Bayesian analysis of canopy transpiration models: A test of posterior parameter means against measurements

D. Scott Mackay a,⁎, Brent E. Ewers b,c, Michael M. Loranty d, Eric L. Kruger e, Sudeep Samanta d

a Department of Geography, State University of New York at Buffalo, 105 Wilkeson Quadrangle, Buffalo, NY 14261, USA
b Department of Botany, University of Wyoming, Laramie, WY 82071, USA
c Program in Ecology, University of Wyoming, Laramie, WY 82071, USA
d Woods Hole Research Center, 145 Woods Hole Rd., Falmouth, MA 02540, USA
e Department of Forest and Wildlife Ecology, University of Wisconsin–Madison, 1630 Linden Dr., Madison, WI 53706, USA

ARTICLE INFO

Article history:
Received 27 January 2011
Received in revised form 31 January 2012
Accepted 8 February 2012
Available online 16 February 2012

This manuscript was handled by Laurent Charlet, Editor-in-Chief, with the assistance of Jose D. Salas, Associate Editor

Keywords:
Bayesian analysis
Canopy transpiration
Canopy gaps
Model data starvation

SUMMARY

Big-leaf models of transpiration are based on the hypothesis that structural heterogeneity within forest canopies can be ignored at stand or larger scales. However, the adoption of big-leaf models is de facto rather than de jure, as forests are never structurally or functionally homogeneous. We tested big-leaf models both with and without modification to include canopy gaps, in a heterogeneous quaking aspen stand having a range of canopy densities. Leaf area index (L) and canopy closure were obtained from biometric data, stomatal conductance parameters were obtained from sap flux measurements, while leaf gas exchange data provided photosynthetic parameters. We then rigorously tested model-data consistency by incrementally starving the models of these measured parameters and using Bayesian Markov Chain Monte Carlo simulation to retrieve the withheld parameters. Model acceptability was quantified with Deviance Information Criterion (DIC), which penalized model accuracy by the number of retrieved parameters. Big-leaf models overestimated canopy transpiration with increasing error as canopy density declined, but models that included gaps had minimal error regardless of canopy density. When models used measured L, the other parameters were retrieved with minimal bias. This showed that simple canopy models could predict transpiration in data scarce regions where only L was measured. Models that had L withheld had the lowest DIC values suggesting that they were the most acceptable models. However, these models failed to retrieve unbiased parameter estimates indicating a mismatch between model structure and data. By quantifying model structure and data requirements this new approach to evaluating model-data fusion has advanced the understanding of canopy transpiration.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Evapotranspiration represents roughly half the water budget of forest ecosystems, and in closed canopies most of this is via transpiration under the control of stomatal conductance. Stomatal conductance is thought to be a coordinated activity that balances the water supply/demand and the demand for carbon by photosynthesis (Brodribb et al., 2002; Katul et al., 2009). The former can be expressed in terms of a function of atmospheric vapor pressure deficit (D) while the latter is explained by a function of photosynthetically active radiation (Q). In most aerodynamically well-mixed forest canopies the vertical and horizontal heterogeneity of stomatal conductance is dominated by changes in Q as influenced by leaf clumping at various scales (Chen and Leblanc, 1997; Kucharik et al., 1999; Ni et al., 1997; Yang et al., 2001). Leaf clumping can be explicitly dealt with in three-dimensional, multi-layer models, but these require data that has generally been expensive to obtain over large spatial domains. All transpiration models, simple or complex, ultimately must rely on empirical prescription of stomatal conductance based on known environmental drivers (Raupach and Finnigan, 1988). Recently, Ewers et al. (2007b) demonstrated that whole-stem hydraulics dampens the heterogeneous responses to Q at leaf and branch levels, and so three-dimensional models may not improve predictions of canopy transpiration over simpler models.

Many canopy transpiration models operating at scales larger than individual crowns or where detailed three-dimensional data are not available use simplified radiative transfer schemes that assume the canopy is a well-mixed turgid medium and treat it as a “big leaf”. The errors associated with the big leaf assumption in models are well known. Law et al. (2001) showed that the effects of light distribution within canopies had the greatest impact on modeled energy exchange at high leaf area index (L), as low L
was mostly sunlit. Recently, it was shown that tree-scale model predictions of transpiration improved when spatial gradients of leaf temperature within the canopy were included (Bauerle et al., 2009). This has implications for the vertical distribution of energy exchange through the canopy, which can be modeled using the big leaf assumption with sun and shade elements (dePury and Farquhar, 1997), and the spatial distribution of leaf area, for which “gappy” models have been developed (Song et al., 2009; Song and Band, 2004) to separately deal with vertical radiative transfer in gaps and canopy areas. Since simplified canopy transpiration models from big leaf to gappy are likely to remain in widespread use at scales ranging from stands to regions we asked the following question: Can these models be parameterized to accurately reproduce transpiration rates and predict parameter values that would be observable in the field? We address this question at the scale of a forest stand.

Canopy transpiration models, such as those that use the Penman–Monteith equation (Monteith, 1965), depend on a small number of key inputs: (1) total energy exchanging leaf area, (2) canopy stomatal conductance ($G_S$), and for gappy models (3) some measure of canopy gap fraction. $G_S$ is reduced from some value under ideal environmental conditions (e.g., reference canopy stomatal conductance, $G_{surf}$) as a function of water loss rate and photosynthetic light limitation such as that provided by phenomenological models (Jarvis, 1976). The daily fractional interception of photosynthetically active radiation in a forest canopy can be computed as (Campbell and Norman, 1998):

$$f = 1 - \exp(-K_d(\psi, L_c) \cdot L_c)$$

(1)

where $K_d(\psi, L_c)$ is the extinction coefficient for diffuse light in the canopy at zenith angle $\psi$, and $L_c = L/P_C$, is leaf area index of the canopy covered areas. $P_c$ (Mackay et al., 2010) is the proportion of canopy closure at the plot scale (m$^2$ maximum crown area m$^{-2}$ ground area). To illustrate the effect of this big leaf assumption on a typical aspen canopy (Campbell and Norman, 1998) consider Eq. (1) applied to a spatially uniform canopy with $L_c$ of 1 m$^2$ m$^{-2}$ and $P_c = 1$ yielding $K_d(\psi) = 0.71$ and $f_{unif orm} = 0.51$. Suppose the same total leaf area was clumped spatially so that half the total area consisted of forest patches with $L_c = 2$ and $P_c = 0.5$, such that there are gaps with $L_{gap} = 0$. In this case, $f_{clumped} = 0.5(f_{patch} + f_{gap})$. For the forest patch, $K_d(\psi, L_c) = 0.65$ and $f_{patch} = 0.73$, and for the gap, $f_{gap}$ approaches zero as $L_{gap}$ approaches zero. Clearly, $f_{clumped} = 0.5 f_{patch} + f_{uniform}$. Both land surface areas have the same total forest leaf surface area, but the latter has a reduced total energy absorbed in the forest canopy and hence a reduced transpiration. A big leaf model could potentially match the transpiration predictions of a gappy model by compensating for the higher energy absorbed with a lower stomatal conductance. Consequently, we hypothesized that (1) big leaf models applied to gappy canopies would fail to accurately predict transpiration using parameter values close to those observed. Indeed, this would shift the modeling burden from observations to parameterization, which should increase model uncertainty. While the debate about correct versus useful models (Raupach and Finnigan, 1998) continues, one diagnostic of a useful model is that the modeled flux responses, obtained using inputs that can (or are) measured, match the observed responses at a level that is consistent with the model’s intended purpose. We also hypothesized that (2) gappy canopy transpiration models that rely more on measured inputs and less on parameterization would be quantitatively more acceptable models.

We test these hypotheses with our canopy transpiration model (Ewers et al., 2008; Loranty et al., 2010b; Mackay et al., 2003; Samanta et al., 2008, 2007). For this analysis we need an extensive sap flux data set supported by detailed canopy structural data, and so we use an established data set (Loranty et al., 2008; Mackay et al., 2010; Traver et al., 2010). Our approach differs from past work because we are using “starvation” analysis and explicitly testing how well the posterior distributions match measured values.

2. Methods

2.1. Study site

Data for this study was obtained in 2005 in a 20-year old regenerating quaking aspen (Populus tremuloides) stand located approximately 800 m southeast of the WLEF eddy flux tower (Bakwin et al., 1998; Davis et al., 2003), in the Chequamegon-Nicolet National Forest (45.95 N, 90.27 W) near the town of Park Falls, Wisconsin, USA. The area lies at the interface between northern temperate and southern boreal ecosystems. It is part of the northern highlands physiographic province, consisting of Precambrian igneous and metamorphic bedrock under 8–90 m of glacial outwash, pitted outwash, and moraines. The climate is characterized by long winters and a short growing season with mean January and July temperatures of $–12$°C and 19°C, respectively (Fassnacht and Gower, 1997).

Details on the data collection are given in previous papers (Loranty et al., 2008; Mackay et al., 2010; Traver et al., 2010). Briefly, we had an inventory of 108 trees instrumented with Granier-type 20-mm length constant-heat sap flux sensor (Granier, 1987) and 752 total P. tremuloides trees. One sensor pair was placed on the north side of each tree at 1.3 m height. Intra tree scaling followed Ewers et al. (2002) to account for bark thickness, sapwood depth, and radial and circumferential trends in sap flux. We included nighttime transpiration by using baselines determined from maximum temperature difference over a week-long period (Oishi et al., 2008). For every tree we recorded location, height, and basal area. Mean per-tree basal areas for the sap flux and other P. tremuloides trees were 80.3 (±5.3 standard error) cm$^2$ and 70.5 (±2.1 standard error) cm$^2$, respectively. Tree heights were measured using a laser rangefinder and clinometer with standard triangulation methods (Bolzan, 2004). Leaf area index and canopy closure (or total projected crown area divided by stand area) were determined through allometric relationships developed within the stand (Mackay et al., 2010).

Leaf-level gas exchange data (LI-6400) was collected from P. tremuloides canopies spanning the wetland to upland positions using portable canopy access scaffolding. All gas exchange measurements were made during mid-June 2005. Maximum carboxylation rate of 47.4 µmol m$^{-2}$ s$^{-1}$, maximum electron transport rate of 80.6 µmol m$^{-2}$ s$^{-1}$, and quantum yield ($\phi$) of 0.08 mol CO$_2$ mol$^{-1}$ photons were obtained from gas exchange measurements (Loranty et al., 2010a). These parameters did not differ between trees growing in the wetland versus the upland or between sun-lit and shaded leaves.

2.2. Model description

Here we used the Terrestrial Regional Ecosystem Simulator (TREES) (Ewers et al., 2008; Loranty et al., 2010b; Mackay et al., 2003; Samanta et al., 2008, 2007) to simulate tree transpiration. Most of the details of the model can be found in these previous manuscripts, and so we describe just the parts of the model needed to interpret the results from the current study, and show how we modified it to account for gappy canopies. TREES employs a simple whole-plant hydraulic function (Oren et al., 1999) to calculate canopy average stomatal conductance prior to photosynthetic limitation:

$$G_S = G_{surf} - m \ln D$$

(2)
where $G_{\text{ref}}$ is reference canopy stomatal conductance (mol m$^{-2}$ s$^{-1}$) at $D = 1$ kPa in absence of photosynthetic light limitation, and $m = - \delta G_{\text{ref}} \delta \ln D$ is the sensitivity of $G_{\text{ref}}$ to $D$. $G_{\text{ref}}$ is used to calculate net $C_3$ photosynthesis ($A_p$) (dePury and Farquhar, 1997; Farquhar et al., 1980) and equilibrium intercellular $CO_2$ ($C_i$) (Katul et al., 2003) in sun and shade canopy elements. $A_p$ and $C_i$ are used in Fick's Law to calculate stomatal conductance to vapor for both sun and shade elements in parallel (Loranty et al., 2010b). We note that this model does not incorporate the response of mesophyll conductance to environmental conditions. However, research shows quite variable responses of mesophyll conductance to environmental stressors (Flexas et al., 2009), and modeled effects of mesophyll conductance on transpiration were inferable only under low soil moisture conditions (Keenan et al., 2010). For the present study soil moisture was not limiting (Loranty et al., 2008; Traver et al., 2010).

Sunlit leaf area of the canopy ($L_{\text{sun}}$) was defined (Campbell and Norman, 1998) as:

$$L_{\text{sun}} = 1 - \exp\left[-K_{\text{le}}(\psi, L_c) \cdot L_c\right]$$

where $K_{\text{le}}(\psi, L_c)$ is the extinction coefficient for light in the canopy at zenith angle $\psi$, calculated using leaf angle distribution and clumping factors (Chen and Cihlar, 1995) appropriate for P. tremuloides (Campbell and Norman, 1998) at the sub-crown scale. $K_{\text{le}}(\psi, L_c)$ does not strictly account for gaps between crowns at a plot scale (plot size $>$ crown size), and so we use $P_{\text{CC}}$ to calculate the gap-based $L_c$ from $L$. Medhurst and Beadle (2002) employed a similar technique to scale transpiration among thinned stands of different stocking densities. Shaded leaf area ($L_{\text{shad}}$) was calculated by subtracting $L_{\text{sun}}$ from $L_c$. When $P_{\text{CC}} = 1$ the model operated in a big-leaf mode with sun and shade elements. We acknowledge that this simple approach only accounts for first-order effects of gaps between crowns on radiative transfer (Ni et al., 1997), but consider it to be acceptable given our modest amount of supporting field data and starvation approach to testing the model.

Simulated canopy transpiration was based on the Penman–Monteith combination equation (Monteith, 1965) as formulated by Campbell and Norman (1998), weighted by the relative amounts of sun and shade element transpiration:

$$E_{\text{Can}} = x^{-1} P_{\text{CC}} \cdot \sum_{k} \left( S \cdot R_{sh} + \gamma_k \frac{g_k (e_l - e_k)}{L_k D/P_k} \right)$$

where $x$ is the latent heat of vaporization, $s$ is the slope of the saturation vapor pressure - temperature relation, $R_{sh}$ is net radiation in the $k$th canopy element, $\gamma_k$ is the apparent psychrometric constant, $S$ is stomatal conductance to vapor for the $k$th canopy element (i.e., sun or shade) taking into account hydraulic, aerodynamic, and photosynthetic factors (Loranty et al., 2010a,b), and $P_k$ is atmospheric pressure.

### 2.3. Model parameterization

#### 2.3.1. Parameters from observations

To test model predictions of canopy transpiration we conducted two experiments. First, we ran the model at each of six canopy density levels yielding six sets of $L$, $P_{\text{CC}}$, $G_{\text{ref}}$, and $m$, to quantify the absolute error of the big leaf and gap-based models. To obtain different canopy densities from the same spatial data set we employed the sap flux plot scaling methodology of Mackay et al. (2010) in which the forest was sampled using a range of plot sizes to exclude (small plots) or include (large plots) canopy gaps ($L = 4.92$ m$^{-2}$ and $P_{\text{CC}} = 0.84$ for 2.5 m plots; $L = 3.83$ m$^{-2}$ and $P_{\text{CC}} = 0.75$ for 3 m plots; $L = 3.28$ m$^{-2}$ and $P_{\text{CC}} = 0.58$ for 4 m plots; $L = 2.81$ m$^{-2}$ and $P_{\text{CC}} = 0.49$ for 5 m plots; $L = 2.13$ m$^{-2}$ and $P_{\text{CC}} = 0.39$ for 7 m plots; and $L = 1.86$ m$^{-2}$ and $P_{\text{CC}} = 0.34$ for 9 m plots). Details on how we quantified $L$ and $P_{\text{CC}}$ from allometric equations developed within our stand are described in Mackay et al. (2010). $G_{\text{ref}}$ was calculated by inverting a form of the Penman–Monteith equation (Monteith and Unsworth, 1990). $G_{\text{ref}}$ and $m$ were estimated from boundary line analysis of $G_{\text{ref}}$ (Ewers et al., 2007a). This data-based technique and process models of stomatal conductance agree on calculated parameters for stomatal conductance (Mackay et al., 2003). Values for $\phi$ were the same for all sets.

#### 2.3.2. Parameters from stochastic modeling

Second, we tested the relative importance of $L$, $P_{\text{CC}}$, $G_{\text{ref}}$, $m$, and $\phi$ using Bayesian analysis while selectively replacing these observation-based inputs with model parameters that were sampled using the Markov Chain Monte Carlo (MCMC) method. Parameter ranges are given in Table 1. These are not quite non-informative priors, as we set upper limits to each parameter based on maximum values we have observed within Northern Wisconsin across a wide range of species, including aspen. $G_{\text{ref}}$ values have not been found to exceed 250 mmol m$^{-2}$ s$^{-1}$ (Ewers et al., 2007a) and so we used 300 as a maximum. For $L$ no value in Northern Wisconsin was found to exceed 6 m$^{-2}$ (Burrows et al., 2002) and this is beyond the values found for aspen, and so we used this as a maximum value. The quantum yield value is the theoretically highest value that can be obtained based on the biochemistry of C3 photosynthesis (Singh et al., 2001). We note that based on our previous work with individual trees in the study site (Loranty et al., 2010a,b) we could have chosen narrower ranges for the prior distributions. However, for this site the current study was our first effort at running TREES at the whole-stand scale, and we did not want to force the model to fit stand-scale priors based on knowledge of parameters at the tree scale.

To test the relative importance of parameters in modeling canopy transpiration we used a Bayesian modeling framework. A complete explanation of the MCMC approach we use is given in two previous papers (Samanta et al., 2008, 2007). Here models were designed in such a way that each withheld one or more observation-based parameters from among $L$, $P_{\text{CC}}$, $G_{\text{ref}}$, $m$, and $\phi$ chosen from the 9 m radius plot scale. We chose this scale because it was most representative of the mean $L$, $P_{\text{CC}}$, and observed transpiration of the stand (Mackay et al., 2010). While it is true that the big leaf model would perform better at higher canopy closure values a true test of its skill should be made at a representative canopy density for the study area. $G_{\text{ref}}$ and $m$ were always included or withheld together because they are mechanistically linked (Oren et al., 1999), and so there were four independent parameters leading to 15 parameter combinations. We also ran one set of chains for a big leaf model form in which $G_{\text{ref}}$, $m$, and $\phi$ were withheld. Withheld parameters were sampled using a multivariate MCMC; otherwise, the parameters were obtained from the data as described above. Letter combinations denote parameters withheld (e.g., model LCCMP had $L$, $P_{\text{CC}}$, $G_{\text{ref}}$, $m$, and $\phi$ withheld, while model $P$ had just $\phi$ withheld and all other parameters measured).

#### Table 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L$</td>
<td>m$^2$ m$^{-2}$</td>
<td>0.1</td>
<td>6.0</td>
</tr>
<tr>
<td>$P_{\text{CC}}$</td>
<td>Percentage</td>
<td>0.01</td>
<td>1.0</td>
</tr>
<tr>
<td>$G_{\text{ref}}$</td>
<td>mmol m$^{-2}$ s$^{-1}$</td>
<td>10.0</td>
<td>300.0</td>
</tr>
<tr>
<td>$m$</td>
<td>mmol m$^{-2}$ s$^{-1}$</td>
<td>10.0</td>
<td>200.0</td>
</tr>
<tr>
<td>$\phi$</td>
<td>mol CO$_2$ mol$^{-1}$ photons</td>
<td>0.0025</td>
<td>0.125</td>
</tr>
</tbody>
</table>
Quantitative comparisons of models were made using Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002) as described in Samanta et al. (2008):

$$DIC = D(\bar{\theta}) + p_D$$

where $D(\bar{\theta})$ is mean deviance, $p_D = D(\bar{\theta}) - D(\tilde{\theta})$ is model complexity, which is defined as the effective number of parameters, and $D(\tilde{\theta})$ is deviance at the posterior means of the parameters. A difference of 7 in DIC values between models was considered significant (Spiegelhalter et al., 2002).

Following MCMC we iteratively simulated a sequence (or chain) of draws from the joint posterior distribution. A candidate parameter value was drawn randomly from a proposal distribution, and a decision was made to accept the parameter based on the ratio of posterior density at the candidate parameter value to that at the current parameter value (Samanta et al., 2008, 2007). This process was continued until a sufficient number of iterations for the chain converged to a stationary posterior distribution, which was then sampled to obtain parameter distributions. For each of the 15 model forms we ran 4 chains of 200,000 iterations. For each chain we initialized the MCMC with a parameter covariance matrix based on trial runs. The covariance matrix was updated every 1000 iterations using either the last 20, 40, or 60 accepted parameter values as needed to converge on a target acceptance rate of about 20% with five parameters (Gelman et al., 2004), but a higher rate with fewer parameters. We discarded the first half of each chain as a “burn-in” period, then visually evaluated and quantitatively assessed for chain convergence using the potential scale reduction factor ($\hat{R}$) (Gelman et al., 2004). In every case ($\hat{R}$) values varied from 0.999 to 1.01, well below the 1.2 maximum recommended by Gelman et al. (2004). To reduce autocorrelation among posterior samples parameter, variances, and model deviance for each iteration $[D(\theta)]$, the second half of each chain was sub-sampled at a gap of 50 simulations (Samanta et al., 2008, 2007).

As a test of the ability of the MCMC algorithm to retrieve the known, measured parameters we also produced synthetic data (e.g., Braswell et al., 2005; Thompson and Katul, 2011) by simulating $\varepsilon_C$ using all of the parameters obtained from the data (i.e., $L = 1.86$, $P_{CC} = 0.34$, $G_{ref} = 79.9 \text{ mmol m}^{-2} \text{s}^{-1}$, $m = 47.4 \text{ mmol m}^{-2} \text{s}^{-1}$, and $\phi = 0.08 \text{ mol CO}_2 \text{ mol}^{-1} \text{ photons}$). The synthetic data was then used to develop MCMC chains for a representative subset of the model forms. In each case the MCMC accurately retrieved all measured parameters (see Supplement Table S1). We also note that the code was tested manually, but we realize that there are recent algorithmic methods (Cook et al., 2006) that could have simplified this process.

3. Results

3.1. Hypothesis 1: Big leaf versus gap-based model

Fig. 1 shows diurnal plots of predicted and observed canopy transpiration ($E_C$) at each of the respective stand densities. The big leaf model was significantly different from observed ($p < 0.01$) for plots of 3.0–9.0 m radii. No gappy model was significantly different from observed. The magnitude of the error of the predictions made with the big leaf model increased as $L$ decreased. Moreover, the relative error in predictions of $E_C$ with the big leaf model was higher as $D$ increased (Fig. 2) indicating that errors were greatest when fluxes were highest and hydrologically most important. By contrast the relative errors in gappy model predictions of $E_C$ showed no relationship with $L$ (or $P_{CC}$) and were never more than 12% of the observed mean flux. However, the gappy model showed a negative bias at low $D$, but also low $E_C$.

3.2. Hypothesis 2: Data richness and model parsimony

The results of Bayesian analysis for the 15 gappy and 1 big leaf model forms are shown in Table 2 ranked from lowest to highest DIC. Model complexity or effective number of parameters ($p_D$) increased with the number of parameters. There were relatively small variations in DIC among the gappy models, but as a guide to our analysis we formed breaks between LCGMP and LCGM, as well as between LGM and L ($\Delta \text{DIC} > 7$ among models). The model

![Fig. 1. Canopy transpiration ($E_C$) using big leaf (solid black lines) and gap-fraction based models (dashed black lines) versus observed $E_C$ (gray lines) for leaf area index ($L$) of (a) 4.92, (b) 3.83, (c) 3.28, (d) 2.81, (e) 2.13, and (f) 1.86. Note the differences in scales on the y-axis among the rows of plots. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).](image-url)
groups tend to highlight common properties among the models that fall within each respective group. The parameters for the group with the lowest DIC values are distinct from those of the second and third groups. The big leaf model is shown as a separate group since it cannot be compared to the gappy models using DIC without knowledge of saturated deviance (Spiegelhalter et al., 2002, p. 606). We note that other measures, such as Bayes Factor, would be more useful if one wanted to use the different models for averaging purposes (Qian et al., 2005), which was not done here.

Models LGMP, LCGMP, LCGM, and LGM all overestimated $G_{\text{Sref}}$ and $m$, and underestimated $L$. The first two models also predicted very low values for $\phi$. The remaining models closely predicted the observed $G_{\text{Sref}}$, $m$, $L$, but also underestimated $\phi$. $G_{\text{Sref}}$ for these models was at most 5 mmol m$^{-2}$ s$^{-1}$ lower than estimated from boundary line analysis from sap flux data, and this was within a standard error (about 6 mmol m$^{-2}$ s$^{-1}$) of these estimates (Mackay et al., 2010). Among gappy models the range in $L_{G_{\text{Sref}}}$ was relatively small (133.3–154.2 mmol m$^{-2}$ s$^{-1}$), which is expected given the constraint to match total $E_C$. $P_{\text{CC}}$ showed very little variation among models. The proportionality between $m$ and $G_{\text{Sref}}$ was approximately 0.6 as predicted by plant hydraulic theory (Oren et al., 1999; Katul et al., 2009) for all gappy models that had mean posterior $G_{\text{Sref}}$ values close to observed, while the other models (LGMP, LCGMP, LCGM, and LGM) had $m/G_{\text{Sref}}$ proportionalities of 0.7, which do disagree with plant hydraulic theory. Given that we had a relatively small standard error on $G_{\text{Sref}}$ compared to the absolute values retrieved for these four models, this ratio can be considered significantly higher than 0.6. By contrast, the big leaf model had a higher DIC, lower $\phi$, and lower $L_{G_{\text{Sref}}}$ than any of the gappy models.

Posterior parameter distributions for models GM, GMP, CGMP, and LCGMP, and the big leaf model are shown in Figs. 3a–d and 4, respectively. We note that parameter posteriors happened to follow Gaussian distributions, except those for $\phi$, although such a finding was neither anticipated nor necessary. Posterior parameters for $m$ and $P_{\text{CC}}$ closely matched their respective measured values when measured $L$ was used. When we also withheld measured $L$ (Fig. 3b) posterior distributions for all parameters, except for $\phi$, remained Gaussian despite having distributions that excluded (except for $P_{\text{CC}}$) the respective measured values. Posterior parameter distributions for the big leaf model (Fig. 4) were narrower than their respective gappy model counterparts. We are pragmatic and so a narrower spread is a desirable outcome. However, here the narrowly spread posterior parameter values underestimate our measured $G_{\text{Sref}}$, $m$, and $\phi$ values, although the $m/G_{\text{Sref}}$ was similar to that of the gappy models (see also Table 2). The big leaf model is clearly not consistent with the data, and so it would be prudent to reject it.

Posterior predictions of $E_C$ by the gappy (GMP) and big leaf models are shown in Fig. 5. For both models the 95% posterior interval bounded 95% of the observations. Also shown are 75% posterior intervals for both models. The big leaf model is clearly not consistent with the data, and so it would be prudent to reject it.

**Fig. 2.** Relative error in modeled canopy transpiration ($E_C$) [(modeled-measured)/measured - 100%] with changes in leaf area index (L) and canopy closure ($P_{\text{CC}}$) was lower than their respective observed values as a means to compensate for the higher absorption of solar radiation by the big leaf model ($L = 1.86$, $P_{\text{CC}} = 100\%$, $f_{\text{uniform}} = 0.7$ using Eq. (1)) compared to the gappy models ($L = 5.47$ in forest patches, $P_{\text{CC}} = 34\%$, $f_{\text{stumped}} = 0.32$ using Eq. (1)).

**Table 2**

Parameters derived from observations1 or from median of the posterior distributions for the parameters from Bayesian analysis. Models are denoted by letter representing the respective parameters that were derived by Bayesian analysis: L for leaf area index (L), C for canopy closure ($P_{\text{CC}}$), G for reference canopy average stomatal conductance ($G_{\text{Sref}}$), M for sensitivity of stomatal closure to vapor pressure deficit (m), and P for reference quantum yield (greek letter phi, $\phi$). The model, big leaf, denotes GMP run with a big leaf model. Breaks in ADIC are indicated by row gaps as a guide to identifying groups of models.

<table>
<thead>
<tr>
<th>Model</th>
<th>$P_C$</th>
<th>DIC</th>
<th>ADIC</th>
<th>$P_{\text{CC}}$ (%)</th>
<th>$G_{\text{Sref}}$</th>
<th>$m$</th>
<th>$L$</th>
<th>$\phi$</th>
<th>$L_{G_{\text{Sref}}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LGMP</td>
<td>4.43</td>
<td>-8265.71</td>
<td>0.00</td>
<td>34.00</td>
<td>206.16</td>
<td>146.14</td>
<td>0.71</td>
<td>0.017</td>
<td>145.4</td>
</tr>
<tr>
<td>LGMP</td>
<td>5.33</td>
<td>-8264.25</td>
<td>1.46</td>
<td>31.29</td>
<td>214.09</td>
<td>153.32</td>
<td>0.72</td>
<td>0.019</td>
<td>154.2</td>
</tr>
<tr>
<td>LCGM</td>
<td>4.31</td>
<td>-8252.31</td>
<td>13.40</td>
<td>27.85</td>
<td>178.32</td>
<td>127.41</td>
<td>0.82</td>
<td>0.080</td>
<td>145.3</td>
</tr>
<tr>
<td>LGM</td>
<td>3.40</td>
<td>-8247.07</td>
<td>18.64</td>
<td>34.00</td>
<td>148.44</td>
<td>101.06</td>
<td>0.90</td>
<td>0.080</td>
<td>133.3</td>
</tr>
<tr>
<td>L</td>
<td>1.43</td>
<td>-8237.80</td>
<td>27.91</td>
<td>34.00</td>
<td>79.95</td>
<td>47.43</td>
<td>1.72</td>
<td>0.080</td>
<td>137.3</td>
</tr>
<tr>
<td>LP</td>
<td>2.68</td>
<td>-8237.24</td>
<td>28.47</td>
<td>34.00</td>
<td>79.95</td>
<td>47.43</td>
<td>1.74</td>
<td>0.058</td>
<td>138.9</td>
</tr>
<tr>
<td>LC</td>
<td>2.32</td>
<td>-8236.30</td>
<td>29.41</td>
<td>33.50</td>
<td>79.95</td>
<td>47.43</td>
<td>1.73</td>
<td>0.080</td>
<td>138.7</td>
</tr>
<tr>
<td>LCP</td>
<td>3.36</td>
<td>-8235.91</td>
<td>29.80</td>
<td>34.08</td>
<td>79.95</td>
<td>47.43</td>
<td>1.74</td>
<td>0.059</td>
<td>138.8</td>
</tr>
<tr>
<td>C</td>
<td>1.41</td>
<td>-8235.00</td>
<td>30.71</td>
<td>31.40</td>
<td>79.95</td>
<td>47.43</td>
<td>1.86</td>
<td>0.080</td>
<td>148.7</td>
</tr>
<tr>
<td>CP</td>
<td>2.72</td>
<td>-8234.94</td>
<td>30.77</td>
<td>31.99</td>
<td>75.40</td>
<td>45.71</td>
<td>1.86</td>
<td>0.056</td>
<td>140.2</td>
</tr>
<tr>
<td>GM</td>
<td>2.37</td>
<td>-8234.63</td>
<td>31.08</td>
<td>34.00</td>
<td>75.40</td>
<td>46.68</td>
<td>1.86</td>
<td>0.080</td>
<td>140.2</td>
</tr>
<tr>
<td>GMP</td>
<td>3.68</td>
<td>-8233.91</td>
<td>31.80</td>
<td>34.00</td>
<td>74.74</td>
<td>45.12</td>
<td>1.86</td>
<td>0.059</td>
<td>139.0</td>
</tr>
<tr>
<td>CGM</td>
<td>3.44</td>
<td>-8232.79</td>
<td>32.92</td>
<td>33.06</td>
<td>76.92</td>
<td>46.94</td>
<td>1.86</td>
<td>0.080</td>
<td>143.1</td>
</tr>
<tr>
<td>CGMP</td>
<td>4.63</td>
<td>-8232.13</td>
<td>33.58</td>
<td>33.40</td>
<td>77.21</td>
<td>46.84</td>
<td>1.86</td>
<td>0.059</td>
<td>143.0</td>
</tr>
<tr>
<td>P</td>
<td>1.04</td>
<td>-8229.05</td>
<td>36.66</td>
<td>34.00</td>
<td>79.95</td>
<td>47.43</td>
<td>1.86</td>
<td>0.055</td>
<td>148.7</td>
</tr>
<tr>
<td>Big leaf</td>
<td>3.66</td>
<td>-8165.31</td>
<td>100.40</td>
<td>100.00</td>
<td>60.13</td>
<td>31.10</td>
<td>1.86</td>
<td>0.017</td>
<td>119.1</td>
</tr>
</tbody>
</table>

1 Parameter values for $G_{\text{Sref}}$, $m$, $L$, and $\phi$ shown in italics are values obtained from field measurements.
4. Discussion

4.1. Hypotheses

We failed to reject the hypothesis that the big leaf model assumption would be subject to increasing error as the stand leaf area became more clumped. Indeed, error in $E_{\text{Csim}}$ was 35% of observed $E_C$ (Fig. 2) as $L$ approached its stand mean value. At low $D$ the gappy models underestimated $E_C$ by 7–12%, but this error showed no trend with $L$. Based on just over 100 sap flux trees we estimate a standard error on the estimate of observed $E_C$ to be about 6%. The underestimation of $E_C$ at low $D$ by the gappy models could be attributed to underestimation of radiation absorbed by the canopy at low sun angles by our simple radiative transfer model, inappropriate response of $G_S$ at low $D$, or simply an underestimation of photosynthesis at low $D$. This was not an issue at high $D$ and hence high sun angles, and only resulted in a significant deviation in $E_{\text{Csim}}$ from $E_C$ at $L<2\, m^2/m^2$. Moreover, all gappy models had lower DIC values than the big leaf model (Table 2).

We note that we could anticipate the parameter values that would be obtained if the big leaf model had been run over smaller plot sizes. At higher $L$ and $P_{\text{CC}}$ values for $E_C$ would be proportionally higher (Mackay et al., 2010), and so $G_{\text{Sref}}$ and $m$ remain approximately the same at all plot radii as they are proportional to $E_C/L$. Moreover, non-representative values for $L$ and $P_{\text{CC}}$ would not improve the posterior estimates of $\phi$. As a further test of the spatial validity of our approach we repeated the Bayesian analysis in three 9 m plots chosen from wetland, transitional, and upland areas of the stand, with corresponding low, intermediate, and high $L$. The posterior parameters made the adjustments as stated above, but the ranking of models based on DIC did not change.

We rejected the hypothesis that more acceptable models in terms of DIC would necessarily be associated with increased data inputs because the four models with the lowest DIC were the worst models in terms of posterior prediction of withheld parameter val-
Fig. 5. 95% (Dark gray) and 75% (light gray) posterior prediction intervals and posterior median (solid line) for modeled canopy transpiration ($E_c$) from scaled sap flux (open circles) for (a) gap and (b) big leaf models. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.2. Implications for ecosystem modeling at stand scale

Previous studies have shown that transpiration rates increase from relatively closed stand centers to open edges (Cienciala et al., 2002; Giambelluca et al., 2003; Herbst et al., 2007; Taylor et al., 2001) in response to changes in advection (Taylor et al., 2001) and effects of light reduction on stomatal conductance (Loranty et al., 2010a). These results indirectly account for the effects of the spatial heterogeneity of radiation absorbed by the canopy. Our results show that inferences made with big leaf models with sun and shade components would likely accrue increasing error in stands with larger gap sizes regardless of the mean leaf area index. This is a corollary to previous results showing that inferences made with big leaf models were not apparent in the gas exchange data. Alternatively, Kurpius et al. (2003) employed the same method for both thinned and un-thinned plantation stands, which may be reasonable given that plantations are designed to be relatively homogeneous with respect to canopy structure. For forests stands with more heterogeneous spatial distribution of $L$ we have shown that parameter adjustments may only partially offset the error associated with the big leaf assumption, but at the cost of having parameter values that are essentially meaningless.

More complex canopy models have been employed to reduce the need for physically unrealistic parameterizations, including multi-layer models (Bernier et al., 2002; Leuning et al., 1995; Raulier et al., 2000) and single-layer models with sun and shade elements (dePury and Farquhar, 1997; Samanta et al., 2007; Schymanski et al., 2007; Wang and Leuning, 1998). However, these do not explicitly consider canopy gaps, and so our results suggest that these models would suffer at least some of the illnesses of pure big leaf models when applied to stands with spatially heterogeneous leaf area. Medhurst et al. (2002) showed that transpiration was accurately predicted when leaf area was divided by projected crown area, and then calculated transpiration was multiplied by projected crown area. Loranty et al. (2010a) used a ray-tracing radiative transfer model to deliver filtered above-canopy radiation on a tree-by-tree basis using three-dimensional data on canopy architecture for two broadleaf species. Such data-intensive methods potentially serve as benchmarks for future canopy transpiration and stomatal conductance modeling efforts, but are impractical at large scales by virtue of the amount of data that must be collected in the field. Alternatively, by representing canopy gaps probabilistically the SVP model (Song et al., 2009; Song...
4.3. Broader implication for large-scale models

Our results are consistent with a broad range of land surface modeling work in which leaf area index is considered key (Davi et al., 2009). We have shown that the other parameters in our study, \( G_{\text{ref}} \), \( m \), \( P_{\text{co}} \), and \( \phi \) can be derived through calibration without loss of model fidelity, but that the loss of \( L \) is critical. However, for large-scale modeling efforts, observations of \( E_{\text{c}} \) for model calibration are not available. Recent work suggests that at least some of the parameters needed may be derivable from canopy structural properties. For example, Novick et al. (2009) demonstrated a relationship between \( G_{\text{ref}} \) and canopy height in temperate forests, and Loranty et al. (2010b) demonstrated for \( P. tremuloides \) a strong relationship between \( G_{\text{ref}} \) and competition for light among crowns. Recent advances in high-resolution remote sensing of vegetation structure, such as from LiDAR (Brandtberg et al., 2003; Houldcroft et al., 2005), and ecosystem models capable of utilizing such information (Hurtt et al., 2004) suggest a practical way forward for the use of three-dimensional models as benchmarks for development of simpler models.

5. Conclusions

Big-leaf models of canopy transpiration are routinely employed by de facto. A major shortcoming of these models is that they ignore canopy structural heterogeneity, which suggests that they would be particularly problematic in forests disturbed by fire, insects, drought, or direct anthropogenic forcing. Alternatively, gap-based, or gappy, models are more complex and data intensive, but should be more sensitive to canopy heterogeneity. Big-leaf models can produce tolerable accuracies when calibrated against known flux measurements, which has been justification for their use over more costly gappy models. Here we used a novel data collection protocol that yields enough detail to fully describe parameters for both big-leaf and gappy models, and a Bayesian model comparison approach where the need for prior observation-based constraints on parameter values is gradually eliminated. We conclude that big-leaf models reproduce observed transpiration using incorrect parameter values whereas gappy models can produce accurate flux estimates using the correct parameter values. We also conclude that gappy models are not more data intensive than big-leaf models, as it was possible to retrieve through parameterization the correct observed parameters for canopy closure, reference canopy stomatal conductance, and quantum yield of photosynthesis as long as leaf area index was well constrained a priori. We have demonstrated that considerable information about minimum data and modeling requirements can be obtained with our approach. We suggest that our approach can, and should be, replicated in other flux studies on water, carbon, energy, and nutrients, especially given the growing network of data-intensive sites.

Acknowledgements

Funding for this study was from the National Science Foundation, Hydrologic Sciences Program (EAR-0405306 to D.S.M., EAR-0405381 to B.E.E., and EAR-0405318 to E.L.K.). The NSF IGERT program also supported M.M.L. D.S.M. also acknowledges funding from the Department of Energy (DOE) Office of Biological and Environmental Research, National Institute for Climatic Change Research (NICCCR) Midwestern region sub-agreement 050516220. The statements made in this manuscript reflect the views of the authors and do not necessarily reflect the views of NSF or DOE. We are grateful to personnel at Kemp Natural Resources Station, and to Elizabeth Traver and David Roberts for their assistance with fieldwork. We are especially grateful for the insightful comments made by the anonymous reviewers and the Associate Editor.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jhydrol.2012.02.019.

References
