Sensitivity analysis of a land surface scheme using multicriteria methods

L. A. Bastidas, H. V. Gupta, S. Sorooshian, W. J. Shuttleworth, and Z. L. Yang
Department of Hydrology and Water Resources, University of Arizona, Tucson

Abstract. Attempts to model surface-atmosphere interactions with greater physical realism have resulted in complex land surface schemes (LSS) with large numbers of parameters. A companion paper describes a multicriteria calibration procedure for extracting plot-scale estimates of the preferred ranges of these parameters from the various observational data sets that are now available. A complementary procedure is presented in this paper that provides an objective determination of the multicriteria sensitivity of the modeled variables to the parameters, thereby allowing the number of calibration parameters and hence the computational effort to be reduced. Two case studies are reported for the BATS model using data sets of typical quality but very different location and climatological regime (ARM-CART and Tucson). The sensitivity results were found to be consistent with the physical properties of the different environments, thereby supporting the reasonableness of the model formulation. Further, when the insensitive parameters are omitted from the calibration process, there is little degradation in the quality of the model description and little change in the preferred range of the remaining parameters.

1. Introduction and Scope

This paper is one of three that discuss the usefulness of multicriteria methods for the evaluation and improvement of land surface schemes (LSS). A companion paper [Gupta et al., this issue] shows how multicriteria methods can be used to improve the estimates of LSS parameters by simultaneously constraining the model to measurements of several observed system responses such as heat fluxes, ground temperature, and surface soil moisture. This paper develops that study by introducing a robust multicriteria approach to parameter sensitivity analysis for LSS models and by showing how the methodology provides a way to reduce the dimensionality of the parameter estimation problem. In both papers, the methodology is illustrated using the Biosphere-Atmosphere Transfer Scheme (BATS) [Dickinson et al., 1993] and two data sets, one from the ARM-CART grassland site and the other from a semiarid site in the Sonoran Desert, Arizona. A third paper (in preparation) will discuss the power and applicability of multicriteria methods for the evaluation of model performance and for model intercomparison.

The paper is organized as follows: Section 2 discusses the background and context for this work and presents a review of the literature. Section 3 introduces the theoretical and practical basis for applying multicriteria methods to the generalized sensitivity analysis of LSS model parameters. In Section 4, the sensitive parameters of the BATS model for two study sites are identified, and it is shown that only the reduced set of sensitive parameters need to be calibrated to obtain good model performance. Section 5 discusses the results and future extensions.
must be specified. Many of these are “conceptual” parameters that are not directly measurable at the GCM-LSS scales and for which all that might reasonably be specified are the approximate ranges for the parameter values, based on some approximate understanding of the regional hydrogeology. If the input-state-output response of the LSS is not essentially sensitive to parameter variations within these ranges, it will be reasonable to use some nominal estimates (such as the midpoints of the parameter uncertainty ranges). However, if the LSS input-state-output response is sensitive to finer specification of these parameters, the only remaining recourse is to adjust the parameter so that model responses are constrained to closely match available observations via calibration [Gupta et al., this issue]. It is important therefore to identify which of the parameters within the model are sensitive (for each hydroclimatic regime) and should be refined further via calibration.

A number of different methods have been proposed in the literature for the sensitivity analysis of LSS schemes. These include the most commonly used “one parameter at a time” approach [e.g., Wilson et al., 1987a, b; Pitman, 1994; Gao et al., 1996], the factorial method [e.g., Henderson-Sellers, 1992; Lettenmaier et al., 1996], the Fourier amplitude sensitivity test (FAST) [Collins and Avissar, 1994], and the regionalized sensitivity analysis (RSA) methodology [Franks et al., 1997]. Most studies assumed parameter independence and explored the individual impacts of each parameter on each system response taken one-at-a-time. None of the methods accounts for the joint multiparameter and multiresponse interactions. Although the factorial method allows (in principle) the exploration of multiparameter interactions, a simple two-level factorial experiment for a 25 parameter model would require an unreasonable number of model runs ($2^{25} = 33,554,432$). While this number can be reduced by a fractional factorial experiment that only considers second- or third-order interdependence, the effects of large numbers of possible interactions would then be ignored. The FAST method would only require about 6000 model runs for the same problem (extrapolating from Cukier et al. [1978]), but it again ignores interdependence among the parameters or requires an a priori estimate of the parameter covariance structure.

Another common feature of past studies is that they were directed not only at finding the influence of the model parameters but also at how this influence varies under different climatic forcing conditions. The studies were therefore conducted using artificially constructed atmospheric forcing data for different environments [e.g., Henderson-Sellers, 1992; Gao et al., 1996]. Further, the impacts were typically evaluated in terms of integrated response statistics such as average or cumulative values over monthly or yearly time spans (e.g., total evapotranspiration, total runoff, and minimum upper layer soil temperature for a year [Henderson-Sellers, 1992]; mean of the annual heat fluxes, ground temperature, and soil moisture [Pitman, 1994]; annual cumulative monthly runoff, annual cumulative monthly accumulated absolute differences between soil moisture and porosity for perturbed and nominal parameters [Lettenmaier et al., 1996]; and annual average variation of model responses, [Gao et al., 1996]). To our knowledge, only Franks et al. [1997] used real system response data for their sensitivity analysis. However, they studied a time span of only 10 days and only considered the latent heat flux.

In general, these studies have reported that the latent and sensible heat fluxes are sensitive to the same parameters. One parameter that seems to be universally important for different LSS formulations is that related to the stomatal resistance of the vegetation. However, the results of different studies using BATS are inconsistent; for example, Lettenmaier et al. [1996] report high sensitivities to the Clapp and Hornberger parameter B, while Gao et al. [1996] and Henderson-Sellers [1992] did not. An important finding by Gao et al. [1996] is that the influence of errors in the initial soil layer water conditions may persist for several years beyond the end of the spin-up period.

The work presented here builds on the multicriteria framework presented by Gupta et al. [1998] and Gupta et al. [this issue]. The major aim of this paper is to demonstrate a robust, practical, and efficient approach for the parameter sensitivity analysis of LSS models, in a manner that properly accounts for the joint multiparameter/multiresponse interactions, and to illustrate the consequent benefits that can accrue. We show that the approach provides a parameter sensitivity classification which is consistent with physical understanding. We also show how this classification can be used effectively to reduce the dimensionality of the parameter estimation problem, thereby significantly reducing the computer time required to conduct a model calibration.

3. Multicriteria Sensitivity Analysis

3.1. Problem Formulation

A theoretical basis for the application of multicriteria theory to the calibration of conceptual multiparameter physically based models was presented by Gupta et al. [1998], and this is extended to LSS models by Gupta et al. [this issue]. The characteristic of LSS models is that they have a relatively large number ($p$) of parameters $\theta = \{\theta_1, \cdots, \theta_p\}$, which influence the time evolution behavior of several ($m$) different modeled variables $(Z_j(\theta, t_j), t_j = ta_j, \cdots, tb_j, j = 1, \cdots, m)$. We define a single criterion $f_j(\theta)$ for each separate ($j$th) model response to measure the distance between the set of modeled responses $Z_j(\theta)$ and some benchmark responses $X_j$ (Notes: (1) it is trivial to extend the methodology to allow more than one complementary criterion to be defined on each modeled response; (2) without loss of generality, we assume that $f_j(\theta) > 0$). The benchmark responses could either be model simulations $Z_j(\theta^*)$ performed at some nominal point $\theta^*$ in the parameter space or a set of observations $O_j$ made at a study site. The specification of the mathematical form of these criteria depends on the problem and the goals of the user; in this work we adopt the commonly used root-mean-square error measure of the deviation between two time series ($\text{RMSE}_j(\theta) = \text{sqt} \{(1/n) \sum_{i=1}^n (Z_j(\theta, t_i) - X_j(t_i))^2\}$). The multicriteria response of the model as a function of the parameters can then be defined on the feasible parameter space $\Theta$ as

$$F(\theta) = \{f_1(\theta), \cdots, f_m(\theta)\} \quad (1)$$

where $\Theta \subset \Theta$. The primary objective of parameter sensitivity analysis is to determine which of the parameters $\theta$ give rise to significant variations in the multicriteria response function $F(\theta)$ as they are allowed to vary jointly over $\Theta$ and to determine the nature of those variations. A secondary objective is to classify the sensitivity of the parameters into some order of relative importance.

The above-stated goal of deriving multicriteria parameter sensitivity can be addressed by extending and improving the “regional sensitivity analysis” (RSA) single-criterion method
minima. Instead, it is common to have a set of solutions, with the property that moving from one solution to another results in the improvement of one criterion while causing deterioration of another. Figure 1a illustrates a simple case with two parameters \((\theta_1, \theta_2)\) and a two-criteria response function \(\{f_1, f_2\}\). Figure 1a shows the feasible parameter space \(\Theta\), and Figure 1b shows the projection of the parameter space into the function space. The points \(\alpha\) and \(\beta\) indicate the solutions that minimize each of the individual criteria \(f_1\) and \(f_2\). The thick line indicates the set \(S\) of multicriteria minimizing points to the function \(\{f_1, f_2\}\); \(\gamma\) is an element of the solution set \(S\) which is superior in the multicriteria sense to any point \(\delta\) not contained in it.

[Spear and Hornberger, 1980; Hornberger and Spear, 1981]. The latter method begins with a uniform random sampling of \(n\) distinct points in the feasible parameter space \(\Theta\) (we shall call this set of points \(R\)) and proceeds by partitioning \(\Theta\) into a region \(B\) (called the “behavioral” region), which is characterized by a desirable modeled behavior, and its complement \(\bar{B}\) (called the “nonbehavioral” region). The behavioral region can be defined either in terms of the parameter space that produces desired trajectories of the modeled responses or state variables [Spear and Hornberger, 1980], or in terms of the (subjectively assessed) desirable values of a summarizing criterion [Hornberger et al., 1985; Beven and Binley, 1992; Chang and Delleur, 1992; Spear et al., 1994; Ohte and Bales, 1995; Franks et al., 1997]. In the context of (1) the behavioral region must be defined in terms of a multicriteria threshold applied to the function \(F(\theta)\). A simple way to do this is to select a \(T = \{T_1, \cdots, T_m\}\) and define \(B(T)\) such that \(F(\theta) < T\) (where \(<\) denotes component-by-component superiority). However, this requires the subjective selection of \(m\) threshold values and, further, does not guarantee multicriteria superiority [Bastidas, 1998]. An alternative approach is to transform the multicriteria formulation (equation (1)) into a single-criterion form by some mapping such as

\[
g(\theta) = \sum_{i=1}^{m} w_f i(\theta) \tag{2}
\]

where the \(w_f\) are user-selected weights defining the relative importance of the individual criteria and to apply a single threshold to \(g(\theta)\). However, this still requires the subjective selection of the \(m\) weights. These problems can be avoided largely by adopting the notion of Pareto Ranking, this being a scale-independent and truly multiobjective way to achieve a partition of the feasible space.

### 3.2. Pareto Ranking

A defining characteristic of the multicriteria response function \(F(\theta)\) is that there is not, in general, a unique minimizing point in the parameter space. That is, because of errors in the model structure and other possible sources, it is usually impossible to find a single point \(\theta\) at which all the criteria have their
The Tucson data were collected at 20 min time intervals and cover an entire year from May 1993 to April 1994. The measured atmospheric forcings are net radiation, incoming radiation, air temperature, precipitation, specific humidity, and wind speed. The observed model response variables are sensible heat ($H$), latent heat ($\lambda E$), and ground temperature ($T_o$).

The measured heat fluxes were flagged at certain times by Unland et al. [1996] as being of dubious quality, and these data were therefore not used in this study. The remaining observational data were available at 5219 time steps.

The ARM-CART data set covers the 5 month period from April 1 to August 25, 1995, with a sampling interval of 30 min. The observed forcing values are net radiation ($R_n$ in W/m$^2$), surface temperature ($T_a$ in Kelvins (K)), atmospheric pressure ($p_a$ in kPa), relative humidity ($r_h$ in percent), wind velocity ($V_w$ in m/s), and precipitation ($P$ in millimeters (mm)).

Several gaps in the data were filled via interpolation by taking account of the diurnal cycle and the occurrence of precipitation [see Gupta et al., this issue]. The data set also contains time series observations for four variables that correspond with two model outputs, sensible heat ($H$ in watts per square meter (W/m$^2$)) and latent heat ($\lambda E$ in W/m$^2$), and two model state variables (ground temperature ($T_g$ in K) and soil moisture ($S_w$ in mm)).

Because of problems with the Bowen ratio system measurements, all $H$ and $\lambda E$ data corresponding to Bowen ratio between 0.75 and 1.25 were discarded. The resulting number of time steps at which values of $H$, $\lambda E$, $T_o$, and $S_w$ are available is 4237.

### 4.1. Sensitivity Analysis Results

The MOGSA methodology was used to estimate the parameter sensitivity of BATS for each site using the RMSE between the observed and the model-simulated time series values of each observed system response as a different system response criterion: {$H$}, {$\lambda E$}, {$T_o$}, and {$S_w$} for ARM-CART and $H$, $\lambda E$, and $T_o$ for Tucson. The parameters were allowed to vary between minimum and maximum values corresponding approximately to the full range of possible values for all locations on the Earth, except for the parameters rough and displa that do not include the rainforest vegetation type and parameter xmoohyd for which the range was somewhat more restricted (see Table 1). Note that proper feasibility constraints were imposed on the allowable values for the parameters to preserve the physical realism of the parameterization. For example, the thicknesses of the shallower soil layers are constrained to be less than the deeper soil layers, the initial water contents are less than the corresponding soil layer depths, and the season variability of vegetation cover and leaf area index are smaller than their maximum values.

The algorithm stabilized for a minimum sample size of 750 for the Tucson data and for a minimum sample size of 3000 for the ARM-CART data. The Pareto rank 10 was chosen as the threshold value $R^*$ because (1) it was found to achieve a higher number of sensitive parameters with a smaller sample size (for both the mean and the median of the K-S probability value and for both sensitivity thresholds 0.01 and 0.05); and (2) because it gave stable results regardless of sample sizes. The results are presented in Figure 3 for the Tucson site and Figure 4 for the ARM-CART site. The top subplots show the results for the 16 vegetation parameters and the bottom subplots show the re-

---

**Figure 2.** Flowchart of the MOGSA parameter sensitivity analysis method.
Table 1. Parameters Considered in BATS Sensitivity Analysis

<table>
<thead>
<tr>
<th>Name</th>
<th>Description (units)</th>
<th>Reasonable Range Tucson</th>
<th>Reasonable Range ARM-CART</th>
</tr>
</thead>
<tbody>
<tr>
<td>vecg</td>
<td>vegetation cover</td>
<td>0.0–0.95</td>
<td>0.10–0.70</td>
</tr>
<tr>
<td>seasf</td>
<td>difference between vecg and fractional cover at 269 K</td>
<td>0.0–0.60</td>
<td>0.00–0.60</td>
</tr>
<tr>
<td>rough</td>
<td>aerodynamic roughness length (m)^2</td>
<td>0.0024–1.0</td>
<td>0.01–0.84</td>
</tr>
<tr>
<td>displa</td>
<td>displacement height (m)^2</td>
<td>0.0–5.0</td>
<td>0.05–1.50</td>
</tr>
<tr>
<td>rsmin</td>
<td>minimum stomatal resistance (s/m)</td>
<td>5.0–200.0</td>
<td>50.0–200.0</td>
</tr>
<tr>
<td>xla</td>
<td>maximum leaf area index</td>
<td>0.0–6.0</td>
<td>0.05–3.00</td>
</tr>
<tr>
<td>xla10</td>
<td>minimum leaf area index</td>
<td>0.0–5.0</td>
<td>0.05–3.00</td>
</tr>
<tr>
<td>sa1</td>
<td>stem area index</td>
<td>0.5–4.0</td>
<td>1.00–4.00</td>
</tr>
<tr>
<td>sqrdi</td>
<td>inverse sqrt of leaf dimension (mm ^ -0.5)</td>
<td>5.0–10.00</td>
<td>5.00–10.00</td>
</tr>
<tr>
<td>fc</td>
<td>light dependence of stomatal resistance (m^2/w)</td>
<td>0.02–0.06</td>
<td>0.02–0.06</td>
</tr>
<tr>
<td>depuv</td>
<td>depth of top soil layer (m)</td>
<td>0.05–0.5</td>
<td>0.01–0.20</td>
</tr>
<tr>
<td>deprev</td>
<td>depth of root zone layer (m)</td>
<td>0.5–2.0</td>
<td>0.50–2.00</td>
</tr>
<tr>
<td>deptv</td>
<td>depth of total zone layer (m)</td>
<td>5.0–10.00</td>
<td>5.00–10.00</td>
</tr>
<tr>
<td>albvg</td>
<td>vegetation albedo for shortwave &lt; 0.7 µm</td>
<td>0.04–0.20</td>
<td>0.10–0.20</td>
</tr>
<tr>
<td>albvgl</td>
<td>vegetation albedo for longwave &gt; 0.7 µm</td>
<td>0.18–0.40</td>
<td>0.20–0.40</td>
</tr>
<tr>
<td>rootf</td>
<td>ratio of roots in upper layer to roots in root layer</td>
<td>0.30–0.90</td>
<td>0.10–0.90</td>
</tr>
<tr>
<td>xmo1</td>
<td>porosity</td>
<td>0.33–0.66</td>
<td>0.33–0.66</td>
</tr>
<tr>
<td>xmosuc</td>
<td>minimum soil suction (mm)</td>
<td>30.0–200.0</td>
<td>30.0–200.0</td>
</tr>
<tr>
<td>xmo1ycd</td>
<td>maximum hydraulic conductivity (mm/s)^c</td>
<td>0.0008–0.01</td>
<td>0.0008–0.01</td>
</tr>
<tr>
<td>xmo1wt</td>
<td>water content at which permanent wilting point occurs</td>
<td>0.0088–0.542</td>
<td>0.0005–0.01</td>
</tr>
<tr>
<td>xmo1cf</td>
<td>ratio of field capacity to saturated water content</td>
<td>0.404–0.866</td>
<td>0.71–0.71</td>
</tr>
<tr>
<td>bee</td>
<td>Clapp and Hornberger &quot;b&quot; parameter</td>
<td>3.5–10.8</td>
<td>3.5–10.8</td>
</tr>
<tr>
<td>skrat</td>
<td>ratio of soil thermal conductivity to that of loam</td>
<td>0.7–1.7</td>
<td>0.7–1.7</td>
</tr>
<tr>
<td>solour</td>
<td>soil albedo for different colored soils</td>
<td>0.05–0.12</td>
<td>0.05–0.12</td>
</tr>
<tr>
<td>ssw</td>
<td>surface zone water content (m)</td>
<td>0.0–deprev</td>
<td>0.0–0.2</td>
</tr>
<tr>
<td>rsw</td>
<td>root zone water content (m)</td>
<td>0.0–deprev</td>
<td>0.0–5.0</td>
</tr>
<tr>
<td>tsw</td>
<td>total zone water content (m)</td>
<td>0.0–deptv</td>
<td>0.0–10.00</td>
</tr>
</tbody>
</table>

Global and reasonable ranges.
^Range does not include tropical rain forest.
^Range restricted.
^Parameters not considered in the analysis.

sults for the eight soil parameters and the three initial soil-moisture conditions.

The top part of each subplot shows the “global” sensitivity of the model, i.e., the sensitivity obtained using the Pareto ranking 10 as threshold. The bottom part of each subplot shows the single-criteria sensitivity of the model to each modeled variable. For consistency, the single-criterion thresholds were chosen to partition the sample populations into fractions similar to those achieved by the Pareto ranking threshold. In the bottom part of each subplot, each group of vertical bars corresponds to the sensitivity of a different parameter; the longer the bar, the more sensitive the parameter. Also indicated in the plots are the 1 and 5% significance levels. If a sensitivity bar crosses the 1% level, it is considered to have high sensitivity; if it crosses the 5% level, it is considered to have medium sensitivity. Otherwise, it is considered to be relatively insensitive.

Considering first the Tucson site (Figure 3), the joint multicriteria analysis (top bars) indicates 16 sensitive parameters, specifically 14 parameters with high sensitivity, of which 8 are related to vegetation (vegc, seasf, rsmin, xla, xla10, depuv, deprev, and rootf), 4 are related to soil (xmosuc, xmo1ycd, bee, and skrat), and 1 is the initial water content of the root soil layer (rsw); and 2 parameters with medium sensitivity, 1 related to soil porosity (xmo1por), and 1 being the initial water content of the root layer (rsfw). Note that the initial total soil layer water content (tsw) might also be considered of “medium” sensitivity. The single-criterion analyses (bottom bars in Figure 4) reveal that the heat fluxes are very sensitive to stomatal resistance but that the state variables (ground temperature and soil moisture) are not. Further, in contrast with the Tucson site, the modeled variables are all sensitive to the seasonal variation of the vegetation cover (seasf) and the heat fluxes are both sensitive to minimum leaf area index (xla10). This is plausible,
considering that the period for which data are available is from April to August, and the vegetation over which measurements are made is irrigated crop.

Thus both sites show relatively high sensitivity to 11 parameters, 5 related to vegetation (vegc, rsmin, xla, depuv, and deprv), 6 related to soil (xmopor, xmosuc, xmohyd, bee, and skrat), and 1 being the initial water content of the surface soil layer (ssw). Both sites show relatively low sensitivity to 6 of the vegetation parameters, 3 related to stem and leaf properties (sai, sqrtdi, and fc), the depth of the total soil layer (deptv), and the vegetation albedos (albvgs, and albvgl).

However, the vegetation-related parameters (rough and displa) and the soil color parameter (solour) show high sensitivity at the Tucson site but low sensitivity at the ARM-CART site, while the vegetation-related parameters (seasf, xlaio, and rootf) and the soil parameter (xmotc) show low sensitivity at the Tucson site but high sensitivity for ARM-CART. Further, the vegetation-related roughness length (rough) is the most sensitive parameter for the Tucson site but is the least sensitive parameter for the ARM-CART site. In contrast, the difference between vegetation and fractional cover (seasf) is a very highly sensitive parameter for the ARM-CART site but a parameter with very low sensitivity at the Tucson site.

The sensitivity results for both sites are consistent with the model formulation and the physical properties of the different environments. For the semiarid Tucson site, 6 of the 8 most sensitive parameters are related to the soil storages and texture, while for the ARM-CART site, 5 of the 8 most sensitive parameters are related to vegetation properties. Moreover, the results generally agree with Gao et al. [1996], Wilson et al. [1987a, b], and Lettenmaier et al. [1996]. Note that at both sites in this study, there is low sensitivity to the light dependence of stomatal resistance (fc) and to the vegetation albedos (albvgs, albvgl) in marked contrast to the results of Henderson-Sellers [1992] for tropical forest.

**4.2. Application of the Sensitivity Results to Model Calibration**

An important aim of the parameter sensitivity analysis is to allow the possible reduction in the number of parameters that must be estimated, thereby reducing the computational time required for model calibration. Gupta et al. [this issue] conducted a multicriteria calibration run for the Tucson site in which they estimated 25 parameters using the latent heat ($H$), sensible heat ($\lambda E$), and ground temperature ($T_g$) system responses. This 25 parameter calibration run required ~20,000 iterations (model simulations) to converge to a solution. To illustrate the usefulness of parameter sensitivity analysis in reducing the computational effort associated with calibration, a similar multicriteria calibration run was made in which only 18 parameters were estimated, including the 15 sensitive parameters (vegc, rough, displa, rsmin, xla, depuv, deprv, xmopor, xmosuc, xmohyd, bee, skrat, solour, ssw, and tsw) and the three marginally sensitive parameters (fc, rootf, and rsw). (Note that in BATS the parameter xmowil is derived from other parameters and was therefore not included). In this run, the insensitive parameters (seasf, xlaio, sai, sqrtdi, deptv, albvgs, and albvgl) were prescribed to be at the BATS default values for semiarid regions [see Gupta et al., this issue, Table 1]. The goal is to demonstrate that if the parameters diagnosed as being insensitive using the MOGSA methodology are not calibrated, there is only a marginal degradation in the quality of the calibrated model performance, and the preferred range of parameters is changed only slightly.

The results of these two runs are compared in Figure 5. The results shown in light shade correspond to the Pareto set ob-
tained by the 25 parameter multicriteria calibration run and the results shown in black correspond to the Pareto set obtained by the 18 parameter multicriteria calibration run. The parameter estimates are shown on a normalized parameter plot (Figure 5a); the parameters and the three initial soil moisture contents are listed along the x axis, and the y axis corresponds to the parameter values plotted on a range normalized by the allowable maximum and minimum values for the parameters (Table 1). Each line from left to right across the plot corresponds to a different parameter estimate. The lightly shaded area in the background corresponds to the parameter ranges considered to be reasonable for the Sonoran Desert semiarid Tucson site (see Table 1); these reasonable ranges were not used to constrain the calibration runs but were instead used as postcalibration checks on the reliability and power of the calibration procedure: that is, a successful calibration run should converge to estimates that fall within these ranges. Figure 5b shows the trade-offs between the different criteria in the multicriteria space; the three response criteria are listed along the x axis, while the y axis corresponds to the criterion values (RMSEs) with preferred values toward the bottom of the plot. Each line going from left to right across the plot corresponds to a calibration result (and hence a different set of parameter values). Note that if a calibration result plots as a line, which falls entirely below (or above) that of a different result, the former can be said to be absolutely superior (inferior) in a multicriteria sense. However, if the two lines cross each other, then the results are noninferior to each other in a multicriteria sense.

It is important to note that all the solutions for both the 25 and the 18 parameter calibration runs fall within the reasonable bounds (Figure 5a). In the case of the 25 parameter calibration, some of the insensitive parameters (sqrtdi, deptv, albvgs, and albvgl) tend toward different locations from the fixed default values used in the 18 parameter calibration, and to be poorly constrained (i.e., they have a relatively wide spread). Highly sensitive parameters (e.g., vegc and xmpor) tend to be tightly constrained and have similar values in both calibration runs, while the less sensitive parameters (e.g., bee and fc) tend to be less constrained and vary more. This is entirely consistent with the expectation that the less sensitive parameters will have a less well defined optimal region.

Figure 5b shows that the 25 parameter calibration run is superior to the 18 parameter calibration run in a multicriteria sense; but only marginally so, the improvement in the RMSEs for the three system responses is less than 10%. However, the 18 parameter calibration run required only 4000 iterations (model simulations) to converge to a solution, a fivefold reduction from the 25 parameter calibration. These results suggest that the parameter sensitivity analysis using MOGSA is a sensible and viable way to reduce the dimensionality of the parameter estimation problem, thereby reducing the computational time required for calibration. The left column of Figure 6 shows a typical 10 day portion of the time series match obtained by the 18 parameter calibration run. For each system response, the dots indicate the observed data and the lightly shaded area represents the range of variation at each time step corresponding to the Pareto set of parameter estimates. The relative uncertainty of the daytime \( \Delta E \) trajectory is a consequence of the high degree of noise in the data. The statistics computed for each of the responses, as well as the scatterplots, are defined for the middle point of the trajectory set and are calculated for the entire year of data. The correlation between the simulated and the observed values is gratifying.

5. Summary and Conclusions

A companion paper [Gupta et al., this issue] describes the development of a multicriteria calibration procedure for determining the preferred range of the many parameters required in a complex, present-day LSS and illustrates the application of that procedure using the BATS model with data from two field sites. In practice, it is desirable to restrict the number of parameters to be calibrated in order to reduce the
computational effort. Further, the ability to make objective
determination of the relative sensitivity of modeled variables to
the specific parameters used in an LSS at a particular site has
a broader value as a basis for specifying the relative importance
of the physical processes represented in the model. Thus a
methodology is required which is consistent with multicriteria
parameter estimation methods and which is capable of deter-
mining the extent to which the modeled variables given by an
LSS are sensitive to the value of the individual parameters
used in it.

This need is addressed directly in this paper, by describing
the development of a novel, multicriteria approach to evaluate
parameter sensitivity, the MOGSA algorithm, which extends
and improves the regional sensitivity analysis single-criterion
method of Spear and Hornberger [1980] to multicriteria by
sorting the possible parameter sets via the notion of Pareto
ranking and by applying bootstrapping and sequential sam-
pling to ensure statistical robustness of the results. The method
enables the parameters to be ranked in terms of both absolute
and relative sensitivity. This procedure is then applied in the
context of the parameter estimation studies described in the
companion paper. First, the parameters that have significant
influence on the values of the modeled variables at the two
sites are determined and are distinguished from the param-
eters that have little influence and which might therefore be
safely omitted from the calibration run. The consequences of
omitting these parameters from the multicriteria calibration
were then investigated.

For both field sites considered in this study, the MOGSA
algorithm was effective in specifying the sensitivity of BATS-
calculated variables to the parameters used in the model. The
results provide insight into important parameters and associ-
ated processes in BATS at the two field sites. In both cases the
modeled variables were found to be sensitive to 12 parameters.
Five are related to vegetation, specifically vegetation cover,
aerodynamic roughness, minimum stomatal resistance, depth
of upper soil layer, and depth of root zone soil layer; six are
related to soil, specifically porosity, minimum soil suction, max-
imum hydraulic conductivity, Clapp and Hornberger “b” pa-
parameter, and ratio of soil thermal conductivity to that of loam;
and one is the initial water content of the surface layer. In both
cases the modeled variables were found to be insensitive to six
vegetation-related parameters, specifically stem area index, in-
verse square root of leaf dimension, light dependence of sto-
matal resistance, depth of the total soil layer, and the vegeta-
tion albedos for both shortwaves and longwaves.

However, important differences were found in the model
sensitivities at the two sites. These presumably reflect the prev-
alent climate and the nature of the vegetation at the sites.
Specifically, the parameters aerodynamic roughness, displace-
ment height, and soil color show high sensitivity at the semiarid
Tucson site but low sensitivity at the ARM-CART grassland
site, while the parameters difference between vegc and frac-
tional cover at 269 K, minimum leaf area index, ratio of roots
in the upper layer to roots in the lower layer, and ratio of field
capacity to saturated water content show low sensitivity at the
Tucson site but high sensitivity for ARM-CART. In general,
the sensitivity results for both sites are consistent with the

Figure 5. Results showing the benefits of using multicriteria parameter sensitivity analysis to reduce pa-
rameter dimensionality before calibration of the BATS model to the Tucson semiarid site. The full 25
parameter calibration results are shown in black and the reduced 18 parameter calibration results are shown
in light shade. (a) Parameter estimates plot: each horizontal line represents a member of the Pareto solution
set. The medium shade area indicates the reasonable range of parameter variability for semiarid regions. (b)
Criterion values plot: each horizontal line represents a member of the Pareto solution set. Lines closer to the
base of the plot indicate superior solutions.
model formulation and the physical properties of the different environments. For the Tucson site, six of the eight most sensitive parameters are related to soil properties, while for the ARM-CART site, five of the eight most sensitive parameters are related to vegetation properties.

Comparison between the results of calibration runs for the Tucson semiarid site, in which all the parameters were optimized and in which only a subset of sensitive parameters were optimized, was consistent with expectations. There was only limited degradation in the quality of the correspondence between modeled and observed values (the RMSE for the three criteria increased by less than 10%), and the ranges of preferred values for the parameters that were optimized in both runs were mutually consistent. However, the calibration run for the sensitive parameter subset required only 4000 iterations (model simulations) to converge to a solution, a fivefold reduction over the full parameter calibration.

Thus in summary, the multicriteria methodology MOGSA developed in this study proved successful when applied to evaluate the sensitivity of model-calculated variables to individual parameters for a complex LSS at two distinctly different field sites. The sensitivity results for both sites were found to be consistent with the physical properties of the different environments, thereby supporting the reasonableness of the model formulation. Further, the analysis provided evidence on the importance of parameters (and the associated processes) operating at the two sites, and when the parameters identified as insensitive were omitted from the optimization process, there was little degradation in the quality of the model description and little change in the preferred range of the remaining parameters. The MOGSA algorithm has been written to be generally applicable to any model, and the code is available from the authors on request (e-mail: hoshin@hwr.arizona.edu).

Figure 6. Model performance with the 18 parameter multicriteria calibration parameter estimates compared to the observed data. The time series plots on the left show a typical 10 day period; the lightly shaded region represents the range of output simulations corresponding to the range of the Pareto parameter estimates, and the dots represent the observed data. The scatterplots on the right compare the entire period of data with the midpoints of the model-simulated ranges.
Acknowledgments. Partial financial support for this research was provided by the National Science Foundation (grants EAR-9415347 and EAR-9418147), the Hydrologic Research Laboratory of the National Weather Service (grants NA47WH04008 and NA57WH0575), and the National Aeronautics and Space Administration (NASA-EOS grant NAGW2425). The financial support provided to L. Bastidas by the Escuela Politecnica Nacional (Quito-Ecuador) and the Fulbright-Laspau Program are gratefully acknowledged. Special thanks are due to Helene Unland and Jim Washburne for providing the data, to Jim Broermann and Dan Brathwaite for providing assistance with computational resources, and to Corrie Thies for proofreading.

References


L. A. Bastidas, H. V. Gupta, W. J. Shuttleworth, S. Sorooshian, and Z. L. Yang, Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721. (hoshin@hwr.arizona.edu).

(Received September 2, 1998; revised February 8, 1999; accepted February 22, 1999.)