15

Multi-Objective Calibration of SWMM for Improved Simulation of the Hydrologic Regime

Mauricio Herrera, Isobel Heathcote, William James and
Andrea Bradford

This chapter presents a multi-objective calibration of the Storm Water
Management Model (SWMM) using the Non-dominated Sorting Genetic
Algorithm (NSGA-II) developed by Deb et al. (2001). The effects of model
calibration on the representation of various hydrologic characteristics with
ecologic and geomorphic relevance are studied. Results indicate that there
are modeling conflicts between low flows, medium to bank-full flows, and
high flows. As a consequence, calibration improvements of minimum water
quality maintenance flows, decreases the agreement between computed and
observed flows above bankfull elevations. The presence of these trade-offs
should be acknowledged in model-based watershed management strategies
in order to minimize the uncertainty bias towards certain characteristics of
the flow regime. Results show the effects of such trade-offs in the model
accuracy to represent different hydrologic quantities, such as mean monthly
flows, peak monthly flow, minimum monthly flows, and flow durations and
exceedance volumes for different flow ranges and months.

15.1 Introduction

To a great extent, the calibration of hydrologic models defines the
effectiveness of model-based watershed management strategies. Within any

model structure, the selection of a set of parameters that most accurately describe the variation of an observed characteristic of the flow regime, for instance high flows or minimum daily flows; is known as model calibration. Ideally, a calibrated model would perfectly match the field measurements. However, this condition is hard to achieve due to data input errors, limitations in the model structure, statistical errors associated with spatial heterogeneity of the watershed, model discretization, and temporal aggregation errors. Thus, the objective of any calibration procedure is to minimize the difference, as much as possible, between computed and observed results. Yet, there is not a unique way to define this difference. Most manual calibration processes rely on a visual comparison of the two values to assess the success of the calibration. In addition, numerous evaluation functions are available can be also used to quantify the model error.

Regardless of the comparison approach adopted, model calibration often requires the evaluation of numerous parameter combinations before a reasonable agreement is reached. Hreiche et al. (2002) found that 1.5 million simulations were required for the four-parameter model MEDOR to systematically explore the complete response surface of the model. Assuming that each simulation requires one second to complete, the total running time would be over 17 days. Evidently, a calibration exercise would not require this many evaluations; however, this example illustrates the complexity of the search space even for a simple model.

Automatic calibration has received a great deal of attention over the last three decades as an efficient method to narrow down the number of parameter combinations (Johnston and Pilgrim, 1976; Pickup, 1977; Sorooshian et al., 1983; Duan et al., 1992; Beven and Binley, 1992; Yapo et al., 1996; Gupta et al, 2003). Perhaps one of the most commonly used approaches is to apply systems optimization theory concepts to the problem of model calibration. Typically, an optimization problem is expressed as the minimization or maximization of a function or group of functions, subject to certain restrictions that delimit the space of feasible alternatives. In addition, both objective functions and constraints depend on the decision variables, which are also constrained to a certain domain. Mathematically, an optimization problem can be defined as follows:

Minimize/Maximize $h_m(\boldsymbol{s})$, $m = 1,2 \ldots M$;

Subject to $f_j(\boldsymbol{s}) \geq 0$, $j = 1, 2, \ldots J$;

$$Q_w(s) = 0, \qquad\qquad w = 1, 2, \ldots W;$$

$$s_i^{(L)} \le s_i \le s_i^{(U)}, \qquad\qquad i = 1, 2, \ldots n;$$

where:

h_m = m^{th} objective function,
M = maximum number of objective functions,
s: = array of n decision variables, $s = (s_1, s_2 \ldots s_n)$,
$s_i^{(L)}$ = low boundary of decision variable s_i,
$s_i^{(U)}$: = upper boundary of decision variable s_i
$f_j(s)$ and $Q_w(s)$ = different constraint functions,
J and W = maximum number for each set of constraints.

In the context of hydrologic model calibration, the objective function h_m represents a mathematical formulation of the difference between observed and simulated values applied to a hydrologic characteristic (i.e. peak flow). Decision variables (*s*) are the set of parameters of the hydrologic model that cannot be defined a priori and are therefore subject of calibration.

Recent advances in automatic calibration indicate that the use of evolutionary multi-objective optimization (EMOA), based on the Pareto optimality or non-dominance concept, notably improves the calibration of hydrologic models (Yapo, 1996; Yapo et al., 1996; Gupta et al., 1998; Madsen, 2000; Gupta et al., 2003). Gupta et al. (2003) argues that the main advantage of multi-objective calibration is that it emulates the expert-driven process in which the several performance measures are evaluated before a calibration is recommended.

The concept of Pareto or non-dominance states that given a population of solutions, under the presence of multiple conflicting objectives, there is a sub-group of solutions, often called the Pareto front or the non-dominated set, that is better than the rest of the population when all objectives are considered. To understand the Pareto optimality concept, consider a minimization problem with two objectives *f(i)* and *g(i)* with a solution array *U(u)*. A subset of solutions *V(v)*∈ *U(u)* is said to be non-dominated, non-inferior or the Pareto optimal set if ∀ *u*∈ {*U*≠*V*} there is at least one solution *v* ∈ *V* for which the following condition is true:

$$f(v) < f(u) \text{ and } g(v) < g(u) \qquad\qquad (15.1)$$

Also, any two solutions $\{a, b\} \in V$ are non-dominated each is better than the other in at least one objective:

$$f(a) < f(b) \text{ and } g(b) < g(a) \tag{15.2}$$

The concept of Pareto dominance is also illustrated in Figure 15.1 using a hypothetical calibration problem with two objectives: minimization of high and low flow model errors. The Pareto front is composed by solutions {**a,b,c,d,e,f,g,h**} while {**i,j,k,l,m,n,o,p**} are dominated solutions. Note that, although **k < f** in terms of low flows, **e < k** in both objectives, so **k** is a dominated solution. When comparing solution **e** with solution **f** we note that they are non-dominated as neither one is better than the other considering both objectives.
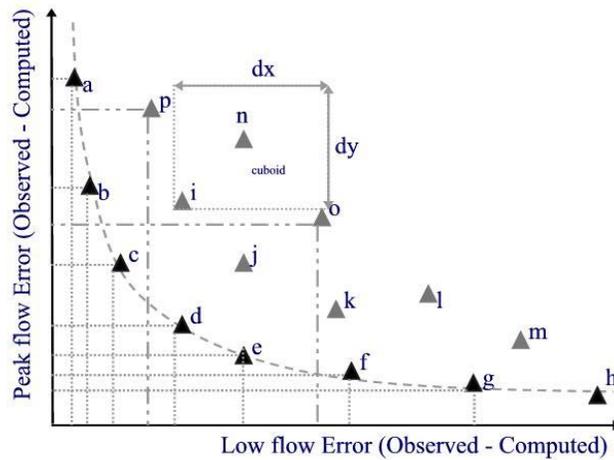

Figure 15.1 Set of Pareto optimal group of solutions for a two-objective hypothetical calibration.

The set of non-dominated or Pareto set of solutions can be interpreted as the minimum uncertainty in the parameters that can be reached, without defining a subjective preference for minimizing one specific objective at the expense of another (Vrugt et al., 2003). As such, the two main objectives of an EMOA are:

    1. to find a Pareto front as close as possible to the 'true' Pareto curve; and

2. to maintain a diverse set of solutions in the non-dominated front in order to include as much variety as possible for each optimized objective.

Like genetic algorithms, EMOAs create a population of solution candidates, which is subjected to a selection process and then is manipulated by genetic operators, usually crossover and mutation (Deb, 2000). However, unlike single objective optimization in which the fitness function is almost the same as the objective function, in a multi-objective optimization the fitness function depends on some other factors such as a global measure of the values of each objective function and also a measure of the diversity of the population.

Madsen (2003) identified three different objective groups for the calibration of hydrologic models:

1. multi-variable measurements (i.e. groundwater levels, stream flow);
2. multi-site measurements (i.e. groundwater levels at different locations in the catchment); and
3. multi-response modes (i.e. different characteristics of the hydrologic regime such as peak and low flows).

Multi-objective calibration provides further insights into the model's ability to reproduce various characteristics of the simulated system. Calibrating the NWSRFS-SMA conceptual rainfall-runoff flood-forecasting model for the US National Weather Service, Yapo et al. (1996) found that low or high flows were better simulated depending on the objective function selected. Also, Madsen (2003) using a multi-objective approach to calibrate the MIKE SHE model, showed significant trade-offs between the performances of the groundwater level simulations and the catchment runoff.

These findings have important implications in modeling applications that need to reproduce two or more characteristics of the hydrologic regime. Often, watershed managers have to deal with various water demands that in many cases are in conflict. One such example is presented in Figure 15.2 that illustrates some of the possible water demands in a stream channel. If a model is, say, very good simulating flow values around bankfull conditions, but presents serious deficiencies simulating longitudinal connectivity maintenance flows (this is the minimum flow to maintain connection of pools along the channel and perhaps also enough water to maintain fish migration depths over the riffles), or water quality maintenance flows; it is possible that management decisions that rely on the model representation of those two targets, might be biased and possibly misdirected.
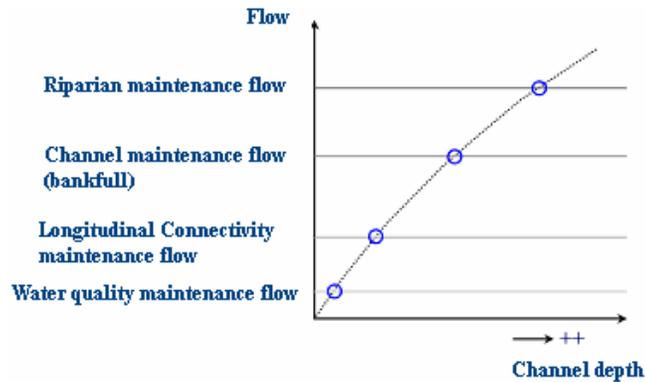
Figure 15.2 Typical in-stream water flow demands.

This situation is further complicated by the complex interrelationships between different characteristics of the hydrologic regime and ecologic and morphologic attributes. In fact, the availability and distribution of energy, particulate matter, dissolved substances and organisms highly depend on the natural variation of the hydrologic regime (Richter et al., 1997). Different authors (Poff and Ward, 1989; Richter et al., 1996; Richter et al., 1997; Roesner and Bledsoe, 2001; Bledsoe, 2001; Instream Flow Council, 2002) agree that annual and intra-annual variations of hydrological conditions are essential to support natural changes of the physical habitat. There are numerous streamflow characteristics that can be associated with riverine ecosystems, biological diversity and geomorphic dynamics.

Richter et al. (1997) argued that streamflow data are excellent indicators for assessing river ecosystems because: (i) many other abiotic characteristics vary with streamflow conditions, such as dissolved oxygen and temperature; (ii) channel morphology is shaped by fluvial processes driven by streamflow; and (iii) the availability of long-term daily records of streamflow provides insights into the history of human perturbations of a watershed. Poff et al (1997) identified five critical components of the flow regime that are of utmost relevance to sustain ecological processes in stream ecosystems: magnitude, frequency, duration, timing and rate of change of flows. Table 15.1 provides examples of hydrological characteristics with physical habitat processes. Examples of hydrograph characteristics of these five components are shown in Table 15.2.

Table 15.1 Example of relations between hydrological characteristics with physical habitat (Richter et al., 1997; Roesner and Bledsoe, 2001).

Flow Characteristics	Availability and effectiveness of habitat refugia	Life cycle requirements	Removal thresholds for substrate, banks, vegetation	Abundance and composition of macroinvertebrates	Recovery of equilibrium conditions	Re-suspension of fines during storm hydrographs	Flood plain Inundation	Erosion	Ability of plants to maintain contact with phreatic water supplies
Magnitude	X			X			X	X	
Frequency		X	X		X	X		X	
Duration		X						X	
Timing	X	X	X						
Rate of change									X

Table 15.2 Examples of flow characteristics relevant for ecological, processes (Richter et al., 1997).

Flow Categories	Mean monthly flow	Annual minima 1-day means	Annual Maximum 7-day means	Julian date of each annual maximum	Number of high pulses each year	Means of all positive differences between daily consecutive daily means
Magnitude	X	X	X			
Frequency					X	X
Duration		X	X		X	
Timing	X			X		
Rate of change						X

Konrad and Booth (2002) selected four hydrological parameters, namely annual mean discharge, annual maximum discharge, 7-day low flow, and the

fraction of the year the mean annual flow was exceeded (T_{QM}) to analyze the effects of urban development in the Puget Basin (western Washington). While the first three parameters are 'magnitude' values, the last one relates to the 'duration' of flow values above the mean annual. Each one of these parameters has associated influences in different ecological processes. For instance, an increase in T_{QM} is likely to represent changes in streambed and aquatic habitat (Konrad and Booth, 2002).

Calculation of these hydrologic values heavily relies on the availability of long-term flow registers. However, generation of data to complement limited measurements, and the prediction of future watershed conditions, for management purposes, highly depends on hydrologic models. Clearly, a biased 'calibrated' model would likely underestimate or overestimate certain hydrologic characteristics.

In this chapter a multi-objective calibration tool is built, linking the NSGA-II (Deb et al., 2000) and the Stormwater Management Model (SWMM) to evaluate and analyze the presence of modeling trade-offs between different hydrologic response modes, relevant for ecologic and morphologic processes. Also, the consequences of selecting a particular model from the Pareto set are evaluated in terms of the model error applied to different hydrologic characteristics.

15.2 Methodology

15.2.1 Non-dominated Sorting Genetic Algorithm (NSGA-II)

Proposed by Deb et al (2000) NSGA-II is an evolutionary multi-objective algorithm with complexity $O(MN^2)$, where M is the number of objectives and N is the population size. It includes both elitism and a crowding procedure for the preservation of diversity in the solutions. The elitism incorporated in NSGA-II, keeps the best solutions throughout the generations to avoid the loss of optimal solutions found prematurely. First, a random population is created, which is subsequently evolved, using genetic operators, into an offspring population. Both parent and children populations are combined and sorted using a fast non-dominated ranking procedure of complexity $O(2N log(2N))$. Then a new empty population is filled using the non-dominated fronts one at a time. Filling starts with the best fronts until the new population of size N is filled. If the size of the last front is larger than the available space, a crowding procedure is applied to select the

solutions located on a less crowded space. NSGA-II introduced a crowding distance concept that is calculated based on the average length of the cuboid formed by adjacent solutions in the objective space (i.e. the sum of the horizontal and vertical distances between any solution and its neighbors in Figure 15.1). This new population is consequently evolved and the process is repeated until the stopping criterion is achieved. The general NSGA-II algorithm is outlined below:

1. random generation of a population *P* of size *N;*
2. non-dominated sorting of *P* and fitness assignment based on the ranking of the individual;
3. binary tournament selection, recombination and mutation operators to create offspring population *Q* of size *N;*
4. combination of *P* and *Q*;
5. perform non-dominated ranking of *P* ∪ *Q* (size 2*N*);
6. fill new population *P'* successively with the best-ranked elements of *P* ∪ *Q*, until size *N* is reached;
7. if *P'* cannot accommodate all individuals from a given rank, perform crowding-distance assignment;
8. individuals with higher distance or density are incorporated into *P'* until size *N* is reached; and
9. create children *Q'* population and go to 5 until maximum number of generations or stopping criterion is reached.

One of the main advantages of NSGA-II is that it does not include extra algorithmic parameters to account for population diversity. This factor is particularly appealing to the automatic calibration process as it reduces its complexity to the specification only of the hydrological parameters.

15.2.2 Stormwater Management Model (SWMM)

SWMM is a comprehensive watershed model originally developed to simulate hydrologic, hydraulic and water quality processes for separate and combined sewer systems (Huber and Dickinson, 1992). Its capabilities, however, enable it to simulate complex watershed settings including rural areas. SWMM simulates rainfall-runoff generation processes, including infiltration, depression storage and snowmelt; water quality generation, including pollutant build-up and washoff processes; groundwater routing and contributions to the flow in the conduits, and, flow routing through a

wide variety of conveyances. SWMM is composed of various modules that are selected according to the user's modeling needs. The RUNOFF module performs flow and pollutant generation routines as describe above, and also, it has some limited flow routing capabilities. The TRANSPORT and EXTRAN modules perform hydraulic routing routines with different levels of complexity; the former uses the kinematic wave routing approach, while the latter solves the complete Saint-Venant equations.

For calibration purposes, the parameters in the RUNOFF module are the only that are modified during the calibration. For this reason, only these module and the relevant processes are described. A general diagram of the hydrologic processes included in the RUNOFF module is depicted in Figure 15.3.
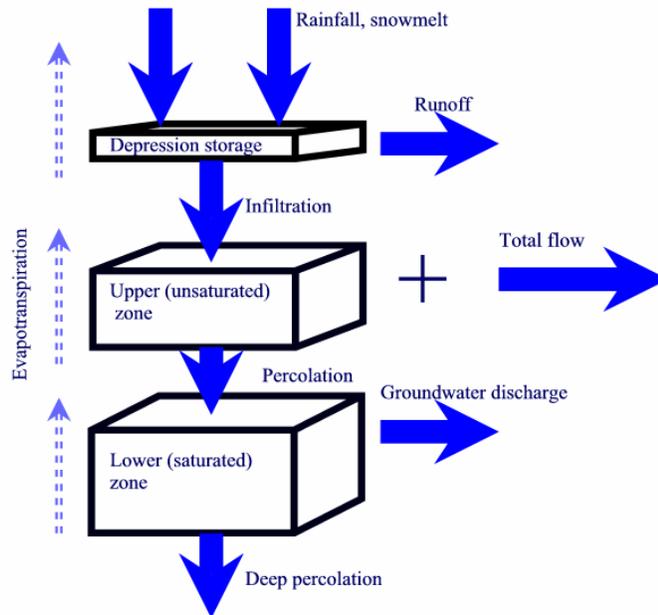

Figure 15.3. Diagram of hydrologic processes in SWMM

RUNOFF reads the meteorological information, such as rainfall, snowfall, temperature, and winds to produce the corresponding runoff. After the depression storage is filled up, the model converts the rainfall excess into surface runoff using the Continuity and Manning equations as shown in Equation 15.3.

$$\frac{dV}{dt} = A\frac{dd}{dt} = A \cdot i * -Q \qquad (15.3)$$

where::

A = sub-catchment area
V = $A \cdot d$ = volume of water on the sub-catchment (mm)
d = water depth, (mm)
t = time, (sec)
A = catchment area (m²)
Q = outflow rate (mm/sec)
I = precipitation (mm/sec)

$$Q = W \cdot \frac{1.49}{n}\left(d - d_p\right)^{5/3} S^{1/2} \qquad (15.4)$$

where:

W = catchment's width (m)
n = Manning's number for catchment
d_p = depression storage (mm)

For pervious areas the program calculates the amount of water infiltration using one of three alternatives: 1) the Horton model, 2) the Green-Ampt equation or 3) the SCS method. For the present research the Green-Ampt method was selected. Equation 15.5 represents the volume of water that infiltrates before the surface is saturated and Equation 15.6 predicts infiltration from this point on.

$$F_s = \frac{SUCT \cdot IMD}{\dfrac{i}{HYDCON} - 1} \qquad (15.5)$$

$$f_p = HYDCON\left(1 + \frac{SUCT \cdot IMD}{F}\right) \qquad (15.6)$$

where:

f_p = infiltration capacity, mm/sec,
i = rainfall intensity, mm/sec,
F = cumulative infiltration volume, this event, mm,

F_s = cumulative infiltration volume required to cause surface saturation, mm,

$SUCT$ = Average capillary suction of water (mm)

$HYDCON$ = Saturated hydraulic conductivity of soil (mm/hr)

IMD = Initial moisture deficit for soil, volume air/volume voids (fraction)

Subsequently, infiltrated water enters the groundwater calculations. The groundwater subroutine performs the simulation for two zones: an upper unsaturated and a lower saturated zone. Water is transferred from the upper to the lower zone through percolation (15.7) and evapotranspiration is allowed for both zones. The program performs water balance calculations for the two zones each time step, updating the moisture content, water depth and water flow to the receiving channel. The transfer of flows from the groundwater to the channel is performed using 15.8.

$$Percolation = HKTH \cdot \left(1 + PCO \cdot \frac{(TH - FC)}{\dfrac{DWT1}{2}} \right) \tag{15.7}$$

$$GWFLW = A1 \cdot (D1 - BC)^{B1} - TWFLW + A3 \cdot D1 \cdot TW \tag{15.8}$$

where:

PCO = average slope of tension versus soil moisture curve (m/fraction)

$DWT1$ = beginning of time step upper zone depth (m)

BC = elevation of bottom of channel (m)

$TWFLW$ = channel water influence flow rate (per subcatchment area)

$\geq$ BC, average elevation of water in channel (m)

< 0 (e.g., -1) channel water influence will be determined by depth in channel at the end of the previous time step.

$A3$ = Coefficient for cross product (mm/hr-m²)

$D1$ = Beginning of time step lower zone depth (m)

TW = Elevation of water in channel (m)

$A1$ = Groundwater flow coefficient (mm-hr/m)

$B1$ = Groundwater flow exponent (dimensionless)

$HKTH$ = Hydraulic conductivity versus moisture content curve-fitting parameter (dimensionless)

TH = Initial upper zone moisture expressed as a fraction
FC = Field capacity expressed as a fraction

For this calibration exercise, the following parameters were considered as the optimization variables. The initial parameter uncertainty shown in Table 15.3 is expressed in terms of an uncertainty factor, as suggested by James (2002), which is randomly generated in each run and multiplies the parameter value suggested as the initial value.

Table 15.3. Calibration parameters and prior uncertainty factor ranges applied in the automatic calibration.

Parameter	Description	Uncertainty factor (James, 2002)
SUCT	Avg. Capillarity Suction ()	0.1 – 2.0
HYDCON	Saturated Hydraulic Conductivity (Green &Ampt)	0.1 – 2.0
IMD	Moisture Deficit (Green &Ampt)	0.1 – 2.0
A1	Groundwater Flow Coefficient	0.1 – 2.0
B1	Groundwater Flow Exponent	0.1 – 2.0
PR	Porosity	0.5 – 1.5
WP	Wilting point	0.5 – 1.5
FC	Field Capacity	0.5 – 1.5
TH	Initial Moisture Content	0.1 – 2.0
HKTH	Hyd. Cond. Vs. Moist. Cont. factor	0.1 – 2.0
DET	Transpiration depth	0.5 – 1.5
WW1	Width	0.1 – 2.0
WW6	Manning's N. Pervious	0.1 – 2.0
WSTORE2	Pervious depression storage	0.1 – 2.0

15.2.3 Building the Multi-objective Calibration Tool

NSGA-II was linked to the RUNOFF and EXTRAN modules as shown in Figure 15.3. New subroutines, written in C language, were added to the original NSGA-II code to allow for the interactions SWMM-NSGA-II and to calculate various hydrologic characteristics from the computed and observed

flow time series. The program first generates a random population of parameter values according to the feasible search space. These parameters are written into the RUNOFF input data file and SWMM is subsequently run for each set of parameters, transferring each time the generated runoff to EXTRAN to perform the hydrograph routing. The program extracts the routed flows to calculate the values of the objective functions. It is worth noting that the program can be easily modified to include the comparison of other variables such as groundwater level readings. The stopping criterion used in this exercise was related to a maximum number of generations, however, this can be replaced with a minimum tolerance limit.
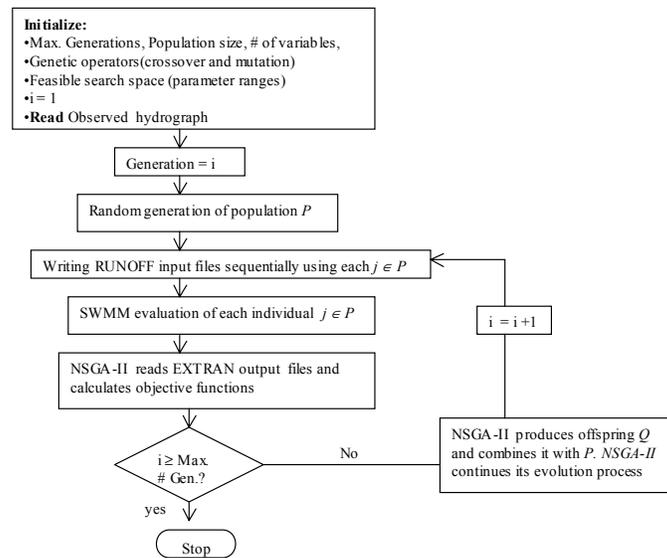
Figure 15.3. Multi-objective calibration algorithm of the RUNOFF-SWMM using the NSGA-II.

15.3 Application

15.3.1 Development of the hydrologic model and data generation

The proposed calibration tool was applied to a sub-watershed of the Eramosa River watershed, located in Southern Ontario in the vicinity of Guelph, as

shown in Figure 15.4. The sub-watershed has an approximate area of 1,324 ha and soils that are primarily sands and gravels. Groundwater is considered to be an important source of stream flow (Schroeter and Boyd, 1987).

A SWMM model was built using existing rainfall registers from a nearby raingauge, local soil characteristics, transects of the Eramosa River, and topographic contours of the watershed. The RUNOFF model was arranged to simulate infiltration, percolation and groundwater contributions to the stream flow, as described above. Flows were routed downstream using the EXTRAN module. The model is assumed to capture the most relevant hydrologic and hydraulic processes that occur in the sub-watershed.
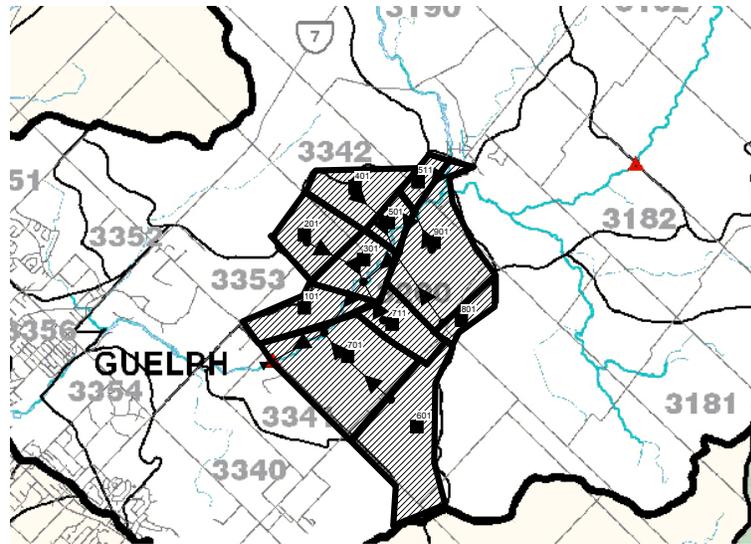
Figure 15.4 Study area near Guelph, Ontario.

Instead of real flow measurements for the calibration, the model was used to generate a four-month flow time series using a 15-minute time step, from June to September of 1986. This series is assumed to be the observed flow record as shown in Figure 15.5.

A second, SWMM model of the same area was built with a smaller number of subcatchments than the previous model. This was done to increase the spatial parameter uncertainty while minimizing the model structural uncertainty.

It is frequently assumed that the inability of hydrologic models to represent different characteristics of the hydrologic regime is mostly due to deficiencies in the model structure. This approach was adopted in order to identify whether the modeling trade-offs that have identified in past studies can be associated to some extent to the spatial parameter uncertainty.

15.3.2 Calibration Objectives

Three objective functions were constructed for the calibration. Each one compares different response modes of the 'observed' hydrograph (i.e. low, medium and high flows) with the 'modeled' results. As shown in Table 15.4, different flow magnitudes were identified from the observed hydrograph as relevant for different ecologic and morphologic processes. For example, minimum flows during the complete 4-month period are considered to be important for maintenance of aquatic habitat and water quality; medium flows are assumed to be relevant for sediment transport and channel forming activities; and high flows are important for riparian corridor activities. The objective functions were therefore built around these flow values. In order to provide a comparative basis, flows were parameterized according to the bankfull flow. Table 15.5 presents the three different flow response modes that were developed to build the objective functions.

Table 15.4 Flow categories.

FLOW (m³/s)	HYDROLOGIC ATTRIBUTE	RELEVANCE
0.15	MINIMUM OBSERVED FLOW	PHYSICAL HABITAT MAINTENANCE FLOW, LOW FLOWS/WATER QUALITY
0.38	MEAN FLOW	GEOMORPHICALLY SIGNIFICANT FLOW
1.53	BANKFULL FLOW (BF)	GEOMORPHICALLY SIGNIFICANT FLOW
2.4	MAXIMUM OBSERVED FLOW	RIPARIAN MAINTENANCE

Table 15.5 Flow response modes parameterized with the bankfull flow.

Flow categories	Description
$0.10 < Q/Q_{BANKFULL} < 0.25$	Low flows
$0.25 < Q/Q_{BANKFULL} < 1.0$	Medium flows
$1.0 < Q_{BANKFULL} < 1.6$	High flows

To perform the mathematical comparison between 'observed' and 'computed' flow response modes, two evaluation or fitness functions were used to compare, the Simple Least Square errors (SLS) and the Root Mean Square errors (RMS) presented in Table 15.6.

Table 15.6 Evaluation functions used for the multi-objective calibration.

Description	Formula
Simple Least Squares (SLS)	$SLS = \sum \left(Q_{comp} - Q_{obs} \right)^2$
Root Mean Square Errors (RMS)	$RMS = \sqrt{\dfrac{1}{n} \sum \left(Q_{comp} - Q_{obs} \right)^2}$

15.3.3 Description of the Experiments

Three numerical experiments were run. For each try the maximum number of generations was set to 100 and the population size to 50. The first calibration experiment was run with three objectives, using the SLS as the evaluation function applied to compare 'observed' vs. 'computed' low, medium and high flows. The second run applied the RMS to the same three flow characteristics, and, the last calibration run was implemented with two objectives. For the latter, both SLS and RMS were applied to the complete flow record, comparing each value of the 'computed' and 'observed' hydrographs. A summary of the experiments is presented in Table 15.7.

15.4 Results and Discussion

15.4.1 Automatic Calibration

For all three experiments, the automatic calibration tool successfully achieved an optimal set of calibrated solutions. In general, after the optimization is completed, the algorithm produces one table for every generation. Each table contains as many solutions as defined by the maximum population, and each solution contains its associated parameter and objective values. The size of the tables depends on the maximum

population (number of rows), and the combined number of parameters and objectives (number of columns). From these results, both parameter and objective function values were extracted and analyzed.

Table 15.7 Summary of the objective functions used for each experiment.

Experiment #	Objective functions
1	1) $SLS_{LOWER\ FLOWS}$ 2) $SLS_{MEDIUM\ FLOWS}$ 3) $SLS_{HIGH\ FLOWS}$
2	1) $RMS_{LOWER\ FLOWS}$ 2) $RMS_{MEDIUM\ FLOWS}$ 3) $RMS_{HIGH\ FLOWS}$
3	1) $SLS_{COMPLETE\ HYDROGRAPH}$ 2) $RMS_{COMPLETE\ HYDROGRAPH}$

Table 15.8 presents the posterior uncertainty ranges obtained with the first two calibration exercises, and the set of calibrated parameters for the last experiment. Because SLS and RMS are non-conflicting objectives, the last experiment converged to a single set of parameters. Parameter values obtained with the latter lie within, or very close, to the range provided for either two of the previous calibrations. This indicates that the solution obtained using the complete flow record is contained in the range obtained when the flow series is divided in different modes, but it does not inform the user of the possible modeling conflicts that may exist between different flow response modes.

It is observed that the ranges of the uncertainty factors, for most parameters, were significantly reduced after the calibration. In spite of the similarity of SLS and RMS, some differences are observed. For example, the upper boundary for parameter SUCT is significantly higher when using SLS than when using RMS. This behaviour indicates that even similar objective functions may direct the search optimization towards slightly different sectors of the search space, and therefore can extract different information from the same observed data set.

From this table it is also noted that some parameters are more readily identifiable than others. For instance, the posterior range of the uncertainty factor obtained for the HYDCON parameter is not as wide as that obtained for other parameters such as IMD, A1, B1, WP and FC.

It is also observed that some of the parameters that present a bigger range of posterior uncertainty are related to the release of groundwater flow to the channel. Due to the magnitude of the groundwater contributions in this watershed, it is expected that the modeling trade-offs between the different objectives could be attributed, to a great extent, to the uncertainty of these parameters.

Table 15.8 Posterior uncertainty factors after the calibration is completed.

Parameter	Objective function		
	SLS	RMS	SLS&RMS
SUCT	0.39-1.43	0.3-0.53	0.39
HYDCON	0.42-0.55	0.47-0.56	0.47
IMD	0.43-1.45	0.11-1.52	0.18
A1	0.96-1.85	0.98-1.29	1.07
B1	0.85-1.84	0.75-1.87	0.83
PR	0.67-0.98	0.69-0.86	0.72
WP	0.61-1.34	0.59-1.36	0.54
FC	0.54-1.13	0.53-0.83	0.59
TH	0.41-0.65	0.5-0.93	0.56
HKTH	0.48-0.56	1.1-1.5	1.68
DET	0.53-1.43	0.5-0.83	0.54
WW1	0.26-1.82	0.24-1.86	1.09
WW6	0.24-1.64	0.2-1.15	0.28
WSTORE2	0.15-0.81	0.11-0.24	0.14

Figure 15.6 illustrates a typical evolution, through each generation, of the population towards the Pareto or non-dominated front. In this case, medium flows $(0.25<Q/Q_{bankfull}<1.0)$ are compared with high flows $(1.0<Q/Q_{bankfull}<1.6)$. As the number of generations increases, the population is replenished with fitter individuals. As explained before, NSGA-II keeps the elites unmodified throughout the evolution. Solutions move from the right to the left until the Pareto front, obtained in the last generation, is located in the far left of the graph. Similar results were obtained when comparing other objective functions for the first two experiments. However, The sharpness of the Pareto curve varied according to which objectives were compared, indicating that different levels of conflicts or trade-offs between objectives.

Figures 15.7 to 15.9 illustrate the corresponding Pareto front obtained in the last generation for different combination of objectives when using the

SLS. The Pareto fronts illustrate the existing trade-offs for each two objective graph. It is evident that improvement in either one of the objectives deteriorates the model performance in terms of the other. For instance in Figure 15.7, moving to the right side of the graph improves high flows (riparian maintenance flows) but decreases the modeling performance for low flows (important for water quality maintenance). Therefore, selection of any 'calibrated' model from the Pareto set is likely to present an uncertainty bias towards any of the two objectives.
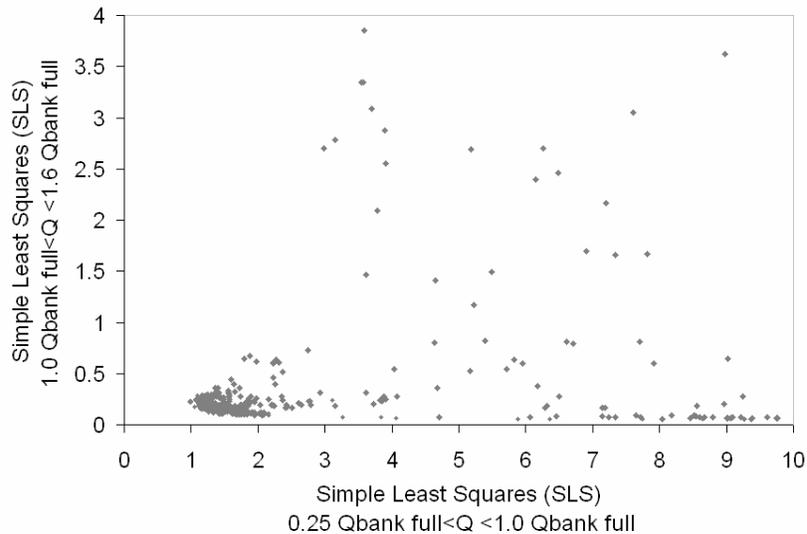
Figure 15.6 Pareto solutions for two objectives: low flows vs. high flows.

Stronger trade-offs were found between low and high flows (Figure 15.7) and between medium and high flows (Figure 15.8) than between low and medium flows (Figure 15.9). Limited trade-offs indicate that the objectives do not conflict for the most part of the search space. Therefore, it is possible to achieve simultaneous improvements in both objectives up to a certain point. Depending on the importance of the trade-offs, it could be possible to aggregate various semi-conflicting objectives (i.e. using a weighted factor function) to speed up the optimization process.
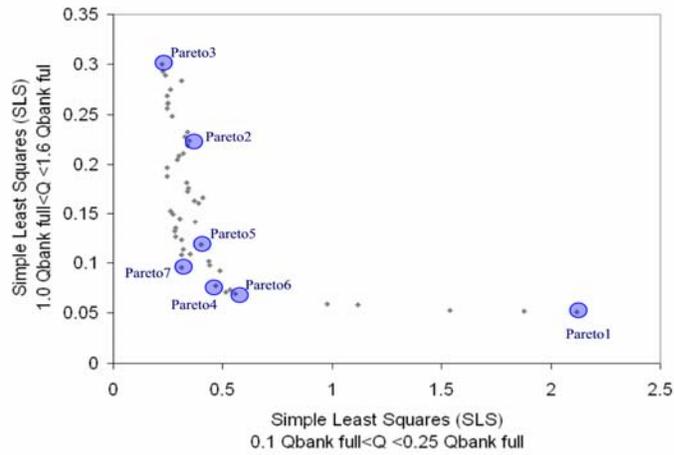
Figure 15.7 Pareto solutions for two objectives: low flows vs. high flows.
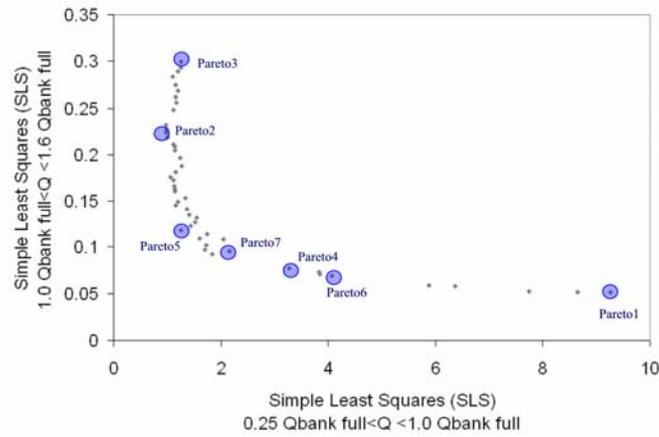
Figure 15.8 Pareto solutions for two objectives: medium vs. high flows.
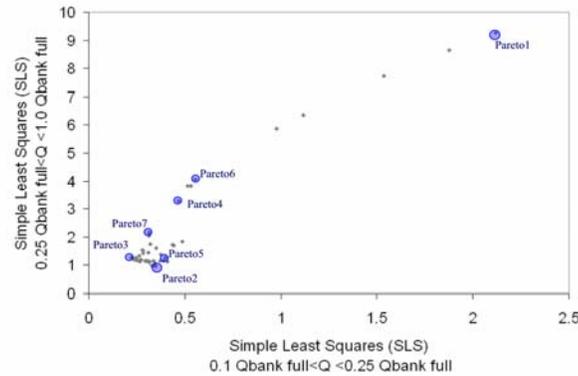
Figure 15.9   Pareto solutions for two objectives: low flows vs. medium flows.

15.4.2. Model Representation of Management Targets

In order to evaluate the performance of the calibrated models in terms of ecologically meaningful parts of the hydrologic regime, various models were sampled, covering different sectors of the Pareto sets, as shown in Figures 15.7, to 15.9. Selected models are described in Table 15.9 and include the best models for each objective and some compromised solutions. Figure 15.12 illustrates the performance of these calibrated models compared with the observed hydrograph.

Table 15.9. Description of selected models

Model name	Description	Evaluation function
Pareto1	Best model for high flows	SLS
Pareto2	Best model for medium flows	SLS
Pareto3	Best model for low flows	SLS
Pareto4	Compromised model	SLS
Pareto5	Compromised model	SLS
Pareto6	Compromised model	SLS
Pareto7	Compromised model	SLS

Table 15.10 presents a summary of the hydrological values and their properties, which are assumed to be relevant targets for this application and are therefore evaluated.

Table 15.10 Summary of ecologically sensitive hydrologic parameters.

	Flow period		Flow range					Ecologic, management relevance						
	Summer Period	Monthly values	Complete hydrograph	$0.1<Q/Q_{BANKFULL}<0.25$	$0.25<Q/Q_{BANKFULL}<1.0$	$1.0<Q/Q_{BANKFULL}<1.6$	$Q/Q_{BANKFULL}>1.6$	Water quality flow	Geomorphic processes	Habitat availability	Environmental disturbance	Wastewater dilution	Riparian corridor maintenance	Flooding protection
MMF (m³/s)		X	X							X				
PMF (m³/s)		X	X						X					X
MIMF (m³/s)		X	X					X		X		X		
TMV mm		X	X							X				
DLF (hr.)		X		X				X		X	X	X		
DMF (hr.)		X			X				X	X	X			
DHF (hr)		X				X				X	X		X	
DFF (hr)		X					X			X			X	X
LFV(mm)		X		X				X		X		X		
MFV mm		X			X				X	X				
HFV mm)		X				X				X			X	
FFV(mm)		X					X			X				X
MSF m³/s	X		X						X	X				
VAMF (mm)	X		X						X	X				
VABF (mm)	X		X										X	X

Legend: MMF = mean monthly flow; PMF = peak monthly flow; MIMF = minimum monthly flow; TMV = total monthly volume; DLF = duration of low flows; DMF = duration of medium flows; DHF = duration of high flows; DFF = duration of flooding flows; LFV = low flows volume; MFV = medium flows volume; HFV = high flows volume; FFV = flooding flows volume; MSF = mean summer flow; VAMF = volume above mean flow; VABF = volume above bankfull flow.

Results are presented from Figure 15.10 to Figure 15.24. In general, model performance varies, depending on the month and hydrologic characteristic considered. There is not a unique model that performs better compared to the others. Instead, different models are found to perform better for certain characteristics, and also, a model that produces the best results for a given month, may not the be best for the other three months. The graphs show that the errors can be positive or negative, indicating underestimation

and overestimation, respectively. In the long term, the sign of the error could be important in terms of direction of a given management strategy. For instance, a model that underestimates flows associated with sediment transport may indicate a reduced risk of channel and ecosystems degradation, thus leading to erroneous or ineffective management strategies.

In terms of monthly mean flows (Figure 15.10), most of the models present mean flow errors between –0.1 and 0.1, except for June, indicating an overall good agreement of the calibrated models for these variables. Regarding peak monthly flow errors, a variety of results are observed in Figure 15.11. While *Pareto 7*, present the best results for June; *Pareto 2* and *Pareto 6* are the best models for July; *Pareto 7* present the best results for September.

Minimum monthly flow errors (Figure 15.12) are better represented for July, August and September; also, from a visual inspection, *Pareto 1* presents an overall better performance than the other models. However, all other models perform better than *Pareto 1* in at least one month of the summer. Total monthly flow errors are presented in Figure 15.13. *Pareto 1* is the best for June; for July, August and September.

Monthly flow duration errors are presented from Figure 15.14 to Figure 15.17. As shown in Figure 15.14, most of the models underestimate the duration of low flows ($0.1 < Q/Q_{BANKFULL} < 0.25$) for the month of June, except *Pareto 1*. *Pareto 2* and *Pareto 5* provide the minimum flow error for July, with the latter being the smallest. *Pareto 2* and *Pareto 5* are the best for August, and *Pareto 1* is the best for September. On the other hand, the duration error of medium flows ($0.25 < Q/Q_{BANKFULL} < 1.0$), shown in Figure 15.15, is overestimated in June by all models, except by *Pareto 1*. For July, most models present a similar behaviour, with the worst models being *Pareto 1*. Regarding duration error for flows between bankfull condition and riparian flooding conditions ($1.0 < Q/Q_{BANKFULL} < 1.6$), and high ($Q/Q_{BANKFULL} > 1.6$), it is observed that only August presents this type of flows, with the best models being, and *Pareto 1*, *Pareto 6*, *Pareto 7*, From Figure 15.18 to Figure 15.21 present the results for volume error for the different flow categories. It is observed that the behaviour is quite similar to that of the flow duration error for each flow response mode. Figure 15.22 presents the mean flow error for the complete summer record indicating that the best model is *Pareto 7*, followed by *Pareto 1* and *Pareto 5*. Figure 15.23 presents the volume above mean flow error for the complete summer as well. The plot identifies *Pareto 2* and *Pareto 7* to be the best model. Finally, the volume error for flows above bankfull for the summer is presented in Figure 15.24.

Clearly, selection of a compromised calibrated model is a difficult task. Different methods can be used to select one model from all the possible options. Some authors (di Pierro et al., 2004; Khu and Madsen, 2005) have used a ranking procedure, based on the Preference Ordering (Das, 1998) to select a reduced set of compromised solutions. This method, basically selects solutions that are Pareto-optimal for all combinations of objective functions (Khu and Madsen, 2005). Another option could be the ranking of models according to their individual performance on individual objectives, adding up the ranks of each solution, and selecting the one with the lowest aggregated rank. However, regardless of the method used for model screening, selection of a preferred model alone neglects the valuable information extracted by the calibration and reflected in the Pareto set. All models that belong to the Pareto set are 'specialists' in terms of certain hydrologic characteristics.

Depending on the specific application, a model can be selected by further discussion between involved parties. If conflicting interests were involved, it would be better to take advantage of the Pareto set as a whole, using a compromised model and checking the results using other 'specialists'. For instance, if the model is used to design stormwater control structures, a compromised model can be used to produce the design, which would be subsequently evaluated by other members of the Pareto set. The advantage of this approach is that it provides an uncertainty assessment of the application, based on specific knowledge extracted from the observed data.
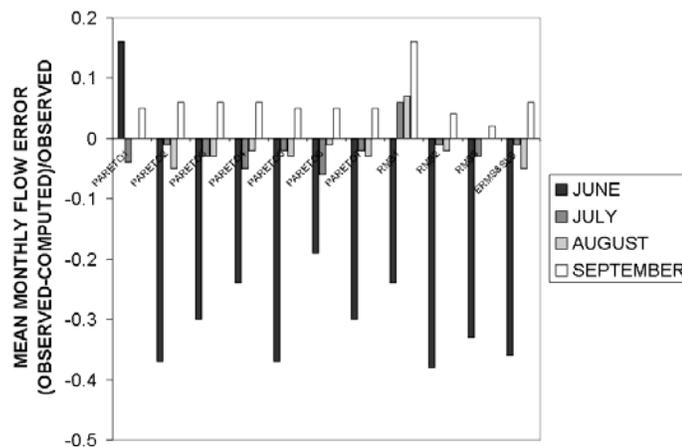
Figure 15.10 Mean monthly flow error.
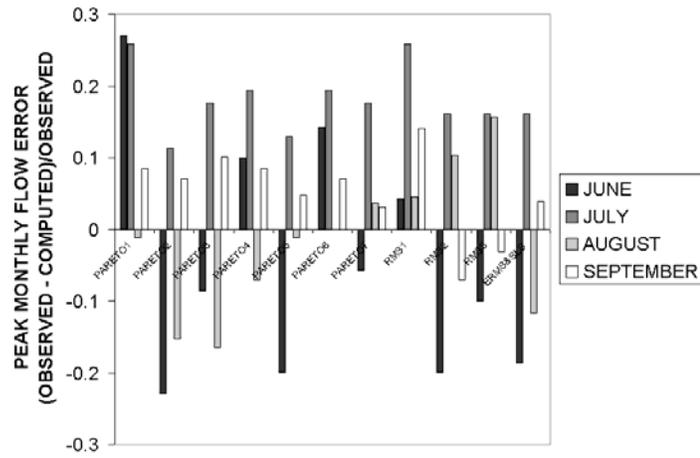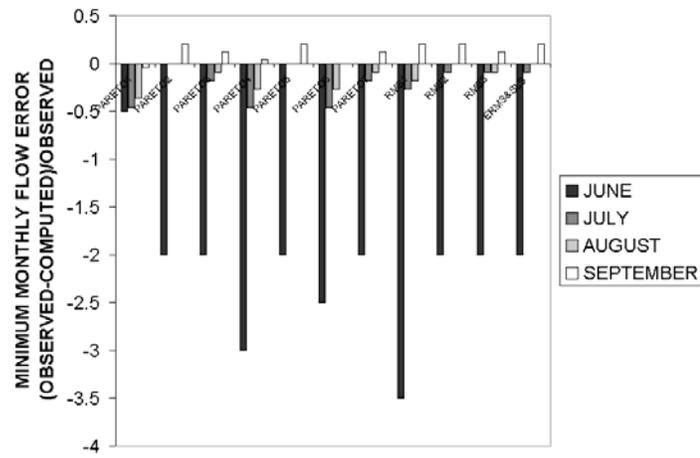
Figure 15.11. Peak monthly flow error

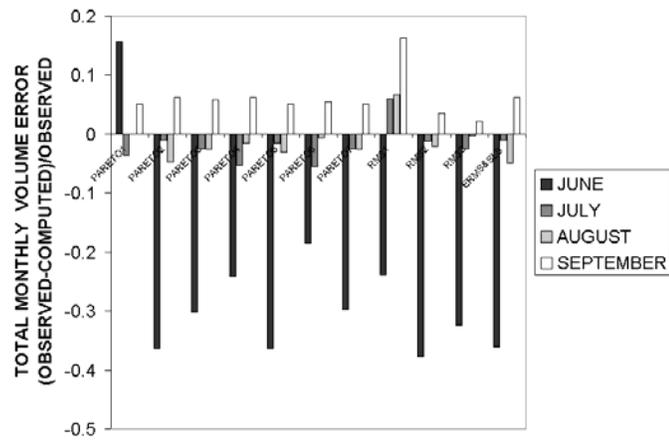

Figure 15.12 Minimum monthly flow error.

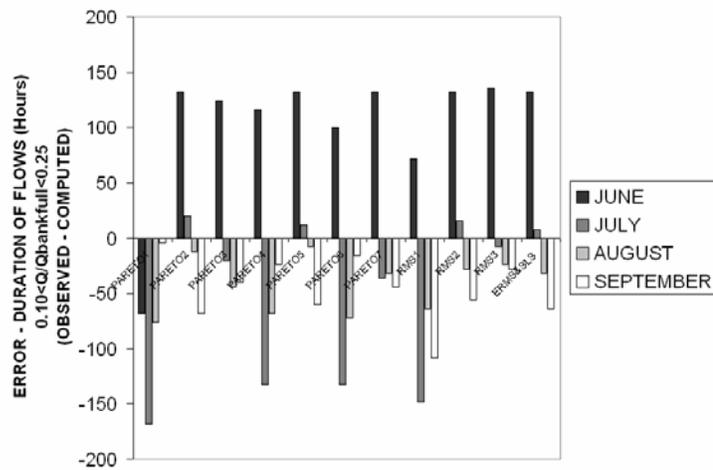**Figure 15.13** Total monthly volume error.

Figure 15.14 Duration of low flow error.
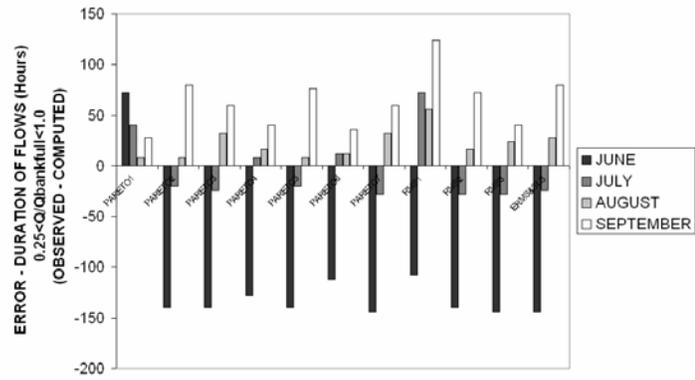
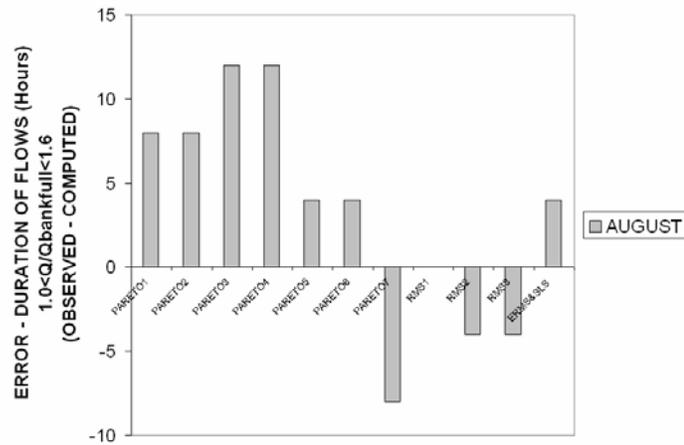**Figure 15.15** Duration of medium flows error.

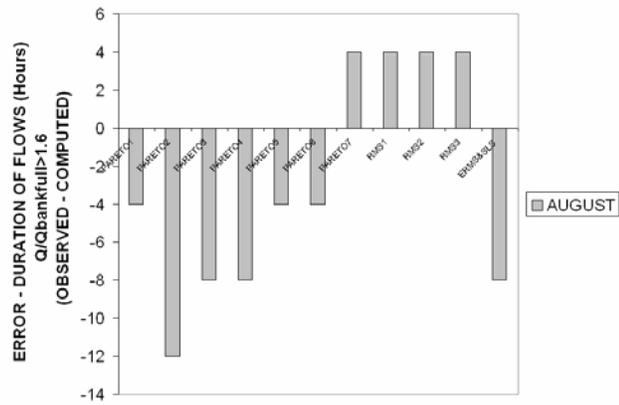**Figure 15.16** Duration of high flows error.
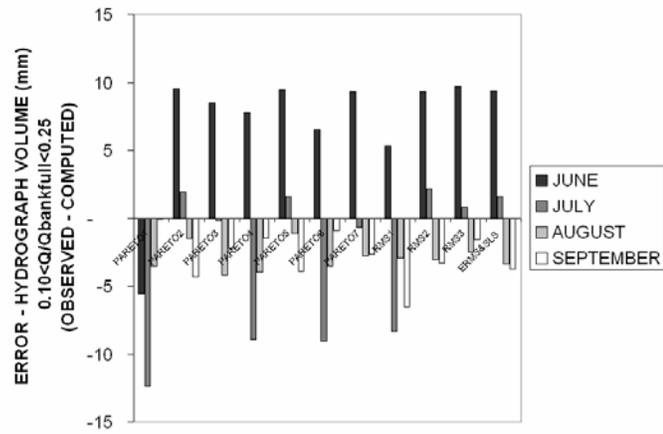
Figure 15.17 Duration of flooding flows error.
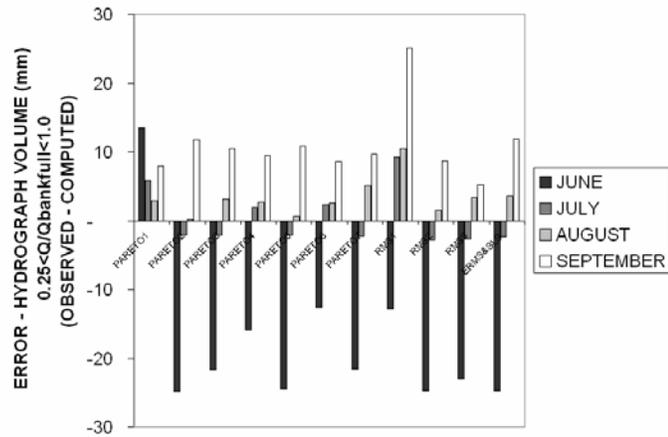

Figure 15.18 Low flows volume error.
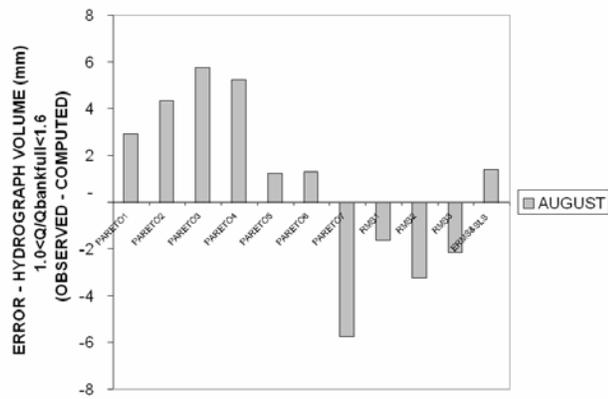
Figure 15.19 Medium flows volume error.

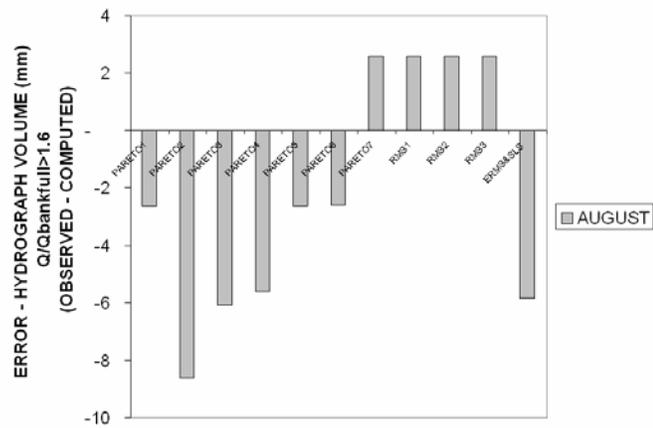**Figure 15.20** High flows volume error.

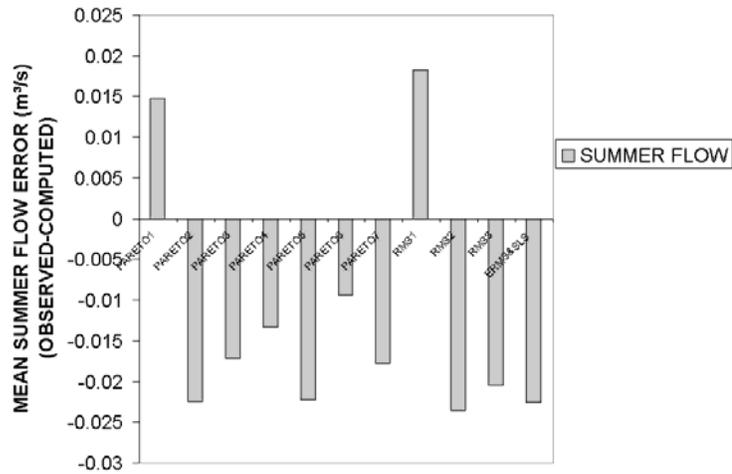**Figure 15.21** Flooding flows volume error.

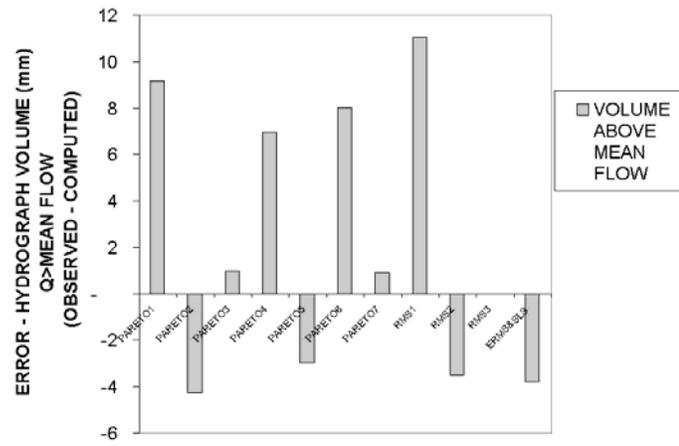**Figure 15.22** Mean summer flow error.

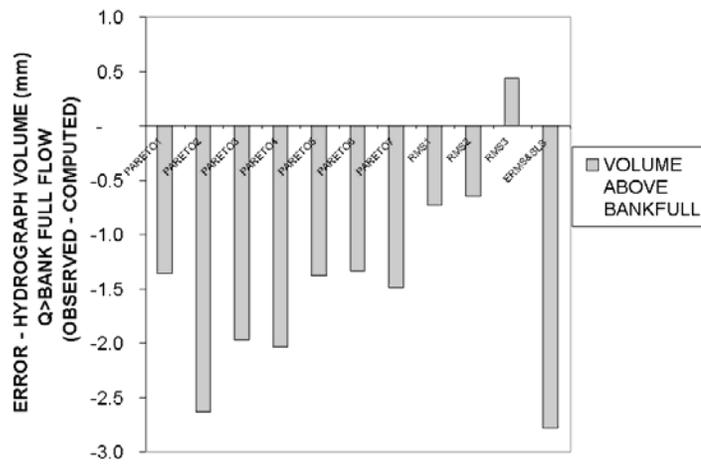**Figure 15.23** Hydrograph volume above mean flow error.

Figure 15.24 Hydrograph volume above bankfull flow error.

15.5 Conclusions

A multi-objective calibration tool was successfully developed combining the NSGA-II (Deb et al., 2000) and the Stormwater Management Model (SWMM). Numerical tests support previous findings (Gupta et al., 2003) that indicate the existence of trade-offs between different calibration objectives.

The results presented in this chapter are important because, even having a 'perfect' model that represent all the process as they occur in nature, spatial parameter uncertainty might induce modeling conflicts between different model performances and reflected in model bias in the simulation of different flow modes or characteristics.

Presenting the calibrated parameters, as an uncertainty range is advantageous because it suggests which parameters would require an improved initial estimate from field or laboratory experiments in order to minimize modeling conflicts.

Calibration of hydrologic models has an important significance on modeling the effects of human activities on ecologic and morphologic processes that cannot be overlooked. Trade-offs between relevant characteristics such as flow durations, hydrograph volumes, minimum flows, bankfull flows and riparian flows should be evaluated before recommending a model.

Screening of models from the optimal set is a difficult task that depends on the specific interest of the modeling group. Different methods are available, to select a model from the Pareto set, however, it would be helpful to consider using the complete set in order to provide a better understanding of the uncertainties involved in a specific modeling application.

Finally, this work underscores the importance of adaptive environmental management approaches that involve the use of models as tools to select targets and designing management strategies.

References

Beven, K. and Binley, A. (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*. Vol. 6. Pages 279 – 298.

Bledsoe, Brian P. (2001) Relationships of stream responses to hydrological changes. Linking Stormwater BMP Designs and Performance to Receiving Water Impact Mitigation. Edited By Ben Urbonas. ASCE, UEF, EWRI.

Das, I. (1998). A preference ordering among various Pareto Optimal Alternatives *Structural Optimization*. Vol. 18. No. 1.

Deb, K.; Pratap, A.S.; Agarwal, S. and Meyarivan, T. (2000). A Fast and Elitist Multi-Objective Genetic Algorithm-NSGA-II. KanGAL Report Number 2000001.

di Pierro, F.; Djordjevic, S.; Kapelan, Z.; Khu, S.T.; Savic, D.; Walters, G.A. (2004) Automatic calibration of urban drainage model using a novel multi-objective genetic algorithm *Proceedings 6th International Conference on Urban Drainage Modeling (UDM'04)*. Dresden, Germany.

Duan, Q. ; Sorooshian, S. and Gupta, V. (1992). Effective and Efficient Global Optimization for Conceptual Rainfall Runoff Models. *Water Resources Research*. Vol. 28, No 4, Pages 1015-1031.

Gupta, H.V.; Sorooshian, S. and Yapo, P. (1998) Toward improved calibration of hydrological models: multiple and non-commensurable measures of information *Water Resources Research*. Vol. 34. No. 4. Pages 751 - 763.

Gupta, H.V.; Sorooshian, S.; Hogue, T. and Boyle, D. (2003) Advances in automatic calibration of watershed models. In Calibration of Watershed Models, Water Sci. Appli. Ser. Vol. 6, edited by Duan, Q.; Gupta, H.; Sorooshian, S.; Rousseau, A. and Turcotte, R. pp. 125– 132, AGU, Washington, D. C., 2003.

Hreiche, A.; Bocquillon, C. ; Dezetter, A. Servat, E. and Najem, W. (2002). Parallel processing for a better understanding of equifinality in hydrologic models. *Proceedings EMS Integrated Assessment and Decision Support*.

Huber, W.C.; Dickinson, R.E. 1992. Stormwater Management Model, version 4: User's Manual. US Environmental Protection Agency, Athens, Georgia.

Instream Flow Council (2002). Instream flows for riverine resource stewardship ISBN 0-9716743-0-2(pbk).

James, W. (2002). Rules for Responsible Modeling. Third Edition. CHI Publications.

Johnston, P.R. and Pilgrim, D.H. (1976). Parameter optimization for watershed models *Water Resources Research.* Vol. 12, No 3. Pages 477 – 486.

Khu, S.T. and Madsen, H. (2005) Multi-objective calibration with Pareto preference ordering: An application to rainfall runoff model calibration. *Water Resources Research,* Vol. 41.

Konrad, C.P.; Booth, D.B. (2002) Hydrologic trends associated with urban development for selected streams in the Puget Sound basin, Western Washington US Geological Survey. Water Resources Investigations Report No 02-4040.

Madsen, H. (2000). Automatic calibration of a conceptual rainfall runoff model using multiple objectives Journal of Hydrology. Vol. 235. Pages 276 – 288.

Madsen, H. (2003). Parameter estimation in distributed hydrological catchment modeling using automatic calibration with multiple objectives. *Advances in Water Resources*. Vol. 26. 205 – 206.

Pickup, G. (1977) Testing the efficiencies of algorithms and strategies for automatic calibration of rainfall runoff models. *Hydrological Sciences Bulletin*. Vol. 22, No 2. Pages 257 – 274.

Poff, N.L. and Ward, J.V. (1989). Implications of streamflow variability and predictability for lotic community structure: a regional analysis of streamflow patterns Can. J. Aquat. Sci., Vol. 46. Pages 1805–1818.

Poff, N.L.; Allan, J.D.; Bain, M.B.; Karr, J.R.; Prestegaard, K.L.; Richter, B.D.; Sparks, R.E.; Stromberg, J.C. (1997) The natural flow regime: a paradigm for river conservation and restoration. Bioscience. Vol. 47. No 11. Pages 769 – 784.

Richter, B.; Baumgartner, J.; Powell, J.; Braun, D. (1996). A Method for Assessing Hydrological Alteration within Ecosystems. Conservation Biology. Vol. 10, No 4. p. 1163 – 1174

Richter, B.; Baumgartner, J.; Wiginton, R. and Braun, D. (1997). How much water does a river need? Freshwater Biology. 37, 231-249.

Roesner, L. and Bledsoe, B. (2001) Effects of wet weather discharges on the physical character of aquatic habitat *WERF Report 000-WSN-4*. Draft Report.

Schroeter, H. and Boyd, D. (1998) Eramosa watershed study hydrology technical appendix Final Draft.

Soorooshian, S.; Gupta, V.K. and Fulton, J.Lloyd (1983). Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall runoff models: the influence of calibration data variability and length on model credibility *Water Resources Research*. Vol. 19, No 1. Pages 251– 259.

Vrugt, J.; Gupta, H.V.; Bouten, W. and Sooroshian, S. 2003. A shuffled complex evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resources Research*. Vol. 39, No 8.

Yapo, P. (1996). A Multi-objective global optimization algorithm with application to calibration of hydrologic models *Ph.D. Dissertation*. Department of Systems and Industrial Engineering, University of Arizona.

Yapo, P.; Gupta, H.V.; Sorooshian, S. (1996) Automatic calibration of conceptual rainfall – runoff models: sensitivity to calibration data. Journal of Hydrology. (181) 23- 48.