):328–338
Khu ST, Madsen H (2005) Multiobjective calibration with pareto preference ordering: an application to rainfall-runoff model calibration. Water Resour Res 41(3):W03004
Leavesley GH (1994) Modeling the effects of climate change on water resources—a review. Clim Change 28:159–177
Liang GC, O'Connor KM, Kachroo RK (1994) A mulitiple-input, single-output, variable gain-factor model. J Hydrol 155:185–198
Liu Y, Yang W (2007) An interface of drainage division for modeling wetlands and riparian buffers in agricultural watersheds. J Spat Hydro 7(1):66–80
Madsen H (2003) Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives. Adv Water Resources 26(2):205–216
Marler RT, Arora JS (2004) Survey of multi-objective optimization methods for engineering. Struct Multidiscipl Optim 26(6):369–395
Miyamoto T, Kaneda K, Noguchi S, Yamashita H (2006) A technique for selecting an optimal solution from among pareto-optima of multi-purposed electromagnetic apparatus design based on game theory. In: Electromagnetic field computation, 12th biennial IEEE conference, p 121
Nazemi A, Yao X, Chan A (2006) Extracting a set of robust pareto-optimal parameters for hydrologic models using nsga-ii and scem. In: 2006 IEEE congress on evolutionary computation, pp 1901–1908. doi:10.1109/CEC.2006.1688539
Neitsch SL, Arnold JG, Kiniry JR, Williams JR (2001) Soil and water assessment tool: theoretical documentation, version 2000. Soil and Water Research Service, Temple, Texas

Rakesh K, Chandan G (2005) Multi-objective optimization of a batch copoly (ethylene-polyoxyethylene terephthalate) reactor using different adaptations of nondominated sorting genetic algorithm. Macromol Theory Simul 14(6):358–373

Sarkar D, Modak YM (2006) Optimal design of multiproduct batch chemical plant using nsga-ii. Asia-Pac J Chem Eng 1(1–2):13–20

Shafii M, Smedt FD (2009) Multi-objective calibration of a distributed hydrological model (wetspa) using a genetic algorithm. Hydrol Earth Syst Sci Discussions 6(1):243–271

Taboada H, Coit D (2006) Data mining techniques to facilitate the analysis of the pareto-optimal set for multiple objective problems. In: Proceedings of the industrial engineering research conference, IERC 2006. IERC, Orlando, Florida

Tabrizi MHN, Said SE, Badr AW, Mashor Y, Billings SA (1998) Nonlinear modelling and prediction of a river flow system. J Am Water Resour Assoc 34:1333–1339

Tang Y, Reed P, Wagener T (2006) How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration? Hydrol Earth Syst Sci 10:289–307

Tappeta RV, Renaud JE, Messac A, Sundararaj GJ (2000) Interactive physical programming: tradeoff analysis and decision making in multidisciplinary optimization. AIAA J 38(5):917–926

Thorndike RL (1953) Who belong in the family? Psychometrika 18(4):267–276

Tolson BA, Shoemaker CA (2007) Cannonsville reservoir watershed swat2000 model development, calibration and validation. J Hydrol 337(1–2):68–86

Vamvakeridou-Lyroudia LS, Savic DA, Walters GA (2006) Fuzzy hierarchical decision support system for water distribution network optimization. Civ Eng Environ Syst 23(3):237–261

Weinert K, Zabel A, Kersting P, Michelitsch T, Wagner T (2009) On the use of problem-specific candidate generators for the hybrid optimization of multi-objective production engineering problems. Evol Comput 17(4):527–544

Wöhling T, Vrugt JA, Barkle GF (2008) Comparison of three multiobjective optimization algorithms for inverse modeling of vadose zone hydraulic properties. Soil Sci Soc Am J 72(2):305–319

Xuebin L (2009) Study of multi-objective optimization and multi-attribute decision-making for dynamic economic emission dispatch. Electr Power Compon Syst 37:1133–1148

Yapo PO, Gupta HV, Sorooshian S (1998) Multi-objective global optimization for hydrologic models. J Hydrol 204:83–97

Young P, Beven KJ (1994) Data-based mechanistic modelling and the rainfall-flow nonlinearity. Environmetrics 5(3):335–363