

The Impact of Error Accounting in a Bayesian Approach to Calibrating Modeled Turbulent Fluxes in an Open-Canopy Forest[✉]

TONY E. WONG^a

*Department of Applied Mathematics, and Cooperative Institute for Research in Environmental Sciences,
University of Colorado Boulder, Boulder, Colorado*

WILLIAM KLEIBER

Department of Applied Mathematics, University of Colorado Boulder, Boulder, Colorado

DAVID C. NOONE^b

*Cooperative Institute for Research in Environmental Sciences, and Department of Atmospheric
and Oceanic Sciences, University of Colorado Boulder, Boulder, Colorado*

(Manuscript received 17 February 2017, in final form 2 May 2017)

ABSTRACT

Land surface models are notorious for containing many parameters that control the exchange of heat and moisture between land and atmosphere. Properly modeling the partitioning of total evapotranspiration (ET) between transpiration and evaporation is critical for accurate hydrological modeling, but depends heavily on the treatment of turbulence within and above canopies. Previous work has constrained estimates of evapotranspiration and its partitioning using statistical approaches that calibrate land surface model parameters by assimilating in situ measurements. These studies, however, are silent on the impacts of the accounting of uncertainty within the statistical calibration framework. The present study calibrates the aerodynamic, leaf boundary layer, and stomatal resistance parameters, which partially control canopy turbulent exchange and thus the evapotranspiration flux partitioning. Using an adaptive Metropolis–Hastings algorithm to construct a Markov chain of draws from the joint posterior distribution of these resistance parameters, an ensemble of model realizations is generated, in which latent and sensible heat fluxes and top soil layer temperature are optimized. A set of five calibration experiments demonstrate that model performance is sensitive to the accounting of various sources of uncertainty in the field observations and model output and that it is critical to account for model structural uncertainty. After calibration, the modeled fluxes and top soil layer temperature are largely free from bias, and this calibration approach successfully informs and characterizes uncertainty in these parameters, which is essential for model improvement and development. The key points of this paper are 1) a Markov chain Monte Carlo calibration approach successfully improves modeled turbulent fluxes; 2) ET partitioning estimates hinge on the representation of uncertainties in the model and data; and 3) despite these inherent uncertainties, constrained posterior estimates of ET partitioning emerge.

^a Current affiliation: Earth and Environmental Systems Institute, The Pennsylvania State University, University Park, Pennsylvania.

^b Current affiliation: College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon.

[✉] Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JHM-D-17-0030.s1>.

Corresponding author: Tony Wong, twong@psu.edu

1. Introduction

The partitioning of total evapotranspiration (ET) into contributions from surface evaporation and plant transpiration provides acute insight into the hydrological and biogeochemical coupling and behavior of ecosystems, but is notoriously difficult to constrain. Several methods exist for ET partitioning, but they can be challenging to implement in the field. Total ET is most commonly determined via micrometeorological methods, such as eddy covariance (Baldocchi et al. 1988). Transpiration T may

be determined by chamber gas exchange or sap flow methods, then divided by total ET to obtain the transpiration fraction T/ET . In practice, however, these methods are applied on plant- or plot-level scales, so obtaining results representative of the ecosystem as a whole is limited by surface and vegetation heterogeneity (Schaeffer et al. 2000; Jarvis 1995; Ehleringer et al. 1993). Soil evaporation can be measured using soil weighing lysimeters, but limitations include difficult implementation and potentially poor spatial representation (Dunin 1991).

ET partitioning quantifies the relative use of the pathways moisture may follow from the soil water to above-canopy water vapor, which is directly related to the surface fluxes. These, in turn, draw from and contribute to the model state variables, the temperature and water content of soil, vegetation, and air. The transport of moisture and heat throughout an ecosystem is commonly modeled as an analogy of a series of electrical resistors (e.g., Oleson et al. 2010). In light of the known uncertainty in parameterizations for these resistances, it is hypothesized that model realizations that best match observations of surface fluxes and state variables are those where the resistance terms are correctly modeled (Landsberg and Powell 1973; Baldocchi et al. 1987; Dausat et al. 2001; Liu et al. 2007). Previous modeling studies for ET partitioning have found the canopy resistance to be a particularly influential parameter (Song et al. 2016; Zhu et al. 2013). These studies point to Bayesian calibration methods as a fruitful approach to constrain model predictions for the transpiration fraction of total ET, or T/ET .

Estimates of T/ET for a region, however, may be inconsistent and poorly constrained. For example, for similar sites in the Sonoran Desert, estimates of transpiration fraction range from 7% to 80% (Sammis and Gay 1979; Liu et al. 2005). Another example at an open-canopy forest site (which will be the subject of this study) has yielded estimated distributions of T/ET that are bimodal and span the entire 0–1 range (Berkelhammer et al. 2016). Models driven by meteorological fields can provide constraint on T/ET by simulating individual evaporation and transpiration streams. However, these models depend on parameterizations that are not easily validated or calibrated beyond a small number of idealized cases. Therefore, the difficulty in acquiring direct measurements of T/ET places high demand on developing formal calibration frameworks through which limited observations may be combined with physical models (Abbaspour et al. 2015). Even within a statistical calibration framework, however, different plausible estimates of uncertainty and different calibration approaches may lead to differing predicted model outcomes (Schmidt et al. 2016; Morton et al. 2013; Jeremiah et al. 2011). Previous studies remain

silent on the impacts of error accounting within a statistical model on resulting estimates for T/ET .

In the present study, we advance previous work by employing a Bayesian calibration framework to evaluate the impacts of uncertainty accounting on estimates of T/ET . In particular, for the case of a 145-day field campaign at a site in central Colorado, United States, we calibrate uncertain model parameters pertaining to the relative pathways by which moisture and heat are exchanged between land and atmosphere and assess the impacts on model estimates of T/ET . It is shown that 1) the assimilation of tower meteorological data with model output within a Bayesian calibration framework provides constraint on the network of ecosystem resistances to moisture and heat transport and 2) the efficacy of the calibration approach relies on properly accounting for uncertainties in both the model and observations. As a consequence of the first point, the informed estimates of the network of resistances lead to an estimate of the modeled T/ET , and practical estimates of the posterior uncertainties in model parameters, as well as T/ET , emerge.

2. Methods

a. Observational data

The observational data consist of meteorological and hydrological data spanning 145 days from May to September 2011 in a semiarid open-canopy forest northwest of Colorado Springs, Colorado. Measurements were made from a 27.1-m-tall tower in the Manitou Experimental Forest Observatory (MEF; 39°06′02″ N, 105°06′05″ W, 2286 m elevation). This is a ponderosa pine forest with minimal undergrowth, a canopy height of roughly 18.5 m, and leaf area index (LAI) of 1.2 (Ortega et al. 2014; DiGangi et al. 2011; Kim et al. 2010; Berkelhammer et al. 2016). The MEF site was selected for this study because of the availability of high-quality previously validated hydrological and flux data. Additionally, a recent study employed two different flux partitioning approaches to determine T/ET at MEF, which provides a baseline against which the results of the present study may be compared (Berkelhammer et al. 2016).

Air temperature (°C), humidity (%), and pressure (mbar) were measured by a series of Vaisala Humidity and Temperature Transmitter (HMT337) and Multi-Weather Sensor (WXT520) probes located at each inlet. CO₂ and humidity concentrations ($\mu\text{mol mol}^{-1}$ and mmol mol^{-1} , respectively) were measured using a LiCor Li6262. Wind speed (m s^{-1}) was measured using a Gill R3-50 sonic anemometer. CO₂ and H₂O

fluxes ($\mu\text{mol m}^{-2}\text{s}^{-1}$ and $\text{mmol m}^{-2}\text{s}^{-1}$, respectively) were measured using a LiCor Li7000 infrared gas analyzer and Campbell CSAT3 sonic anemometer. Eddy covariance methods were used to determine latent and sensible heat fluxes (W m^{-2}). A suite of Kipp and Zonen CNR1 net radiometers were used to measure upward and downward longwave and shortwave radiation (W m^{-2}). Observations of soil temperature (K) were made at depths of 5, 50, 70, 100, and 150 cm using Campbell Scientific T107 thermistors. Further information regarding the experimental setup can be found elsewhere (Berkelhammer et al. 2013; DiGangi et al. 2011; Ortega et al. 2014).

b. Land surface model

The model used is the National Center for Atmospheric Research Community Land Model, version 4 (CLM4; Oleson et al. 2010). CLM4 is a one-dimensional land surface model of the energy, momentum, water, and CO_2 exchanges between land and atmosphere, and it accounts for ecosystem dynamics, biophysical processes, hydrological processes, and biogeochemical processes. The Community Land Model is the land component of the Community Earth System Model (CESM; Hurrell et al. 2013). Data fields needed as inputs to force CLM4 include incident longwave and solar radiation (diffuse and direct components), incident precipitation, atmospheric humidity, wind speed, and pressure and temperature (measured at the top of the tower; 27.1 m). Surface data configured for MEF are soil texture, soil color, monthly LAI and stem area index (SAI), vegetation composition, and albedo (Vertenstein et al. 2010).

For the calibration data, daily averages of latent heat flux, sensible heat flux, and top soil layer temperature are used. Latent and sensible heat fluxes are key quantities in the exchange of moisture and heat between the land surface and atmosphere and are readily observable by eddy covariance techniques (Baldocchi et al. 1988). Soil temperature is an attractive option as calibration data because it is a widely observed state variable and characterizes an important part of the energy state of the ecosystem. We note that the top soil layer temperature is distinct from the radiative temperature of the soil surface.

We account for uncertainty in ecosystem turbulence characteristics by calibrating the aerodynamic resistances to moisture and heat transport 1) between surface and vegetation air f_g , 2) between vegetation and above-canopy air f_a , 3) through vegetation stomata f_s , and 4) through the leaf boundary layer f_b . The calibration parameters are applied as multiplicative factors in the default resistance parameterizations as is practiced in other studies (e.g., Song et al. 2016). That is, in Eqs. (3)–(6), the calibration parameters f_s , f_b , f_a , and f_g would all be

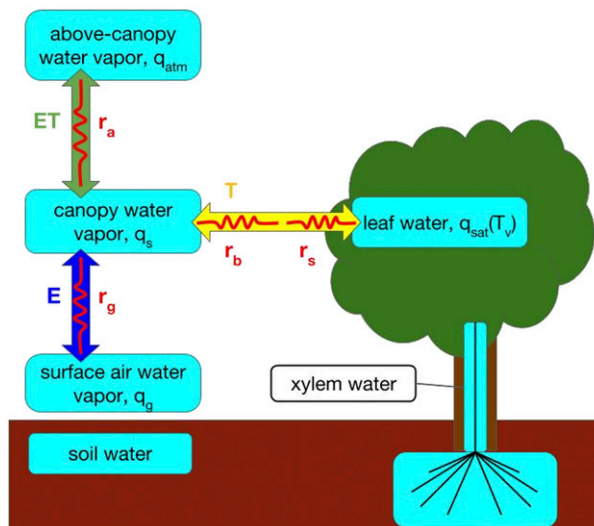


FIG. 1. Schematic depicting the evaporation and transpiration pathways contributing to total evapotranspiration in the model. Adapted from Wong et al. (2017).

equal to one in the underlying CLM4 model. The following section provides an overview of the parameterizations in CLM4 most relevant to the present work. The interested reader is directed to Oleson et al. (2010) for additional details and to Fig. 1 for a depiction of the evaporation, transpiration, and total evapotranspiration fluxes; the water pools on which they act; and the ecosystem resistances that facilitate these transfers of moisture and heat.

Total evapotranspiration is calculated as described in Oleson et al. (2010, their Eq. 5.86):

$$ET = -\rho_{\text{atm}} \frac{q_{\text{atm}} - q_s}{r_a}, \tag{1}$$

where r_a is the aerodynamic resistance between the vegetation canopy and above-canopy air spaces (s m^{-1}), ρ_{atm} (kg m^{-3}) is the atmospheric air density, q_{atm} (kg kg^{-1}) is the atmospheric specific humidity, and q_s (kg kg^{-1}) is the canopy air space specific humidity. Transpiration is calculated as in Oleson et al. (2010, their Eq. 5.87):

$$T = -\rho_{\text{atm}} \frac{q_s - q_{\text{sat}}(T_v)}{r_s + r_b}, \tag{2}$$

where $q_{\text{sat}}(T_v)$ is the saturation specific humidity (kg kg^{-1}) at the temperature of the vegetation T_v , r_s is the stomatal resistance ($\text{m}^2 \text{s } \mu\text{mol}$), and r_b is the leaf boundary layer resistance ($\text{m}^2 \text{s } \mu\text{mol}$). We note that Eq. (2) is an approximation of the true physics of CLM4, which decompose the vegetative flux T into contributions from sunlit and shaded leaves, as well as from evaporation of canopy-intercepted water.

Stomatal resistance ($\text{m}^2 \text{s} \mu\text{mol}$) is calculated as

$$r_s = f_s \left(m \frac{A}{c_s} \frac{e_s}{e_i} P_{\text{atm}} + b \right)^{-1}, \quad (3)$$

where m is a parameter that depends on vegetation type, A is leaf photosynthesis ($\mu\text{mol CO}_2 \text{m}^{-2} \text{s}^{-1}$), c_s is the CO_2 partial pressure at the leaf surface (Pa), e_s is the vapor pressure at the leaf surface (Pa), e_i is the saturation vapor pressure (Pa) inside the leaf at the temperature of the vegetation, P_{atm} is the atmospheric pressure (Pa), and $b = 2000$ is the minimum stomatal conductance ($\mu\text{mol m}^{-2} \text{s}^{-1}$) when $A = 0$ (Collatz et al. 1991; Oleson et al. 2010). The factors A , c_s , e_s , and e_i are all modeled by CLM4, P_{atm} is provided as atmospheric forcing, and m and b are parameters that depend on vegetation type; all have associated uncertainties. Parameters m and b , in particular, are known from laboratory experiments for various leaf types (Ball 1988; Ball et al. 1987), though it remains less clear how these parameters scale up to the ecosystem scale considered by CLM4. Figure 2 from Collatz et al. (1991) showed that the overall match between measured and modeled stomatal conductance (the inverse of r_s) is reasonable, but parametric uncertainty in Eq. (3) persists at the leaf scale, which will aggregate when integrated to ecosystem scale.

Leaf boundary layer resistance ($\text{m}^2 \text{s} \mu\text{mol}$) is parameterized as

$$r_b = f_b \frac{1}{C_v} \left(\frac{U}{d_{\text{leaf}}} \right)^{-1/2}, \quad (4)$$

where $C_v = 0.01 \text{ m s}^{-1/2}$ is the turbulent transfer coefficient between the vegetation surface and canopy air, U (m s^{-1}) is the magnitude of the wind velocity incident on the leaves at a reference level, and d_{leaf} is the characteristic dimension of the leaves in the direction of the wind flow (taken throughout CLM4 to be 0.04 m; Oleson et al. 2010). Parameter U is modeled by CLM4 and C_v and d_{leaf} are parameters that CLM4 prescribes to all vegetation types.

The aerodynamic resistance to moisture and heat transfer between ground and canopy air (s m^{-1}) is as follows:

$$r_g = f_g \frac{1}{C_s U}, \quad (5)$$

where C_s is the turbulent transfer coefficient between the underlying soil and the canopy air, which is interpolated between values for bare soil and for dense canopy (Zeng et al. 2005; Oleson et al. 2010). While the values for r_g calculated using Eq. (5) agree with observational data, it

has been noted that there are few measurements available that can constrain these values (Zeng et al. 2005). This lack of direct constraint on r_g suggests a need to constrain this parameter indirectly, through observations of surface fluxes and state variables.

The aerodynamic resistance between the vegetation canopy and above-canopy air spaces (s m^{-1}) is given by

$$r_a = f_a \frac{q_{\text{atm}} - q_s}{q_* u_*}, \quad (6)$$

where u_* (m s^{-1}) is the friction velocity and q_* (kg kg^{-1}) is the moisture scale (Oleson et al. 2010). Parameters u_* and q_* are derived from Monin–Obukhov similarity theory. Their parameterizations were developed by separate groups using different data in different parameter regimes and are intended for global modeling studies, with typical grid spacing much larger than the extent of the MEF tower footprint (Zeng et al. 1998). Therefore, it is not expected that these gridcell-scale parameters will compare favorably with the site-level data at MEF, further suggesting calibration is needed.

The resistances r_g , r_s , r_a , and r_b are exchange parameters that define the way in which the energy and water balances at the surface behave, yet the theory behind them is not well established. For example, Liu et al. (2007) showed that realistic errors in the wind speed used to determine the aerodynamic resistances can lead to errors as large as 22% in modeled surface fluxes. Therefore, estimating these parameters is important to accurately model the transfer processes. The evapotranspiration flux partitioning (i.e., T/ET) is a quantity that characterizes the hydrological and biogeochemical behaviors of an ecosystem, and the resistance parameters selected here partly control the balance of evaporation and transpiration fluxes in the model. Therefore, the effects of the error accounting on the modeled seasonal average T/ET are analyzed to assess the impacts of the calibration approaches.

c. Model biases and the need for a statistical approach

In the control model, CLM4 output displays substantial errors and biases relative to the observational data, aggregated to daily time scale (Fig. 2). The gray shaded regions denote a plus/minus one observational standard error region around the observations. The model represents these physical processes qualitatively well, as peaks and troughs in the data are present in the modeled values. However, the error distributions are not centered at zero, indicative of model structural errors (Figs. 2d–f). The low latent heat flux bias at MEF (-25 W m^{-2}) is consistent with the generally low bias found in northern midlatitude sites in a global analysis

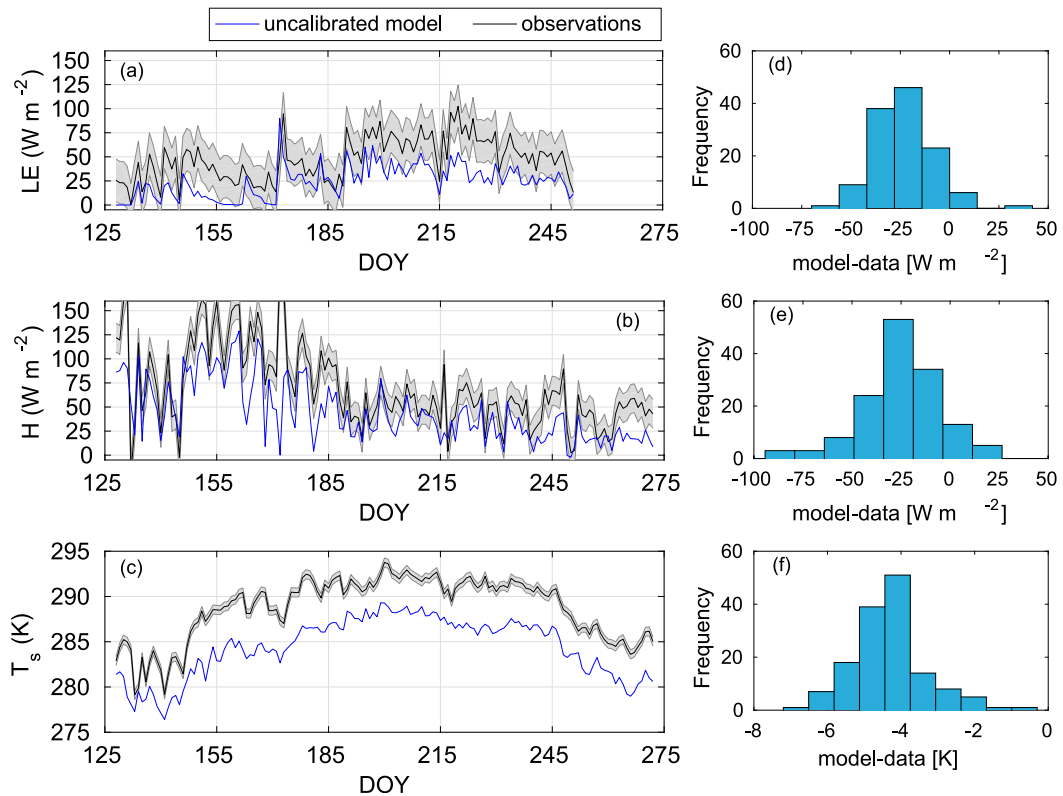


FIG. 2. Uncalibrated model results, compared to observational data. Time series of (a) latent heat flux, (b) sensible heat flux, and (c) top soil layer temperature. (d)–(f) Histograms of the errors in these quantities, reported as model observations. The gray shaded region denotes plus/minus one observational standard error around the observations.

(Bonan et al. 2012). The fact that both the latent heat flux and sensible heat flux exhibit negative biases may seem surprising because this points to a substantial total bias in the surface energy budget. However, both surface fluxes are facilitated by the network of aerodynamic resistances (cf. Fig. 1), so it is reasonable for the fluxes to be impacted similarly by uncertainties in these resistance parameters. It is unclear to what degree these biases are attributable to model parameter choice or are an inherent limitation of CLM4 that stems from the use of approximate equations, thus motivating the need to account for these model biases.

d. Statistical calibration framework

1) ALGORITHM

Within a statistical framework, the CLM4 model output is related to various types of observational data. Using a Bayesian approach, prior distributions on the model parameters are joined with the CLM4 solution and assimilated field data, leading to posterior distributions that represent uncertainty in the parameters. The calibration algorithm proceeds by exploring the

parameter space to discover high-probability regions of the posterior distribution. In this manner, the calibration framework enables model simulation results that match the observational data in light of the model biases and other uncertainties.

Let $\boldsymbol{\eta}(t, \boldsymbol{\theta}) \in \mathbf{R}^3$ refer to the vector of model output for the calibration fields (latent heat flux, sensible heat flux, and top soil temperature) at calibration parameter values $\boldsymbol{\theta} = (f_g, f_s, f_a, f_b)^T$ on day t and let $\mathbf{y}(t) \in \mathbf{R}^3$ be the corresponding observations of these fields. Let $\boldsymbol{\alpha} = (\alpha_{LE}, \alpha_H, \alpha_T)^T$ refer to the model additive bias, where α_{LE} is the additive bias in the modeled latent heat flux, α_H is the bias in the sensible heat flux, and α_T is the temperature bias. We make the simplifying assumption $\alpha_F = \alpha_{LE} = \alpha_H$, where α_F is a bias assumed to be present in both the modeled latent and sensible heat fluxes. In preliminary experiments (see Fig. 2), the biases in the sensible and latent heat fluxes were similar enough to warrant this assumption and thereby reduce the size of the parameter space the calibration approach must explore. The impacts of this assumption will be explored later. The bias corrections α_F and α_T may be considered as accounting for model structural error. These biases

are assumed to be stationary, which is a reasonable assumption for the single growing season dataset for MEF. This assumption should be carefully examined if the present method is to be applied to multiseason observational datasets. Finally, denote the optimal, unknown, model parameters by θ^* and let α^* be the additive model bias that optimizes the match between modeled and observed fluxes and soil temperature. The statistical model is

$$\mathbf{y}(t) = \alpha^* + \boldsymbol{\eta}(t, \theta^*) + \boldsymbol{\varepsilon}(t). \quad (7)$$

It is simple to generalize Eq. (7) to include a multiplicative bias term as well, but based on exploratory analyses such flexibility is neither required nor considered in the present application. For the sake of brevity, throughout the remainder of this work the term “bias” refers only to the additive bias α .

It is convenient to think of $\boldsymbol{\zeta}(t) = \alpha^* + \boldsymbol{\eta}(t, \theta^*)$ as representing the physical process the model $\boldsymbol{\eta}$ seeks to capture. In Eq. (7), $\boldsymbol{\varepsilon}(t) \in \mathbf{R}^3$ accounts for model and observational errors and is modeled as a multivariate mean zero Gaussian process that is independent between components and across time points. The impacts of these assumptions will later be evaluated. Denote by $\boldsymbol{\Sigma}$ the covariance matrix of $\boldsymbol{\varepsilon}(t)$. We assume independent and uncorrelated errors, so it suffices at each time t to specify $\boldsymbol{\Sigma} = \text{diag}(\sigma_{LE}^2, \sigma_H^2, \sigma_T^2)$, where σ_k is the empirical standard deviation of the set of observations of field k , where $k = LE, H$, or T , for latent heat flux, sensible heat flux, and top soil layer temperature, respectively. An experiment will be conducted to assess the impacts of the assumption of uncorrelated errors [E5; described in section 2d(2)]. The specific form of $\boldsymbol{\Sigma}$ is addressed in the “Error covariance” section below.

The likelihood function $L(\mathbf{y} | \boldsymbol{\theta}, \boldsymbol{\alpha})$ of $\mathbf{y} = [y(1), \dots, y(145)]^T$ (where there are 145 days of data) is then a product of univariate normal likelihoods. At each time $t = 1, 2, \dots, 145$, the normal likelihood is centered at the observation $\mathbf{y}(t)$ with covariance $\boldsymbol{\Sigma}$. Given a joint prior distribution $\pi(\boldsymbol{\theta}, \boldsymbol{\alpha})$ on the calibration parameters and biases, the joint posterior probability is

$$\pi(\boldsymbol{\theta}, \boldsymbol{\alpha} | \mathbf{y}) \propto L(\mathbf{y} | \boldsymbol{\theta}, \boldsymbol{\alpha}) \pi(\boldsymbol{\theta}, \boldsymbol{\alpha}). \quad (8)$$

This posterior density contains all uncertainty information regarding the model parameters and biases, and optimal values can be derived directly from the posterior (e.g., a posterior median minimizes absolute Bayes loss).

Prior distributions are required for the model parameters and bias parameters and are assumed to be independent between parameters. Each individual component of $\boldsymbol{\theta}$ is given a uniform prior with bounds $[1/10, 10]$. These bounds are chosen to allow an order of magnitude deviation

above or below the control model behavior and to prevent nonphysical negative parameter values. The model flux and temperature biases are assigned normal priors with mean zero and standard deviation 100 W m^{-2} for α_F and 100 K for α_T , representing a relatively uninformative prior.

A Markov chain of draws from the joint posterior distribution of $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}$ [Eq. (8)] is constructed using the Metropolis–Hastings algorithm (Metropolis et al. 1953; Hastings 1970). This approach follows the detailed outline presented by Higdon et al. (2004) for model calibration, and a brief overview of the algorithm is given here:

- 1) Initial values for the Monte Carlo iteration are selected ($\boldsymbol{\theta}_1$ and $\boldsymbol{\alpha}_1$).
- 2) The Metropolis–Hastings algorithm (below) is iterated for $j = 2, \dots, N$:
 - (i) Parameter $\boldsymbol{\theta}_{\text{new}}$ is drawn from a normal distribution with mean $\boldsymbol{\theta}_{j-1}$ and variance $\mathbf{s}_{\boldsymbol{\theta}}^2$. Similarly, $\boldsymbol{\alpha}_{\text{new}}$ is drawn from a normal distribution with mean $\boldsymbol{\alpha}_{j-1}$ and variance $\mathbf{s}_{\boldsymbol{\alpha}}^2$.
 - (ii) The posterior probability of these new iterates, $\pi(\boldsymbol{\theta}_{\text{new}}, \boldsymbol{\alpha}_{\text{new}} | \mathbf{y})$, is calculated according to Eq. (8).
 - (iii) The acceptance probability p is calculated as $p = \min[1, \pi(\boldsymbol{\theta}_{\text{new}}, \boldsymbol{\alpha}_{\text{new}} | \mathbf{y}) / \pi(\boldsymbol{\theta}_{j-1}, \boldsymbol{\alpha}_{j-1} | \mathbf{y})]$.
 - (iv) Select the next Markov chain Monte Carlo (MCMC) iterate $\boldsymbol{\theta}_j$:
 - $\boldsymbol{\theta}_j = \boldsymbol{\theta}_{\text{new}}, \boldsymbol{\alpha}_{\text{new}}$ with probability p or
 - $\boldsymbol{\theta}_j = \boldsymbol{\theta}_{j-1}, \boldsymbol{\alpha}_{j-1}$ with probability $1 - p$.

We initialize each component of the proposal density variances at $\mathbf{s}_{\boldsymbol{\theta}}^2$ and $\mathbf{s}_{\boldsymbol{\alpha}}^2$ at 1. After 500 iterations, an adaptive Metropolis algorithm is implemented (Haario et al. 2001). The covariance matrix used to propose new calibration parameters is the covariance matrix of the previous iterates, scaled by $s_d = 2.38/N^{1/2}$ to optimally explore the parameter space, where N is the dimension of the parameter space (Gelman et al. 1996).

The above Markov chain model generates samples from the posterior distribution of model parameters $\boldsymbol{\theta}$ and biases $\boldsymbol{\alpha}$. Initial testing using a Latin hypercube analysis of the six-dimensional parameter space (four calibration parameters and two additive biases) highlighted a unimodal region of high likelihood near $\boldsymbol{\theta} = (1.50, 0.85, 3.80, 1.00)^T$; Markov chains for the calibration parameters are therefore initialized at these values. The additive biases $\boldsymbol{\alpha}_1$ are initialized at 24 W m^{-2} and 4.0 K , in light of the biases from the control model behavior (Fig. 2). Preliminary experiments indicated that initialization at dispersed initial conditions converges to the same posterior modes (see supplemental material), but is not used for analysis based on slower speed of convergence relative to using the preliminary Latin hypercube exploration to inform the initial parameter estimates.

Theoretically, the Metropolis–Hastings algorithm converges to a stationary distribution of the calibration parameters and biases. Gelman–Rubin diagnostics and inspection are used to assess convergence of our sampling chains (Gelman and Rubin 1992). Six parallel runs of the adaptive Metropolis–Hastings algorithm outlined above are each iterated 30 000 times. The first half of each run is discarded for burn-in and the remaining values are used for analysis. No thinning is done (Link and Eaton 2012).

2) ERROR COVARIANCE

Following Kennedy and O’Hagan (2001), four distinct error terms compose the total uncertainty: $\sigma_k^2 = \sigma_{\text{par},k}^2 + \sigma_{\text{obs},k}^2 + \sigma_{\text{rep},k}^2$ and the calibrated additive bias term α . Each of these component uncertainties is described below.

Structural uncertainty (i.e., α) is the inherent disagreement between any physical model and the process it seeks to model, due to the use of approximate equations of nature. We formally define this as the difference between the model response at the true values of the calibration parameters and the mean of the physical process being modeled (Kennedy and O’Hagan 2001), and it is assumed here to result in a constant bias rather than contribute to random uncertainty. Because these optimal parameters are unknown, structural uncertainty may only be estimated. In the present work, the bias calibration terms account for structural uncertainty, and the biases are obtained from the posterior.

Observational uncertainty $\sigma_{\text{obs},k}$ is error due to imperfect measurement systems. The values used are $\sigma_{\text{obs,LE}} = 22 \text{ W m}^{-2}$, $\sigma_{\text{obs,H}} = 15 \text{ W m}^{-2}$ (Twine et al. 2000), and $\sigma_{\text{obs,T}} = 0.5 \text{ K}$ (supplied by manufacturer). Eddy covariance methods rely on measurements of wind speed, temperature, and humidity originating from the MEF tower observatory, which represents the entire tower footprint. Thus, defining the spatial scale of interest to be that represented by eddy covariance fluxes, $\sigma_{\text{obs,H}}$ and $\sigma_{\text{obs,LE}}$ incorporate representation uncertainty as well. The flux uncertainties for MEF (semiarid, open-canopy forest) may differ from the estimates based on observations from a grassland (Twine et al. 2000), although these estimates serve as a basis from which to assess the impacts of other sources of uncertainty.

Representation uncertainty $\sigma_{\text{rep},k}$ results from heterogeneity in the site or region selected to model. Based on the estimates of subgrid-scale variability in surface temperature found from high-resolution modeling, we set the representation uncertainty for the top soil layer temperature at $\sigma_{\text{rep,T}} = 2 \text{ K}$ (Essery et al. 2003). The estimates of flux observational uncertainties $\sigma_{\text{obs,LE}}$ and

$\sigma_{\text{obs,H}}$ are calculated on a spatial scale representative of the flux tower footprint. Thus, the representation uncertainties $\sigma_{\text{rep,LE}}$ and $\sigma_{\text{rep,H}}$ are implicitly incorporated into $\sigma_{\text{obs,LE}}$ and $\sigma_{\text{obs,H}}$.

Parametric uncertainty $\sigma_{\text{par},k}$ results from uncertainty in model parameters. While the purpose of this assimilation calculation is to reduce parametric uncertainty (that associated with the ecosystem resistance terms), CLM4 has many other parameters that attempt to best represent the physical attributes and ecosystem behavior of the observation site. Among these parameters are surface albedo, LAI, soil texture (percentage of soil that is sand, silt, or clay), and percentages of the vegetation that are bare soil, trees, grass, or shrubs. A one-at-a-time sensitivity analysis is performed to inform an estimate of parametric error (e.g., Abbaspour et al. 2015; Das et al. 2008). First, a series of 11 model simulations are produced: a control run, two runs with $\pm 20\%$ sand fraction in estimating soil hydraulic conductivity, two runs with $\pm 20\%$ LAI, two runs with $\pm 20\%$ albedo, and four runs with $+20\%$ bare soil, trees, grass, and shrubs individually. For each simulation, the root-mean-square (RMS) deviations between the control modeled and the experiment modeled latent heat flux, sensible heat flux, and top soil layer temperature are calculated. For each of these fields, the parametric uncertainty is estimated as the Euclidean distance of the 10-dimensional RMS deviation point from the sensitivity tests to the origin. These result in parametric uncertainty estimates of $\sigma_{\text{par,LE}} = 3.86 \text{ W m}^{-2}$, $\sigma_{\text{par,H}} = 3.09 \text{ W m}^{-2}$, and $\sigma_{\text{par,T}} = 0.15 \text{ K}$. These estimates are the maximum deviation considering fairly large (20%) perturbations on each parameter. Thus, the true uncertainty due to these parameters is likely lower than these estimates. Conversely, uncertainty due to feedbacks when multiple parameters are perturbed simultaneously and uncertainty due to parameters not considered here conspire to increase the parametric uncertainty beyond this estimate. While CLM4 has other parameters and parameterizations that would need to be accounted for to fully quantify parametric uncertainty, we assume that these experiments provide a sufficient estimate of $\sigma_{\text{par},k}$. Furthermore, the contribution of $\sigma_{\text{obs},k}$ to σ_k is much larger than any reasonable estimate of $\sigma_{\text{par},k}$.

We conduct a series of five experiments to assess the importance of accounting for different uncertainty terms in this land surface model calibration. Table 1 provides a summary of these experiments. We note that other methods for statistical model calibration besides the adaptive Metropolis approach outlined above may be useful in practice (Jeremiah et al. 2011). However, the experiments presented here focus on the impacts of uncertainty quantification on the calibration results, as

TABLE 1. Summary of error accounting in the five calibration experiments.

Uncertainty	<i>E1</i>	<i>E2</i>	<i>E3</i>	<i>E4</i>	<i>E5</i>
Observational	X	X	X	X	X
Representation	X	X	X	X	X
Structural		X	X	X	X
Parametric			X	X	X
Distinct flux biases				X	
Sample covariance					X

opposed to the calibration algorithm itself. In the first experiment *E1*, the structural and parametric uncertainties in the statistical model are neglected. Thus, the error covariance matrix Σ has entries given by $\sigma_{\text{obs},k}^2 + \sigma_{\text{rep},k}^2$. The second experiment *E2* incorporates the structural uncertainty (additive biases) into the statistical model. The third experiment *E3* makes use of the full uncertainty accounting outlined above, incorporating the additive bias terms into the model calibration, and the entries of the error covariance matrix Σ are given by $\sigma_k^2 = \sigma_{\text{par},k}^2 + \sigma_{\text{obs},k}^2 + \sigma_{\text{rep},k}^2$. The fourth experiment *E4* examines the impacts of assuming a shared bias for both latent heat and sensible heat fluxes α_F . In *E4*, uncertainties are estimated as in *E3*, and separate flux biases are estimated for latent heat (i.e., α_{LE}) and sensible heat flux (i.e., α_H). The fifth experiment *E5* examines the impacts of assuming fixed observational uncertainties in the error covariance matrix

Σ and neglecting correlation in the error structure. In *E5*, uncertainties σ_{LE}^2 , σ_H^2 , and σ_T^2 are estimated as in *E3*, and we sample Σ from its conjugate prior, an inverse Wishart distribution centered at $\text{diag}(\sigma_{LE}^2, \sigma_H^2, \sigma_T^2)$ (see supplemental material).

3. Results

a. Posterior inference and the importance of error accounting

The calibration technique results in tight constraints on the model parameters and bias corrections that vary depending on the method for accounting of uncertainty (Fig. 3). The most drastic improvement comes from accounting for model structural uncertainty (experiments *E2*–*E5* vs *E1*). When model structural uncertainty is neglected, parameter values accumulate near the boundaries of the prior ranges (Figs. 3a,d). This indicates that the model is attempting to compensate for lack of a key physical mechanism (the bias terms, α_F and α_T) by selecting extreme calibration parameters. It is unlikely that the theory governing transfer processes is in error by more than an order of magnitude; thus, this result highlights that accounting for model structural uncertainty is critical.

We find that accounting for parametric uncertainty has only a minor impact on the resulting posterior estimates for calibration parameters (Fig. 3, *E2* vs *E3*). Across all experiments, the exchange between canopy

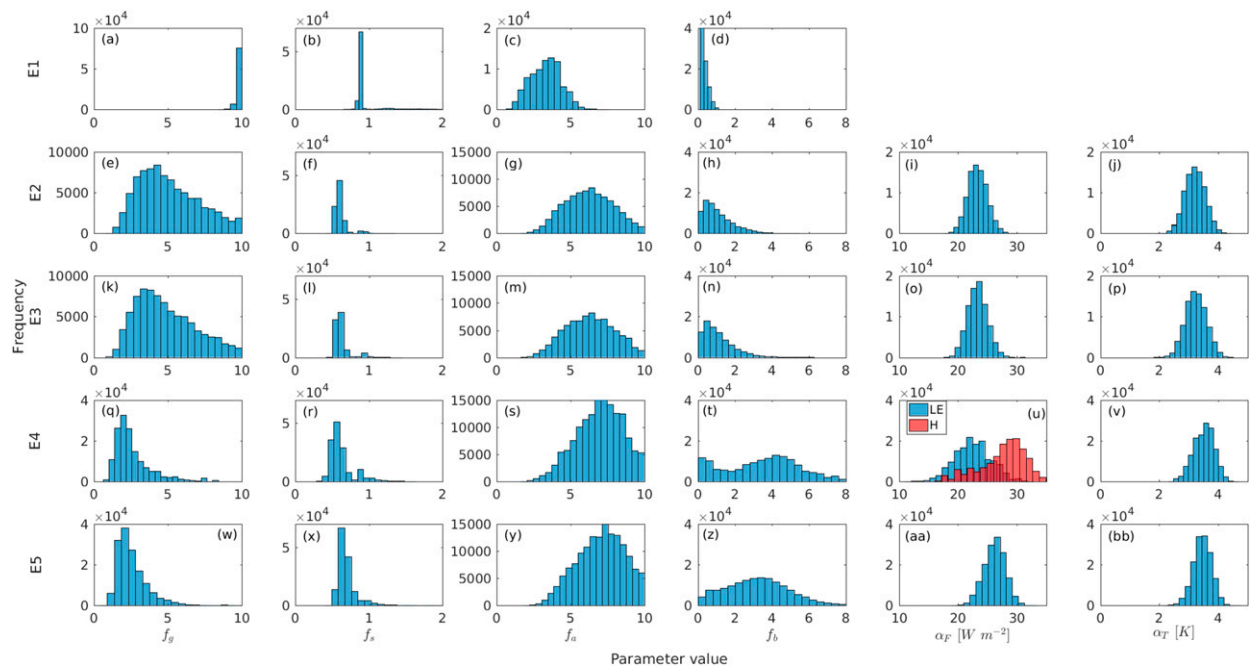


FIG. 3. Frequency distributions of the calibrated model parameters and biases for (a)–(d) *E1*, (e)–(j) *E2*, (k)–(p) *E3*, (q)–(v) *E4*, and (w)–(bb) *E5*.

and above-canopy air spaces (f_a) and the structural uncertainties (α_F and α_T) are tightly constrained. When the statistical model is given, the flexibility of separate bias corrections for latent and sensible heat fluxes ($E3$ vs $E4$) or to sample uncertainty in the observational covariance matrix ($E3$ vs $E5$), the constraint on surface-to-canopy exchange (f_g) increases, but the constraint on leaf boundary layer exchange decreases (f_b). With total ET fixed by calibrating to latent heat flux data, lower f_g and higher f_b both serve to decrease the relative contribution of transpiration to the total ET.

This is reflected in the calibrated estimates of T/ET in experiments $E4$ and $E5$ relative to $E3$ (Fig. 4). Table 2 provides the 95% credible intervals and medians for the parameters, biases, and T/ET for each calibration experiment. As a wider representation of uncertainty is incorporated into the calibration framework, the resulting distribution of calibrated estimates of T/ET also widens (Table 2, Fig. 4). Interestingly, as the accounting of uncertainty becomes more complete (generally, moving from left to right along the columns of Table 1), the estimates of T/ET decrease. A recent study estimates T/ET at MEF to be 0.49 ± 0.23 (Berkelhammer et al. 2016). By incorporating uncertainty and correlated errors in the error covariance matrix Σ , experiment $E5$ yields the best agreement between our calibrated estimates for T/ET and these previous estimates. While our estimates for T/ET are notably higher than those of Berkelhammer et al. (2016), all of the experiments presented here overlap within the uncertainties in both previous estimates of T/ET and those presented here.

b. Observationally constrained model performance

Posterior ensembles of model realizations are generated by drawing 100 random samples from the posterior estimates of the resistance parameters and corresponding biases. The calibration technique yields considerable improvement in the agreement between the modeled and observed latent heat flux, sensible heat flux, and soil temperature (Fig. 5). The poor performance of experiment $E1$ highlights the importance of accounting for structural uncertainties (Figs. 5a–c).

Calibration yields substantial improvement in model performance (Table 3, Fig. 5), as compared to the uncalibrated model (Fig. 2). These improvements are particularly notable in the simulation of latent heat flux in experiments $E2$ and $E3$. When uncertainty in the ecosystem resistance terms is accounted for, modeled top soil layer temperatures and sensible heat fluxes also agree well with observations, given the spread in these ensemble members. The posterior model ensemble results for latent heat flux agree well with observations for mid- and late season [day of year (DOY) 155 and later],

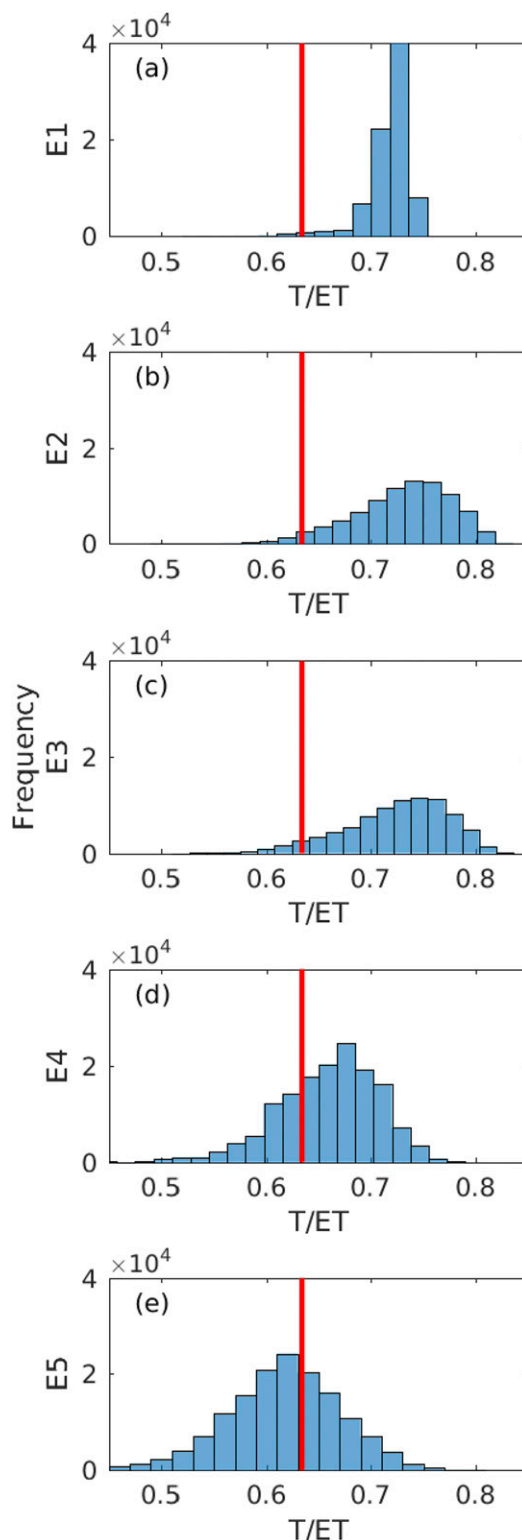


FIG. 4. Frequency distributions of the calibrated transpiration fraction in (a) $E1$ (observational and representation uncertainty only), (b) $E2$ (add structural uncertainty to $E1$), (c) $E3$ (add parametric uncertainty to $E2$), (d) $E4$ (add separate structural error terms for latent and sensible heat fluxes to $E3$), and (e) $E5$ (estimate error covariance matrix in $E3$). The vertical red line denotes the control model ET partitioning.

TABLE 2. Median and 2.5%–97.5% credible intervals for model parameters and T/ET .

	$E1$	$E2$	$E3$	$E4$	$E5$
f_g	9.88 (9.37–10.0)	4.74 (2.00–9.48)	4.49 (1.92–9.22)	2.21 (1.18–6.09)	2.29 (1.27–5.12)
f_s	0.89 (0.85–1.63)	0.59 (0.52–0.92)	0.59 (0.52–0.94)	0.57 (0.43–1.06)	0.66 (0.55–1.04)
f_a	3.35 (1.41–5.23)	6.23 (3.23–9.37)	6.30 (3.21–9.37)	7.08 (3.74–9.68)	7.14 (3.85–9.72)
f_b	0.29 (0.11–0.80)	0.90 (0.14–3.16)	0.92 (0.14–3.27)	3.66 (0.15–7.60)	3.18 (0.38–6.69)
α_{LE}	—	23.2 (20.2–26.5)	23.3 (20.1–26.7)	22.4 (16.7–28.3)	26.3 (22.6–29.6)
α_H	—	—	—	28.4 (18.3–33.7)	—
α_T	—	3.20 (2.54–3.87)	3.22 (2.56–3.93)	3.53 (2.81–4.11)	3.46 (2.83–4.03)
Σ_{11}	—	—	—	—	498 (493–504)
Σ_{12}	—	—	—	—	–0.31 (–2.78–2.11)
Σ_{13}	—	—	—	—	0.00 (–0.34–0.33)
Σ_{22}	—	—	—	—	236 (233–238)
Σ_{23}	—	—	—	—	0.01 (–0.22–0.24)
Σ_{33}	—	—	—	—	4.27 (4.23–4.31)
T/ET	0.72 (0.65–0.74)	0.74 (0.63–0.80)	0.73 (0.61–0.80)	0.66 (0.55–0.74)	0.62 (0.50–0.72)

but yield a poor representation of early season latent heat flux (DOY 125–155). This is attributed largely to the increase in modeled soil temperature that was a result of improving the match between modeled and observed soil temperature (note that the uncalibrated

model produces soil temperatures that are too cool; cf. Fig. 5f).

While calibration in experiments $E4$ and $E5$ did yield substantially different estimates of T/ET (Fig. 4), these experiments did not improve model performance

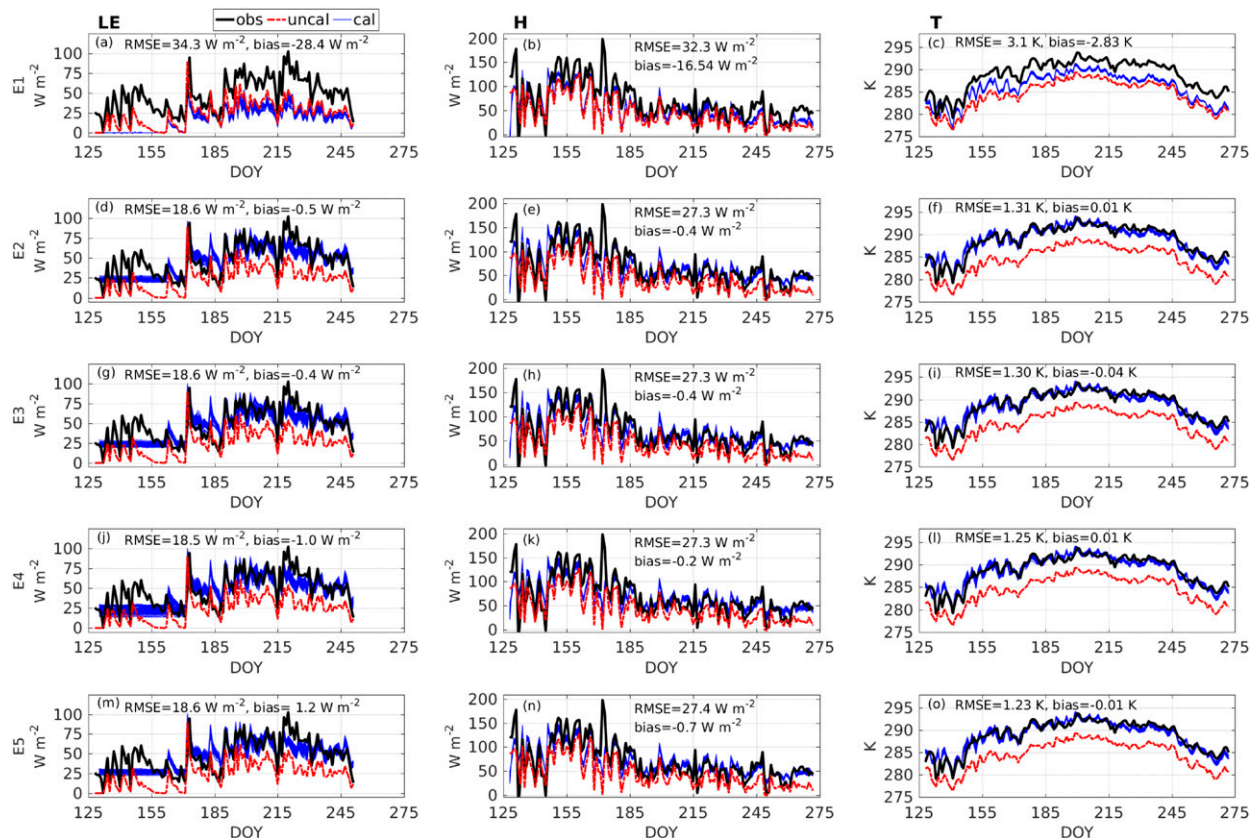


FIG. 5. Modeled fluxes of (left) latent heat, (center) sensible heat, and (right) temperature, using calibration parameters drawn from the Markov chain results. Thick black lines are the observations, thin blue lines are the posterior ensemble results, and the dot-dashed red lines are uncalibrated model results for the five experiments (a)–(c) $E1$, (d)–(f) $E2$, (g)–(i) $E3$, (j)–(l) $E4$, and (m)–(o) $E5$.

TABLE 3. RMSE (and mean model bias, in parentheses) for the uncalibrated model and calibrated model experimental ensembles.

	Latent heat flux (W m^{-2})	Sensible heat flux (W m^{-2})	Soil temperature (K)
Uncalibrated	27.3 (−23.3)	35.9 (−25.1)	4.43 (−4.31)
<i>E1</i>	34.3 (−28.4)	32.3 (−16.6)	3.09 (−2.83)
<i>E2</i>	18.6 (−0.5)	27.3 (−0.4)	1.31 (0.01)
<i>E3</i>	18.6 (−0.4)	27.3 (−0.4)	1.30 (−0.04)
<i>E4</i>	18.5 (−1.0)	28.6 (−6.0)	1.25 (0.01)
<i>E5</i>	18.6 (1.2)	27.4 (−0.7)	1.23 (−0.01)

relative to *E3* (Table 3). Calibration (*E3*) decreased model bias from -23.3 W m^{-2} , -25.1 W m^{-2} , and -4.31 K to -0.4 W m^{-2} , -0.4 W m^{-2} , and -0.04 K for latent heat flux, sensible heat flux, and soil temperature, respectively. RMSE decreased from 27.3 W m^{-2} , 35.9 W m^{-2} , and 4.43 K for uncalibrated latent heat flux, sensible heat flux, and top soil layer temperature to 18.6 W m^{-2} , 27.3 W m^{-2} , and 1.30 K in *E3* (Table 3). The reductions in RMSE correspond to percent improvements of 32%, 24%, and 71% for latent heat flux, sensible heat flux, and top soil layer temperature, respectively. These posterior RMSE (and bias) values are calculated as the mean of the RMSE (bias) from each of the 100 ensemble members. Note that the improvements in RMSE and bias are not only due to the accounting of structural uncertainty (i.e., the model additive bias terms), but also due to providing the calibration framework with the flexibility needed to converge to more realistic posterior parameter estimates (Fig. 3).

4. Discussion

We have outlined and implemented a Bayesian approach for the calibration of land surface model parameters. This calculation demonstrates that the challenges posed by having only data with large total errors and models with large intrinsic biases may be overcome by adequately accounting for uncertainties and biases within a statistical calibration framework. Our approach focused on the goal of improving estimates of T/ET and therefore emphasized calibration of aerodynamic and vegetation resistances that facilitate the transfer of heat and moisture between soils, leaves, and the overlying atmosphere. We have demonstrated the ability of this calibration approach to constrain the posterior distributions of aerodynamic and vegetation resistance parameters in CLM4. The advantages of this method include 1) quantification of uncertainty in model structure, model parameters, and model output (Tables 2, 3) and 2) calibrating model parameters such that model output agrees with observational data, given the uncertainties inherent in both the model and observations. We find calibrated estimates of T/ET that agree with previous estimates for MEF (Berkelhammer et al. 2016).

A soil moisture data assimilation experiment was conducted to assess the extent to which the imperfect simulation of soil moisture in the ensemble affected the modeled latent and sensible heat fluxes and T/ET . Soil moisture as modeled by CLM4 was “nudged” toward observations of soil moisture using a Kalman filter (Kalman 1960). Soil moisture RMSE was reduced from $0.13 \text{ mm}^3 \text{ mm}^{-3}$ (control model) to $0.06 \text{ mm}^3 \text{ mm}^{-3}$. This data-assimilated model, matching the soil moisture data well, yielded only slight changes in the surface fluxes. The mean differences between the control and moisture-assimilated latent and sensible heat fluxes were -1.70 and 1.63 W m^{-2} , respectively. These changes are much lower than the uncertainties in these quantities. The modeled T/ET changed from 0.622 in the control model to 0.624 when soil moisture was assimilated. The low sensitivity to the simulated soil moisture is attributed in part to the fact that this simple data assimilation approach does not strictly conserve water mass (a water mass imbalance is calculated, based on the updated soil moisture profile). Thus, while the true sensitivity of T/ET to soil moisture is likely higher, this simple experiment suggests that the results of this study were not substantially skewed by an imperfect simulation of soil moisture. This result reaffirms an earlier model evaluation using a modified version of CLM4 (Wong et al. 2017).

The varying levels of constraint provided by the calibration method on the model parameters suggests limitations in the characterization of turbulent regimes within the complex canopy of the MEF site. This result is likely applicable to other types of landscapes. Thus, for any given site, it is reasonable to expect to calibrate the model to account for this representative discrepancy (Abbaspour et al. 2015). The results presented here focus on a single site and should be verified in a variety of ecosystem settings. Performing these experiments for a network of sites would permit the interpolation of the benefits of this calibration approach to the regions between sites within the network, and cautious extrapolation to the areas outside of the network, where observations are lacking. This offers a method to improve model hydrological performance on a global scale, despite the lack

of readily available global datasets against which to calibrate (Das et al. 2008).

The ecosystem resistance parameters chosen in the present study are clearly not exhaustive, yet serve well to demonstrate the potential for this calibration method to be applied to a wide range of land surface modeling applications. Furthermore, the sensitivity analysis described in error covariance was a representative, but not complete, treatment of the parametric uncertainty in CLM4. The model calibration results are conditional on the data employed (Schmidt et al. 2016); thus, important extensions of this work include incorporating additional model parameters as calibration parameters and assimilating other observational data. Observations of soil moisture or stable water isotope ratios in ecosystem water pools, for example, would offer further constraint on the moisture fluxes (Yan et al. 2015). The assimilation of CO₂ fluxes would offer constraint on stomatal resistance, independent of transpiration rates. Indeed, this type of statistical approach has been used widely with success in carbon cycle applications (Hararuk et al. 2014; Xu et al. 2006; Schmidt et al. 2016; Ricciuto et al. 2008).

LAI was considered in the sensitivity analysis presented in section 2d(2), but it is an attractive parameter to incorporate into the calibration itself. The focus of this work was on the exchange parameters themselves, so calibration of LAI directly is beyond the scope of this study. A recent study found an empirical relationship between LAI and T/ET of $T/ET = 0.71(LAI)^{0.14}$ (Wu et al. 2017). Using a site- and season-averaged LAI for MEF of $1.2 \text{ m}^2 \text{ m}^{-2}$ (Berkelhammer et al. 2016), this relationship gives an estimated T/ET of 0.73, which is within agreement with the 95% credible intervals from $E1$ to $E4$. It is unclear, however, how this relationship between LAI and T/ET extends to the open-canopy forest of MEF, from the maize field from which it was derived (Wu et al. 2016).

5. Conclusions

These caveats notwithstanding, we have presented a comparison of calibrated model estimates for ET partitioning in a land surface model. We have shown that the model estimates for T/ET hinge on a sound characterization of the uncertainties inherent in the models and data employed. Our results indicate that underestimating or misrepresenting uncertainty may lead to estimates of T/ET that vary by more than 0.10. Furthermore, the 95% credible range of estimated partitioning narrows as more uncertainties are neglected (Table 2). This indicates that neglecting key uncertainties may result in overconfident estimates of T/ET .

The single-column land surface model calibration experiment presented here may be reproduced for a representative network of sites scattered on a regional scale to create calibrated regional surface datasets and calibrated model predictions. These projections and datasets would prove useful in other hydrological applications, such as that presented here. Such experiments will facilitate the use of widely available high-quality tower data for the validation and calibration of land surface models (Keller et al. 2008).

Acknowledgments. We thank David Gochis, Andrew Turnipseed, Peter Harley (NCAR), and Max Berkelhammer (U. of Illinois) for assistance in interpreting the meteorological, soil moisture, and vegetation data from Manitou. We thank Jesse Nusbaumer, Aleya Kaushik, and Bronwen Konecky for fruitful conversations. Wong and Noone's portion of this research is supported by the NSF Paleo Perspectives and Climate Change Program (AGS-1203928) and Climate and Large-scale Dynamics (AGS-0955841) as part of the CAREER program. Kleiber's portion was supported by NSF DMS-1417724. There are no real or perceived conflicts of interest. All codes and data are available from the corresponding author upon request. The CLM4 model is available through the National Center for Atmospheric Research code repository. The analysis was done in MATLAB.

REFERENCES

- Abbaspour, K. C., E. Rouholahnejad, S. Vaghefi, R. Srinivasan, H. Yang, and B. Kløve, 2015: A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. *J. Hydrol.*, **524**, 733–752, doi:10.1016/j.jhydrol.2015.03.027.
- Baldocchi, D. D., B. B. Hicks, and P. Camara, 1987: A canopy stomatal resistance model for gaseous deposition to vegetated surfaces. *Atmos. Environ.*, **21**, 91–101, doi:10.1016/0004-6981(87)90274-5.
- , —, and T. P. Meyers, 1988: Measuring biosphere-atmosphere exchanges of biologically related gases with micrometeorological methods. *Ecology*, **69**, 1331–1340, doi:10.2307/1941631.
- Ball, J., 1988: An analysis of stomatal conductance. Ph.D. thesis, Stanford University, 145 pp.
- , I. E. Woodrow, and J. A. Berry, 1987: A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. *Progress in Photosynthesis Research*, Vol. 4, J. Biggins, Ed., Martinus Nijhoff, 221–224, doi:10.1007/978-94-017-0519-6_48.
- Berkelhammer, M., and Coauthors, 2013: The nocturnal water cycle in an open-canopy forest. *J. Geophys. Res. Atmos.*, **118**, 10 225–10 242, doi:10.1002/jgrd.50701.
- , D. C. Noone, T. E. Wong, S. P. Burns, J. F. Knowles, A. Kaushik, P. D. Blanken, and M. W. Williams, 2016: Convergent approaches to determine an ecosystem's transpiration fraction. *Global Biogeochem. Cycles*, **30**, 933–951, doi:10.1002/2016GB005392.

- Bonan, G. B., K. W. Oleson, R. A. Fisher, G. Lasslop, and M. Reichstein, 2012: Reconciling leaf physiological traits and canopy flux data: Use of the TRY and FLUXNET databases in the Community Land Model version 4. *J. Geophys. Res.*, **117**, G02026, doi:10.1029/2011JG001913.
- Collatz, G. J., J. T. Ball, C. Grivet, and J. A. Berry, 1991: Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: A model that includes a laminar boundary layer. *Agric. For. Meteorol.*, **54**, 107–136, doi:10.1016/0168-1923(91)90002-8.
- Das, N. N., B. P. Mohanty, and E. G. Njoku, 2008: A Markov chain Monte Carlo algorithm for upscaled soil–vegetation–atmosphere-transfer modeling to evaluate satellite-based soil moisture measurements. *Water Resour. Res.*, **44**, W05416, doi:10.1029/2007WR006472.
- Dauzat, J., B. Rapidel, and A. Berger, 2001: Simulation of leaf transpiration and sap flow in virtual plants: Model description and application to a coffee plantation in Costa Rica. *Agric. For. Meteorol.*, **109**, 143–160, doi:10.1016/S0168-1923(01)00236-2.
- DiGangi, J. P., and Coauthors, 2011: First direct measurements of formaldehyde flux via eddy covariance: Implications for missing in-canopy formaldehyde sources. *Atmos. Chem. Phys.*, **11**, 10 565–10 578, doi:10.5194/acp-11-10565-2011.
- Dunin, F. X., 1991: Extrapolation of point measurements of evaporation: Some issues of scale. *Vegetatio*, **91**, 39–47, doi:10.1007/BF00036046.
- Ehleringer, J. R., and C. B. Field, Eds., 1993: *Scaling Physiological Processes: Leaf to Globe*. Academic Press, 388 pp.
- Essery, R. L. H., M. J. Best, R. A. Betts, P. M. Cox, and C. M. Taylor, 2003: Explicit representation of subgrid heterogeneity in a GCM land surface scheme. *J. Hydrometeorol.*, **4**, 530–543, doi:10.1175/1525-7541(2003)004<0530:EROSHI>2.0.CO;2.
- Gelman, A., and D. B. Rubin, 1992: Inference from iterative simulation using multiple sequences. *Stat. Sci.*, **7**, 457–511, doi:10.1214/ss/1177011136.
- , G. Roberts, and W. Gilks, 1996: Efficient Metropolis jumping rules. *Bayesian Stat.*, **5**, 599–608.
- Haario, H., E. Saksman, and J. Tamminen, 2001: An adaptive Metropolis algorithm. *Bernoulli*, **7**, 223–242, doi:10.2307/3318737.
- Hararuk, O., J. Xia, and Y. Luo, 2014: Evaluation and improvement of a global land model against soil carbon data using a Bayesian Markov chain Monte Carlo method. *J. Geophys. Res. Biogeosci.*, **119**, 403–417, doi:10.1002/2013JG002535.
- Hastings, W. K., 1970: Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97–109, doi:10.1093/biomet/57.1.97.
- Higdon, D., M. Kennedy, J. C. Cavendish, J. A. Cafeo, and R. D. Ryne, 2004: Combining field data and computer simulations for calibration and prediction. *SIAM J. Sci. Comput.*, **26**, 448–466, doi:10.1137/S1064827503426693.
- Hurrell, J. W., and Coauthors, 2013: The Community Earth System Model: A framework for collaborative research. *Bull. Amer. Meteor. Soc.*, **94**, 1339–1360, doi:10.1175/BAMS-D-12-00121.1.
- Jarvis, P. G., 1995: Scaling processes and problems. *Plant Cell Environ.*, **18**, 1079–1089, doi:10.1111/j.1365-3040.1995.tb00620.x.
- Jeremiah, E., S. Sisson, L. Marshall, R. Mehrotra, and A. Sharma, 2011: Bayesian calibration and uncertainty analysis of hydrological models: A comparison of adaptive Metropolis and sequential Monte Carlo samplers. *Water Resour. Res.*, **47**, W07547, doi:10.1029/2010WR010217.
- Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. *J. Basic Eng.*, **82**, 35–45, doi:10.1115/1.3662552.
- Keller, M., D. S. Schimel, W. W. Hargrove, and F. M. Hoffman, 2008: A continental strategy for the National Ecological Observatory Network. *Front. Ecol. Environ.*, **6**, 282–284, doi:10.1890/1540-9295(2008)6[282:ACSFTN]2.0.CO;2.
- Kennedy, M. C., and A. O'Hagan, 2001: Bayesian calibration of computer models. *J. Roy. Stat. Soc.*, **63B**, 425–464, doi:10.1111/1467-9868.00294.
- Kim, S., T. Karl, A. Guenther, G. Tyndall, J. Orlando, P. Harley, R. Rasmussen, and E. Apel, 2010: Emissions and ambient distributions of biogenic volatile organic compounds (BVOC) in a ponderosa pine ecosystem: Interpretation of PTR-MS mass spectra. *Atmos. Chem. Phys.*, **10**, 1759–1771, doi:10.5194/acp-10-1759-2010.
- Landsberg, J. J., and D. B. B. Powell, 1973: Surface exchange characteristics of leaves subject to mutual interference. *Agric. Meteorol.*, **12**, 169–184, doi:10.1016/0002-1571(73)90017-4.
- Link, W. A., and M. J. Eaton, 2012: On thinning of chains in MCMC. *Methods Ecol. Evol.*, **3**, 112–115, doi:10.1111/j.2041-210X.2011.00131.x.
- Liu, S., L. Lu, D. Mao, and L. Jia, 2007: Evaluating parameterizations of aerodynamic resistance to heat transfer using field measurements. *Hydrol. Earth Syst. Sci.*, **11**, 769–783, doi:10.5194/hess-11-769-2007.
- Liu, Y., H. V. Gupta, S. Sorooshian, L. A. Bastidas, and W. J. Shuttleworth, 2005: Constraining land surface and atmospheric parameters of a locally coupled model using observational data. *J. Hydrometeorol.*, **6**, 156–172, doi:10.1175/JHM407.1.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, 1953: Equation of state calculations by fast computing machines. *J. Chem. Phys.*, **21**, 1087, doi:10.1063/1.1699114.
- Morton, C. G., J. L. Huntington, G. M. Pohl, R. G. Allen, K. C. Mcgwire, and S. D. Bassett, 2013: Assessing calibration uncertainty and automation for estimating evapotranspiration from agricultural areas using METRIC. *J. Amer. Water Resour. Assoc.*, **49**, 549–562, doi:10.1111/jawr.12054.
- Oleson, K. W., and Coauthors, 2010: Technical description of version 4.0 of the Community Land Model (CLM). NCAR Tech. Note NCAR/TN-478+STR, 257 pp., doi:10.5065/D6FB50WZ.
- Ortega, J., and Coauthors, 2014: Overview of the Manitou Experimental Forest Observatory: Site description and selected science results from 2008 to 2013. *Atmos. Chem. Phys.*, **14**, 6345–6367, doi:10.5194/acp-14-6345-2014.
- Ricciuto, D. M., K. J. Davis, and K. Keller, 2008: A Bayesian calibration of a simple carbon cycle model: The role of observations in estimating and reducing uncertainty. *Global Biogeochem. Cycles*, **22**, GB2030, doi:10.1029/2006GB002908.
- Sammis, T. W., and L. W. Gay, 1979: Evapotranspiration from an arid zone plant community. *J. Arid Environ.*, **2**, 313–321.
- Schaeffer, S. M., D. G. Williams, and D. C. Goodrich, 2000: Transpiration of cottonwood/willow forest estimated from sap flux. *Agric. For. Meteorol.*, **105**, 257–270, doi:10.1016/S0168-1923(00)00186-6.
- Schmidt, A., B. E. Law, M. Gockede, C. Hanson, Z. Yang, and S. Conley, 2016: Bayesian optimization of the Community Land Model simulated biosphere–atmosphere exchange using CO₂ observations from a dense tower network and aircraft campaigns over Oregon. *Earth Interact.*, **20**, doi:10.1175/EI-D-16-0011.1.
- Song, Y., L. Jin, G. Zhu, and M. Ma, 2016: Parameter estimation for a simple two-source evapotranspiration model using Bayesian inference and its application to remotely sensed

- estimations of latent heat flux at the regional scale. *Agric. For. Meteor.*, **230–231**, 20–32, doi:[10.1016/j.agrformet.2016.03.019](https://doi.org/10.1016/j.agrformet.2016.03.019).
- Twine, T. E., and Coauthors, 2000: Correcting eddy-covariance flux underestimates over a grassland. *Agric. For. Meteor.*, **103**, 279–300, doi:[10.1016/S0168-1923\(00\)00123-4](https://doi.org/10.1016/S0168-1923(00)00123-4).
- Vertenstein, M., A. Middleton, D. Feddema, and C. Fischer, 2010: CESM1.0.4's user guide. NCAR, 152 pp. [Available online at http://www.cesm.ucar.edu/models/cesm1.0/cesm/cesm_doc_1_0_4/ug.pdf.]
- Wong, T. E., J. Nusbaumer, and D. C. Noone, 2017: Evaluation of modeled land–atmosphere exchanges with a comprehensive water isotope fractionation scheme in version 4 of the Community Land Model. *J. Adv. Model. Earth Syst.*, doi:[10.1002/2016MS000842](https://doi.org/10.1002/2016MS000842), in press.
- Wu, Y., T. Du, R. Ding, L. Tong, S. Li, and L. Wang, 2017: Multiple methods to partition evapotranspiration in a maize field. *J. Hydrometeor.*, **18**, 139–149, doi:[10.1175/JHM-D-16-0138.1](https://doi.org/10.1175/JHM-D-16-0138.1).
- Xu, T., L. White, D. Hui, and Y. Luo, 2006: Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction. *Global Biogeochem. Cycles*, **20**, GB2007, doi:[10.1029/2005GB002468](https://doi.org/10.1029/2005GB002468).
- Yan, H., C. M. DeChant, and H. Moradkhani, 2015: Improving soil moisture profile prediction with the particle filter-Markov chain Monte Carlo method. *IEEE Trans. Geosci. Remote Sens.*, **53**, 6134–6147, doi:[10.1109/TGRS.2015.2432067](https://doi.org/10.1109/TGRS.2015.2432067).
- Zeng, X., M. Zhao, and R. E. Dickinson, 1998: Intercomparison of bulk aerodynamic algorithms for the computation of sea surface fluxes using TOGA COARE and TAO data. *J. Climate*, **11**, 2628–2644, doi:[10.1175/1520-0442\(1998\)011<2628:IOBAAF>2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011<2628:IOBAAF>2.0.CO;2).
- , M. Barlage, R. E. Dickinson, Y. Dai, G. Wang, and K. Oleson, 2005: Treatment of undercanopy turbulence in land models. *J. Climate*, **18**, 5086–5094, doi:[10.1175/JCLI3595.1](https://doi.org/10.1175/JCLI3595.1).
- Zhu, G., Y. Su, X. Li, K. Zhang, and C. Li, 2013: Estimating actual evapotranspiration from an alpine grassland on Qinghai-Tibetan plateau using a two-source model and parameter uncertainty analysis by Bayesian approach. *J. Hydrol.*, **476**, 42–51, doi:[10.1016/j.jhydrol.2012.10.006](https://doi.org/10.1016/j.jhydrol.2012.10.006).