A fully multiple-criteria implementation of the Sobol′ method for parameter sensitivity analysis

Rafael Rosolem,[1] Hoshin V. Gupta,[1] W. James Shuttleworth,[1] Xubin Zeng,[2] and Luis Gustavo Gonçalves de Gonçalves[3]

Received 3 June 2011; revised 2 February 2012; accepted 4 February 2012; published 6 April 2012.

[1] We present a novel rank-based fully multiple-criteria implementation of the Sobol′ variance-based sensitivity analysis approach that implements an objective strategy to evaluate parameter sensitivity when model evaluation involves several metrics of performance. The method is superior to single-criterion approaches while avoiding the subjectivity observed in "pseudo" multiple-criteria methods. Further, it contributes to our understanding of structural characteristics of a model and simplifies parameter estimation by identifying insensitive parameters that can be fixed to default values during model calibration studies. We illustrate the approach by applying it to the problem of identifying the most influential parameters in the Simple Biosphere 3 (SiB3) model using a network of flux towers in Brazil. We find 27–31 (out of 42) parameters to be influential, most (∼78%) of which are primarily associated with physiology, soil, and carbon properties, and that uncertainties in the physiological properties of the model contribute most to total model uncertainty in regard to energy and carbon fluxes. We also find that the second most important model component contributing to the total output uncertainty varies according to the flux analyzed; whereas morphological properties play an important role in sensible heat flux, soil properties are important for latent heat flux, and carbon properties (mainly associated with the soil respiration submodel) are important for carbon flux (as expected). These distinct sensitivities emphasize the need to account for the multioutput nature of land surface models during sensitivity analysis and parameter estimation. Applied to other similar models, our approach can help to establish which soil-plant-atmosphere processes matter most in land surface models of Amazonia and thereby aid in the design of field campaigns to characterize and measure the associated parameters. The approach can also be used with other sensitivity analysis procedures that compute at least two model performance metrics.

Citation: Rosolem, R., H. V. Gupta, W. J. Shuttleworth, X. Zeng, and L. G. G. de Gonçalves (2012), A fully multiple-criteria implementation of the Sobol′ method for parameter sensitivity analysis, *J. Geophys. Res.*, *117*, D07103, doi:10.1029/2011JD016355.

1. Introduction and Background

[2] Sensitivity analysis (SA) is the study of how uncertainty in the output of a model can be apportioned to different sources of uncertainty in the model input factors (e.g., input forcing data, parameters, etc.) [*Saltelli et al.*, 2010]. When applied to parameter estimation (i.e., model calibration) [*Saltelli et al.*, 1999; *Bastidas et al.*, 1999, 2006; *Demarty et al.*, 2004; *Hall et al.*, 2005; *Tang et al.*, 2007a, 2007b; *van Werkhoven et al.*, 2008; *Prihodko et al.*, 2008; *Rosero et al.*, 2010], the objective of SA becomes to identify the most important parameters that, if optimally determined, would lead to the greatest reduction in the variance of the output of interest [*Saltelli*, 2002a]. Sensitivity Analysis is an important tool used in environmental modeling [*U.S. Environmental Protection Agency* (*EPA*), 2009], contributes to the understanding of structural characteristics of a model, and can also help to identify model parameters having minimal effect on the model output (that can therefore be set to default values during simulation).

[3] According to *Campolongo et al.* [2000], local, one at-a-time (OAT) sensitivity analysis methods, although widely applied among scientists, are not appropriate for nonlinear models because they fail to account for potential interactions among different factors. An approach commonly used to study parameter sensitivity of catchment and land-atmosphere interaction models is regionalized sensitivity

[1]Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA.
[2]Department of Atmospheric Sciences, University of Arizona, Tucson, Arizona, USA.
[3]Centro de Previsão de Tempo e Estudos Climáticos, Instituto Nacional de Pesquisas Espaciais, Cachoeira Paulista, Brazil.

analysis (RSA), which uses Monte Carlo sampling to partition the parameter space into behavioral (acceptable) and nonbehavioral (nonacceptable) regions [*Hornberger and Spear*, 1981; *Bastidas et al.*, 1999, 2006; *Demarty et al.*, 2004; *Rosolem et al.*, 2005; *Prihodko et al.*, 2008]. However, RSA does not quantify the extent to which a parameter affects the variance of model output and relies upon subjective decisions to partition between behavioral and nonbehavioral solutions [*Bastidas et al.*, 1999, 2006; *Demarty et al.*, 2004]; the metric used for partitioning is sometimes graphical [*Wagener and Kollat*, 2007] and sometimes statistical, in the latter case with thresholds predefined by the user [*Bastidas et al.*, 1999; *Prihodko et al.*, 2008]. Moreover, a limitation to RSA methods is that once the partitioning is made, the method ignores potential variations of the output within the class of the acceptable values [*Saltelli*, 2002a]. Attempts to identify influential parameters have also been conducted through the analysis of "signature measures" [*Yilmaz et al.*, 2008] but the approach can also become subjective, in the sense that a standardized list of signature measures is not available for different applications (e.g., hydrological versus land surface models) and for different regions on Earth, driven by distinct climatic forcing [e.g., *Nemani et al.*, 2003; *Jolly et al.*, 2005].

[4] In contrast, variance-based (also known as importance measure) methods have recently become the preferred approach for sensitivity analysis across a wide range of applications [*Saltelli*, 2002a, 2002b; *Saltelli et al.*, 1999, 2006, 2010; *Frey*, 2002; *Frey and Patil*, 2002; *EPA*, 2009], including hydrological and land surface modeling [*Tang et al.*, 2007a, 2007b; *van Werkhoven et al.*, 2008; *Rosero et al.*, 2010]. These methods provide a factor-based decomposition of the output variance, and provide robust quantitative results regardless of the model used, can deal with nonlinearities, and are able to account for interactions among different factors; the price for these properties is that they are computationally demanding. These methods provides not just a rank ordering of key factors, but also a quantitative measure of sensitivity so that it is possible to evaluate how important a given factor is in relation to others. Among these methods, the Sobol′ variance-based analysis [*Sobol′*, 1993] has become one of the most widely used in environmental modeling [*Saltelli*, 2002b; *Saltelli et al.*, 1999, 2006; *Campolongo et al.*, 2000; *Tang et al.*, 2007a, 2007b; *van Werkhoven et al.*, 2008; *Rosero et al.*, 2010].

[5] A major drawback of previous evaluation and sensitivity analysis studies of hydrology and land surface interaction models, regardless of the analysis method applied, was their failure to properly consider the fully multiple-criteria (and often noncommensurable) nature [*Gupta et al.*, 1998] of model performance. While several studies have employed multiple objective functions (model performance criteria), the ultimate sensitivity analysis was typically carried out for each objective function individually, making it difficult to objectively rank the parameters in regard to their relative importance in the model or to decide which ones should be selected for calibration. For example, *Liu et al.* [2004] demonstrate that parameters can show varying sensitivities for different criteria, such that a "globally sensitive parameter" is not necessarily sensitive to all individual criteria. Consequently, the way this is usually addressed is to define sensitivity thresholds on each criterion (for variance-based analysis, "sensitive" is usually chosen to correspond to ≥1% of total variance, and "high sensitivity" corresponds to ≥10% [*Tang et al.*, 2007a, 2007b; *van Werkhoven et al.*, 2008]); however, the selection of these thresholds can sometimes be arbitrary, and introduces subjectivity into the analysis.

[6] Here we present a fully multiple-criteria implementation of the Sobol′ SA method based on the notion of multiple-criteria (Pareto) ranking. We demonstrate its value by analyzing parameter sensitivities of the Simple Biosphere 3 (SiB3) model [*Baker et al.*, 2003, 2008] with regard to sensible and latent heat fluxes, and net ecosystem exchange of CO_2. The results indicate which model parameters are most influential in controlling simulations of these fluxes at eight flux tower sites located mainly in the Amazon basin, and such information can help to define data measurement priorities for future field studies. The results presented here are designed to illustrate the sensitivity analysis methodology. Further investigation of sensitivity for all flux tower sites and subsequent calibration of the SiB3 model parameters are described by *Rosolem et al.* [2012].

2. Variance-Based Sensitivity Analysis

2.1. The Sobol′ Method

[7] Sobol′ analysis estimates, via approximate Monte Carlo integrations, what fraction of the variability of some entity can be attributed to changes in the values of various factors that control the value of that entity [*Sobol′*, 1993; *Saltelli et al.*, 1999, 2006; *Hall et al.*, 2005]. Consider the problem formulation:

$$y = f(\theta) = f(\theta_1, \theta_2, \cdots, \theta_k) \tag{1}$$

where y is the vector of metrics used to evaluate model performance (e.g., the root-mean-square errors of model fit to different model simulated fluxes), and $\theta = \{\theta_1, \theta_2, \cdots, \theta_k\}$ is the vector of k model factors (e.g., parameters) that are believed to control the behavior of the model and hence its performance as indicated by y. Our interest is in knowing how much of the total variance $V(y)$ in y, can be explained by variability in the factors θ. The Sobol′ approach computes this by decomposing the function $f(\theta_1, \theta_2, \ldots, \theta_k)$ into terms of increasing dimensionality, such that each successive dimension represents increasing degrees of interactions among the parameters,

$$f(\theta_1, \theta_2, \ldots, \theta_k) = f_0 + \sum_{i=1}^{k} f_i(\theta_i) + \sum_{1 \le i \le j \le k} f_{ij}(\theta_i, \theta_j) + \ldots + f_{1,2,\cdots,k}(\theta_1, \theta_2, \ldots, \theta_k), \tag{2}$$

on the basis of which the total variance $V(y)$ in y can be shown to be composed of the terms

$$V(y) = \sum_i V_i + \sum_{i<j} V_{ij} + \rightleftharpoons + V_{1,2,\cdots,k} \tag{3}$$

where V_i is the portion of $V(y)$ contributed by factor θ_i, and V_{ij} is the portion contributed by interactions between factors θ_i and θ_j, etc. The fractional contribution of the ith factor to

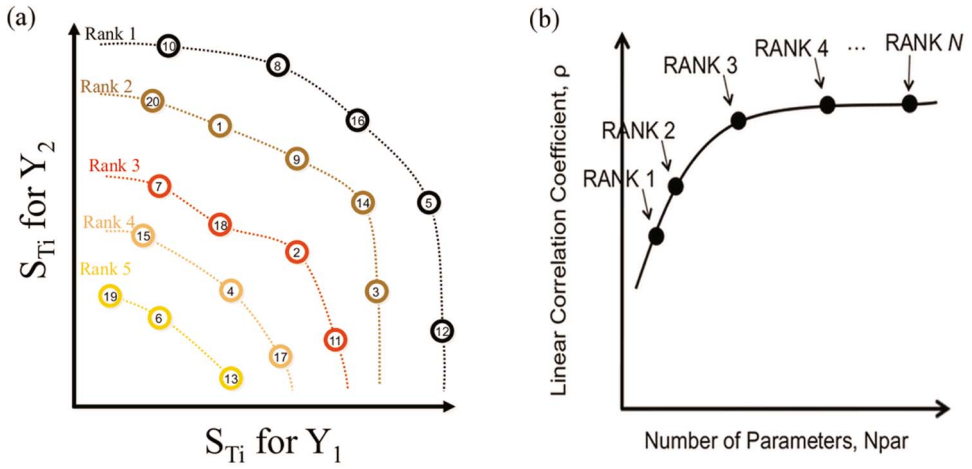

Figure 1. Hypothetical example of the approach used in this study: (a) Pareto rank calculated for a pair of Sobol′'s total order indices (S_{Ti}) associated with objective functions Y_1 and Y_2. The numbers inside circles represent hypothetical factor numbers (i.e., parameters). The color coding indicates the parameters belonging to a particular Pareto rank (for clarity all parameters with same Pareto rank are also linked together by a surrounding dotted line). (b) Relationship between linear correlation coefficient (ρ) and number of identified parameters (N_{par}) based on alternative approach (see section 2.2 for details).

the total output variance $V(y)$ is called the first-order sensitivity index S_i (sometimes called "main effect"):

$$S_i = \frac{V_i}{V(y)}. \tag{4}$$

[8] Similarly, the second-order indices S_{ij}, and total order indices S_{Ti} are defined as

$$S_{ij} = \frac{V_{ij}}{V(y)} \tag{5}$$

$$S_{Ti} = 1 - \frac{V_{\sim i}}{V(y)} \tag{6}$$

where $V_{\sim i}$ is the average variance obtained when all of the factors except for θ_i are allowed to vary (θ_i is kept fixed at one value). The total order index S_{Ti} (equation (6)) therefore represents the total contribution of factor θ_i to the total variance $V(y)$ through both direct and indirect effects.

[9] Software to perform the computations is available at http://simlab.jrc.ec.europa.eu/. To compute these indices, the model performance vector y must be assessed at a sufficiently large number of randomly selected parameter locations in the feasible parameter space for the estimates to be stable (note also that the random factor variations must be uniformly distributed independently of each other). The main disadvantage of the Sobol′ approach (as with other variance-based methods) is its large computational requirement.

[10] In summary, the Sobol′ method is typically used to produce two matrices for each metric y^o in the model performance vector y; one matrix (a vector) specifies the total order sensitivity index S^o_{Ti} and the other specifies the first-order sensitivity indices S^o_i. The value of each index can vary from 0 to 1. When there is only one model performance metric the factors can easily be ranked in order of sensitivity by examining S_i and S_{Ti}, while the strength of interactions can be assessed by examining the matrix of the total order sensitivity indices when subtracting the magnitudes of the first-order sensitivity indices [*Saltelli et al.*, 2005]. To select the "influential" factors, the user then specifies a cutoff threshold (usually applied to S_{Ti}) below which the contribution of a factor to variability in the model performance metric is judged to be inconsequential. However, when there are several model performance metrics, the typical approach is to perform such analysis on each performance metric separately; in this case the user must specify cutoff thresholds for each of the metrics (although, usually a cutoff value is defined commonly to all metrics). While the approach then indicates the relative sensitivity of each metric to each factor, when (as is desirable) metrics tend to be orthogonal [*Gupta et al.*, 1998], different metrics will inevitably be sensitive to different factors and specification of the "most and least" influential factors can become somewhat subjective.

2.2. Fully Multiple-Criteria Implementation

[11] To achieve a more objective ranking of multiple-criteria factor sensitivity, we propose a multiple-criteria screening approach based on the Pareto ranking concept [*Goldberg*, 1989]. This enables definition of groups of factors having relatively stronger and weaker contributions to overall model performance; this can help, for example, in establishing which factors to simultaneously optimize regardless of their relative sensitivity to individual model performance metrics. The approach assigns a "group rank" to each factor based on a multiple-criteria assessment of its individual total order sensitivities. This group ranking then permits a more objective analysis of how many of factors contribute significantly to the overall variability in model performance.

[12] The approach is described below; for simplicity we discuss a hypothetical case of only two performance metrics (Y_1 and Y_2) for a problem having 20 factors (Figure 1). Of course, more than two metrics can be easily handled.

2.2.1. Step 1: Assign Pareto Rank Group Numbers to Each Factor

[13] For each factor, assign a Pareto rank r [see *Gupta et al.*, 1998] on the basis of simultaneous maximization of the total order sensitivity indices S_{T1} and S_{T2} for Y_1 and Y_2, respectively. The maximum Pareto rank $R_{\max}$ will of course be less than or equal to the number of factors k.

[14] By way of our example, the color coding in Figure 1a shows that each of the 20 factors occupies a location in the S_{T1} and S_{T2} space (the number in each circle indicates the factor number). Factors having the strongest influence on metric Y_1 plot closer to the right and on metric Y_2 plot closer to the top, while those strongly influencing both plot toward the top right. Factors having the same Pareto rank are shown linked by a dotted line. A factor is assigned rank 1 if there is no other factor plotting in the upper right quadrant to it. Factors are assigned rank 2 if they become Pareto optimal when all rank 1 factors are removed from the plot, and so on, until all factors have been assigned a rank group number.

2.2.2. Step 2. Create a Control Sequence

[15] Create a "control sequence" of N_C points randomly sampled in the total factor space. Ideally one should use a stratified sampling method, such as Latin hypercube sampling (LHS) [*McKay et al.*, 1979] to ensure uniform coverage. Compute the vector of performance metrics y^C at each sampled location.

2.2.3. Step 3: Create Rank-to-r Factor Sensitivity Groups

[16] Define a rank-to-r factor sensitivity group as containing all factors belonging to rank groups up to and including rank r (i.e., ranks 1, 2, …, r); hence, as r is increased, the number of included factors also increases till all of them are accounted for.

2.2.4. Step 4: Create Rank-to-r Factor Sensitivity Samples

[17] For each rank-to-r factor sensitivity group create a sequence of N_C points randomly sampled in the total factor space such that all the factors belonging to that group are set to their values in the control sequence while the remaining (out of group) factors are fixed to their default values. For each group, compute the corresponding vector y^r of performance metrics at each sampled location.

[18] The consequence of this is a single control sequence and $R_{\max}$ sequences of rank-to-r factor sensitivity samples.

2.2.5. Step 5: Select a Pareto Rank Threshold to Determine the Influential Factors

[19] To determine which factors should be deemed influential, proceed as follows. For each rank-to-r factor sensitivity group compute the linear correlation coefficient ρ^r between y^r and y^C for each performance metric, and then compute the minimum $\rho^r_{\min}$ across all of the metrics. This coefficient $\rho^r_{\min}$ indicates how much of the variability in the control sequence is explained by variability in all factors belonging to groups having rank r or lower. Generate a plot of $\rho^r_{\min}$ versus r (see Figure 1b); note that as r varies from 1 to $R_{\max}$ the number of included factors increases and $\rho^r_{\min}$ approaches 1.0. However, if some of the parameters do not significantly contribute to variability in model performance, $\rho^r_{\min}$ will effectively become 1 for some $r = r^* < R_{\max}$. Notice that this value can be selected by a test of significance for $\rho^r_{\min} \sim 1$, considering the sample size N_C used.

[20] Of course, one can also examine each of the "rank correlation lines" generated for each model performance metric taken separately, thereby obtaining detailed analysis about different aspects of model behavior (e.g., different output fluxes); while this approach is analogous to the one-metric-at-a-time analysis applied by others, it allows for a more robust and objective statistical evaluation as indicated above.

3. Case Study: Land Surface Modeling of Flux Sites in the Amazon Basin

[21] The influence on global climate on the Amazon region selected for this study is unquestionable [*Nobre et al.*, 1991; *Cox et al.*, 2000; *Betts et al.*, 2004; *Malhi et al.*, 2009], and therefore, an understanding of ecological and hydrometeorological interactions in this region is of great importance. Because of its spatial extent (on the order of ~5.4 million km^2), the Amazon basin contributes substantially to regional and global hydrological cycles [*Salati*, 1987; *Shuttleworth*, 1988; *Brubaker et al.*, 1993; *Eltahir and Bras*, 1994; *Werth and Avissar*, 2004] and is a major participant in the global carbon budget [*Malhi and Grace*, 2000; *Houghton et al.*, 2001; *Malhi and Davidson*, 2009].

[22] In general circulation models the soil-vegetation-atmosphere interface is described by land surface parameterization schemes (LSPs) [*Sellers et al.*, 1997; *Pitman*, 2003; *Sakaguchi et al.*, 2011]. Some of these LSPs are highly complex model containing tens of parameters to be specified by the user. As part of the Large-Scale Biosphere-Atmosphere Experiment in Amazonia (LBA) [*LBA Science Planning Group*, 1996; *Avissar and Nobre*, 2002; *Keller et al.*, 2004, 2009], the Data-Model Intercomparison Project (LBA-DMIP; http://www.climatemodeling.org/lba-mip/) seeks to improve the description of land surface processes in this region by analyzing the performance of several LSPs relative to observations. Here we apply the multiple-criteria sensitivity analysis approach described above to evaluate the relative importance of various parameters in the SiB3 model with regard to the strength of their influence on the behaviors of sensible and latent heat fluxes, and net ecosystem exchange of CO_2 simulated by the model.

3.1. The Simple Biosphere 3 (SiB3) Land Surface Model

[23] The land surface parameterization scheme used in this study is a version of the third generation [*Baker et al.*, 2003] of the SiB land surface model [*Sellers et al.*, 1986, 1996a, 1996b] that includes a prognostic representation of canopy air space properties [*Vidale and Stöckli*, 2005], a user-determined number of soil layers, soil representation based on the Community Land Model [*Dai et al.*, 2003], and a new soil water stress parameterization [*Baker et al.*, 2008]. The model is driven by meteorological data (air temperature and specific humidity, horizontal wind speed, barometric surface pressure, downward longwave and shortwave radiation, and rainfall), and simulates various aspects of biosphere-atmosphere interaction (momentum, heat, water exchange, biogeochemistry, etc.) at the interface between the land surface and the atmosphere [*Randall et al.*, 1996]. Our version has been slightly modified to incorporate an alternative soil respiration submodel and a more realistic

Table 1. Geographical and Main Meteorological Characteristics of the LBA-DMIP Sites Used in This Study[a]

Site	Site Name	Latitude	Longitude	Elevation (m)	Reference Height (m)	Data Period	Annual Rainfall (m)	Annual Temperature (°C)	Dry Season (Length)	Biome Type
K34	Manaus Km34	2°36′33″S	60°12′33″W	130	50	2002–2005	2.4	25.6	Jul–Sep (3 months)	EBF
K67	Santarém Km67	2°51′24″S	54°57′32″W	130	63	2002–2004	1.6	25.3	Jul–Nov (5 months)	EBF
K83	Santarém Km83	3°01′05″S	54°58′17″W	130	64	2001–2003	1.7	25.9	Jul–Nov (5 months)	EBF
RJA	Reserva Jarú	10°04′59″S	61°55′51″W	191	60	2000–2002	2.3	25.1	Jun–Oct (5 months)	EBF
BAN	Bananal Island	9°49′28″S	50°09′33″W	120	40	2004–2006	1.7	26.4	May–Sep (5 months)	DBF
PDG	Reserva Pé de Gigante	21°37′10″S	47°39′00″W	690	21	2002–2003	1.3	22.5	Apr–Sep (6 months)	SAV
K77	Santarém Km77	3°01′11″S	54°53′40″W	130	18	2001–2005	1.6	26.3	Jul–Nov (5 months)	CROP
FNS	Fazenda Nossa Senhora	10°45′54″S	62°21′13″W	306	8.5	1999–2001	1.7	24.7	Jun–Oct (5 months)	CROP

[a]Biome types: evergreen broadleaf forest (EBF), deciduous broadleaf forest (DBF), savanna (SAV), and cropland/pastureland (CROP). References: K34 [*Araújo et al.*, 2002], K67 and K83 [*Saleska et al.*, 2003; *Miller et al.*, 2004; *Hutyra et al.*, 2007], RJA and FNS [*von Randow et al.*, 2004], BAN [*Borma et al.*, 2009], PDG [*da Rocha et al.*, 2002], and K77 [*Sakai et al.*, 2004; R. K. Sakai et al., Relating carbon and water exchange over an intensive agriculture field in the Amazon crop to locally and remotely sensed phenological measurements, submitted to *Agricultural and Forest Meteorology*, 2010].

physically based estimate of cloud fraction (based on original work by *Rosolem et al.* [2010] and applied to Amazonian conditions). It also includes a factor to correct for systematic bias in nighttime NEE.

3.2. LBA Study Sites and Data

[24] Continuous hourly meteorological forcing data for eight sites in Brazil were provided by the LBA-DMIP project, and validation data comprising sensible heat flux (H), latent heat flux (λE) and net ecosystem exchange of CO_2 (NEE) were obtained from the principal investigators responsible for each site. The sites are part of *Brasilflux*, the eddy covariance flux tower network in Brazil (N. Restrepo-Coupe et al., What drives the seasonality of productivity across the Amazon basin? A cross-site analysis of eddy flux tower measurements from the Brasil flux network, submitted to *Agricultural and Forest Meteorology*, 2010) and include four evergreen broadleaf forests (EBF), a deciduous broadleaf forest (DBF), a savanna biome (SAV), and two cropland/pasture biomes (CROP) following IGBP classification. Table 1 summarizes the main characteristics of each field site and includes individual site references; for details on climatological characteristics, see *Rosolem et al.* [2008] and *da Rocha et al.* [2009]. For most sites data are available for three consecutive years between 1999 and 2005, but the period of collection varies.

[25] To reduce uncertainties in the measurements, simple quality control procedures were applied to the validation data (fluxes). A visual inspection was first made to remove unrealistic outliers in the fluxes. Hourly NEE values were then calculated by adding the eddy flux (F_{CO_2}) and storage (S_{CO_2}) observations [*Wofsy et al.*, 1993], except at the PDG and FNS sites where no storage measurements were made and NEE was necessarily set equal to eddy flux alone. At sites where carbon storage measurements were made there were sometimes periods with missing data. When this occurred, the missing value of S_{CO_2} at a given hour of the day was estimated to be the monthly average hourly value of S_{CO_2} at the same hour. If the hourly average value of S_{CO_2} for that month could not be calculated (no measurements available), the equivalent hourly value in the mean annual diurnal cycle was used. This approach is similar to the one used by *Hutyra et al.* [2008]. Nighttime NEE values were also filtered for reliability at some sites on the basis of thresholds of friction velocity (u*). For the Manaus K34 site, nighttime NEE observations were considered unreliable when $u^* < 0.15$ m s^{-1}, and at Tapajós sites (K67 and K83) when $u^* < 0.22$ m s^{-1}. The selection of the u* threshold values for these sites were based on previous site-specific studies [*Araújo et al.*, 2002; *Kruijt et al.*, 2004; *Saleska et al.*, 2003].

3.3. Time-Varying Inputs

[26] Phenology information, such as leaf area index (LAI) and fraction of the absorbed photosynthetically active radiation (FPAR), was derived from a prognostic phenology model [*Stöckli et al.*, 2008] that assimilates Moderate Resolution Imaging Spectroradiometer (MODIS) LAI/FPAR products at a monthly time scale. Monthly LAI was provided by LBA-DMIP and monthly FPAR by R. Stöckli (unpublished data, 2009). Additional time-varying inputs, such as green fraction of vegetation, were acquired from the global 1982–2001 European Fourier-adjusted and interpolated normalized difference vegetation index (EFAI-NDVI; 10 day temporal and 0.1° spatial resolution) derived from advanced very high resolution radiometer Pathfinder NDVI data. Biophysical land surface inputs were derived from EFAI-NDVI by applying simple empirical relationships between satellite radiometry and vegetation physiology; see *Stöckli and Vidale* [2004]. For these additional time-varying inputs, a climatological monthly mean was used in this study.

3.4. Parameters Used as Factors in the Sensitivity Analysis

[27] Our modified version of SiB3 has 42 parameters (see Table 2), all of which were used as factors in this study (time-varying inputs, such as LAI and FPAR, were not included in the sensitivity analysis). Parameters ranges highlighted in bold are specified in the form of a multiplier as opposed to actual parameter value; these correspond to parameters whose ranges depend on the values of other parameters. This procedure helps to ensure that the independence assumption of the Sobol′ methodology is met; *Prihodko et al.* [2008] applied a similar approach to SiB2.5. Note that for those parameters, we test the sensitivity of the multiplier rather than that of the actual parameter. Seven such relationships exist as indicated in Table 2, where the '*M*' terms indicate the multipliers applied in each case.

[28] Feasible ranges for each parameter (minimum and maximum values) were selected to encompass values for a

Table 2. Summary of SiB3 Parameters and Feasible Ranges[a]

Parameter	Description	Units	Minimum	Maximum
	Morphological Properties			
z2	Canopy top height	m	0.3	45.0
z1* [$z_1 = M_1 \cdot z_2$]	Canopy bottom height	m	**0.025**	**0.800**
z0d* [$z_{0d} = M_2 \cdot z_2$]	Canopy roughness length	m	**0.05**	**0.20**
zp_disp* [$zp\_disp = M_3 \cdot z_2$]	Canopy zero plane displacement height	m	**0.65**	**0.85**
vcovr	Vegetation cover fraction		0.4	1.0
chil	Leaf-angle distribution factor		−0.3	0.3
	Optical Properties			
tran(1,1)	Live leaf transmittance visible band		0.0400	0.0800
tran(1,2)	Dead leaf transmittance visible band		0.0005	0.2500
tran(2,1)	Live leaf transmittance near infrared band		0.0100	0.4000
tran(2,2)	Dead leaf transmittance near infrared band		0.0005	0.4000
ref(1,1)	Live leaf reflectance visible band		0.0500	0.1500
ref(1,2)	Dead leaf reflectance visible band		0.1000	0.4000
ref(2,1)	Live leaf reflectance near infrared band		0.4000	0.6000
ref(2,2)	Dead leaf reflectance near infrared band		0.3500	0.6000
soref(1)	Soil reflectance visible		0.05	0.15
soref(2)* [$soref(2) = M_4 \cdot soref(1)$]	Soil reflectance near infrared		**1.1**	**2.5**
	Physiological Properties			
vmax0	Maximum Rubisco capacity at canopy top	mol m^{-2} s^{-1}	2.50E-5	1.50E-4
effcon	Intrinsic quantum efficiency	mol mol^{-1}	0.03	0.15
gradm	Stomatal slope factor		3	18
binter	Minimum stomatal conductance	mol m^{-2} s^{-1}	1.00E-8	0.04
atheta	Photosynthesis coupling coefficient		0.5	1.0
btheta	Photosynthesis coupling coefficient		0.5	1.0
trda	Slope of high temperature inhibition function (leaf respiration)	K^{-1}	0.1	2.0
trdm* [$trdm = M_5 \cdot trop$]	One half point of high temperature inhibition function (leaf respiration)	K	**1.01**	**1.20**
trop	Temperature coefficient in photosynthesis-conductance submodel	K	283.16	308.16
respcp	Respiration fraction of Vm		0.005	0.030
slti	Slope of low-temperature inhibition function (Vm)	K^{-1}	0.1	2.0
shti	Slope of high-temperature inhibition function (Vm)	K^{-1}	0.1	2.0
hlti	One half point of low-temperature inhibition function (Vm)	K	273.16	298.16
hhti* [$hhti = M_6 \cdot hlti$]	One half point of high-temperature inhibition function (Vm)	K	**1.01**	**1.20**
	Soil Properties			
bee	Soil wetness exponent		2.5	19.0
phsat	Soil water potential at saturation	m	−0.80	−0.03
satco	Soil hydraulic conductivity at saturation	m s^{-1}	1.00E-6	1.00E-4
poros	Soil water content at saturation (porosity)	m^3 m^{-3}	0.2	0.8
wssp	Soil water stress curvature parameter		0	1
kroot* [$kroot = M_7/scalez$]	Root density extinction coefficient		**0.01**	**0.20**
scalez	Factor to determine exponential distribution of soil depths		0.01	0.10
	Parameters Associated with CO_2 Components			
wopt	Soil moisture percent at maximum soil respiration rate	%	30	80
wsat	Parameter for soil respiration at saturation		0.3	0.8
zm	Skewness exponent for soil respiration submodel		−2	1
respref	Reference soil respiration rate	μmol m^{-2} s^{-1}	1	10
alpha	Multiplier to account for systematic bias on nighttime NEE components		0	1

[a]The sensitivity of parameters followed by an asterisk is analyzed on the basis of its multipliers (i.e., the *M* term in each parameter's equation). Although description and units refer to actual parameters in SiB3, parameter values in bold represent the multiplier values (instead of actual parameter).

broad range of biome types and defined following *Prihodko et al.* [2008] and recognizing the ranges proposed by the Land Data Assimilation System for these biome types (http://ldas.gsfc.nasa.gov/gldas/GLDASmapveg.php). Note that the parameter ranges defined in Table 2 are common to all biomes. This choice is deliberate since our interest in this study is to compute the sensitivity of a given parameter across all biome types; hence the parameter ranges encompass the natural ecosystems for the study region.

3.5. Soil Moisture Initialization

[29] To maintain consistency in the model simulations, careful soil moisture initialization was carried out for each parameter set sampled. Testing showed very little difference between a strategy of initializing soil moisture by using a single year of data repeatedly, versus using the entire time series of forcing data repeatedly. Consequently, we used the former approach to "spin up" the model, resulting in significant computational savings. For each site, the spin-up year was selected (from those for which forcing data were available) to be the one most climatologically similar to the year prior to the period for which model simulations were to be conducted [*Rosolem et al.*, 2008]. Consistent with LBA-DMIP protocol, the spin-up year was applied repeatedly until the difference in total column soil moisture from year to year changed by less than 0.1%.

3.6. Sensitivity Analysis Procedure

[30] The 42 parameters defined in Table 2 were used as factors in the sensitivity analysis. To evaluate model performance, we computed three metrics at each field site, the root-mean-square error (RMSE) between model simulated and observed values of H, λE, and NEE, where RMSE is defined as

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{m}\left(F_{s,t} - F_{o,t}\right)^2}{m}} \tag{7}$$

and $F_{s,t}$ and $F_{o,t}$ are the simulated and observed hourly fluxes of energy H, λE, or NEE calculated over the entire time series (when observations available). The method was applied to each site, with the parameters sampled randomly at ~45,000 locations uniformly distributed throughout the feasible space. *Saltelli et al.* [2005] shows that the pairs of first and total order indices can be computed at a cost of $(k + 2)N$ model simulations, where k is the number of factors ($k = 42$ in our case), and N should typically be between 500 and 1000. Hence, our sample size may be considered conservative ($N \sim 1000$). For the occasional situation where (because of truncation errors) slightly negative indices were obtained, these values were set to zero [see *Tang et al.*, 2007b].

[31] To compute the linear correlation coefficient ρ^r used to establish which parameters are to be deemed sensitive in the multiple criteria sense we used a sample size $N_C = 1000$; we also tested the effect of using larger sample sizes ($N_C = 2000$ and $N_C = 4000$) and found no significant impact on the results.

4. Results

4.1. Total Parameter Sensitivities and the Contributions of Direct and Indirect Effects

[32] Sobol′ total order indices (S_{Ti}) calculated for each SiB3 parameter at each site are summarized in Figure 2. Assuming the conventional variance cutoff values of 1% and 10%, the values indicate that the simulated fluxes show significant sensitivity to 13–20 and 3–8 of the model parameters tested, respectively (with level of sensitivity varying from site to site). Similar results were reported for SiB2.5 at a temperate site [*Prihodko et al.*, 2008] and for SiB2 at the K83 site [*Rosolem et al.*, 2005] using the RSA approach. The fact that so many of the parameters show lack of sensitivity in this and other studies suggests that the SiB model may be overparameterized [*Saltelli et al.*, 2006] with respect to the simulation of energy and carbon fluxes in the region (the parameters may, of course, be important for other aspects of model performance). To check this fact, we performed a similar application of the Sobol′ approach using runoff, ground heat flux, and canopy heat storage flux, for which observational data are not available, as individual criteria; to enable this, the RMSE was calculated with respect to the output values provided by a control simulation using the default parameter values. In summary, most of soil property parameters were found to be highly sensitive to runoff at all sites (especially, *bee*, *satco*, *poros*, and *scalez*), and *z*2 and *vcovr* (plus *poros*, for ground heat flux analysis) were found to be the most influential parameters for both ground heat flux and canopy heat storage flux (results not shown).

[33] In general the parameters that dominate the list in terms of influence are related to physiology, soil and carbon properties. This includes the physiology parameters *vmax0* (maximum Rubisco enzyme capacity at canopy top), *gradm* and *binter* (slope and intercept coefficients in the Ball-Berry-Collatz equation), and *hlti* and *hhti* (low- and high-temperature half-inhibition points for internal temperature stress). Similarly, while the sensitivity of soil properties is not high, the influential parameters include *bee* (Clapp and Hornberger coefficient), *phsat* (soil water potential at saturation), *poros* (soil porosity), and *scalez* (which controls the exponential distribution of soil layers in the model).

[34] Not unexpectedly, the carbon-related parameters (related to soil respiration and possible systematic bias in nighttime NEE flux) show sensitivity only when computing carbon fluxes. On the basis of the cutoff used (1%), almost all carbon-related parameters are sensitive at all sites. A surprising exception is the parameter associated with potential systematic bias in observed nighttime carbon flux (*alpha*), which shows no sensitivity at almost all forest sites. This result supports the study of *Hutyra et al.* [2008], which shows that most of the systematic bias can be reduced after careful screening using the friction velocity (u*) filter. In their study, the magnitudes of CO_2 flux obtained using the u* filter match closely with their best estimate of CO_2 flux at K67 site. We assume their results can be extrapolated to other tall canopy sites in our study; therefore, we would expect a reduction in systematic biases due to proper u* filtering also at K83 and K34 sites. Finally, notice that all other sites that did not contain appropriate screening for turbulent conditions suggest some sensitivity to the parameter *alpha*, regardless of canopy height.

[35] From Figure 2 (computed for S_{Ti}) and Figure 3 (computed for S_i), an interesting pattern emerges that sensitive parameters (high total order sensitivity) make their contributions mainly through direct effects (e.g., *z*2, *vcovr*, *hhti*, and *respref*), while those with low total order sensitivity make their contributions mainly through interactions or a combination of direct and interaction effects (e.g., *effcon* for NEE fluxes, *hlti*, *wssp*, and *wopt*). However, it is notable that for a few high sensitivity parameters (e.g., *gradm*, *hhti*, *poros*, for SAV and CROP sites) the parameter interaction effects dominate. We mention this because *Rosero et al.* [2010] suggest that physical realism of models should go hand in hand with fewer undesirable (unrealistic) parameter interactions and that the Noah model did not correspond well with this principle. While we are not sure whether their argument should hold in every case, it is interesting that most of the flux variability in this model originates from direct parameters contributions. For completeness, Figure S1 in the auxiliary material shows the results for the indirect effect index, S_{INDi}, computed as the difference $S_{Ti} - S_i$ between total and direct sensitivity.[1]

[36] In equation (4), V_i describes the so-called "additive" part of a model [*Saltelli et al.*, 2005]. In simple words, additive models have the property that the combined

[1] Auxiliary materials are available in the HTML. doi:10.1029/2011JD016355.

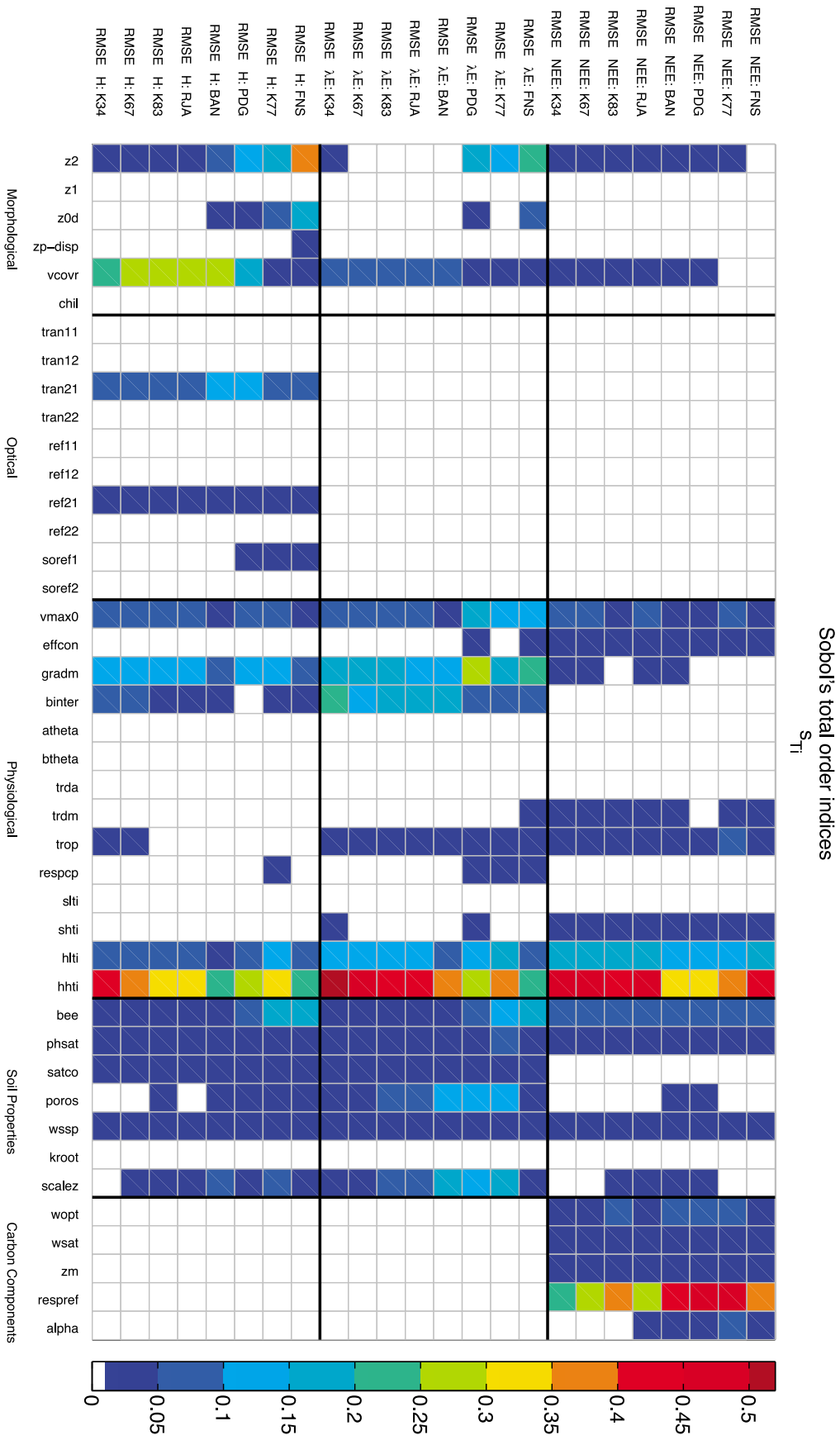
RMSE NEE: FNS
RMSE NEE: K77
RMSE NEE: PDG
RMSE NEE: BAN
RMSE NEE: RJA
RMSE NEE: K83
RMSE NEE: K67
RMSE NEE: K34
RMSE λE: FNS
RMSE λE: K77
RMSE λE: PDG
RMSE λE: BAN
RMSE λE: RJA
RMSE λE: K83
RMSE λE: K67
RMSE λE: K34
RMSE H: FNS
RMSE H: K77
RMSE H: PDG
RMSE H: BAN
RMSE H: RJA
RMSE H: K83
RMSE H: K67
RMSE H: K34
z2
z1
z0d
zp–disp
vcovr
chil
tran11
tran12
tran21
tran22
ref11
ref12
ref21
ref22
soref1
soref2
vmax0
effcon
gradm
binter
atheta
btheta
trda
trdm
trop
respcp
slti
shti
hlti
hhti
bee
phsat
satco
poros
wssp
kroot
scalez
wopt
wsat
zm
respref
alpha
Morphological
Optical
Physiological
Soil Properties
Carbon Components
Sobol's total order indices
S_{Ti}
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5

Figure 2

response of interest can be defined by the sum of their separate effects. In turn, additive models are defined as those for which $\sum_i S_i = 1$. In general, as a rule of thumb, if $\sum_i S_i > 0.6$ then the analysis is considered to be successful [*Saltelli et al.*, 2005; *Liepmann and Stephanopoulos*, 1985], and the model can be considered nearly additive if $\sum_i S_i > 0.8$ [*Campolongo et al.*, 2000]. Figure 4a shows that while the sensitivity analysis of SiB3 produced satisfactory results, SiB3 does not behave as a nearly additive model. Although the "additive" part of the model can be high for some site/flux combinations (e.g., BAN and PDG for H flux, and almost all sites for NEE flux), in general $\sum_i S_i$ for SiB3 falls below 0.8, implying that interaction effects need to be considered in the analysis. Therefore, we focus our analysis on the total order indices (S_{Ti}). With S_{Ti}, we no longer have to limit the analysis to additive models [*Saltelli et al.*, 2005].

[37] Another important aspect of variance-based sensitivity analysis methods (such as Sobol′) is the ability to group inputs as if they were a single factor corresponding directly to the idea of importance measures [*Frey*, 2002; *Saltelli*, 2002a]. In our case, grouping the individual factor (based on Sobol′ total order indices) contributions to the output variance of a given flux can provide valuable information about the SiB3 model (Figures 4b–4d). Our SiB3 results (evaluated for all LBA sites as a whole) can be summarized as follows: (1) "Physiological properties" in the model are the most important measure that if tackled appropriately can provide the largest reduction in SiB3 uncertainty of all fluxes for the Amazon region, and (2) if all fluxes are considered equally important (in terms of contribution to model development), it is not clear which model component is second most important (after the physiological properties). For sensible heat flux (Figure 4b), the physiological component of the model corresponds to almost 50% of the total output variance, followed by "morphological properties" which correspond to about 25%. On the other hand, physiological component explains about 65% of the latent heat flux variance with "soil properties" being the second most importance measure, representing about 25% (Figure 4c). Notice that "carbon properties" make no contribution to the variance of these fluxes but are, not surprisingly, important for carbon fluxes in SiB3, corresponding to about 35% of the output variance, while the physiological component of the model suggests approximate 50% contribution. The simple analysis presented above reinforces the need to account for the multiobjective nature of the system, an aspect that is sometimes ignored when dealing with hydrological and land surface models. This is the main motivation for our proposed multiple-criteria method whose results are described below.

4.2. Multiple-Criteria Analysis of Parameter Sensitivities

[38] The main purpose of Figure 1a is to provide a relatively simple and direct example of the methodology so that potential users can be easily guided through each of the steps. Figure 5 shows examples of a two-dimensional Pareto ranking analysis conducted for the BAN and K83 sites. For ease of illustration, Figure 5 shows 2D scatterplots, although our ultimate goal is to analyze the Pareto ranking in the 3D criteria space (i.e., S_{Ti} for H, λE, and NEE simultaneously). As expected, actual results will not necessarily resemble our simplest case depicted in Figure 1a, although several Pareto rank fronts are easily observed (the number inside each circle is the index number of a parameter following the order presented in Figure 2 and Table 2, i.e., *z*2 = parameter 1, *z*1 = parameter 2, *z*0*d* = parameter 3, and so on). For instance, at the BAN site (Figure 5a), parameters *vcovr*, *hhti*, and *respref* (parameters 5, 30, and 41, respectively) belong to the same group of factors (i.e., rank 1) on the basis of Pareto ranking. However, their individual contributions to the total output variance in SiB3 are different depending on the simulated fluxes analyzed. The parameter *respref* is highly influential to NEE while it exerts minimal (or no) influence on the simulated H. On the other hand, *vcovr* (parameter 5) has strong influence in simulating H but low influence in NEE. Finally, *hhti* shows a good balance of influence in both of these fluxes. Similar analyses can be performed for subsequent Pareto ranks in this site. Figure 5b shows that *hhti* is identified as the only Pareto rank 1 factor because of its strong influence on both λE, and NEE fluxes simulated by SiB3 (S_{Ti} for both criteria are above 40%). Pareto rank 2 factors include *respref* (parameter 41) highly influential for NEE, *hlti* (parameter 29) with contributions to both λE, and NEE, and surprisingly the parameters associated to the photosynthesis-conductance formulation in SiB3 (*grad* and *binter*, parameters 19 and 20, respectively) which show relative higher influence on λE but low influence on NEE. Examples from other sites are somewhat similar to those presented in Figure 5 but are not shown because they can be easily reproduced with the information provided in Figure 2.

[39] To illustrate the shortcomings of a single-criterion-based sensitivity test and for comparison with our multiple-criteria approach, we show the ρ versus rank plots for site K67 in Figure 6. The rank correlation lines obtained when using only one of the H, λE, and NEE model performance metrics at a time are shown in Figures 6a, 6b, and 6c, respectively. The 95% confidence interval of each correlation coefficient (ρ) is estimated using the Fisher transform (also known as r-to-z transform) method [*Fisher*, 1915, 1921]; and the shaded area indicates the lowest and highest bounds of any computed ρ for these analyses. This graphical approach allows us to interpret the results in terms of both at what rank ($r = r^* < R_{\max}$) the value of $\rho_{\min}$ effectively becomes one, and by examining at what rank the confidence interval $\rho - \delta < \rho < \rho + \delta$ of the estimate overlaps the optimal value of $\rho = 1$ ($\rho \pm \delta$ indicates the confidence interval of ρ). The results show a tendency for simulations of

Figure 2. Sobol′'s total order indices (S_{Ti}) calculated for each SiB3 parameter and associated with each site–objective function combination. Each column corresponds to one of 42 parameters used in SiB3, and the columns (parameters) are grouped by model features they control. The rows correspond to each of the eight sites and are grouped by the three objective functions; from top to bottom, the groups are root-mean-square error (RMSE) of NEE, λE, and H. Each cell is color coded to indicate the magnitude of S_{Ti} computed for each combination of parameter and objective function at each site. Larger values of S_{Ti} indicate a greater contribution by parameter i to the total variance and hence greater sensitivity of the modeled response to this parameter. Consistent with the literature, a minimum threshold of $S_{Ti} \geq 0.01$ was used so that colors are only shown for parameters and sites that contribute at least 1% to the total variance.

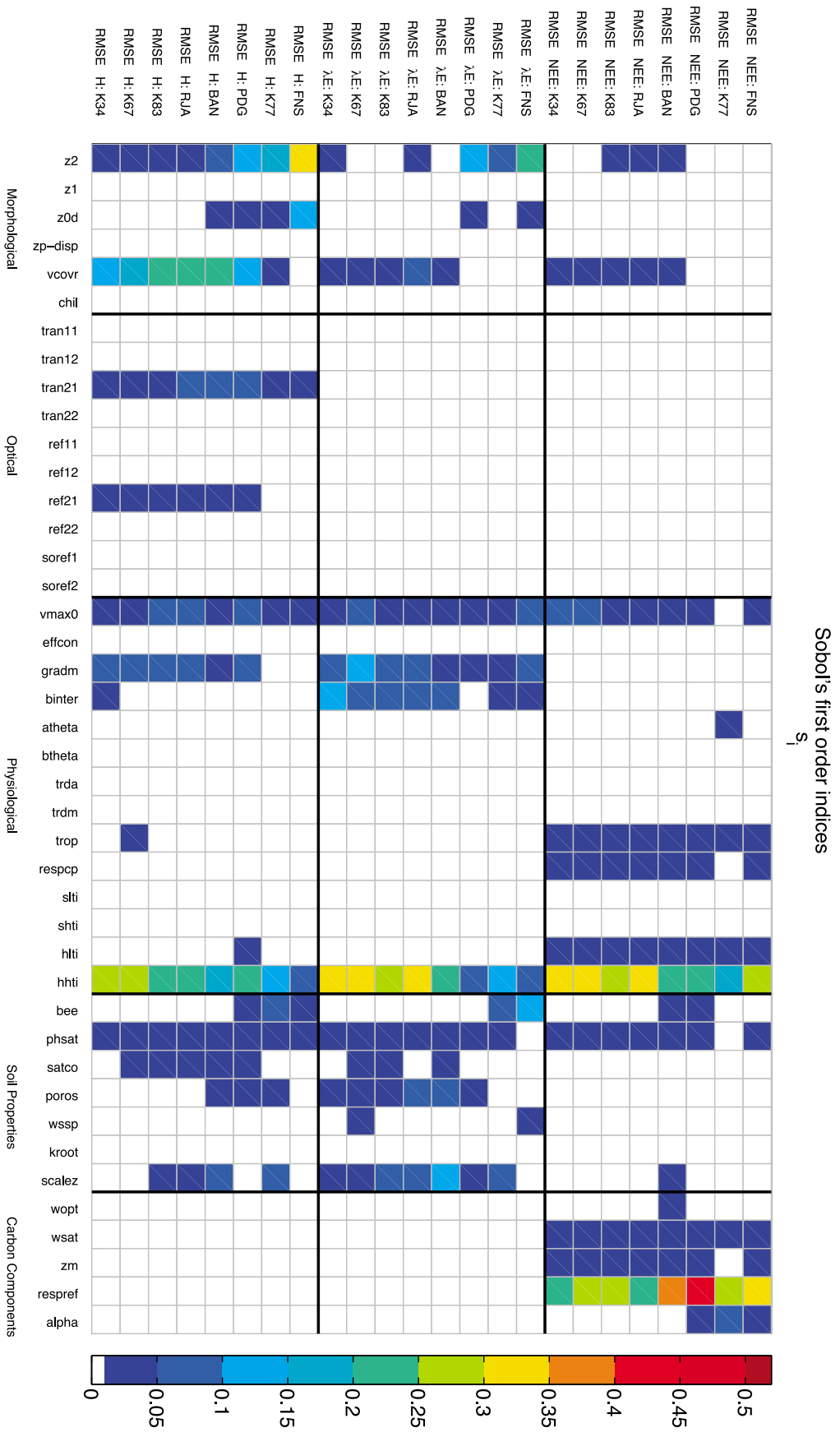

Figure 3. The same as Figure 2, but for Sobol′'s first-order indices (S_i).

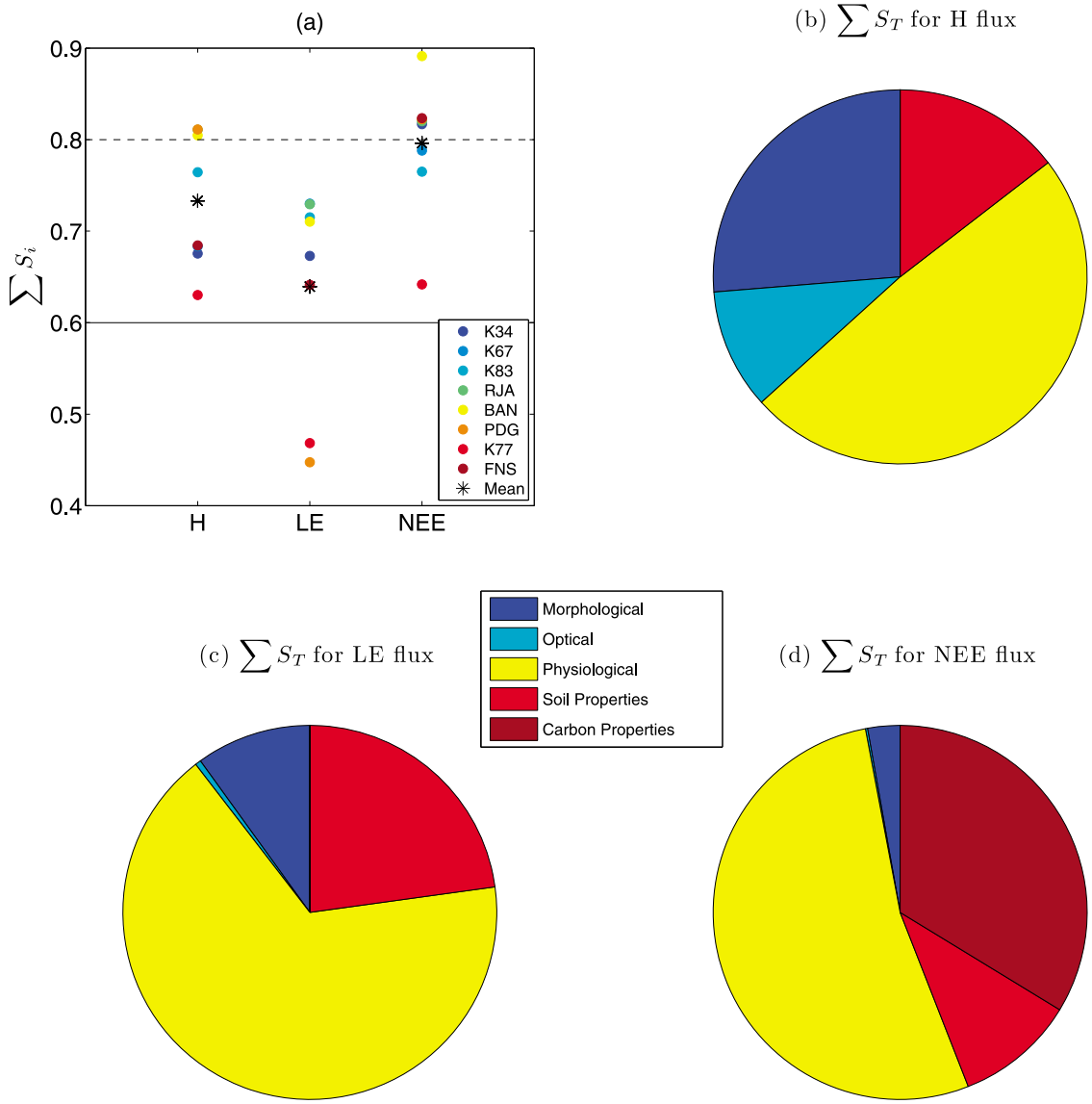

Figure 4. (a) Sum of individual Sobol′ first-order indices ($\sum_i S_i$) computed at each site (color coded) and for each SiB3 flux. The black asterisks correspond to average values calculated for each flux. In general, when $\sum_i S_i$ is greater than 0.6 (i.e., solid black line), the analysis is considered to be satisfactory. When $\sum_i S_i$, is greater than 0.8, the model can be considered nearly additive (see section 4.1 for details). Sum of Sobol′ total order indices (S_T) grouped for each model component (color-coded) for (b) sensible heat flux (H), (c) latent heat flux (λE), and (d) net ecosystem exchange of CO_2 (NEE).

both energy fluxes to improve simultaneously as more parameters are included in the analysis. However, there is a marked difference between these energy fluxes and the flux associated with CO_2 exchange in SiB3 model. The reason is that some parameters having strong influence on both H and LE are ranked as having very low influence on NEE (Figure 2). For instance, parameters *binter* and *scalez* contribute 6% and 2% of the total variance of RMSE for H and 11% and 4% of RMSE for λE; they are ranked 7th and 12th for H flux and 4th and 7th for λE flux. On the other hand, these two parameters are only ranked 26th and 20th for NEE flux, contributing marginally to the total variance of RMSE. Conversely, parameters *respref* and *wopt* contribute approximately 28% and 4% to the total variance associated with NEE flux (ranked 2nd and 6th, respectively). However, these parameters have essentially no contribution to the variance associated with H and λE errors, and are located near the bottom of each individual ranks. As a result, the linear correlation coefficient (either taken as $\rho_{\min}$ or as $\rho \pm \delta$) effectively becomes one (for all metrics) only when almost all parameters are taken into account.

[40] The corresponding results for our multiple-criteria approach are shown in Figure 6d. Clearly, as more ranks (and hence parameters) are included, the explanation of variability in all three fluxes increases, so that the correlation coefficients approach unit. Of interest is that the increases in ρ occur almost simultaneously for all three of the flux criteria so that no metric is "left behind." Consequently, the

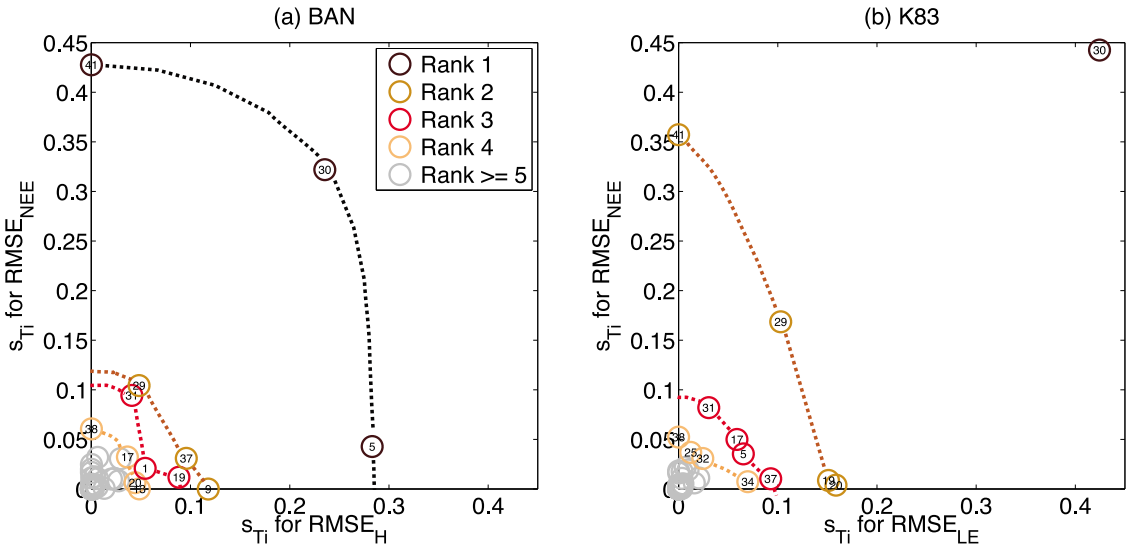

Figure 5. Two-dimensional Pareto ranking results for Sobol′ total order indices (S_{Ti}) computed for (a) RMSE of sensible heat flux (H) and net ecosystem exchange of CO_2 (NEE) at the Bananal Island site (BAN) and (b) RMSE of latent heat flux (λE) and net ecosystem exchange of CO_2 (NEE) at the Santarém K83 site (K83). The number inside each circle represents a parameter following the order presented in Figure 2 and Table 2 (i.e., z2 = parameter 1, z1 = parameter 2, z0d = parameter 3, and so on). The color coding indicates the parameters belonging to a particular Pareto rank (for clarity all parameters with same Pareto rank are also linked together by a dotted line).

overall lower and upper bounds are better constrained, improving the robustness of the analysis and making it easier to identify the most influential parameters and separate those that exert no influence at all to these fluxes. If deemed necessary to identify a smaller set of parameters, one can reduce the target value for ρ to a lower value than unity (such as 0.95). In our case, it is clear that a rank-to-N equal to 6 (i.e., *Npar* = 20 for $\rho_{\min} \sim 0.95$) is sufficient to explain most of the system performance, whereas the single-criterion approach indicates that at least 30 parameters must be included to achieve the same target value for all fluxes. Further, we now find that a rank number of eight (i.e., *Npar* = 28) is sufficient to account for all of variability in the RMSE calculated for all fluxes (i.e., $\rho_{\min}$ reaches unit).

[41] This analysis shows that 14 of the model parameters have little or no effect on the simulation of energy and carbon fluxes, and can therefore be fixed to their nominal values during calibration of SiB3 to these fluxes at site K67. Note that "freezing" unimportant factors contributes to model simplification and was indeed the original motivation of the work of *Sobol′* [1993]. Other sites (Figure 7) show similar results. Overall, we find that approximately 27 to 31 parameters influence the simulation of H, λE, and NEE at these sites. The smaller number is for broadleaf forests while the larger number is for savannah and cropland/pastureland biomes.

[42] For completeness, we also compare our fully multiple-criteria approach with three partial multiple-criteria approaches (here called "pseudo" multiple objective methods) applied to the K67 site (Figure 8). The first approach (pseudo multiobjective method 1) identifies the common set of parameters that simultaneously have Sobol′'s total order indices for all flux metrics greater or equal than the threshold subjectively chosen to be between 1% and 10% (we focus our analysis on parameters whose $S_{Ti} > 1\%$ but it becomes obvious that the subjective choice of the cutoff value impacts largely these results) (Figure 8a). Only nine parameters were found to be "sensitive" (i.e., *vcovr*, *vmax0*, *gradm*, *trop*, *hlti*, *hhti*, *bee*, *phsat*, and *wssp*), and the corresponding $\rho_{\min}$ was found to be approximately 0.57 (Table 3). Notice that this approach may be considered too conservative allowing only those parameters whose contributions are above the specified cutoff value for all fluxes simultaneously.

[43] The second approach (pseudo multiobjective method 2) computes the average S_T from all individual indices (i.e., $S_T = (S_{T1} + S_{T2} + S_{T3})/3$, where 1, 2, 3 indices correspond to H, λE, and NEE, respectively). This method can be interpreted as an attempt to reduce a three-dimensional optimization space to a one-dimensional problem. The approach is sometimes referred to as the weighting method, originally proposed by *Zadeh* [1963] and mentioned by others [*Gupta et al.*, 1998; *Bastidas et al.*, 1999]. In optimization, the weighting coefficients may be adjusted until the results are satisfactory [*Yan and Haan*, 1991a, 1991b], but the initial assumption used is that all of the objectives contribute equally. Although optimization has not been applied in this study, we use this first assumption in our analysis. Once again, we tested cutoff values from 1% to 10% and found that a cutoff value of 1% indicates 20 influential parameters with a corresponding $\rho_{\min} \sim 0.96$ (Figure 8b and Table 3). The remaining contribution to the variability of model output (indicated by $\rho_{\min} < 1$) is because this method does not include additional parameters (assumed to be influential in our Pareto ranking method) that contribute to about 2% to 6% of the total output variance of the simulated fluxes as a whole. Interestingly, an increase of only 1% in the cutoff

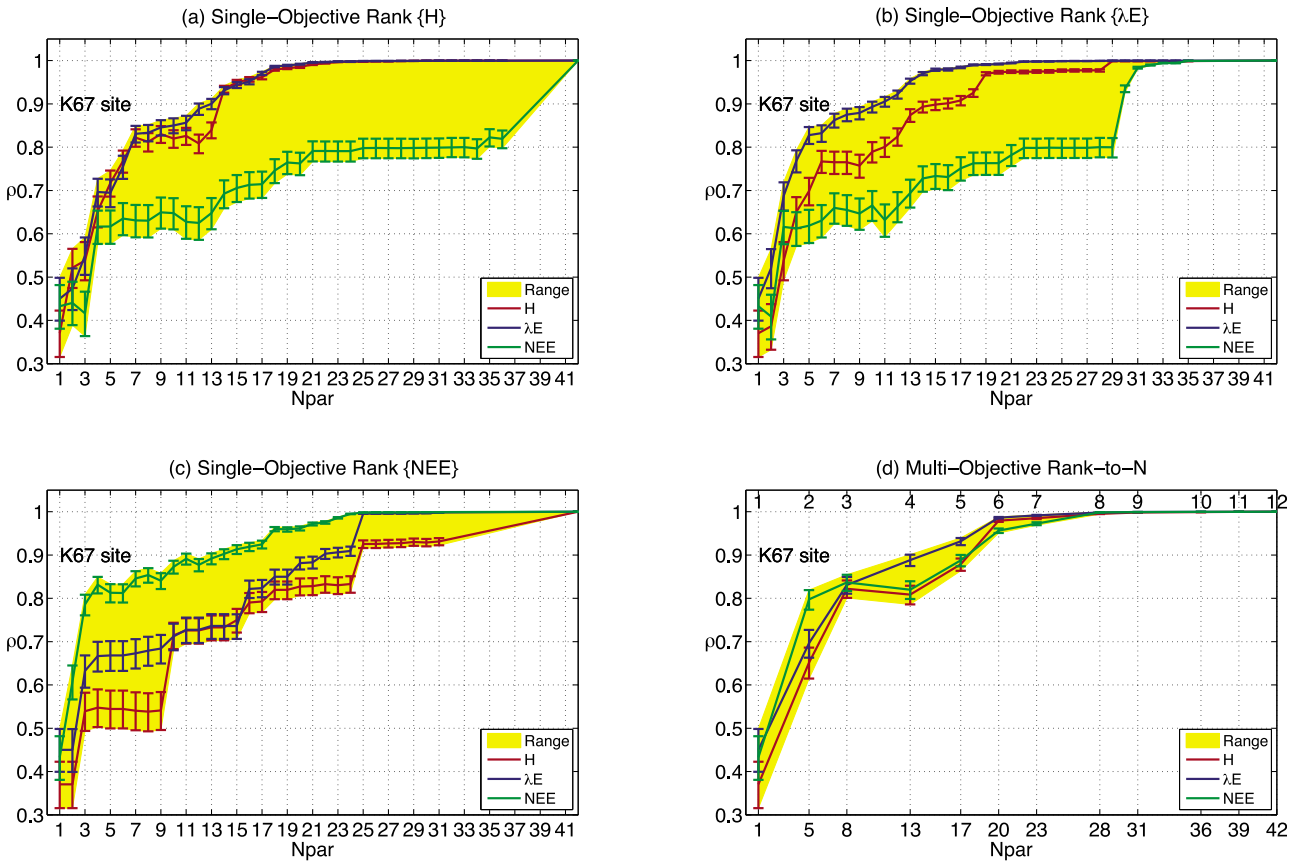

Figure 6. Linear correlation coefficient (ρ) versus number of parameters (*Npar*) calculated on the basis of a single-criterion-based test for (a) H, (b) λE, and (c) NEE fluxes and compared to (d) our newly proposed multiple-criteria approach. Results are for the K67 site, and individual rank correlation lines in each plot are shown as red, blue, and green for H, λE, and NEE, respectively. The 95% confidence interval of each correlation coefficient (error bars) and the shaded area, indicating the lowest and highest bounds of any computed ρ for these analyses, are also shown. Rank-to-N values for the multiple-criteria approach (Figure 6d) are shown on the top *x* axis.

value (i.e., 2% threshold) affects the results dramatically (reducing the number of influential parameters to 12 and ρ_{min} to ~0.85). Notice also that, unlike pseudo multiobjective method 1 and 3 (explained in the next paragraph), this approach does allow for ranking the parameters since the three-dimensional problem becomes one-dimensional.

[44] The third method (pseudo multiobjective method 3) was proposed during review of this manuscript; it may perhaps be viewed as the most sensible strategy of the three. Upon selection of a cutoff value, the approach accepts any parameter whose total order index is larger than the prespecified threshold, regardless of which objective function is being analyzed. In our opinion, this approach is weak in that no information is exchanged among individual criteria; it is essentially equivalent to three single-criterion analyses applied one at a time (in contrast, for example, to method 1 where we have a three-dimensional objective analysis applied once). For a threshold of 1%, the method identifies 23 parameters as influential and yet ρ_{min} does not reach unity ($\rho_{min} \sim 0.97$) (Figure 8c). About 4% of the total output variance of the simulated fluxes is still not captured by this method (Table 3). These results also confirm that the subjectivity inherent in pseudo multiple-criteria approaches can dramatically influence the outcome of the analysis, with a small 1% increase in the threshold cutoff value reducing the number of influential parameters to 18 and ρ_{min} to ~0.89.

5. Discussion and Conclusions

[45] We present a novel rank-based multiple criteria implementation of the Sobol′ sensitivity analysis approach and apply it to the problem of identifying the most influential parameters in the SiB3 model for a network of flux towers in Brazil. The approach implements an objective fully multiple-criteria strategy to evaluate parameter sensitivity. We show that it is superior to single-criterion approaches while avoiding the high subjectivity involved in pseudo multiple-criteria methods. While robust and simple to apply, the Pareto ranking method applied to the Sobol′ sensitivity analysis does require large numbers of model simulations because variance-based are known to be computationally expensive. In addition, we recognize the importance of traditional analyses of parameter sensitivity such as those presented in Figures 2, 3, and 4; therefore the

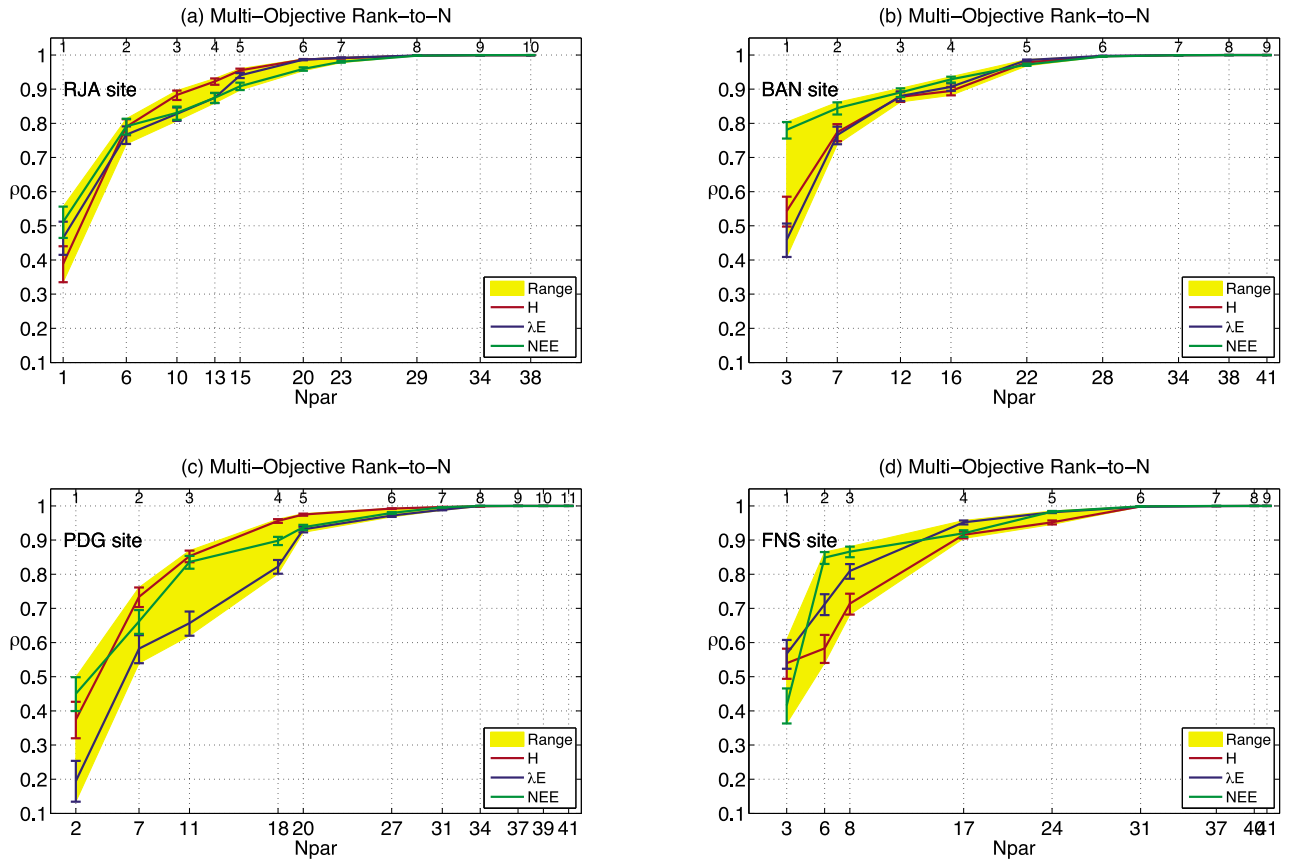

Figure 7. Linear correlation coefficient (ρ) versus number of parameters (*Npar*) calculated on the basis of our newly proposed multiple-criteria approach for the (a) RJA, (b) BAN, (c) PDG, and (d) FNS sites. Individual rank correlation lines in each plot are shown as red, blue, and green for H, λE, and NEE, respectively. The 95% confidence interval of each correlation coefficient (error bars) and the shaded area, indicating the lowest and highest bounds of any computed ρ for these analyses, are also shown. Rank-to-N values for the multiple-criteria approach (Figure 7d) are shown on the top *x* axis.

rank-based strategy we are proposing is intended to complement (arguably, in a more objective way) rather than substitute these analyses. While certain parameters might not show up in a rank-based strategy, we certainly believe that for diagnostic reasons it is very valuable to understand whether they make a contribution to the overall variance or not.

[46] For SiB3, we found 27 to 31 parameters to be influential with the lower value associated with broadleaf forests and the higher value associated with savannah and cropland/pastureland biomes. The majority of these parameters (~78%) are associated mainly with physiology, or soil and carbon properties; optical properties were found to have little influence on surface fluxes taken together. This result can be used to constrain the dimension of the parameter space during model calibration studies [see *Rosolem et al.*, 2005]. Note that freezing unimportant parameters was a major motivation for the development of the Sobol′ variance-based method in the early 1990s.

[47] The lack of sensitivity observed in some parameters could suggest that the SiB3 model may be over-parameterized with respect to the simulation of energy and carbon fluxes in the region, but nothing can be said about the overall behavior of the model in simulating other components, in other regions of the world. The sensitivity of SiB3 parameters was tested under meteorological conditions assumed to be representative of the long-term climatology in the region [*Rosolem et al.*, 2008]. Whether the sensitivity of these parameters might change under future climate or during extreme events (e.g., droughts of 2005 and 2010) needs further investigation.

[48] Although the application discussed in this study is focused on the Amazon region and for a single model, we encourage testing of the method for other regions and recommend its use for evaluating the sensitivity of parameters in other land surface parameterization schemes. Applied to several land surface schemes, the approach can help in establishing which aspects of soil-plant-atmosphere processes matter most in land surface models of Amazonia and other regions, while accounting for the different aspects represented by those models. Such studies can be helpful in informing the design of field campaigns that aim to characterize and measure these parameters. Note that the multiple-criteria ranking and parameter selection methodology is general and can be applied to other numerical models (e.g., hydrological models) and/or with other sensitivity analysis

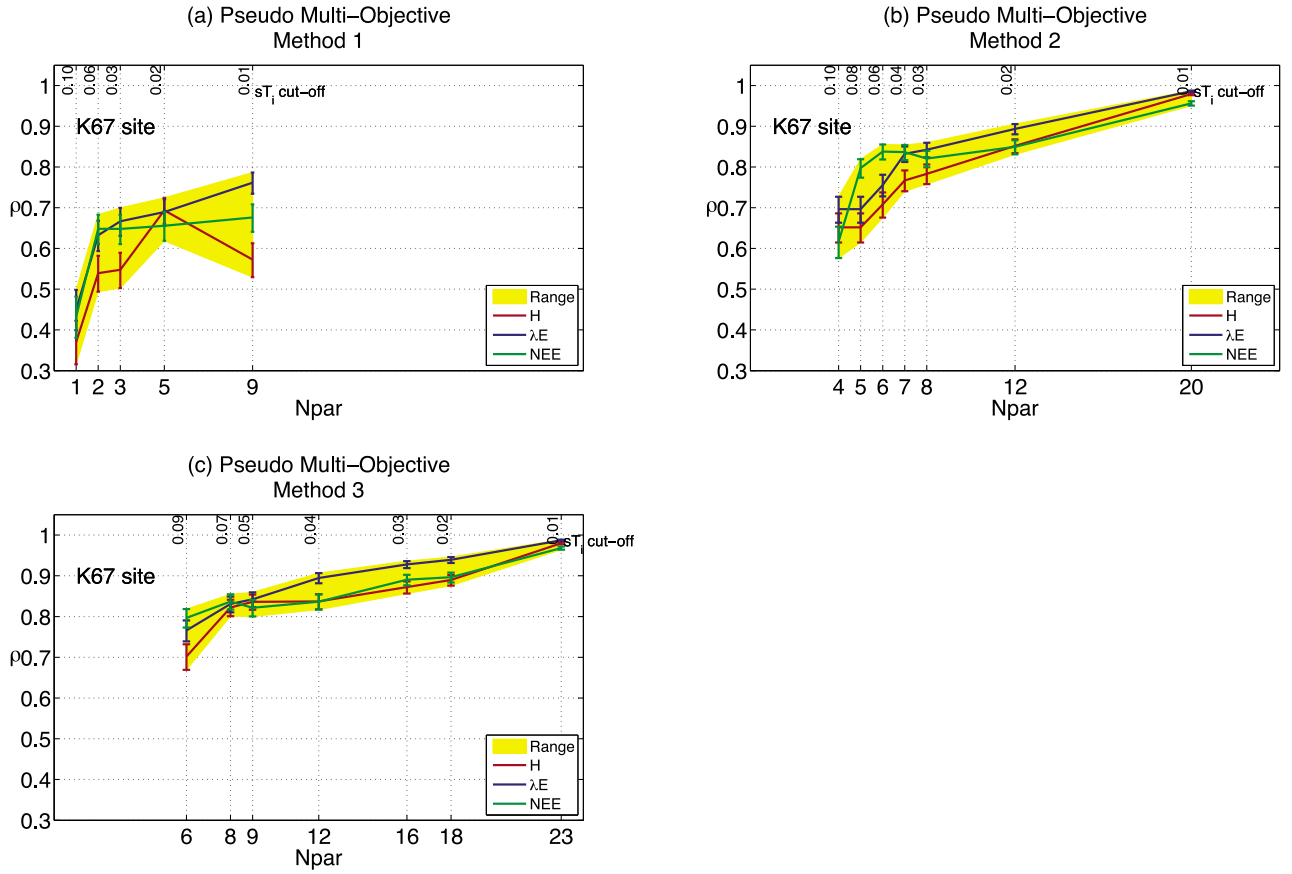

Figure 8. Linear correlation coefficient (ρ) versus number of parameters (*Npar*) calculated on the basis of three pseudo multiple-criteria approaches (see section 4.2 for details). (a) Method 1 identifies the common set of parameters that simultaneously have Sobol''s total order indices for all flux metrics greater than or equal to the threshold subjectively chosen. (b) Method 2 computes the average S_T from all individual indices (i.e., $S_T = (S_{T1} + S_{T2} + S_{T3})/3$, where indices 1, 2, 3 correspond to H, λE, and NEE, respectively) and selects those parameters whose S_T is above a prespecified threshold. (c) Method 3 accepts any parameter whose total order index is larger than the prespecified threshold, regardless of which objective function is being analyzed. The degree of subjectivity of those methods is shown by allowing the thresholds to be defined between 1% and 10%. Individual rank correlation lines in each plot are shown as red, blue, and green for H, λE, and NEE, respectively. The 95% confidence interval of each correlation coefficient (error bars) and the shaded area, indicating the lowest and highest bounds of any computed ρ for these analyses, are also shown.

Table 3. List of Influential Parameters Obtained From Three Different Pseudo Multiobjective Methods and Our Proposed Multiobjective Pareto Ranking Approach[a]

	Pseudo Multiobjective Method 1	Pseudo Multiobjective Method 2	Pseudo Multiobjective Method 3	Multiobjective Pareto Ranking Method
Parameters	**vcovr**, **vmax0**, gradm, trop, **hlti**, **hhti**, bee, **phsat**, wssp	z2 (15), satco (18), **vcovr** (3), poros (13), **tran21** (9), wopt (17), ref21 (19), wssp (14), **vmax0** (6), scalez (11), effcon (16), respref (5), **gradm** (4), **binter** (7), **trop** (12), shti (20), **hlti** (2), **hhti** (1), **bee** (8), **phsat** (10)	**z2**, **phsat**, **vcovr**, satco, **tran21**, **poros**, **ref21**, **wopt**, **vmax0**, wsat, **effcon**, zm, **gradm**, **wssp**, **binter**, **scalez**, trdm, **respref**, **trop**, shti, **hlti**, **hhti**, **bee**	z2 (4), shti (6), z0d (8), hlti (2), vcovr (2), hhti (1), tran21 (3), bee (4), ref21 (4), phsat (4), vmax0 (3), satco (6), effcon (6), poros (5), gradm (2), wopt (5), binter (3), wsat (7), atheta (8), zm (8), btheta (8), wssp (5), trdm (7), scalez (4), trop (5), respref (2), respcp (7), alpha (8)

	Pseudo Multiobjective Method 1 S_{Ti} Threshold		Pseudo Multiobjective Method 2 S_{Ti} Threshold		Pseudo Multiobjective Method 3 S_{Ti} Threshold		Multiobjective Pareto Ranking Method[b]
	1%	2%	1%	2%	1%	2%	
$\rho_{\min}$	0.57	0.66	0.96	0.85	0.97	0.89	1.00
Number of Parameters	9	5	20	12	23	18	28
Uncertainty Remaining	34	53	4	16	4	6	0

[a]Two S_{Ti} thresholds are used (S_{Ti} larger than 1% and 2%) to show how the results are affected by these subject choices. Here $\rho_{\min}$ is the minimum linear correlation coefficient computed across all of the metrics (see section 2.2). The uncertainty remaining corresponds to the sum of Sobol′ total order indices (S_{Ti}) computed for those parameters that were not identified as influential in either of the two cases. Notice that the parameters are ordered on the basis of Table 2 and Figures 2 and 3. Pseudo multiobjective method 2 and our new multiobjective Pareto ranking method are the only two methods that allow for further ranking of these parameters. (Each rank number is shown within parentheses.) Notice that these ranks have different meaning for these two distinct methods (see section 4.2 for details). The influential parameters obtained with an S_{Ti} threshold of 2% are bold.
[b]No need to specify threshold.

procedures provided a measurable sensitivity index (or statistics) is computed for at least two objective functions. Furthermore, note that we have computed all Pareto rank groups for the sake of clarity of our new proposed approach (Figures 6d and 7). One can easily automate this method to stop when $\rho_{\min}$ becomes statistically close to unity (1.0) without the need to compute the remaining rank groups. If this condition had been used in our case, it would have avoided 23,000 additional model simulations. As always, we welcome dialog on these and other ideas related to land surface model calibration. Code for the multiple-criteria sensitivity analysis approach is available from the first author by request.

[49] **Acknowledgments.** This study was supported by NASA Earth and Science Fellowship under grant NNX09AO33H with additional support from NSF Amazon-PIRE under grant 0730305, B2-Earthscience, and the NSF Center for Sustainability of semi-Arid Hydrology and Riparian Areas (SAHRA) under the STC Program of the National Science Foundation, agreement EAR-9876800 and NSF award DEB-0415977. Additional support was also provided by the COsmic-ray Soil Moisture Observing System project under US National Science Foundation grant AGS-0838491. The authors acknowledge Ian T. Baker and A. Scott Denning for providing the original version of the SiB3 model, S. Yatheendradas and P. Pokhrel for technical discussions on Sobol′ sensitivity analysis, R. Stöckli for providing time-varying model inputs relevant to the Amazon region, N. Restrepo-Coupe for providing quality-controlled forcing data for all sites studied here, and the entire LBA-DMIP community for discussions on model improvements. The authors also thank all LBA investigators who provided the flux tower data used in this study: Alessandro C. Araújo and Antônio O. Manzi for the Manaus K34 site; Steve C. Wofsy and Scott R. Saleska for the Santarém K67 site; Michael L. Goulden, Scott D. Miller, and Humberto R. da Rocha for the Santarém K83 site; Celso von Randow, Antônio O. Manzi, and Renata Aguiar for the Rondônia RJA and FNS sites; Humberto R. da Rocha for the Tocantins BAN and São Paulo PDG sites; and David R. Fitzjarrald, Rodrigo da Silva, and Ricardo K. Sakai for the Santarém K77 site. Finally, we acknowledge Steven C. Wofsy for additional comments regarding systematic biases in nighttime fluxes and Santarém K67 eddy flux tower characteristics and James Broermann, Nate Bryant, and Michael Leuthold for technical support during model simulations.

References

Araújo, A. C., et al. (2002), Comparative measurements of carbon dioxide fluxes from two nearby towers in a central Amazonian rainforest: The Manaus LBA site, *J. Geophys. Res.*, *107*(D20), 8090, doi:10.1029/2001JD000676.

Avissar, R., and C. A. Nobre (2002), Preface to special issue on the Large-Scale Biosphere-Atmosphere Experiment in Amazonia (LBA), *J. Geophys. Res.*, *107*(D20), 8034, doi:10.1029/2002JD002507.

Baker, I., A. S. Denning, N. Hanan, L. Prihodko, M. Uliasz, P. L. Vidale, K. Davis, and P. Bakwin (2003), Simulated and observed fluxes of sensible and latent heat and CO_2 at the WLEF-TV tower using SiB2.5, *Global Change Biol.*, *9*(9), 1262–1277, doi:10.1046/j.1365-2486.2003.00671.x.

Baker, I. T., L. Prihodko, A. S. Denning, M. Goulden, S. Miller, and H. R. da Rocha (2008), Seasonal drought stress in the Amazon: Reconciling models and observations, *J. Geophys. Res.*, *113*, G00B01, doi:10.1029/2007JG000644.

Bastidas, L. A., H. V. Gupta, S. Sorooshian, W. J. Shuttleworth, and Z. L. Yang (1999), Sensitivity analysis of a land surface scheme using multicriteria methods, *J. Geophys. Res.*, *104*(D16), 19,481–19,490, doi:10.1029/1999JD900155.

Bastidas, L. A., T. S. Hogue, S. Sorooshian, H. V. Gupta, and W. J. Shuttleworth (2006), Parameter sensitivity analysis for different complexity land surface models using multicriteria methods, *J. Geophys. Res.*, *111*, D20101, doi:10.1029/2005JD006377.

Betts, R. A., P. M. Cox, M. Collins, P. P. Harris, C. Huntingford, and C. D. Jones (2004), The role of ecosystem-atmosphere interactions in simulated Amazonian precipitation decrease and forest dieback under global climate warming, *Theor. Appl. Climatol.*, *78*(1–3), 157–175.

Borma, L. S., et al. (2009), Atmosphere and hydrological controls of the evapotranspiration over a floodplain forest in the Bananal Island region, Amazonia, *J. Geophys. Res.*, *114*, G01003, doi:10.1029/2007JG000641.

Brubaker, K. L., D. Entekhabi, and P. S. Eagleson (1993), Estimation of continental precipitation recycling, *J. Clim.*, *6*(6), 1077–1089, doi:10.1175/1520-0442(1993)006<1077:EOCPR>2.0.CO;2.

Campolongo, F., A. Saltelli, and S. Tarantola (2000), Sensitivity analysis as an ingredient of modeling, *Stat. Sci.*, *15*(4), 377–395, doi:10.1214/ss/1009213004.

Cox, P. M., R. A. Betts, C. D. Jones, S. A. Spall, and I. J. Totterdell (2000), Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model, *Nature*, *408*(6809), 184–187, doi:10.1038/35041539.

Dai, Y. J., et al. (2003), The Common Land Model, *Bull. Am. Meteorol. Soc.*, *84*(8), 1013–1023, doi:10.1175/BAMS-84-8-1013.

da Rocha, H. R., H. C. Freitas, R. Rosolem, R. I. N. Juarez, R. N. Tannus, M. V. Ligo, O. M. R. Cabral, and M. A. F. Silva Dias (2002), Measurements of CO_2 exchange over a woodland savanna (cerrado sensu stricto) in southeast Brasil, *Biota Neotropica*, *2*(1), 1–11. [Available at http://www.biotaneotropica.org.br/v2n1/pt/fullpaper?bn01702012002+en.]

da Rocha, H. R., et al. (2009), Patterns of water and heat flux across a biome gradient from tropical forest to savanna in Brazil, *J. Geophys. Res.*, *114*, G00B12, doi:10.1029/2007JG000640.

Demarty, J., C. Ottle, I. Braud, A. Olioso, J. P. Frangi, L. A. Bastidas, and H. V. Gupta (2004), Using a multiobjective approach to retrieve information on surface properties used in a SVAT model, *J. Hydrol.*, *287*(1–4), 214–236, doi:10.1016/j.jhydrol.2003.10.003.

Eltahir, E. A. B., and R. L. Bras (1994), Precipitation recycling in the Amazon basin, *Q. J. R. Meteorol. Soc.*, *120*(518), 861–880, doi:10.1002/qj.49712051806.

Fisher, R. A. (1915), Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population, *Biometrika*, *10*(4), 507–521.

Fisher, R. A. (1921), On the "probable error" of a coefficient of correlation deduced from a small sample, *Metron*, *1*, 3–32.

Frey, H. C. (2002), Guest editorial: Introduction to special section on sensitivity analysis and summary of NCSU/USDA Workshop on Sensitivity Analysis, *Risk Anal.*, *22*, 539–545, doi:10.1111/0272-4332.00037.

Frey, H. C., and S. R. Patil (2002), Identification and review of sensitivity analysis methods, *Risk Anal.*, *22*(3), 553–578, doi:10.1111/0272-4332.00039.

Goldberg, D. E. (1989), *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, Mass.

Gupta, H. V., S. Sorooshian, and P. O. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, *34*(4), 751–763, doi:10.1029/97WR03495.

Hall, J. W., S. Tarantola, P. D. Bates, and M. S. Horritt (2005), Distributed sensitivity analysis of flood inundation model calibration, *J. Hydraul. Eng.*, *131*(2), 117–126, doi:10.1061/(ASCE)0733-9429(2005)131:2(117).

Hornberger, G. M., and R. C. Spear (1981), An approach to the preliminary analysis of environmental systems, *J. Environ. Manage.*, *12*(1), 7–18.

Houghton, R. A., K. T. Lawrence, J. L. Hackler, and S. Brown (2001), The spatial distribution of forest biomass in the Brazilian Amazon: A comparison of estimates, *Global Change Biol.*, *7*(7), 731–746, doi:10.1046/j.1365-2486.2001.00426.x.

Hutyra, L. R., W. Munger, S. R. Saleska, E. Gottlieb, B. C. Daube, A. L. Dunn, D. F. Amaral, P. B. de Camargo, and S. C. Wofsy (2007), Seasonal controls on the exchange of carbon and water in an Amazonian rain forest, *J. Geophys. Res.*, *112*, G03008, doi:10.1029/2006JG000365.

Hutyra, L. R., J. W. Munger, E. Hammond-Pyle, S. R. Saleska, N. Restrepo-Coupe, B. C. Daube, P. B. de Camargo, and S. C. Wofsy (2008), Resolving systematic errors in estimates of net ecosystem exchange of CO_2 and ecosystem respiration in a tropical forest biome, *Agric. For. Meteorol.*, *148*, 1266–1279, doi:10.1016/j.agrformet.2008.03.007.

Jolly, W. M., R. Nemani, and S. W. Running (2005), A generalized, bioclimatic index to predict foliar phenology in response to climate, *Global Change Biol.*, *11*, 619–632, doi:10.1111/j.1365-2486.2005.00930.x.

Keller, M., et al. (2004), Ecological research in the large-scale biosphere-atmosphere experiment in Amazonia: Early results, *Ecol. Appl.*, *14*(SP4), 3–16, doi:10.1890/03-6003.

Keller, M., M. Bustamante, J. Gash, and P. S. Dias (Eds.) (2009), *Amazonia and Global Change*, *Geophys. Monogr. Ser.*, vol. 186, 576 pp., AGU, Washington, D. C.

Kruijt, B., J. A. Elbers, C. von Randow, A. C. Araújo, P. J. Oliveira, A. Culf, A. O. Manzi, A. D. Nobre, P. Kabat, and E. J. Moors (2004), The robustness of eddy correlation fluxes for Amazon rain forest conditions, *Ecol. Appl.*, *14*(SP4), 101–113, doi:10.1890/02-6004.

LBA Science Planning Group (1996), The Large-Scale Biosphere-Atmosphere Experiment in Amazonia (LBA): Concise experimental plan, 44 pp., Wageningen, Netherlands.

Liepmann, D., and G. Stephanopoulos (1985), Development and global sensitivity analysis of a closed ecosystem model, *Ecol. Modell.*, *30*, 13–47, doi:10.1016/0304-3800(85)90035-3.

Liu, Y. Q., H. V. Gupta, S. Sorooshian, L. A. Bastidas, and W. J. Shuttleworth (2004), Exploring parameter sensitivities of the land surface using a locally coupled land-atmosphere model, *J. Geophys. Res.*, *109*, D21101, doi:10.1029/2004JD004730.

Malhi, Y., and E. A. Davidson (2009), Biogeochemistry and ecology of terrestrial ecosystems of Amazonia, in *Amazonia and Global Change*, *Geophys. Monogr. Ser.*, vol. 186, edited by M. Keller et al., pp. 293–298, AGU, Washington, D. C.

Malhi, Y., and J. Grace (2000), Tropical forests and atmospheric carbon dioxide, *Trends Ecol. Evol.*, *15*(8), 332–337, doi:10.1016/S0169-5347(00)01906-6.

Malhi, Y., S. Saatchi, C. Girardin, and L. E. O. C. Aragão (2009), The production, storage, and flow of carbon in Amazonian forests, in *Amazonia and Global Change*, *Geophys. Monogr. Ser.*, vol. 186, edited by M. Keller et al., pp. 355–372, AGU, Washington, D. C, USA.

McKay, M. D., R. J. Beckman, and W. J. Conover (1979), A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, *21*(2), 239–245, doi:10.2307/1268522.

Miller, S. D., M. L. Goulden, M. C. Menton, H. R. da Rocha, H. C. de Freitas, A. Figueira, and C. A. D. de Sousa (2004), Biometric and micrometeorological measurements of tropical forest carbon balance, *Ecol. Appl.*, *14*(SP4), 114–126, doi:10.1890/02-6005.

Nemani, R. R., C. D. Keeling, H. Hashimoto, W. M. Jolly, S. C. Piper, C. J. Tucker, R. B. Myneni, and S. W. Running (2003), Climate-driven increases in global terrestrial net primary production from 1982 to 1999, *Science*, *300*, 1560–1563, doi:10.1126/science.1082750.

Nobre, C. A., P. J. Sellers, and J. Shukla (1991), Amazonian deforestation and regional climate change, *J. Clim.*, *4*(10), 957–988, doi:10.1175/1520-0442(1991)004<0957:ADARCC>2.0.CO;2.

Pitman, A. J. (2003), The evolution of, and revolution in, land surface schemes designed for climate models, *Int. J. Climatol.*, *23*(5), 479–510, doi:10.1002/joc.893.

Prihodko, L., A. S. Denning, N. P. Hanan, I. Baker, and K. Davis (2008), Sensitivity, uncertainty and time dependence of parameters in a complex land surface model, *Agric. For. Meteorol.*, *148*(2), 268–287, doi:10.1016/j.agrformet.2007.08.006.

Randall, D. A., D. A. Dazlich, C. Zhang, A. S. Denning, P. J. Sellers, C. J. Tucker, L. Bounoua, S. O. Los, C. O. Justice, and I. Fung (1996), A revised land surface parameterization (SiB2) for GCMs. Part III: The greening of the Colorado State University general circulation model, *J. Clim.*, *9*(4), 738–763, doi:10.1175/1520-0442(1996)009<0738:ARLSPF>2.0.CO;2.

Rosero, E., Z. L. Yang, T. Wagener, L. E. Gulden, S. Yatheendradas, and G. Y. Niu (2010), Quantifying parameter sensitivity, interaction, and transferability in hydrologically enhanced versions of the Noah land surface model over transition zones during the warm season, *J. Geophys. Res.*, *115*, D03106, doi:10.1029/2009JD012035.

Rosolem, R., L. A. Bastidas, W. J. Shuttleworth, L. G. G. de Gonçalves, E. J. Burke, H. R. da Rocha, S. D. Miller, and M. L. Goulden (2005), Evaluation of effects of selective logging on energy-water and carbon exchange processes, in *Regional Hydrological Impacts of Climatic Change: Hydroclimatic Variability*, edited by S. Franks et al., pp. 118–125, Int. Assoc. of Hydrol. Sci. Press, Wallingford, U. K.

Rosolem, R., W. J. Shuttleworth, and L. G. G. de Gonçalves (2008), Is the data collection period of the Large-Scale Biosphere-Atmosphere Experiment in Amazonia representative of long-term climatology?, *J. Geophys. Res.*, *113*, G00B09, doi:10.1029/2007JG000628.

Rosolem, R., W. J. Shuttleworth, X. Zeng, S. R. Saleska, and T. E. Huxman (2010), Land surface modeling inside the Biosphere 2 tropical rainforest biome, *J. Geophys. Res.*, *115*, G04035, doi:10.1029/2010JG001443.

Rosolem, R., H. V. Gupta, W. J. Shuttleworth, L. G. G. de Gonçalves, and X. Zeng (2012), Assessing components of uncertainty in parameter estimation of the SiB3 land surface model for Amazon biomes, *Hydrol. Processes*, in press.

Sakaguchi, K., X. Zeng, B. Christoffersen, N. Coupe, S. Saleska, and P. Brando (2011), Natural and drought scenarios in an east central Amazon forest: Fidelity of the Community Land Model 3.5 with three biogeochemical models, *J. Geophys. Res.*, *116*, G01029, doi:10.1029/2010JG001477.

Sakai, R. K., D. R. Fitzjarrald, O. L. L. Moraes, R. M. Staebler, O. C. Acevedo, M. J. Czikowsky, R. Da Silva, E. Brait, and V. Miranda (2004), Land-use change effects on local energy, water, and carbon balances in an Amazonian agricultural field, *Global Change Biol.*, *10*(5), 895–907, doi:10.1111/j.1529-8817.2003.00773.x.

Salati, E. (1987), The forest and the hydrological cycle, in *The Geophysiology of Amazonia: Vegetation and Climate Interactions*, edited by R. E. Dickinson, pp. 273–296, John Wiley, New York.

Saleska, S. R., et al. (2003), Carbon in Amazon forests: Unexpected seasonal fluxes and disturbance-induced losses, *Science*, *302*(5650), 1554–1557, doi:10.1126/science.1091165.

Saltelli, A. (2002a), Sensitivity analysis for importance assessment, *Risk Anal.*, *22*(3), 579–590, doi:10.1111/0272-4332.00040.

Saltelli, A. (2002b), Making best use of model evaluations to compute sensitivity indices, *Comput. Phys. Commun.*, *145*(2), 280–297, doi:10.1016/S0010-4655(02)00280-1.

Saltelli, A., S. Tarantola, and K. P. S. Chan (1999), A quantitative model-independent method for global sensitivity analysis of model output, *Technometrics*, *41*(1), 39–56, doi:10.2307/1270993.

Saltelli, A., M. Ratto, S. Tarantola, and F. Campolongo (2005), Sensitivity analysis for chemical models, *Chem. Rev.*, *105*, 2811–2828, doi:10.1021/cr040659d.

Saltelli, A., M. Ratto, S. Tarantola, F. Campolongo, European Commission, and Joint Research Centre of Ispra (I) (2006), Sensitivity analysis practices: Strategies for model-based inference, *Reliab. Eng. Syst. Saf.*, *91*(10–11), 1109–1125, doi:10.1016/j.ress.2005.11.014.

Saltelli, A., P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola (2010), Variance based sensitivity analysis of model output: Design and estimator for the total sensitivity index, *Comput. Phys. Commun.*, *181*, 259–270, doi:10.1016/j.cpc.2009.09.018.

Sellers, P. J., Y. Mintz, Y. C. Sud, and A. Dalcher (1986), A Simple Biosphere model (SiB) for use within general circulation models, *J. Atmos. Sci.*, *43*(6), 505–531, doi:10.1175/1520-0469(1986)043<0505:ASBMFU>2.0.CO;2.

Sellers, P. J., D. A. Randall, G. J. Collatz, J. A. Berry, C. B. Field, D. A. Dazlich, C. Zhang, G. D. Collelo, and L. Bounoua (1996a), A revised land surface parameterization (SiB2) for atmospheric GCMs. 1. Model formulation, *J. Clim.*, *9*(4), 676–705, doi:10.1175/1520-0442(1996)009<0676:ARLSPF>2.0.CO;2.

Sellers, P. J., S. O. Los, C. J. Tucker, C. O. Justice, D. A. Dazlich, G. J. Collatz, and D. A. Randall (1996b), A revised land surface parameterization (SiB2) for atmospheric GCMs. 2. The generation of global fields of terrestrial biophysical parameters from satellite data, *J. Clim.*, *9*(4), 706–737, doi:10.1175/1520-0442(1996)009<0706:ARLSPF>2.0.CO;2.

Sellers, P. J., et al. (1997), Modeling the exchanges of energy, water, and carbon between continents and the atmosphere, *Science*, *275*(5299), 502–509, doi:10.1126/science.275.5299.502.

Shuttleworth, W. J. (1988), Evaporation from Amazonian rainforest, *Proc. R. Soc. London, Ser. B*, *233*(1272), 321–346, doi:10.1098/rspb.1988.0024.

Sobol′, I. M. (1993), Sensitivity analysis for nonlinear mathematical models, *Math. Modell. Comput. Exp.*, *1*, 407–414.

Stöckli, R., and P. L. Vidale (2004), European plant phenology and climate as seen in a 20-year AVHRR land-surface parameter dataset, *Int. J. Remote Sens.*, *25*(17), 3303–3330, doi:10.1080/01431160310001618149.

Stöckli, R., T. Rutishauser, D. Dragoni, J. O'Keefe, P. E. Thornton, M. Jolly, L. Lu, and A. S. Denning (2008), Remote sensing data assimilation for a prognostic phenology model, *J. Geophys. Res.*, *113*, G04021, doi:10.1029/2008JG000781.

Tang, Y., P. Reed, K. van Werkhoven, and T. Wagener (2007a), Advancing the identification and evaluation of distributed rainfall-runoff models using global sensitivity analysis, *Water Resour. Res.*, *43*, W06415, doi:10.1029/2006WR005813.

Tang, Y., P. Reed, T. Wagener, and K. van Werkhoven (2007b), Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation, *Hydrol. Earth Syst. Sci.*, *11*(2), 793–817, doi:10.5194/hess-11-793-2007.

U.S. Environmental Protection Agency (EPA), (2009), Guidance on the development, evaluation, and application of environmental models, *Rep. EPA/100/K-09/003*, 99 pp., Off. of the Sci. Advisor, Counc. for Regul. Environ. Model., Washington, D. C. [Available at http://www.epa.gov/crem/library/cred_guidance_0309.pdf.]

van Werkhoven, K., T. Wagener, P. Reed, and Y. Tang (2008), Characterization of watershed model behavior across a hydroclimatic gradient, *Water Resour. Res.*, *44*, W01429, doi:10.1029/2007WR006271.

Vidale, P. L., and R. Stöckli (2005), Prognostic canopy air space solutions for land surface exchanges, *Theor. Appl. Climatol.*, *80*(2–4), 245–257, doi:10.1007/s00704-004-0103-2.

von Randow, C., et al. (2004), Comparative measurements and seasonal variations in energy and carbon exchange over forest and pasture in south west Amazonia, *Theor. Appl. Climatol.*, *78*(1–3), 5–26.

Wagener, T., and J. Kollat (2007), Numerical and visual evaluation of hydrological and environmental models using the Monte Carlo analysis toolbox, *Environ. Model. Software*, *22*(7), 1021–1033, doi:10.1016/j.envsoft.2006.06.017.

Werth, D., and R. Avissar (2004), The regional evapotranspiration of the Amazon, *J. Hydrometeorol.*, *5*(1), 100–109, doi:10.1175/1525-7541(2004)005<0100:TREOTA>2.0.CO;2.

Wofsy, S. C., M. L. Goulden, J. W. Munger, S. M. Fan, P. S. Bakwin, B. C. Daube, S. L. Bassow, and F. A. Bazzaz (1993), Net exchange of CO_2 in a midlatitude forest, *Science*, *260*(5112), 1314–1317, doi:10.1126/science.260.5112.1314.

Yan, J., and C. T. Haan (1991a), Multiobjective parameter estimation for hydrologic models—Weighting of errors, *Trans. ASAE*, *34*(1), 135–141.

Yan, J., and C. T. Haan (1991b), Multiobjective parameter estimation for hydrologic models—Multiobjective programing, *Trans. ASAE*, *34*(3), 848–856.

Yilmaz, K. K., H. V. Gupta, and T. Wagener (2008), A process-based diagnostic approach to model evalutation: Application to the NWS distributed hydrologic model, *Water Resour. Res.*, *44*, W09417, doi:10.1029/2007WR006716.

Zadeh, L. A. (1963), Optimality and non-scalar valued performance criteria, *IEEE Trans. Autom. Control*, *8*, 59–60.

L. G. G. de Gonçalves, Centro de Previsão de Tempo e Estudos Climáticos, Instituto Nacional de Pesquisas Espaciais, Rodovia Presidente Dutra, Km 40, SP-RJ Cachoeira Paulista, SP 12630-000, Brazil.

H. V. Gupta, R. Rosolem, and W. J. Shuttleworth, Department of Hydrology and Water Resources, University of Arizona, 1133 East James E. Rogers Way, Tucson, AZ 85721, USA. (rosolem@email.arizona.edu)

X. Zeng, Department of Atmospheric Sciences, University of Arizona, 1118 East 4th St., PO Box 210081, Tucson, AZ 85721, USA.