Computers & Geosciences 83 (2015) 127–138

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Active subspaces for sensitivity analysis and dimension reduction of an integrated hydrologic model

Jennifer L. Jefferson [a,c,d,*], James M. Gilbert [a,c,d], Paul G. Constantine [b], Reed M. Maxwell [a,c,d]

[a] Colorado School of Mines, Hydrologic Science and Engineering Program, 1500 Illinois Street, Golden, CO 80215, USA
[b] Colorado School of Mines, Department of Applied Mathematics and Statistics, 1500 Illinois Street, Golden, CO 80215, USA
[c] Colorado School of Mines, Geology and Geological Engineering Department, 1500 Illinois Street, Golden, CO 80215, USA
[d] Integrated Ground Water Modeling Center, 1500 Illinois Street, Golden, CO 80215, USA

ARTICLE INFO

Article history:
Received 22 December 2014
Received in revised form
2 July 2015
Accepted 6 July 2015
Available online 14 July 2015

Keywords:
Active subspaces
Dimension reduction
Energy flux
Hydrologic model
Sensitivity

ABSTRACT

Integrated hydrologic models coupled to land surface models require several input parameters to characterize the land surface and to estimate energy fluxes. Uncertainty of input parameter values is inherent in any model and the sensitivity of output to these uncertain parameters becomes an important consideration. To better understand these connections in the context of hydrologic models, we use the ParFlow-Common Land Model (PF-CLM) to estimate energy fluxes given variations in 19 vegetation and land surface parameters over a 144-hour period of time. Latent, sensible and ground heat fluxes from bare soil and grass vegetation were estimated using single column and tilted-v domains. Energy flux outputs, along with the corresponding input parameters, from each of the four scenario simulations were evaluated using active subspaces. The active subspace method considers parameter sensitivity by quantifying a weight for each parameter. The method also evaluates the potential for dimension reduction by identifying the input–output relationship through the active variable – a linear combination of input parameters. The aerodynamic roughness length was the most important parameter for bare soil energy fluxes. Multiple parameters were important for energy fluxes from vegetated surfaces and depended on the type of energy flux. Relationships between land surface inputs and output fluxes varied between latent, sensible and ground heat, but were consistent between domain setup (i.e., with or without lateral flow) and vegetation type. A quadratic polynomial was used to describe the input–output relationship for these energy fluxes. The reduced-dimension model of land surface dynamics can be compared to observations or used to solve the inverse problem. Considering this work as a proof-of-concept, the active subspace method can be applied and extended to a range of domain setups, land cover types and time periods to obtain a reduced-form representation of any output of interest, provided that an active subspace exists.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Energy transfer into and out of the land surface is a primary driver of climate and weather patterns. Incoming solar radiation provides energy to change the phase of water through latent heat and heats the land surface and subsurface through sensible and ground heat, respectively. Latent heat fluxes redistribute water between the land and atmosphere through evaporation and transpiration while sensible and ground heat fluxes change the surface temperature and promote atmospheric circulation. Evapotranspiration depends on the surface temperature, is the largest component of the surface energy balance (Trenberth et al., 2009) and is responsible for precipitation that falls around the globe. Components of the surface energy balance are difficult to measure because they vary in both space and time. As a result, hydrologic models emerge as one approach to obtain estimates of these fluxes; other approaches such as remote sensing (Carlson et al., 1995; Mu et al., 2007) also exist. As computing power has increased these models have become more complex, often requiring several input parameters to model interacting systems. If the complexity of models can be reduced while still maintaining accuracy, understanding the distribution of water and energy throughout the hydrologic system becomes more accessible.

Hydrologic models can be used to estimate energy and water fluxes for a given set of environmental conditions. Unlike measurements that capture conditions over a finite period of time, models can simulate a range of scenarios that cannot be observed

* Corresponding author at: Colorado School of Mines, Hydrologic Science and Engineering Program, 1500 Illinois Street, Golden, CO 80215, USA.
E-mail address: jejeffer@mymail.mines.edu (J.L. Jefferson).

http://dx.doi.org/10.1016/j.cageo.2015.07.001
0098-3004/© 2015 Elsevier Ltd. All rights reserved.

but are of interest to researchers and end-users of model output. Representing complex, natural systems in such models can require between 10 and 50 surface and subsurface input parameters (Bastidas et al., 1999). Unfortunately, many input parameters are difficult to measure or obtain, which introduces uncertainty in the precise values required for model simulations (Beven, 1989). It is important to understand how uncertain inputs influence model output.

Recently, integrated hydrologic models have emerged as an option to combine attributes of groundwater and land surface models. Groundwater models handle flow within the saturated zone but often poorly represent the unsaturated zone. Land surface models estimate water and energy fluxes and may simulate the unsaturated zone. Traditionally, land surface models have limited capabilities to account for the saturated zone and do not consider subsurface flow in three dimensions (Beven, 1997; Henderson-Sellers et al., 1995; Liang et al., 1994). Integrated models compute the flow of water in both unsaturated and saturated regions of the subsurface, and account for spatial and temporal variations in water table depth. Subsurface conditions have been shown to influence surface energy fluxes (Ferguson and Maxwell, 2011; Mahmood and Hubbard, 2003; Rihani et al., 2010; Szilagyi et al., 2012), so integrated hydrologic models should provide more realistic representations of natural systems, given that their parameters are well constrained by observations. Historically, integrated models have been less studied but given their growing use it is important to consider the sensitivities of these types of models.

Several studies analyze the sensitivity of energy balance estimates to changes in inputs for land surface models (Table 1). Some studies use a one-at-a-time (OAT) approach to evaluate how output changes as a result of changing one input parameter, while others use methods that account for parameter interactions. Depending on the method, the number of hydrologic model simulations required to conduct a sensitivity analysis (SA) study can vary over five orders of magnitude. Detailed descriptions of the methods included in Table 1 can be found in the respective references and are summarized in Beringer et al. (2002). In addition to differences in methods, SA studies of hydrologic models have varying scopes where some compare findings between multiple models while others focus on one model. The studies in Table 1 identify important model input parameters like leaf area index, aerodynamic roughness length and stomatal resistance, but often lack specific recommendations for how to use this information. Other studies, not included in Table 1, focus on the sensitivity of energy fluxes to meteorological inputs, subsurface properties or the spatial scale of the domain, as opposed to the model inputs used to characterize the land surface (Abramowitz et al., 2008; Chen and Dudhia, 2001; Pau et al., 2014; Rihani et al., 2010).

Only one of the studies included in Table 1 evaluates the sensitivity of an integrated hydrologic model. Srivastava et al. (2014) evaluated the sensitivity of latent heat and streamflow to soil and surface parameters using the ParFlow-Common Land Model (PF-CLM). Single column OAT SA were completed to determine the most important CLM parameters (i.e., leaf area index, field capacity, stem area index, wilting point and aerodynamic roughness length). The sensitivity of latent heat and streamflow estimates were then assessed using the Morris OAT method and a watershed-specific domain of the Santa Fe River Basin located in northern Florida (Srivastava et al., 2014). Vertically-variable subsurface units and spatially-variable surface characteristics were used to represent the domain. Results of the Morris OAT show that latent heat is sensitive to the hydraulic conductivity in both confined and unconfined regions of the domain and that leaf area index (in confined regions) and wilting point (in unconfined regions) are also important input parameters (Srivastava et al., 2014).

In this work we use the recently developed active subspace method (Constantine, 2015) to build a low-dimensional model of the relationship between energy flux predictions from PF-CLM and 19 of its input parameters. The active subspace also identifies the parameters whose perturbations cause the greatest change in predictions, and we compare these important parameters to previous sensitivity analyses. Different from Srivastava et al. (2014), this work focuses only on land surface parameters, evaluates additional energy fluxes, uses a different SA method and presents reduced-dimension models of surface energy fluxes. Furthermore, we study two idealized domains which allows us to more easily interpret the SA and dimension reduction results. The low-

Table 1
Summary of energy flux sensitivity analysis (SA) studies, methods and sensitive land surface parameters.

Reference	Model[a]	SA method(s)[b]	Sensitive land surface parameters	Number of simulations
Bastidas et al. (1999)	BATS	MOGSA	*lai*, *lai0*, *sai*, *z0m*, stomatal resistance, soil and root layer depths, root function[c]	750–3000
Beringer et al. (2002)	NCAR LSM	Reduced form model	*lai*, *z0m*, displacement height	200
Collins and Avissar (1994)	LAID	FAST	*lai*, *z0m*, stomatal resistance	13,000
Franks et al. (1997)	TOPUP	RSA	Surface resistance[c]	20,000
Gao et al. (1996)	BATS	Feature curves	*wp*, *z0m*[c]	>150
Göhler et al. (2013)	CLM3.5	Eigendecomposition	*rl_n*, *rs_n*, soil resistance factor and layer depth, photosynthesis factors[c]	3000
Henderson-Sellers (1993)	BATS	Factorial	*z0m*, photosynthesis factor[c]	96
Hou et al. (2012)	CLM4	MRE		128
Jacquemin and Noilhan (1990)	NP-89	OAT	*lai*, *z0m*, stomatal resistance[c]	>50
Li et al. (2013)	CLM	Local, SOT, MARS, delta, Morris OAT	*dl*, *rl_n*, *tl_n*, *wp*, *z0m*, photosynthesis factor[c]	≈ 400
Liang and Guo (2003)	10 LSMs	Factorial	*lai*, stomatal resistance[c]	>1000
Liu et al. (2004)	NCAR LSM	MOGSA	*dmx*, *rl_n*, *tl_n*, *xl*, canopy and displacement height, photosynthesis factor[c]	500–20,000
Pitman (1994)	BATS	OAT	*lai*, *sc*, *z0m*, root layer depth[c]	≈ 200
Rosero et al. (2010)	Noah LSM	Sobol'	*lai*, stomatal resistance[c]	405,000
Schwinger et al. (2010)	CLM3.5	Linearized model	*lai*, *z0m*[c]	72
Srivastava et al. (2014)	PF-CLM	Morris OAT	*lai*, *wp*[c]	340

[a] Biosphere–Atmosphere Transfer Scheme (BATS), Common Land Model (CLM), Land–Atmosphere Interactive Dynamics (LAID), Land surface model (LSM), National Center for Atmospheric Research (NCAR), Noilhan Planton (NP-89), Variable Infiltration Capacity (VIC).

[b] Fourier amplitude sensitivity test (FAST), Multiobjective generalized sensitivity analysis (MOGSA), Multivariate adaptive regression splines (MARS), Minimum Relative Entropy (MRE), one-at-a-time (OAT), regionalized sensitivity analysis (RSA), sum-of-trees (SOT).

[c] Study included subsurface parameters that were identified as important; see reference for additional details.

dimensional model is used to gain insight into the behavior of latent, sensible and ground heat fluxes for a specific set of meteorological conditions. Active subspaces are novel tools for dimension reduction and have not been previously applied within the field of hydrology.

2. PF-CLM integrated hydrologic model

2.1. Model applications and description

PF-CLM has been used to investigate a range of hydrologic questions in both idealized and watershed-specific domains. Several studies have used PF-CLM to identify and understand relationships between subsurface heterogeneity and energy fluxes (Condon et al., 2013; Kollet, 2009; Maxwell and Kollet, 2008; Rihani et al., 2010). PF-CLM has also been used to compare energy fluxes before and after land cover changes due to the infestation and spread of the mountain pine beetle in the Rocky Mountains of Colorado (Mikkelson et al., 2013). Irrigation and water management decisions in agricultural settings perturb the hydrology of a watershed by altering energy feedbacks between the subsurface and land surface; this has been studied extensively using a PF-CLM model of the Little Washita watershed in Oklahoma (Condon and Maxwell, 2014a, 2014b, 2013; Ferguson and Maxwell, 2011).

Latent, sensible and ground heat fluxes are the three major components of the surface energy balance. The CLM parameterizations used to quantify these fluxes originate from the Biosphere Atmosphere Transfer Scheme (BATS; Dai et al., 2003) but have been modified to incorporate subsurface moisture conditions. Subsurface pressures are approximated in PF using the Richards' equation:

$$S_s S\left(\psi_p\right)\frac{\partial \psi_p}{\partial t} + \phi\frac{\partial S\left(\psi_p\right)}{\partial t} = \nabla\cdot\left[-\mathbf{K_s}(x)k_r\left(\psi_p\right)\cdot\nabla\left(\psi_p - z\right)\right] + q_s \quad (1)$$

where S_s is the specific storage [L^{-1}], S is the relative saturation [dimensionless] which is a function of the pressure head ψ_p [L], t is time [T], ϕ is the porosity of the soil [dimesionless], $\mathbf{K_s}(x)$ is the saturated hydraulic conductivity tensor [LT^{-1}], k_r is the relative permeability [dimensionless] which depends on the pressure head, z is the depth below the surface [L] and q_s is the source/sink term which includes infiltration, evaporation and transpiration [T^{-1}]. PF approximates the subsurface pressure head using a cell-centered finite difference approach in space and an implicit backward Euler scheme in time; the discretized system is solved with a Newton–Krylov method (Jones and Woodward, 2001; Maxwell, 2013; Osei-Kuffuor et al., 2014). This method, along with the use of preconditioners to accelerate the solving process, has been used in several geoscience applications (Dawson et al., 1997; Jenkins et al., 1999; White and Borja, 2011).

The parameterizations of latent, sensible and ground heat require several input parameters to describe vegetation and soil properties (Table 2). Parameters used to describe the root distribution or to compute rates of photosynthesis are not included here. The selected input parameters are connected to the output energy fluxes through a series of intermediate variables (Fig. 1). Latent heat is computed differently depending on whether or not vegetation is present. Bare soil evaporation E_{gr} ($kgm^{-2}s^{-1}$) is computed as

$$E_{gr} = -\beta\rho_a u_* q_* \quad (2)$$

where β (dimensionless) is a soil resistance factor, ρ_a (kgm^{-3}) is the air density, u_* (ms^{-1}) is the friction velocity and q_* (dimensionless) is the humidity scaling parameter. When vegetation is present on the land surface, evapotranspiration is computed as

$$E_{veg} = \left[R_{pp,dry} + L_w\right]L_{SAI}\left[\frac{\rho_a}{r_b}\left(q_{sat} - q_{af}\right)\right] \quad (3)$$

where the air density, boundary resistance factor r_b (sm^{-1}), saturated humidity at the land surface q_{sat} (dimensionless) and humidity within the canopy q_{af} (dimensionless) combine to form the potential evapotranspiration. Potential evapotranspiration is

Table 2
Selected bare soil and grass surface input parameters.

Parameter description	Name	Distribution (Range)[a]	Default value	Units
Maximum dew on canopy	*dmx*	U(0.05, 0.15)	0.1	mm
Leaf area index (maximum)	*lai*[b]	U(1, 4)	2	–
Leaf area index (minimum)	*lai0*	U(0.05, 3)	0.5	–
Stem area index	*sai*	U(0.5, 4)	4	–
Leaf and stem orientation index	*xl*	U(−0.3, 0.4)	−0.3	–
Leaf dimension	*dl*[c]	U(0, 0.16)	0.04	m
Leaf reflectance (visible light)	*rl_v*[d]	U(0.06, 0.13)	0.11	–
Leaf reflectance (near infrared light)	*rl_n*[d]	U(0.35, 0.55)	0.58	–
Stem reflectance (visible light)	*rs_v*[d]	U(0.3, 0.4)	0.36	–
Stem reflectance (near infrared light)	*rs_n*[d]	U(0.5, 0.6)	0.58	–
Leaf transmittance (visible light)	*tl_v*[d]	U(0.01, 0.09)	0.07	–
Leaf transmittance (near infrared light)	*tl_n*[d]	U(0.1, 0.45)	0.25	–
Stem transmittance (visible light)	*ts_v*[d]	U(0.15, 0.25)	0.22	–
Stem transmittance (near infrared light)	*ts_n*[d]	U(0.3, 0.4)	0.38	–
Field capacity	*fc*[e]	U(0.3, 1)	1	–
Wilting point	*wp*[e]	U(0.1, 0.3)	0.1	–
Aerodynamic roughness length (bare soil)	*z0m*[f]	LN(−8.83, 1.47)	0.01	m
Aerodynamic roughness length (grass)	*z0m*[f]	LN(−5.11, 0.52)	0.03	m
Soil color	*sc*	U(1, 8)	2	–
Water table depth (below surface)	*wt*	U(−10, −1)	–	m

[a] Uniform (minimum value, maximum value) and LogNormal (mean, standard deviation).
[b] Friedl et al. (1994).
[c] Smith and Geller (1980).
[d] Asner et al. (1998).
[e] Campbell and Norman (1998).
[f] Park et al. (2010), Wieringa (1993).

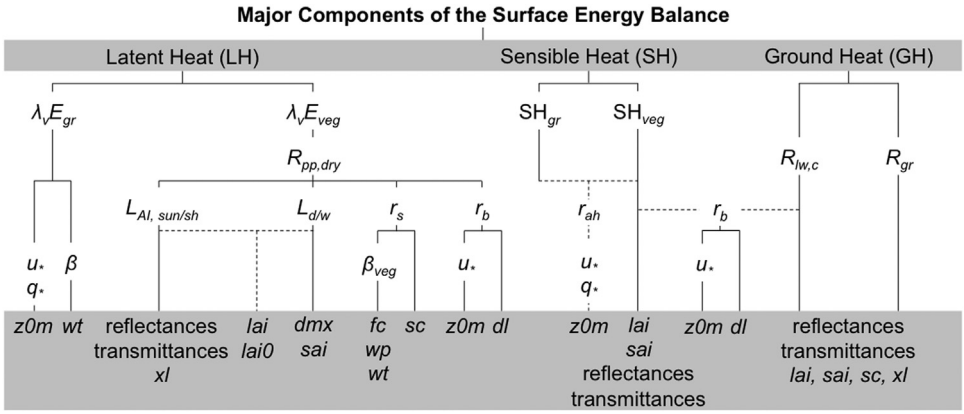

Fig. 1. Major connections between latent, sensible and ground heat output fluxes and input parameters in PF-CLM. Dashed lines indicate that intermediate variables and corresponding input parameters are similar for each of the variables along the solid lines; for example, both SH_{veg} and $R_{lw,c}$ depend on r_b.

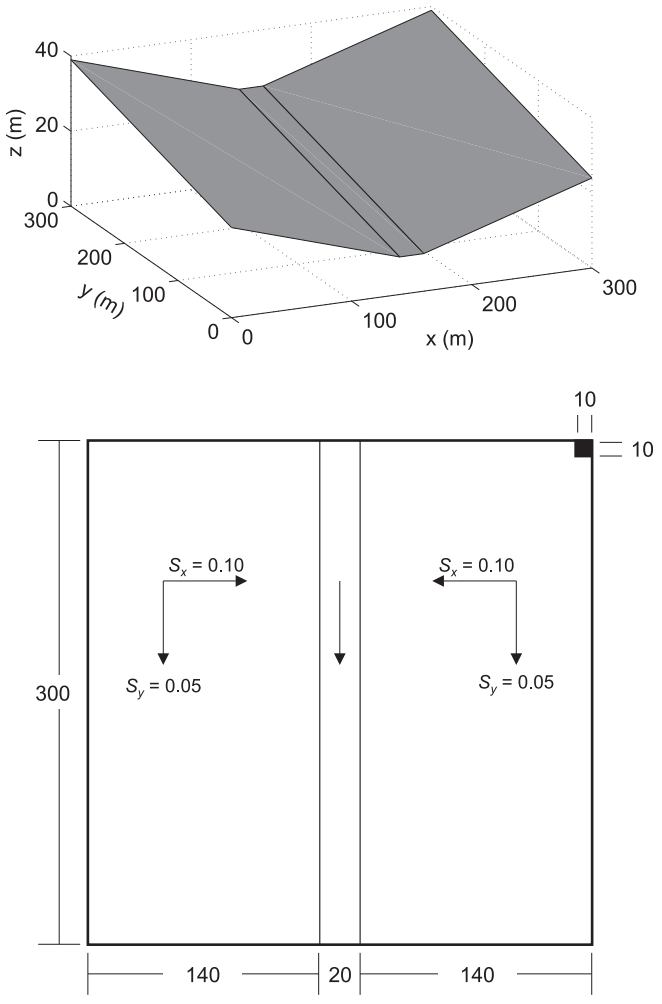

Fig. 2. Three dimensional (top) and plan view (bottom) of the tilted-v domain. Units of plan view are in meters. S_x and S_y are slopes in the x- and y-directions, respectively, and the dark square indicates a resolution of 10 m.

partitioned between transpiration $R_{pp,dry}$ (dimensionless), which depends on the dry fraction of the canopy L_d (dimensionless), and evaporation, which depends on the fraction of foliage covered by the water L_w (dimensionless). The sum of the leaf and stem area indices L_{SAI} (dimensionless) quantifies the total surface from which evaporation can occur. Transpiration only occurs from the leaf

Table 3
Model setup for single column and tilted-v model domains.

Parameter	Single Column	Tilted-V	Units
Dimensions			
Length	10	300	m
Width	10	300	m
Thickness	10	10	m
Δx	10	10	m
Δy	10	10	m
Δz	0.1	0.1	m
Subsurface			
Saturated hydraulic conductivity, K_{sat}	0.04465		mh^{-1}
Specific storage, S_s	1.00E−06		m^{-1}
van Genuchten α	4.0738		m^{-1}
van Genuchten n	2.19		–
Residual saturation, S_{res}	0.11		–
Porosity, Φ	0.512		%
Surface			
Slope (x, y)	5, 0	10, 5	%
Land cover	bare soil, grass		–
Manning's roughness	1E−06		$hm^{-1/3}$
Forcing			
Climate	Plains		–
Event duration	144		h
Δt	1		h

surface, quantified by the leaf area index L_{AI} (dimensionless), and from the dry fraction of the canopy, so $R_{pp,dry}$ is computed as

$$R_{pp,dry} = \frac{L_d r_b}{L_{AI}}\left[\frac{L_{AI,sun}}{r_b + r_{s,sun}} + \frac{L_{AI,sh}}{r_b + r_{s,sh}}\right] \tag{4}$$

where r_s (sm^{-1}) is the stomatal resistance. The leaf area index and stomatal resistance are further partitioned between the sunlit (subscript *sun*) and shaded (subscript *sh*) portions of the canopy. When photosynthesis is not limited by light r_s depends on the subsurface moisture, which is accounted for through

$$\beta_{veg} = \sum_{l=1}^{l=10} f_{root,l}\frac{\phi S_l - \phi S_{wp}}{\phi S_{fc} - \phi S_{wp}} \tag{5}$$

where S_l is the degree of saturation (passed from PF to CLM) and $f_{root,l}$ (dimensionless) is the fraction of roots in soil layer l. PF and CLM exchange data across the 10 layers directly beneath the surface. Bare soil evaporation and evapotranspiration are expressed as

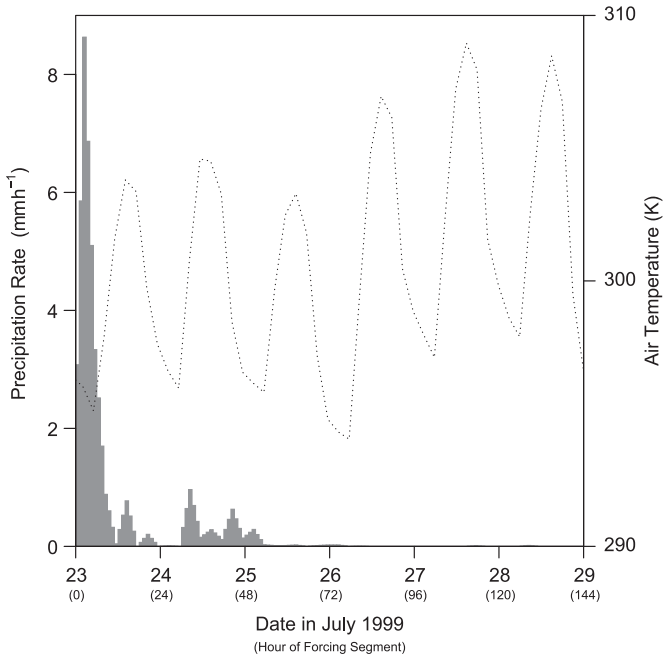

Fig. 3. Precipitation rates and air temperature (dashed line; right axis) for July 23-29, 1998 (144 h) for the Little Washita watershed in Oklahoma.

latent heat fluxes when multiplied by the latent heat of vaporization λ_v (2.5104e–6 Jkg^{-1}).

Sensible heat fluxes also depend on whether or not vegetation is present. Sensible heat from bare soil SH_{gr} (Wm^{-2}) is computed as

$$SH_{gr} = \rho_a c_p \frac{T_{gr} - T_{air,gr}}{r_{ah}} \tag{6}$$

where c_p is the specific heat of dry air (1004.67 Jkg^{-1}K^{-1}), $T_{air,gr}$ is the air temperature at the ground surface (K), T_{gr} is the ground temperature (K), which depends on the subsurface pressure head, and r_{ah} is the aerodynamic resistance factor (sm^{-1}) which accounts for atmospheric stability. The sensible heat flux from a vegetated surface SH_{veg} (Wm^{-2}) is computed as

$$SH_{veg} = \rho_a c_p \frac{L_{SAI}}{r_b}\left(C_{veg} T_{veg} - C_{air,gr} T_{air,gr} - C_{gr} T_{gr}\right) \tag{7}$$

where C_{veg} is the normalized heat conductance of vegetation, T_{veg} is the vegetation temperature, $C_{air,gr}$ is the normalized heat conductance of air at the ground surface and C_{gr} is the normalized heat conductance of the ground.

The ground heat flux GH (Wm^{-2}) depends strongly on longwave radiation and accounts for both latent and sensible heat fluxes from bare soil, if present

$$GH = R_{gr} + R_{lw,c} + \varepsilon_{gr} R_{lw} - \varepsilon_{gr}\sigma T^3_{gr,\,t-1}\left(T_{gr,t-1} + 4\Delta T_{gr}\right) - SH_{gr} - \lambda_v E_{gr} \tag{8}$$

where R_{gr} is the radiation absorbed by the ground surface (Wm^{-2}), $R_{lw,c}$ is the longwave radiation below the canopy (Wm^{-2}), R_{lw} is the downward longwave radiation obtained from the atmospheric forcing (Wm^{-2}), ε_{gr} is the emissivity of the ground taken as 0.96 (dimensionless), σ is the Stefan–Boltzmann constant (5.67e−8 Wm^{-2}K^{-4}), $T_{gr,t-1}$ is the ground temperature at the previous time step and ΔT_{gr} is the difference in soil temperature between the previous and current time step.

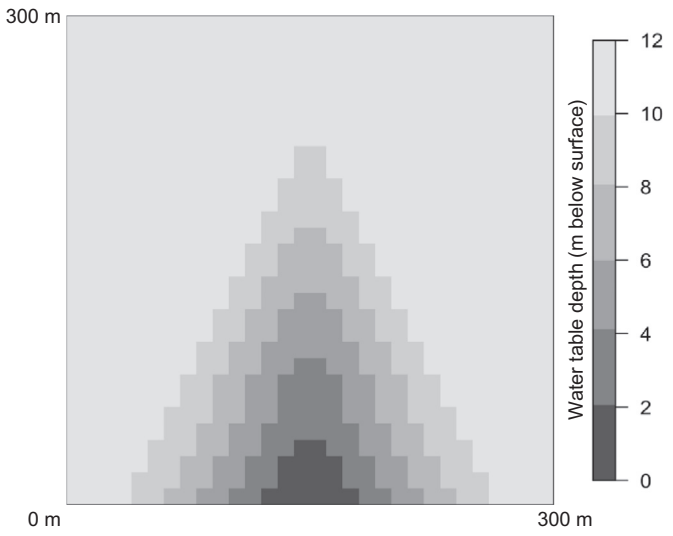

Fig. 4. Initial depth to water table for the tilted-v domain simulations. Orientation of figure is similar to Fig. 2.

2.2. Domain setup

Two domains are used for the sensitivity analysis – a single column and a tilted-v (Fig. 2). The single column represents the simplest hydrologic case with no lateral flow. The tilted-v is a standard test problem in hydrology (Kollet and Maxwell, 2006; Maxwell et al., 2014; Panday and Huyakorn, 2004; Sulis et al., 2010) and represents an idealized watershed with lateral subsurface and overland flow; two hillslopes intersect to form a channel that directs flow out of the domain. The dimensions and discretization of each domain were selected so that the single column domain represents approximately one column of the tilted-v domain (Table 3). Each domain was simulated with uniform land cover, either as bare soil or grass, for a total of four cases. In order to isolate the sensitivity of land surface parameters, subsurface parameters were set to be homogeneous and representative of an average loam soil for all simulations (Table 3; Schaap and Leij, 2000).

Computing energy fluxes in PF-CLM requires meteorological forcing data. Eight atmospheric variables are required for each time step of the model simulation: shortwave radiation, longwave radiation, precipitation rate, air temperature, east–west and north–south wind speeds, atmospheric pressure and specific humidity. A 144-hour segment of meteorological data from the Little Washita watershed in Oklahoma is used for both the single column and tilted-v domains (Fig. 3). In the tilted-v simulations the water table drains by gravity to reach a steady state of subsurface storage before the forcing is applied (Fig. 4); this ensures lateral flow is occurring and that energy fluxes are a result of the precipitation event. The 144-hour segment of meteorological data from July 23–29, 1998 was selected because it captures both wet and dry periods in the late summer when the energy fluxes are expected to be near their maximum values. Evaluating the domains on a seasonal or annual basis is likely of interest for practical applications, but for purposes of this study a short period was selected so that multiple simulations could be completed quickly.

3. Active subspace method

In this hydrologic application active subspaces are used to study the relationship between the PF-CLM input parameters and the predicted quantities of interest – namely, latent heat, sensible

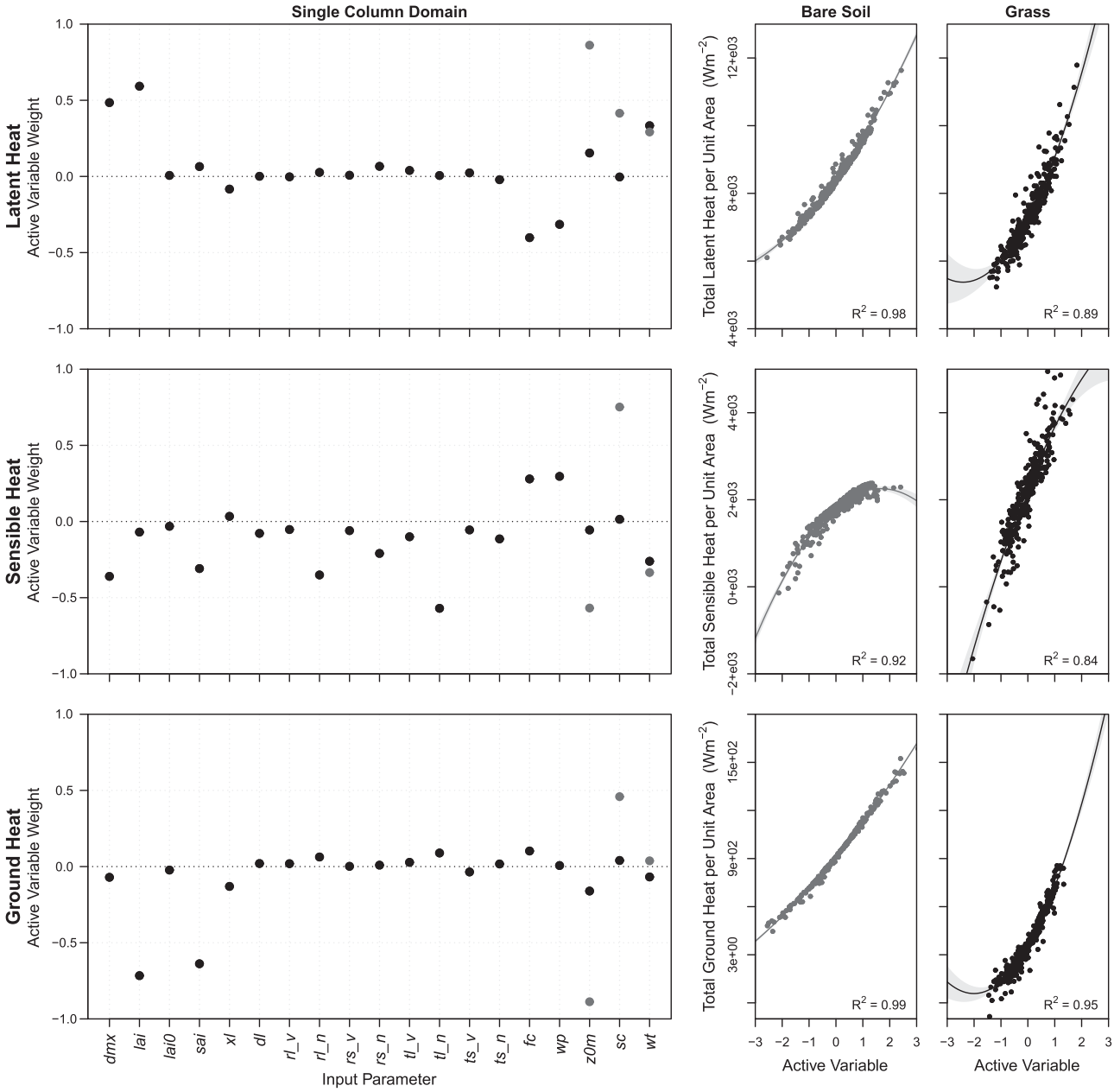

Fig. 5. Input parameter weights (left column) and sufficient summary plots (right columns) for latent, sensible and ground heat fluxes from bare soil (gray) and grass (black) single column domain simulations. The quadratic polynomial (solid line), 95% confidence interval (light gray shading) and R^2 value are shown on the sufficient summary plots for each land cover and flux type.

heat and ground heat. The active subspace is the span of a set of directions. These directions are eigenvectors of a matrix derived from the gradient of the map between model inputs and quantities of interest. Parameter perturbations along these directions change quantities of interest more, on average, than parameter perturbations in orthogonal directions. To compute these directions, one randomly samples the gradient and computes the uncentered principal components (as opposed to centered principle components where the mean is subtracted from the data) from the collection of samples. However, this method requires access to the gradient of the quantity of interest with respect to the parameters, which is not available in PF-CLM. Finite difference approximations are not feasible, since they would require up to 19 runs per gradient sample.

Without gradients, one must estimate gradients with a model. In this case, a linear model is used to approximate the input–output map which is discussed in Chapter 1 of Constantine (2015) and briefly summarized below. The normalized gradient of the linear model defines one direction in the 19-dimensional space of input parameters. This approach is appropriate when (i) the active subspace is one-dimensional and (ii) the quantity of interest is roughly monotonic with respect to each of its input parameters. Both of these conditions can be verified with a sufficient summary plot, which displays the relationship between the modeled quantity of interest and a linear combination of the input parameters (i.e., the active variable); the weights of the linear combination are the components of the linear model's normalized gradient. Each point on the sufficient summary plot corresponds to the inputs and corresponding output from one model realization.

The algorithm is written for a generic function $f(\mathbf{x})$, where f is the quantity of interest, and $\mathbf{x}$ is the vector containing the normalized input parameters; for PF-CLM, $\mathbf{x}$ contains 19 components.

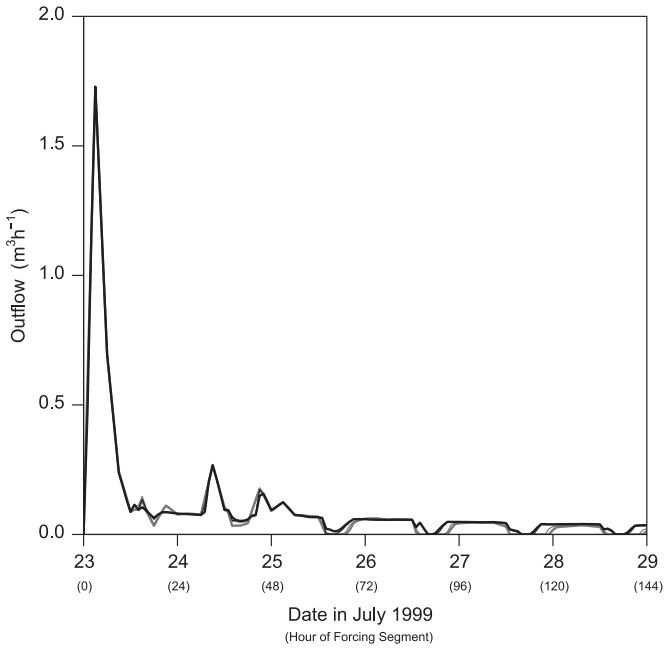

Fig. 6. Outflow from tilted-v domain for bare soil (gray) and grass (black) land cover over the 144-hour forcing segment.

The following procedure is applied to each of the four cases using $M=300$, for a total of 1200 simulations:

1. For $i = 1, \ldots, M$, draw $\mathbf{x}_i$ independently according to the joint density on $\mathbf{x}$.
2. For each $\mathbf{x}_i$, compute $f_i = f(\mathbf{x}_i)$, which involves shifting and scaling $\mathbf{x}_i$ to the PF-CLM input ranges and running PF-CLM.
3. Use least-squares to fit the coefficients $\hat{a}_0$ and $\hat{\mathbf{a}} = \left[\hat{a}_1, \ldots, \hat{a}_M\right]^T$ of a linear model,

$$f_i \approx \hat{a}_0 + \hat{\mathbf{a}}^T \mathbf{x}_i, \tag{9}$$

using the pairs $\mathbf{x}_i$ and f_i.
4. Compute the normalized gradient of the linear model

$$\hat{\mathbf{w}} = \frac{\hat{\mathbf{a}}}{\|\hat{\mathbf{a}}\|}. \tag{10}$$

5. Plot the pairs $\left(\hat{\mathbf{w}}^T \mathbf{x}_i,\ f_i\right)$, which is the sufficient summary plot.

If the sufficient summary plot reveals a strong, univariate relationship between $\hat{\mathbf{w}}^T\mathbf{x}$ and f, then

$$f(\mathbf{x}) = g\left(\hat{\mathbf{w}}^T \mathbf{x}\right) \tag{11}$$

where g is a univariate, scalar-valued function fit with the pairs $\left(\hat{\mathbf{w}}^T \mathbf{x}_i, f_i\right)$ – e.g., linear, quadratic or cubic polynomial. Fitting a function in this manner allows for visual confirmation of the relationship between the inputs and the outputs using the sufficient summary plot. Appendix A.1 contains the derivation of the previous algorithm.

Additionally, the weights of the linear combination (i.e., the components of $\hat{\mathbf{w}}$) identify the input parameters that contribute the most to the one-dimensional active subspace. Parameters with relatively large (in magnitude) weights are the important parameters; perturbations in these parameters change the quantity of interest more than perturbations in other parameters. The sign of the weight indicates the direction moved along the horizontal axis (i.e., the active variable) of the sufficient summary plot as the input parameter value increases or decreases. In other words, the weights allow for analysis of the model's sensitivity to changes in parameter values.

Active subspaces not only provide insight into which parameters are important, but also approximate the relationship between model inputs and outputs with fewer parameters than the number of inputs. While this method has never been applied in the field of hydrology, successful applications to aerospace models include shape optimization (Lukaczyk et al., 2014) and safety engineering (Constantine et al., 2014).

4. Results and discussion

4.1. Single column domain

The weights in Fig. 5 quantify the importance of the parameters for the single column domain; weights with large magnitudes imply that the corresponding parameters cause more change in the heat fluxes than those with small weights. The aerodynamic roughness length (*z0m*) is the most important input parameter for bare soil latent and ground heat energy fluxes. Soil color is also important and influences the bare soil sensible heat flux the most. The magnitudes of soil color weights are greater for bare soil than for grass, where soil color showed little to no importance. The location of the initial water table influences latent and ground heat fluxes from grass more than from bare soil. As the initial water table depth decreases in magnitude (i.e., becomes closer to the surface) the latent heat flux increases whereas the sensible heat decreases.

The weights of the 16 remaining parameters for grass land cover are not consistent between the different fluxes. Latent heat is most influenced by the maximum leaf area index (*lai*) followed by the maximum dew that the canopy can hold (*dmx*), field capacity (*fc*), initial water table depth (*wt*), wilting point (*wp*) and aerodynamic roughness length. Similar to latent heat, the depth of dew, wilting point, field capacity and water table are important for sensible heat, but not as much as the near-infrared leaf transmittance (*tl_n*). Two other parameters, the near-infrared leaf reflectance (*rl_n*) and stem area index (*sai*), have approximately the same weights as the dew depth for sensible heat. Weights of the reflectance and transmittance input parameters for sensible heat have the largest magnitude and most variability compared to the other two fluxes. In contrast, weights for ground heat are the least variable and only two of the 19 parameters show significant influence (*lai* and *sai*).

4.2. Tilted-v domain

The tilted-v domain simulations do not include the initial water table as an input parameter because the spin-up process eliminates the need for a user-selected value. The primary difference between the single column and tilted-v domains is the presence of lateral subsurface flow. The tilted-v domain has flow out of the domain throughout most of the 144-hour forcing period (Fig. 6). Fig. 6 also shows that there is little variability in the predicted outflow between the bare soil and grass simulations. Only small differences (< 0.14) exist between input parameter weights for the single column and tilted-v domains (Fig. 7). The relationship between the input parameters and output fluxes for each domain setup are also similar. These similarities indicate that lateral flow does not change the parameter sensitivity or relationship for this particular simulation. Additional simulations comparing other

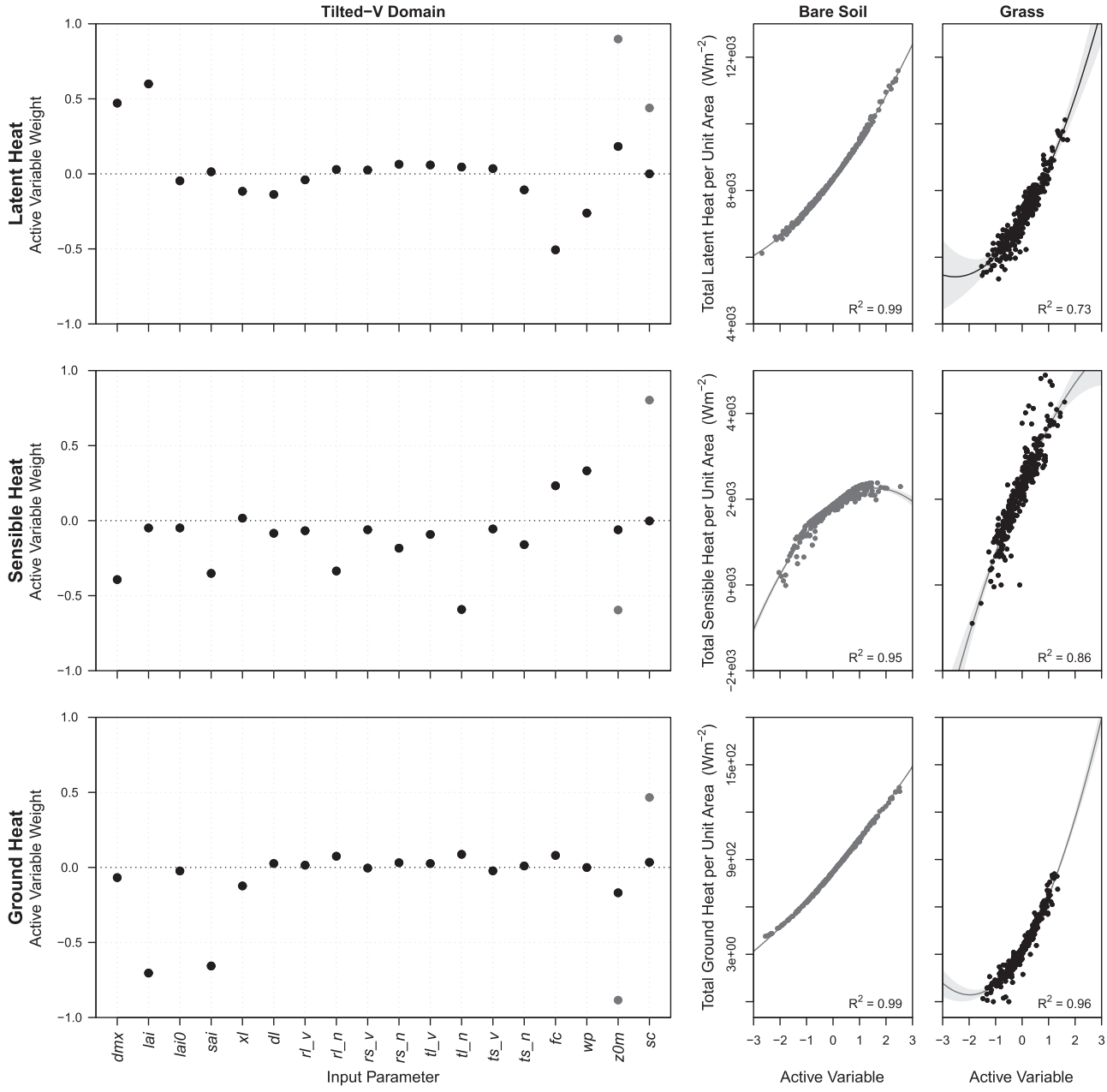

Fig. 7. Input parameter weights (left column) and sufficient summary plots (right columns) for latent, sensible and ground heat fluxes from bare soil (gray) and grass (black) tilted-v domain simulations. The quadratic polynomial (solid line), 95% confidence interval (light gray shading) and R^2 value are shown on the sufficient summary plots for each land cover and flux type.

domain setups and vegetation types are necessary to validate this preliminary conclusion.

4.3. Sufficient summary plots

In both domains, energy fluxes for bare soil and grass exhibit a univariate trend with respect to the linear combination of inputs (i.e. the active variable); see Figs. 5 and 7. The relationship for these scenarios is described by a quadratic polynomial:

$$g(y) \approx C_0 + C_1 y + C_2 y^2 \approx C_0 + C_1\left(\hat{\mathbf{w}}^T\mathbf{x}\right) + C_2\left(\hat{\mathbf{w}}^T\mathbf{x}\right)^2 \quad (12)$$

where

$$x = 2\left(\frac{x' - x'_L}{x'_U - x'_L}\right) - 1. \quad (13)$$

for variables sampled from a uniform distribution with upper x'_U and lower limits x'_L, and

$$x = \frac{\log(x') - \mu}{\sigma}, \quad (14)$$

for variables sampled from a lognormal distribution with a mean μ and standard deviation σ, where x is the normalized input parameter and x' is the actual PF-CLM input value sampled from its respective distribution. The function $g(y)$ returns the total latent, sensible or ground heat per unit area (Wm^{-2}) and y is the active variable – the weighted sum of the scaled input parameters (12). The total energy per unit area is obtained by multiplying the flux

Table 4
Summary of quadratic model coefficients.

Model output	C_0	C_1	C_2	R^2
Single column				
Bare soil				
Latent heat	8.41E+03	1.11E+03	1.04E+02	0.98
Sensible heat	1.81E+03	5.27E+02	−1.57E+02	0.92
Ground heat	9.10E+02	2.05E+02	1.01E+01	0.99
Grass				
Latent heat	7.22E+03	1.53E+03	3.17E+02	0.89
Sensible heat	2.28E+03	1.54E+03	−1.48E+02	0.84
Ground heat	3.54E+02	2.95E+02	7.36E+01	0.95
Tilted-V				
Bare soil				
Latent heat	8.34E+03	1.06E+03	9.74E+01	0.99
Sensible heat	1.86E+03	4.99E+02	−1.56E+02	0.95
Ground heat	8.22E+02	1.95E+02	8.99E+00	0.99
Grass				
Latent heat	7.09E+03	1.33E+03	2.62E+02	0.73
Sensible heat	2.36E+03	1.47E+03	−1.48E+02	0.86
Ground heat	3.19E+02	2.80E+02	7.08E+01	0.96

rate from each CLM tile by the tile area (i.e., 100 m^2) and then normalizing by the entire domain area. Table 4 lists the coefficients C_0, C_1 and C_2 and the R^2 value for each land cover and domain setup. A comparison of R^2 values for a linear versus quadratic function are included in Appendix A.2 (Table A1). In general, the quadratic model fits better for bare soil simulations as compared to grass and for the tilted-v domain as compared to the single column domain (Table 4); better quadratic model fits coincide with circumstances where the water table has less influence. Increased scatter (i.e., lower R^2 values) in the sufficient summary plots of grass simulations could also be due to the larger number of input parameters. Sufficient summary plots for latent and sensible heat fluxes from the single columns have less scatter when the water table is not included as an input parameter (not shown here). From a physical perspective, the water table depth can have a more dramatic effect on energy fluxes than variations in the other parameters because conditions can range between saturated and dry over a short period of time. Soil moisture influences both the ground temperature as well as the amount of water available for evaporation and transpiration. As a result, it is possible that water table variations introduce nonlinear behavior that cannot be captured by a linear model.

The parameter weights and sufficient summary plots are a powerful combination of information that can be used to understand land surface dynamics. The sign of the input parameter weight can predict whether the output parameter will increase or decrease with changes in the input parameter. For example, since the leaf area index weight is positive for latent heat, greater leaf area index increases the latent heat flux. However, because the weight for the leaf area index is negative for ground heat, greater leaf area index decreases the ground heat flux. These relationships agree with physical intuition; more leaf area means evaporation and transpiration can occur from a larger area which results in a greater latent heat flux. Conversely, more leaf area decreases the

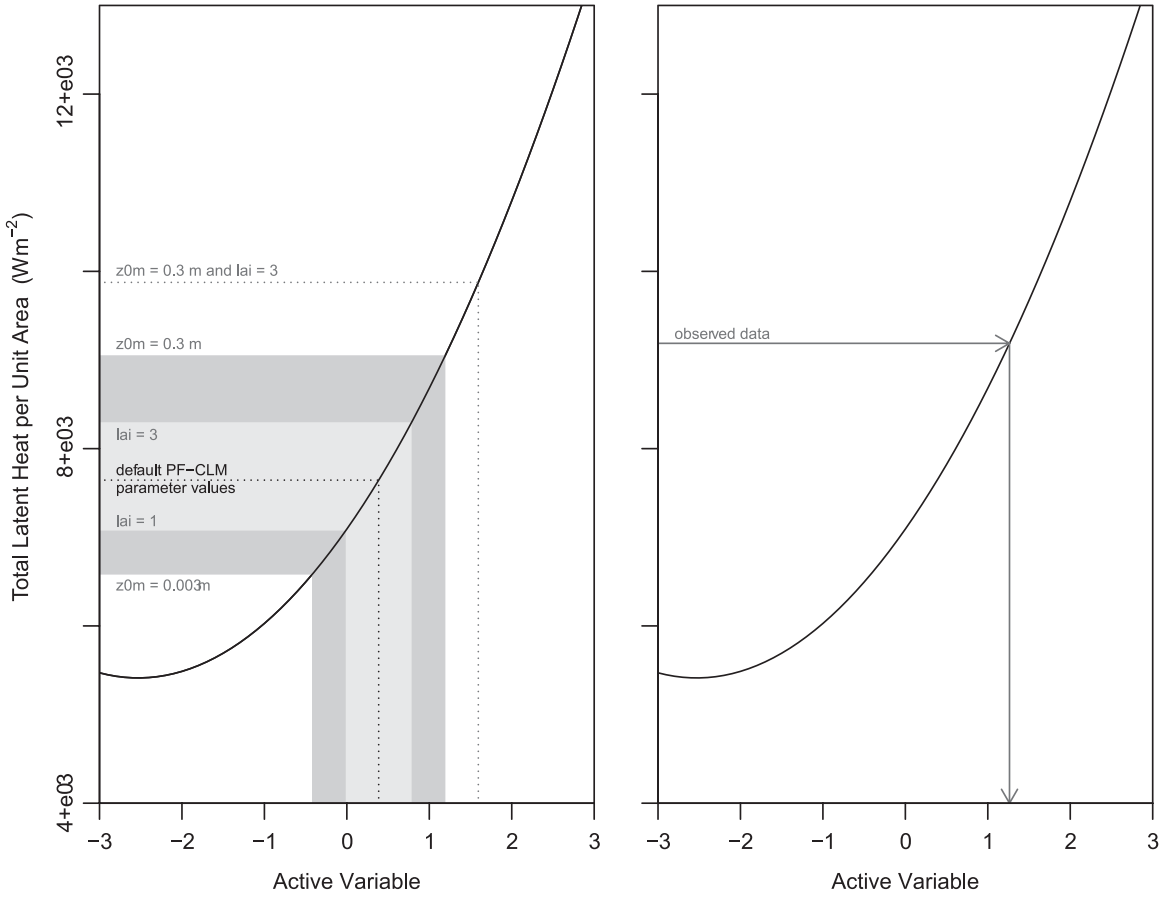

Fig. 8. Two examples of how sufficient summary plots can used (1) to evaluate how estimates of output change with respect to input parameter values (left) and (2) to predict input parameter values using observed data. The black dashed line in the left panel corresponds to the active variable computed using default PF-CLM values; shaded areas show the range of output estimates using the default values except for the parameter noted (e.g., *lai* and/or *z0m*).

amount of radiation reaching the ground surface and therefore reduces the ground heat flux. Similar logic can be used to explain the connection between other input parameters and their influence on the output of interest.

The leaf area index example can also be described quantitatively (Fig. 8). The active variable is 0.39 when all input parameters assume their default values (shown in Table 2). The active variable is simple to compute using (12) and the substitution shown in (11). It is easy to study how output estimates are affected by changes in input parameters. For the case of leaf area index, decreasing *lai* by one unit reduces the latent heat estimates by approximately 7.5% whereas increasing *lai* by one unit results in estimates approximately 8.5% larger (Fig. 8). Fig. 8 includes another example showing the sensitivity bounds as a result of changes in the aerodynamic roughness length. This type of evaluation can be completed for combinations of input parameters as well (e.g., both *lai* and *z0m*).

Univariate sufficient summary plots can also be used to solve the inverse problem. Observed data can be plotted on the *y*-axis and the input–output relationship can be used to identify the active variable which constrains the input parameter values (Fig. 8). There are an infinite number of input combinations that can produce the same active variable value. Choosing from these combinations requires either more data or a choice of regularization. Unfortunately, AmeriFlux energy flux data for Little Washita is not available for the 144-hour period analyzed here. Latent heat flux observations from two years prior are within the *y*-axis range of Fig. 8 which provides confidence that estimates using the simplified domain are the correct order of magnitude.

5. Conclusions

In this work the active subspace method was applied to the PF-CLM integrated hydrologic model to analyze the sensitivity of latent, sensible and ground heat fluxes to 19 land surface parameters. The use of each input parameter depends on the type of energy flux. The active subspace method identifies which of the input parameters are most important and how the combination of inputs relate to the output of interest. Of the three input parameters evaluated for bare soil, the aerodynamic roughness length was important for all energy fluxes. Estimates of energy fluxes from a vegetated surface were sensitive to approximately half of the 19 input parameters: five parameters exerted the most influence on the latent heat flux (*lai*, *dmx*, *fc*, *wp*, *z0m*), 6 on the sensible heat flux (*tl_n*, *dmx*, *rl_n*, *sai*, *wp*, *fc*) and 3 on the ground heat flux (*lai*, *sai*, *z0m*). The initial water table depth was also important for the grass single columns. Sensible heat fluxes were sensitive to a similar number of parameters as the latent heat fluxes and also showed the greatest variability in parameter weight magnitudes. The parameters with large weights are similar to parameters identified in SA studies completed on other hydrologic models (Table 1). Furthermore, the relationship between the input parameters and each output flux could be described using a quadratic function of the active variable for all cases.

In addition to comparing fluxes from bare soil and grass vegetation, two domain setups were evaluated to determine if lateral flow changed the input parameter sensitivity. The weights and relationships shown in the sufficient summary plots are similar for the single column and tilted-v domains suggesting that lateral flow has a negligible effect on the land surface parameter–flux relationship. While lateral flow will influence soil moisture magnitudes and patterns systematically across a domain, it will not change the relationship between soil moisture and ET. Srivastava et al. (2014) effectively assumed that lateral flow was not important when the parameters identified as sensitive from a single column domain were the only parameters varied in their watershed domain. These findings support their assumption, as long as PF-CLM has been simulated for a long enough duration that the subsurface storage is not changing appreciably. The active subspaces method could be applied to a watershed-domain because the required input parameters are the same regardless of the domain configuration, but vary in magnitude based on land cover type. However, the results shown here suggest that a simple, computationally cheaper domain can potentially be used to provide insight into more complex domains. The ability to isolate sensitive parameters without the expense of multiple simulations of a complex domain makes SA more tractable.

Results from this proof-of-concept example show how active subspaces can be used in the context of hydrology. The application of this method to hydrologic models has great potential; this method could be used to derive relationships between any combination of surface or subsurface inputs and outputs, for any climate, soil type or period of interest. We anticipate that this method is applicable during both water- and energy-limited times of the year but that the input parameter weights and input–output relationships will vary. For example, estimates of energy fluxes from a snow-dominated alpine location are likely to be sensitive to different input parameters in the summer (i.e., water-limited) than in the winter (i.e., energy-limited). Including a heterogeneous subsurface will further complicate feedbacks between the subsurface and atmosphere and it is plausible that more than one dimension will be needed to approximate the relationship between the inputs and outputs. For high-dimensional problems the dimension reduction from this method preserves the fine scale processes and physical intuition behind the model and works to average functions not quantities. Furthermore, the usefulness of sufficient summary plots allows for investigation into model response and behavior. Once a reduced-form model is established and validated, expensive simulations may be bypassed altogether. Calibration of input parameters using observations and the simplified system is possible if the simulation setup matches observation locations. Recent research using CLM suggests that reduced-form models can be used to reduce the number of simulations required for parameter optimization (Gong et al., 2015), adjust parameter values used to predict methane emission from wetlands (Müller et al., 2015) and calibrate parameters identified as important for latent heat estimates (Ray et al., 2015).

As hydrologic models continue to be used to estimate hydrologic outputs such as energy fluxes, the need for efficient and accessible means to evaluate the sensitivity and behavior of model output becomes even more important. While active subspaces approximate the physics of a modeled system they can improve our confidence and understanding of processes within the model and also provide ways to reduce the computational demands of completing multiple simulations of expensive domains. The analysis of two hypothetical domains suggests that use of this method can be extended far beyond this proof-of-concept example. Active subspaces have the potential to quantify uncertainty and reduce the dimension of other PF-CLM scenarios as well as be applied to other high-dimension hydrologic models.

Acknowledgments

This work was completed with the support of the National Science Foundation Water Sustainability and Climate Grant WSC-1204787 and Climate Change Water and Society (CCWAS) Integrated Graduate Education and Research Traineeship (IGERT) Program (http:// ccwas.ucdavis.edu/; DGE-1069333).

Appendix

A.1 Derivation of the active subspace method

The active subspace is derived for a general continuously differentiable, nonlinear, scalar-valued function of several variables; complete details are found in (Constantine, 2015). We denote this function by $f(\mathbf{x})$, where $\mathbf{x}$ is the vector of m continuous inputs, and f returns a scalar. The gradient $\nabla f(\mathbf{x})$ is the column-oriented vector of m partial derivatives with respect to the components of $\mathbf{x}$. We assume that the domain of f – i.e., the space of $\mathbf{x}$ – is equipped with a normalized probability density function $\rho(\mathbf{x})$, which enables us to compute weighted averages. In the PF-CLM grass-covered case, f is a heat flux, $\mathbf{x}$ has $m=19$ components, and ρ is the product of a uniform density over 18 parameters times a normal density over the log of roughness length ($z0m$).

The active subspace is defined by the eigenvectors of the following symmetric, positive semidefinite matrix,

$$\mathbf{C} = \int \nabla f \nabla f^T \rho d\mathbf{x} = \mathbf{W}\Lambda\mathbf{W}^T, \tag{A1}$$

where $\mathbf{W}$ is the orthogonal matrix of eigenvectors, and Λ is the diagonal matrix of non-negative eigenvalues in decreasing order. The ith eigenvalue λ_i satisfies

$$\lambda_i = \int \left(\mathbf{w}_i^T \nabla f\right)^2 \rho d\mathbf{x}, \tag{A2}$$

where $\mathbf{w}_i$ is the corresponding eigenvector. In words, this says that the λ_i quantifies the average squared change in the function subject to small perturbations along $\mathbf{w}_i$ – i.e., the directional derivative along $\mathbf{w}_i$. A large gap between large and small eigenvalues identifies an *active subspace*, defined by the eigenvectors corresponding to the large eigenvalues; a small perturbation to inputs along directions in the active subspace changes f more, on average, than a perturbation along the orthogonal inactive subspace.

To estimate the eigenvectors and eigenvalues, we estimate $\mathbf{C}$ using independent random samples of the gradient vector and compute its eigenvalue decomposition,

$$\mathbf{C} \approx \hat{\mathbf{C}} = \frac{1}{M}\sum_{i=1}^{M} \nabla f(\mathbf{x}_i)\nabla f(\mathbf{x}_i)^T = \hat{\mathbf{W}}\hat{\Lambda}\hat{\mathbf{W}}^T, \tag{A3}$$

where $\mathbf{x}_i$ are drawn independently at random according to $\rho(\mathbf{x})$. When gradients of f are not available, as in the case of PF-CLM, we must estimate the gradients from evaluations of f. Finite differences are inappropriate when m is large and f is expensive to evaluate and/or noisy. Another option is to fit a least-squares polynomial model to a set of pairs $(\mathbf{x}_j, f(\mathbf{x}_j))$ and compute the gradient of the polynomial approximation; this is the approach we take in Section 3. When the polynomial approximation is a linear function of $\mathbf{x}$ (i.e., a polynomial of degree at most 1 in each variable), the computation of $\hat{\mathbf{W}}$ reduces dramatically. The gradient of the global linear model is constant for all $\mathbf{x}$,

$$f(\mathbf{x}) \approx \hat{a}_0 + \hat{\mathbf{a}}^T\mathbf{x}, \ \nabla f(\mathbf{x}) \approx \hat{\mathbf{a}}. \tag{A4}$$

In this case, $\hat{\mathbf{C}}$ becomes

$$\hat{\mathbf{C}} \approx \frac{1}{M}\sum_{i=1}^{M} \hat{\mathbf{a}}\hat{\mathbf{a}}^T = \hat{\mathbf{a}}\hat{\mathbf{a}}^T = \hat{\mathbf{w}}\hat{\lambda}\hat{\mathbf{w}}^T, \tag{A5}$$

where $\hat{\lambda} = \|\hat{\mathbf{a}}\|^2$ and $\hat{\mathbf{w}} = \hat{\mathbf{a}}/\|\hat{\mathbf{a}}\|$ as in Eq. (10) in Section 3. This is identical to the algorithm in Section 3. With the linear model, we can identify only the one-dimensional active subspace. Input perturbations along any direction orthogonal to $\hat{\mathbf{w}}$ are deemed inactive. The appropriateness of this approach is validated by the sufficient summary plots as in Section 4.3. The strong univariate trends assure us that a one-dimensional active subspace is sufficient to explain the relationship between the inputs x and the outputs f.

Table A1
Comparison of R^2 values for sufficient summary plot relationships.

Model output	R^2 (quadratic)	R^2 (linear)
Single column		
Bare soil		
Latent heat	0.9846	0.9723
Sensible heat	0.9211	0.8300
Ground heat	0.9900	0.9907
Grass		
Latent heat	0.8881	0.8654
Sensible heat	0.8408	0.8353
Ground heat	0.9456	0.9167
Tilted-V		
Bare soil		
Latent heat	0.9900	0.9850
Sensible heat	0.9453	0.8349
Ground heat	0.9900	0.9945
Grass		
Latent heat	0.7255	0.7065
Sensible heat	0.8580	0.8525
Ground heat	0.9567	0.9268

A.2 Comparison of reduced dimension model fit

Table A1 displays the R^2 values for various polynomial models of the heat fluxes as a function of the active variable. These values suggest that a quadratic polynomial is an appropriate choice for the data shown in Figs. 5 and 7.

References

Abramowitz, G., Leuning, R., Clark, M., Pitman, A., 2008. Evaluating the performance of land surface models. J. Clim. 21, 5468–5481. http://dx.doi.org/10.1175/2008JCLI2378.1.

Asner, G.P., Wessman, C.A., Schimel, D.S., Archer, S., 1998. Variability in leaf and litter optical properties : implications for BRDF model inversions using AVHRR , MODIS, and MISR 257, 243–257.

Bastidas, L.A., Gupta, H.V., Sorooshian, S., Shuttleworth, W.J., Yang, Z.L., 1999. Sensitivity analysis of a land surface scheme using multicriteria methods. J. Geophys. Res. 104, 19481–19490. http://dx.doi.org/10.1029/1999JD900155.

Beringer, J., McIlwain, S., Lynch, A.H., Chapin III, F.S., Bonan, G.B., 2002. The use of a reduced form model to assess the sensitivity of a land surface model to biotic surface parameters. Clim. Dyn. 19, 455–466. http://dx.doi.org/10.1007/s00382-002-0237-9.

Beven, K., 1989. Changing ideas in hydrology-the case of physically-based models. J. Hydrol. 105, 157–172.

Beven, K., 1997. TOPMODEL : a critique. Hydrol. Process. 1085, 1069–1085.

Campbell, G.S., Norman, J.M., 1998. An Introduction to Environmental Biophysics, 2nd ed. Springer-Verlag, New York, Inc., New York.

Carlson, T.N., Taconet, O., Vidal, A., Gillies, R.R., Olioso, A., Humes, K., 1995. An overview of the workshop on thermal remote sensing held at La Londe les Maures, France, September 20–24, 1993. Agric. For. Meteorol. 77, 141–151.

Chen, F., Dudhia, J., 2001. Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. Mon. Weather Rev. 129, 569–585. http://dx.doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Collins, D.C., Avissar, R., 1994. An evaluation with the Fourier Amplitude Sensitivity Test (FAST) of which land-surface parameters are of greatest importance in atmospheric modeling. J. Clim. 7, 681–703.

Condon, L.E., Maxwell, R.M., 2013. Implementation of a linear optimization water allocation algorithm into a fully integrated physical hydrology model. Adv. Water Resour. 60, 135–147. http://dx.doi.org/10.1016/j.advwatres.2013.07.012.

Condon, L.E., Maxwell, R.M., 2014a. Groundwater-fed irrigation impacts spatially distributed temporal scaling behavior of the natural system: a spatio-temporal framework for understanding water management impacts. Environ. Res. Lett. 9, 9. http://dx.doi.org/10.1088/1748-9326/9/3/034009.

Condon, L.E., Maxwell, R.M., 2014b. Feedbacks between managed irrigation and water availability: diagnosing temporal and spatial patterns using an integrated hydrologic model. Water Resour. Res. 50, 17. http://dx.doi.org/10.1002/2013WR014868, Received.

Condon, L.E., Maxwell, R.M., Gangopadhyay, S., 2013. The impact of subsurface conceptualization on land energy fluxes. Adv. Water Resour. 60, 188–203. http://dx.doi.org/10.1016/j.advwatres.2013.08.001.

Constantine, P., Emory, M., Larsson, J., Iaccarino, G., 2014. Exploiting Active Subspaces to Quantify Uncertainty in the Numerical Simulation of the HyShot II Scramjet.

Constantine, P.G., 2015. Active Subspaces: Emerging Ideas in Dimension Reduction for Parameter Studies. SIAM, Philadelphia.

Dai, Y., Zeng, X., Dickinson, R.E., Baker, I., Bonan, G.B., Bosilovich, M.G., Denning, A. S., Dirmeyer, P.A., Houser, P.R., Niu, G., Oleson, K.W., Schlosser, C.A., Yang, Z.-L., 2003. The Common Land Model. Bull. Am. Meteorol. Soc. 84, 1013–1023. http://dx.doi.org/10.1175/BAMS-84-8-1013.

Dawson, C.N., Klie, H., Wheeler, M.F., Woodward, C.S., 1997. A parallel, implicit, cell-centered method for two-phase flow with a preconditioned Newton–Krylov solver. Comput. Geosci. 1, 215–249.

Ferguson, I.M., Maxwell, R.M., 2011. Hydrologic and land–energy feedbacks of agricultural water management practices. Environ. Res. Lett. 6, 7. http://dx.doi.org/10.1088/1748-9326/6/1/014006.

Franks, S.W., Beven, K.J., Quinn, P.F., Wright, I.R., 1997. On the sensitivity of soil-vegetation-atmosphere transfer (SVAT) schemes: equifinality and the problem of robust calibration. Agric. For. Meteorol. 86, 63–75. http://dx.doi.org/10.1016/S0168-1923(96)02421-5.

Friedl, M.A., Michaelsen, J., Davis, F.W., Walker, H., Schimel, D.S., 1994. Estimating grassland biomass and leaf area index using ground and satellite data. Int. J. Remote Sens. 15, 1401–1420.

Gao, X., Sorooshian, S., Gupta, H.V., 1996. Sensitivity analysis of the biosphere–atmosphere scheme. J. Geophys. Res. 101, 7279–7289.

Göhler, M., Mai, J., Cuntz, M., 2013. Use of eigendecomposition in a parameter sensitivity analysis of the Community Land Model. J. Geophys. Res. Biogeosci. 118, 904–921. http://dx.doi.org/10.1002/jgrg20072.

Gong, W., Duan, Q., Li, J., Wang, C., Di, Z., Dai, Y., Ye, A., Miao, C., 2015. Multi-objective parameter optimization of common land model using adaptive surrogate modeling. Hydrol. Earth Syst. Sci. 19, 2409–2425. http://dx.doi.org/10.5194/hess-19-2409-2015.

Henderson-Sellers, A., 1993. A factorial assessment of the sensitivity of the BATS land-surface parameterization scheme. J. Atmos. Sci. 6, 227–247.

Henderson-Sellers, A., Pitman, A.J., Love, P.K., Irannejad, P., Chen, T.H., 1995. The Project for Intercomparison of Land Surface Parameterization Schemes (PILPS): phases 2 and 3. Bull. Am. Meteorol. Soc., 489–503.

Hou, Z., Huang, M., Leung, L.R., Lin, G., Ricciuto, D.M., 2012. Sensitivity of surface flux simulations to hydrologic parameters based on an uncertainty quantification framework applied to the Community Land Model. J. Geophys. Res. 117, D15108. http://dx.doi.org/10.1029/2012JD017521.

Jacquemin, B., Noilhan, J., 1990. Sensitivity study and validation of a land surface parameterizatioin using the Hapex-Mobilhy data set. Bound.-Layer Meterol. 52, 93–134.

Jenkins, E.W., Berger, R.C., Hallberg, J.P., Howington, S.E., Kelley, C.T., Schmidt, J.H., Stagg, A., Tocci, M.D., 1999. Newton–Krylov–Schwarz Methods for Richards' equation.

Jones, J.E., Woodward, C.S., 2001. Newton–Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems. Adv. Water Resour. 24, 763–774.

Kollet, S.J., 2009. Influence of soil heterogeneity on evapotranspiration under shallow water table conditions: transient, stochastic simulations. Environ. Res. Lett. 4, 9. http://dx.doi.org/10.1088/1748-9326/4/3/035007.

Kollet, S.J., Maxwell, R.M., 2006. Integrated surface–groundwater flow modeling: a free-surface overland flow boundary condition in a parallel groundwater flow model. Adv. Water Resour. 29, 945–958. http://dx.doi.org/10.1016/j.advwatres.2005.08.006.

Li, J., Duan, Q.Y., Gong, W., Ye, A., Dai, Y., Miao, C., Di, Z., Tong, C., Sun, Y., 2013. Assessing parameter importance of the Common Land Model based on qualitative and quantitative sensitivity analysis. Hydrol. Earth Syst. Sci. 17, 3279–3293. http://dx.doi.org/10.5194/hess-17-3279-2013.

Liang, X., Guo, J., 2003. Intercomparison of land-surface parameterization schemes: sensitivity of surface energy and water fluxes to model parameters. J. Hydrol. 279, 182–209. http://dx.doi.org/10.1016/S0022-1694(03)00168-9.

Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. J. Geophys. Res. 99, 14415–14428.

Liu, Y., Gupta, H.V., Sorooshian, S., Bastidas, L.A., Shuttleworth, W.J., 2004. Exploring parameter sensitivities of the land surface using a locally coupled land-atmosphere model. J. Geophys. Res. 109, 13. http://dx.doi.org/10.1029/2004JD004730.

Lukaczyk, T.W., Constantine, P., Palacios, F., Alonso, J.J., 2014. Active subspaces for shape optimization. In: Proceedings of the 10th AIAA Multidisciplinary Design Optimization Conference, pp. 1–18. doi: http://dx.doi.org/10.2514/6.2014-1171.

Mahmood, R., Hubbard, K.G., 2003. Simulating sensitivity of soil moisture and evapotranspiration under heterogeneous soils and land uses. J. Hydrol. 280, 72–90. http://dx.doi.org/10.1016/S0022-1694(03)00183-5.

Maxwell, R.M., 2013. A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling. Adv. Water Resour. 53, 109–117. http://dx.doi.org/10.1016/j.advwatres.2012.10.001.

Maxwell, R.M., Kollet, S.J., 2008. Interdependence of groundwater dynamics and land-energy feedbacks under climate change. Nat. Geosci. 1, 665–669. http://dx.doi.org/10.1038/ngeo315.

Maxwell, R.M., Putti, M., Meyerhoff, S., Delfs, J.-O., Ferguson, I.M., Ivanov, V., Kim, J., Kolditz, O., Kollet, S.J., Kumar, M., Lopez, S., Niu, J., Paniconi, C., Park, Y.-J., Phanikumar, M.S., Shen, C., Sudicky, E. a, Sulis, M., 2014. Surface-subsurface model intercomparison: a first set of benchmark results to diagnose integrated hydrology and feedbacks. Water Resour. Res. 50, 1531–1549. http://dx.doi.org/10.1002/2013WR013725.

Mikkelson, K.M., Maxwell, R.M., Ferguson, I., Stednick, J.D., McCray, J.E., Sharp, J.O., 2013. Mountain pine beetle infestation impacts: modeling water and energy budgets at the hill-slope scale. Ecohydrology 6, 64–72. http://dx.doi.org/10.1002/eco.278.

Mu, Q., Heinsch, F.A., Zhao, M., Running, S.W., 2007. Development of a global evaportranspiration algorithm based on MODIS and global meteorology data. Remote Sens. Environ. 111, 519–536. http://dx.doi.org/10.1016/j.rse.2007.04.015.

Müller, J., Paudel, R., Shoemaker, C.A., Woodbury, J., Wang, Y., Mahowald, N., 2015. CH_4 parameter estimation in CLM4.5bgc using surrogate global optimization. Geosci. Model Dev. Discuss. 8, 141–207. http://dx.doi.org/10.5194/gmdd-8-141-2015.

Osei-Kuffuor, D., Maxwell, R.M., Woodward, C.S., 2014. Improved numerical solvers for implicit coupling of subsurface and overland flow. Adv. Water Resour. 74, 185–195. http://dx.doi.org/10.1016/j.advwatres.2014.09.006.

Panday, S., Huyakorn, P.S., 2004. A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. Adv. Water Resour. 27, 361–382. http://dx.doi.org/10.1016/j.advwatres.2004.02.016.

Park, S.-J., Park, S.-U., Ho, C.-H., 2010. Roughness length of water vapor over land surfaces and its influence on latent heat flux. Terr. Atmos. Ocean. Sci. 21, 855–867.

Pau, G.S.H., Bisht, G., Riley, W.J., 2014. A reduced-order modeling approach to represent subgrid-scale hydrological dynamics for land-surface simulations: application in a polygonal tundra landscape. Geosci. Model Dev. 7, 2091–2105. http://dx.doi.org/10.5194/gmd-7-2091-2014.

Pitman, A.J., 1994. Assessing the sensitivity of a land-surface scheme to the parameter values using a single column model. J. Clim. 7, 1856–1869.

Ray, J., Hou, Z., Huang, M., Sargsyan, K., Swiler, L., 2015. Bayesian calibration of the community land model using surrogates. SIAM/ASA J. Uncertain. Quantif. 3; , pp. 199–233.

Rihani, J.F., Maxwell, R.M., Chow, F.K., 2010. Coupling groundwater and land surface processes: idealized simulations to identify effects of terrain and subsurface heterogeneity on land surface energy fluxes. Water Resour. Res. 46, 14. http://dx.doi.org/10.1029/2010WR009111.

Rosero, E., Yang, Z.-L., Wagener, T., Gulden, L.E., Yatheendradas, S., Niu, G.-Y., 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, 21. http://dx.doi.org/10.1029/2009JD012035.

Schaap, M.G., Leij, F.J., 2000. Improved prediction of unsaturated hydraulic conductivity with the Mualem-van Genuchten Model. Soil Sci. Soc. Am. 64, 843–851.

Schwinger, J., Kollet, S.J., Hoppe, C.M., Elbern, H., 2010. Sensitivity of latent heat fluxes to initial values and parameters of a land-surface model. Vadose Zo. J. 9, 984–1001. http://dx.doi.org/10.2136/vzj2009.0190.

Smith, W.K., Geller, G.N., 1980. Leaf and environmental parameters influencing transpiration: theory and field measurements. Oecologia 46, 308–313.

Srivastava, V., Graham, W., Muñoz-Carpena, R., Maxwell, R.M., 2014. Insights on geologic and vegetative controls over hydrologic behavior of a large complex basin – global sensitivity analysis of an integrated parallel hydrologic model. J. Hydrol. 519, 2238–2257. http://dx.doi.org/10.1016/j.jhydrol.2014.10.020.

Sulis, M., Meyerhoff, S.B., Paniconi, C., Maxwell, R.M., Putti, M., Kollet, S.J., 2010. A comparison of two physics-based numerical models for simulating surface water–groundwater interactions. Adv. Water Resour. 33, 456–467. http://dx.doi.org/10.1016/j.advwatres.2010.01.010.

Szilagyi, J., Zlotnik, V.A., Jozsa, J., 2012. Net recharge vs. depth to groundwater relationship in the Platte River Valley of Nebraska, United States. Ground Water 51, 945–951. http://dx.doi.org/10.1111/gwat.12007.

Trenberth, K.E., Fasullo, J.T., Kiehl, J., 2009. Earth's Global Energy Budget. Bull. Am. Meteorol. Soc. 90, 311–323. http://dx.doi.org/10.1175/2008BAMS2634.1.

White, J.A., Borja, R.I., 2011. Block-preconditioned Newton–Krylov solvers for fully coupled flow and geomechanics. Comput. Geosci. 15, 647–659. http://dx.doi.org/10.1007/s10596-011-9233-7.

Wieringa, J., 1993. Representative roughness parameters for homogeneous terrain. Bound. Layer Meterol. 63, 323–363.