Use of FLUXNET in the Community Land Model development

R. Stöckli,1,2,3 D. M. Lawrence,4 G.-Y. Niu,5 K. W. Oleson,4 P. E. Thornton,4 Z.-L. Yang,5 G. B. Bonan,4 A. S. Denning,1 and S. W. Running6

Received 26 July 2007; revised 12 October 2007; accepted 21 December 2007; published 19 March 2008.

[1] The Community Land Model version 3 (CLM3.0) simulates land-atmosphere exchanges in response to climatic forcings. CLM3.0 has known biases in the surface energy partitioning as a result of deficiencies in its hydrological and biophysical parameterizations. Such models, however, need to be robust for multidecadal global climate simulations. FLUXNET now provides an extensive data source of carbon, water and energy exchanges for investigating land processes, and it encompasses a global range of ecosystem-climate interactions. Data from 15 FLUXNET sites are used to identify and improve model deficiencies. Including a prognostic aquifer, a bare soil evaporation resistance formulation and numerous other changes in the model result in a significantly improved soil hydrology and energy partitioning. Terrestrial water storage increased by up to 300 mm in warm climates and decreased in cold climates. Nitrogen control of photosynthesis is revealed as another missing process in the model. These improvements increase the correlation coefficient of hourly and monthly latent heat fluxes from a range of 0.5–0.6 to the range of 0.7–0.9. RMSE of the simulated sensible heat fluxes decrease by 20–50%. Primary production is overestimated during the wet season in mediterranean and tropical ecosystems. This might be related to missing carbon-nitrogen dynamics as well as to site-specific parameters. The new model (CLM3.5) with an improved terrestrial water cycle should lead to more realistic land-atmosphere exchanges in coupled simulations. FLUXNET is found to be a valuable tool to develop and validate land surface models prior to their application in computationally expensive global simulations.


1. Introduction

[2] The land surface provides a lower boundary to the atmosphere for exchanges of radiation, heat, water, momentum and chemical species such as CO2. The importance of these exchanges for the climate system is increasingly being recognized [Betts et al., 2000; Cox et al., 2000; Pielke, 2001; Friedlingstein et al., 2003; Seneviratne et al., 2006; Betts et al., 2007]. Storage of heat and water on land constitutes a significant memory component within the climate system. For instance, soil moisture has strong controls on ecosystem function and boundary layer processes in regions where evapotranspiration as a biophysical process is limited by soil moisture availability [Seneviratne and Stöckli, 2007]. Furthermore the global carbon cycle interacts with soil and vegetation biophysics since carbon assimilation and ecosystem respiration are regulated by the land surface radiation, water and heat balances.

[3] Land surface models for use in global climate models have been developed over the last three decades. They range from simple energy balance parameterizations to complex schemes including the full terrestrial biogeochemical cycle [Sellers et al., 1997; Friedlingstein et al., 2006] and are based on knowledge gained from field and laboratory research in plant physiology, soil science and micrometeorology. However, many model components resulted from relatively few observations and from idealized laboratory experiments. This leads to significant uncertainty in the parameterization of processes which are now employed on a global scale for studying land-climate interaction at seasonal to decadal timescales.

[4] These model uncertainties have been documented in model inter-comparison studies (e.g., PILPS [Henderson-Sellers et al., 1996; Pitman et al., 1999; Nijssen et al., 2003]). Large differences still exist in the simulation of
seasonal and annual evapotranspiration and runoff dynamics [Gedney et al., 2000]. It is not clear how much present climate model predictions are affected by these limitations. For instance, coupling strength between the land surface and the atmosphere varies not only by region but also by the used parameterization [Koster et al., 2004]. A realistic representation of land surface responses to climatic variability as part of global climate simulations is important for future climate impact studies. It is also mandatory in the prediction of the global carbon balance, with regional sinks and sources, which will be part of the next generation earth system models.

[5] The Community Land Model version 3 (CLM3.0) is a community-developed land surface model maintained at NCAR (National Center for Atmospheric Research) and includes a comprehensive set of mechanistic descriptions of soil physical and vegetation biophysical processes [Oleson et al., 2004]. The model can be extended to a full biogeochemical description of the terrestrial carbon-nitrogen interactions [Thornton et al., 2007] based on the BIOME-BGC model and vegetation biogeography with disturbance dynamics [Levis et al., 2004] based on the LPJ model.

[6] Despite being an advanced process-based land surface model, CLM3.0 has known deficiencies in simulating the long term terrestrial hydrological cycle in climate simulations. They can influence the surface climate and vegetation biogeography through plant-soil carbon and water dynamics [Dickinson et al., 2006]. In coupled simulations with many feedback processes, these shortcomings can further amplify errors from the atmospheric model, with unhealthy consequences for the simulated climate system and land-atmosphere interactions [Bonan and Levis, 2006; Hack et al., 2006; Lawrence et al., 2007]. The CLM model development community has proposed a number of improved soil hydrological and plant physiological formulations that represent previously missing processes that appear to be responsible for a damped soil water storage cycle in the tropics and the generally dominating fraction of bare soil evaporation to plant transpiration [see, e.g., Lawrence et al., 2007]. For details about the full set of proposed changes to CLM, see Oleson et al. [2007]. Here, we both evaluate how these changes have improved the model and also elucidate how the use of FLUXNET data has contributed to the identification of deficiencies in the model including the aforementioned missing processes. The subset of changes to the model evaluated in detail here include: (1) a Topmodel-based runoff, infiltration and aquifer model, (2) a bare soil evaporation resistance and, (3) an empirical function for nitrogen control of the photosynthesis-conductance formulation. The aim of this study is to individually implement and evaluate the proposed algorithms and to quantify their impact on the simulated terrestrial carbon and water cycle on hourly to seasonal timescales.

[7] Such a study is difficult for a global land surface model due to a lack of suitable global observations [Henderson-Sellers et al., 2003]. However, long-term ground-based ecosystem observations such as FLUXNET [Baldocchi et al., 2001], the global network of research sites where the eddy covariance technique is used to monitor surface-atmosphere exchanges of carbon, water, and energy, are a unique data source for process-based land surface model development [Running et al., 1999; Canadell et al., 2000; Reichstein et al., 2002; Turner et al., 2004; Stöckli and Vidale, 2005; Boga da et al., 2006; Friend et al., 2007] although it is important to remember that these observations are of local scale and can be subject to potentially large random and systematic errors [Wilson et al., 2002; Foken, 2008]. FLUXNET is probably the most comprehensive terrestrial ecosystem data set today, and uncertainties in radiation, heat, water and carbon flux measurements can be accurately quantified [Falge et al., 2001; Schmid, 2002; Hollinger and Richardson, 2005; Richardson et al., 2006]. Flux tower observations per se only have limited spatial scalability and do not provide a gridded global coverage. They do, however, span a global range of ecosystems where we can exercise land surface models like CLM3.0. The importance of individual processes regulating the heat, water and carbon exchanges varies by climate. Certain processes may only play a role at one end of the multidimensional spectrum of climatic environments.

[8] In this study we use 15 FLUXNET tower sites from the temperate, mediterranean, tropical, boreal and subalpine climate zones to interactively assess the realism of proposed CLM3.0 enhancements during model development. Gap-filled yearly meteorological forcing data sets at the tower sites are used to conduct off-line single-point simulations. In the results section quality-screened heat, water and carbon fluxes as well as soil moisture and soil temperature measurements are compared to simulated equivalents. Several model hydrological deficiencies controlling turbulent surface fluxes, are successively identified and corrected with this study. It is therefore demonstrated how FLUXNET helps to reduce model biases in the simulation of land surface processes and how it can be used as an efficient tool for the reevaluation of land surface models like CLM3.0 during their development stage.

2. Methods

2.1. Model

[9] CLM3.0 (Community Land Model Version 3 [Oleson et al., 2004]) is the land model component of CCSM3 (Community Climate System Model Version 3 [Collins et al., 2006]). It includes mechanistic formulations of physical, biophysical and biogeochemical processes that simulate the terrestrial radiation, heat, water and carbon fluxes in response to climatic forcings. CLM3.0 provides an integrated coupling of photosynthesis, stomatal conductance, and transpiration. Therefore vegetation biophysical processes strongly interact with soil hydrological processes. The CLM3.0 community has proposed a number of model changes as a response to the above discussed deficiencies of the CLM3.0 code. Three of them, in particular, are directly related to simulations of the global hydrological cycle and are summarized here (full documentation in Oleson et al. [2007]):

[10] 1. Infiltration, runoff and groundwater: A Topmodel-based infiltration, saturation and runoff scheme [Beven and Kirkby, 1979; Niu et al., 2005] introduces catchment-scale soil water dynamics from classical hydrological modeling to a land surface model for global applications. Additionally, a prognostic aquifer scheme [Niu et al., 2007] allows for seasonal to inter-annual soil water storage fluctuations which involve soil depths beyond the 3.43 m deep soil of
CLM3.0. The depth of the water table is highly related to subsurface runoff magnitude [Sivapalan et al., 1987; Chen and Kumar, 2001]. During dry periods the aquifer contributes to base-flow and provides a long-term storage for soil water. It is hydraulically connected to the root zone and therefore interacts with vegetation biophysical state and function. During rainfall events or in moist climates the water table can rise into the model soil column, which increases root zone soil moisture and subsurface runoff. It also increases infiltration since soil hydraulic conductivity shows a highly nonlinear dependence on soil water content. In the original CLM3.0 the magnitude of soil water dynamics is constrained to total soil depth, while here the aquifer acts as a buffer with a storage capacity varying by climate, soil, vegetation and topography.

[11] 2. Soil evaporation: In the original CLM3.0 an unreasonably high fraction of evapotranspiration comes from bare soil evaporation [Lawrence et al., 2007]. In addition to the already simulated top soil humidity [Oleson et al., 2007, equations (F1)–(F4)] a new resistance function was implemented, based on work by Sellers et al. [1992]. Equation (F5) in Oleson et al. [2007] is an empirical parameterization of the bare soil evaporation resistance, which was developed on a limited number of FIFE 87 measurements. It had previously been successfully used in SiB 2 and 2.5 (Simple Biosphere Model Versions 2 and 2.5 [Sellers et al., 1996, Vidale and Stöckli, 2005]).

[12] 3. Nitrogen limitation: Initial simulations including the above soil hydrological processes revealed an exaggerated light response of photosynthesis, resulting in too much primary production and slightly overestimated latent heat flux. Apart from soil water, temperature, humidity and radiation, leaf nitrogen content can define the maximum rate of carboxylation in the photosynthesis formulation and therefore stomatal opening. While prognostic nitrogen is part of the separately developed biogeochemistry scheme CLM-CN [Thornton et al., 2007], many applications require the standard CLM. In order to simulate nitrogen control on photosynthesis and therefore stomatal conductance, PFT-dependent factors \( f(N) \) were diagnosed from a simulation employing CLM-CN from a fully spun-up preindustrial state of terrestrial biogeochemistry. \( f(N) \) represents the proportion of potential photosynthesis that is realized in the face of nitrogen limitation, as predicted by CLM-CN. For our simulations \( f(N) \) is imposed on the maximum rate of carboxylation \( V_{\text{max}} \) in a similar manner to, e.g., plant water stress, as described in Oleson et al. [2007, Appendix G]. \( V_{\text{max}} \) then modulates canopy photosynthesis and therefore carbon uptake as well as canopy conductance and therefore transpiration in the model.

### 2.2. Data

[13] FLUXNET is a global network of currently more than 400 flux towers which operate independently or as part of regional networks (CarboEurope, AmeriFlux, LBA, etc.). The off-line single point simulations with CLM3.0 were carried out at 15 FLUXNET sites covering a range of climatic environments listed in Table 1: temperate (5), mediterranean (3), boreal (3), tropical (2), north boreal (1) and subalpine (1). Only towers providing three or more years of continuous driver and validation data as part of the publicly accessible AmeriFlux or CarboEurope standardized Level 2 database have been selected. In order to obtain a balanced set of flux towers, only a few temperate sites could be used. On the other hand arctic and especially more arid sites with multiyear continuous coverage were difficult to find.

#### 2.2.1. Forcing Data

[14] Yearly gap-filled meteorological driver data were created from level 2 flux tower data sets at 30 or 60 min time steps. For off-line simulations the model requires \( R_{\text{G,d}} \) (downwelling short-wave radiation; \( \text{W m}^{-2} \)), \( L_{\text{W,d}} \) (downwelling long-wave radiation; \( \text{W m}^{-2} \)), \( T_{\text{a}} \) (air temperature; K), \( RH_{\text{d}} \) (relative humidity; %), \( u \) (wind speed; m s\(^{-1}\)), \( P_{\text{e}} \) (surface pressure; Pa), \( P \) (rainfall rate; mm s\(^{-1}\)). Measurements of these quantities at the tower reference height...
(Table 1) were used. Outliers which deviated $n \sigma$ times from the median-filtered time series were removed ($\sigma$ is the standard deviation of the original time series and $n = 4$, except for $RH_u$ where $n = 8$; for $u$ where $n = 20$ and for $LW_d$ where $n = 16$). Up to two month long successive gaps were filled by applying a 30 day running mean diurnal cycle forwards and backwards through the yearly time series. Years with more than 2 month of consecutive missing data were not used.

[15] The following exceptions were applied to the above procedure:

[16] 1. $RG_d$ was not median-filtered since most of its variability occurs on diurnal timescale. Instead the potential solar radiation as a function of latitude and local solar time, scaled with the annual maximum observed $RG_d$, provided an upper bound for $RG_d$.

[17] 2. $P$ was neither median-filtered nor gap-filled. Where provided by tower sites, daily precipitation totals from nearby stations were used to replace missing 30 or 60 min tower data. Daily precipitation totals were evenly distributed at night between 00:00–04:00 during days when no daily 30 or 60 min were available; and they were used to augment valid 30 and 60 min data during days with partial missing data periods.

[18] 3. For sites with no $P$, it was estimated by

$$P = P_a e^{-\sigma m},$$

where $P_a$ is the mean sea level pressure (101300 Pa), $M$ is the molecular weight of air (0.029 kg mol$^{-1}$), $g$ is the gravitational acceleration ($9.81$ m s$^{-2}$), $z$ is the tower height above sea level (m) and $R$ is the universal gas constant ($8.314$ J K$^{-1}$ mol$^{-1}$).

[19] 4. For sites with no $LW_d$ (most sites), it was estimated from the surface radiation balance:

$$LW_d = R_n - RG_d + RG_u + \sigma\left(\frac{T_a + T_i}{2}\right)^4,$$

where $R_n$ and $RG_u$ are non-gap-filled net radiation (W m$^{-2}$) and upwelling solar radiation (W m$^{-2}$), $\sigma$ is the Stefan-Boltzmann constant ($5.67 \cdot 10^{-8}$ W m$^{-2}$K$^{-4}$) and $T_i$ is either the canopy temperature or soil surface temperature (K), depending on data availability. As a backup algorithm (any of the right hand side variables in equation (2) missing, most sites, again) downwelling long-wave radiation was estimated by using the clear-sky $LW_d$ parameterization by Idso [1981], modified by an emissivity correction factor as proposed by Gabathuler et al. [2001]:

$$LW_d = \epsilon_c \epsilon_0 \sigma T_a^4,$$

where $\epsilon_0$ is the clear sky atmospheric emissivity as a function of $T_a$ and atmospheric vapor pressure $e_a$ (mb). $\epsilon_c$ adjusts $\epsilon_0$ for cloud cover. It depends on the clearness index $K_0$, which ranges from 0 to 1 (full cloud cover to clear sky). $K_0$ can be approximated by dividing measured by potential downwelling solar radiation, but only during daytime. We replaced all nocturnal $K_0$ values where $RG_d$ was below 50 W m$^{-2}$ with linearly interpolated values. While clear sky $LW_d$ can be reasonably estimated the above formulation for all-sky $LW_d$ is a rough fix in need of some data. Since cloud emissivity depends on, e.g., cloud type, water content and cloud vertical extent an uncertainty of roughly 5–20 W m$^{-2}$ is introduced to the driver data set [Gabathuler et al., 2001] by using this algorithm.

[20] The consistently gap-filled meteorological forcing data from the above 15 sites (and from around 50 additional sites) are available from the authors (upon request also as ALMA-compliant NetCDF files).

### 2.2.2. Validation Data

[21] Turbulent surface fluxes and soil physical state variables from the Level 2 flux tower data sets were used to validate the model during the implementation stage of above-described modifications. None of the validation data were gap-filled since our intention was to look at timing and phase of the seasonal fluxes in response to climatic forcings rather than to match the local-scale heat, water and carbon balance at the end of the year. LE (latent heat flux; W m$^{-2}$), H (sensible heat flux; W m$^{-2}$), and NEE (net ecosystem exchange; $\mu$mol m$^{-2}$ s$^{-1}$) were $u^*$ filtered in order to account for the well documented biases in eddy covariance measurements during periods of low turbulence [Schmid et al., 2003]: comparisons to modeled fluxes were only performed for times when the $u^*$ value was above 0.2 m s$^{-1}$ (in the mean 67.4% of the data). Ideally, the $u^*$ threshold should be site-dependent and would only need to be applied to nocturnal data. Random uncertainties in turbulent surface fluxes [Hollinger and Richardson, 2005] were estimated based on empirical findings by Richardson et al. [2006]. Systematic errors in measured surface fluxes due to failure in energy balance closure were accounted for by multiplying $u^*$-screened surface fluxes with the residual of the energy balance closure (as % of $R_n$) for each site, which was calculated from the regression of hourly observed $R_n$ versus LE and H fluxes [Wilson et al., 2002; Grunwald and Bernhofer, 2007]. In all LE and H plots, the total errors were calculated as the square root of the sum of the squares of random and systematic errors for each analysis time step (e.g., hourly or monthly). Table 2 presents a summary of these uncertainty estimates for each site.

[22] The model in its standard configuration simulates GPP (gross primary productivity; $\mu$mol m$^{-2}$ s$^{-1}$) but not $R_e$ (ecosystem respiration; $\mu$mol m$^{-2}$ s$^{-1}$). In order to calculate NEE (net ecosystem exchange; $\mu$mol m$^{-2}$ s$^{-1}$), which is the difference between two large terms GPP and $R_e$, both terms would need to be accurately prognosed. This requires a mechanistic formulation involving prognostic carbon and nitrogen fluxes and pools as for instance presented by Thornton et al. [2007]. In order to compare modeled carbon uptake to observations, observed estimates of GPP were empirically derived from observed NEE, PAR (photosynthetically active radiation; $\mu$mol m$^{-2}$ s$^{-1}$), and $T_s$ (5 cm soil
temperature; K) using the algorithms by Desai et al. [2005].

Measured volumetric soil moisture was converted to percent saturation by assuming a porosity of 0.48 and by using the model soil layer which was closest to observation depth.

2.3. Experiment

Single point model simulations were performed for each of the 15 flux tower sites. The original model (CLM3.0) was successively modified with the proposed changes:

1. CLM3.0: the original and publicly available release code of CLM3.0.
2. CLMgw: addition of a Topmodel-based infiltration, runoff and aquifer storage formulation to CLM3.0. CLMgw (gw stands for groundwater) further includes all other major updates in the model (e.g., a new canopy integration scheme, canopy interception changes, new frozen soil and plant soil water availability parameterizations) as described in Oleson et al. [2007] which were not part of the original CLM3.0.
3. CLMgw_rsoil: addition of the bare soil evaporation resistance formulation to CLMgw (rsoil stands for soil resistance).
4. CLM3.5: addition of a PFT-dependent nitrogen limitation factor to CLMgw_rsoil. This simulation is equivalent to the public release code of CLM3.5.

2.3.1. Boundary Conditions

Vegetation and soil parameters for each site were derived from the standard CLM3.0 PFT-dependent look-up tables based on vegetation type and soil type (constant vertical profiles of sand/clay fractions for each site) from Table 1. A single PFT was used for each site. Visible and

Table 2. Uncertainty of Observations: % of u* Filtered Data (u*), % of Energy Balance Closure (ebc), Mean Error in LE Due to Failure of Energy Balance Closure (ebc LE), Mean Random Error in LE (ran LE), Mean Error in H Due to Failure of Energy Balance Closure (ebc H), and Mean Random Error in H (ran H)

<table>
<thead>
<tr>
<th>Number</th>
<th>Site</th>
<th>u* %</th>
<th>ebc %</th>
<th>ebc LE W m^-2</th>
<th>ran LE W m^-2</th>
<th>ebc H W m^-2</th>
<th>ran H W m^-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vielsalm</td>
<td>21</td>
<td>73</td>
<td>7.0</td>
<td>20.2</td>
<td>7.5</td>
<td>29.7</td>
</tr>
<tr>
<td>2</td>
<td>Tharandt</td>
<td>44</td>
<td>78</td>
<td>8.4</td>
<td>24.5</td>
<td>7.7</td>
<td>29.8</td>
</tr>
<tr>
<td>3</td>
<td>Castel Porziano</td>
<td>34</td>
<td>83</td>
<td>8.6</td>
<td>25.7</td>
<td>13.9</td>
<td>36.6</td>
</tr>
<tr>
<td>4</td>
<td>Collelongo</td>
<td>41</td>
<td>81</td>
<td>8.8</td>
<td>28.1</td>
<td>12.3</td>
<td>37.1</td>
</tr>
<tr>
<td>5</td>
<td>Kaamanen</td>
<td>56</td>
<td>72</td>
<td>10.4</td>
<td>23.9</td>
<td>3.4</td>
<td>22.9</td>
</tr>
<tr>
<td>6</td>
<td>Hyytiäla</td>
<td>47</td>
<td>72</td>
<td>8.3</td>
<td>21.5</td>
<td>6.2</td>
<td>26.1</td>
</tr>
<tr>
<td>7</td>
<td>El Saler</td>
<td>24</td>
<td>83</td>
<td>7.8</td>
<td>25.6</td>
<td>12.3</td>
<td>39.1</td>
</tr>
<tr>
<td>8</td>
<td>Santarem KM83</td>
<td>47</td>
<td>81</td>
<td>32.7</td>
<td>59.2</td>
<td>8.6</td>
<td>27.7</td>
</tr>
<tr>
<td>9</td>
<td>Tapajos KM67</td>
<td>36</td>
<td>81</td>
<td>24.9</td>
<td>44.9</td>
<td>6.5</td>
<td>26.3</td>
</tr>
<tr>
<td>10</td>
<td>Morgan Monroe</td>
<td>22</td>
<td>65</td>
<td>18.8</td>
<td>28.2</td>
<td>11.0</td>
<td>29.5</td>
</tr>
<tr>
<td>11</td>
<td>Boreas OBS</td>
<td>25</td>
<td>80</td>
<td>5.8</td>
<td>20.7</td>
<td>10.4</td>
<td>31.1</td>
</tr>
<tr>
<td>12</td>
<td>Lethbridge</td>
<td>42</td>
<td>77</td>
<td>7.3</td>
<td>21.8</td>
<td>8.8</td>
<td>34.3</td>
</tr>
<tr>
<td>13</td>
<td>Fort Peck</td>
<td>42</td>
<td>68</td>
<td>13.6</td>
<td>24.6</td>
<td>13.2</td>
<td>30.3</td>
</tr>
<tr>
<td>14</td>
<td>Harvard Forest</td>
<td>17</td>
<td>84</td>
<td>6.9</td>
<td>24.9</td>
<td>6.4</td>
<td>31.1</td>
</tr>
<tr>
<td>15</td>
<td>Niwot Ridge</td>
<td>17</td>
<td>76</td>
<td>12.7</td>
<td>27.6</td>
<td>12.3</td>
<td>40.7</td>
</tr>
</tbody>
</table>

Figure 1. Performance of four model versions at 15 FLUXNET towers (numbers 1–15). Statistics in the Taylor diagram are derived from hourly simulated and observed LE and H fluxes. Legend: CLM3.0: red asterisks; CLMgw: green crosses; CLMgw_rsoil: cyan diamonds; CLM3.5: violet triangles. In CLM3.0 H is off-scale for the two tropical sites 8 and 9 (and therefore not shown).
Table 3. Performance of Simulated LE and H Fluxes in Four CLM Versions (3.0, gw, gw_rsoil, 3.5): R and RMSE (W m$^{-2}$, in Brackets) are Diagnosed on Hourly and Monthly Timescales

<table>
<thead>
<tr>
<th>Number</th>
<th>Site</th>
<th>3.0</th>
<th>gw</th>
<th>gw_rsoil</th>
<th>3.5</th>
<th>LE</th>
<th>3.0</th>
<th>gw</th>
<th>gw_rsoil</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hourly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Vielsalm</td>
<td>0.68(43.0)</td>
<td>0.84(57.1)</td>
<td>0.88(46.1)</td>
<td><strong>0.90</strong> (35.9)</td>
<td>0.79(67.0)</td>
<td>0.71(67.1)</td>
<td>0.74(63.9)</td>
<td><strong>0.83</strong> (54.2)</td>
</tr>
<tr>
<td>2</td>
<td>Tharandt</td>
<td>0.65(44.3)</td>
<td>0.74(51.4)</td>
<td><strong>0.79</strong> (37.2)</td>
<td><strong>0.79</strong> (35.2)</td>
<td>0.80(75.3)</td>
<td>0.75(75.3)</td>
<td>0.82(60.8)</td>
<td><strong>0.86</strong> (55.4)</td>
</tr>
<tr>
<td>3</td>
<td>Castel Porziano</td>
<td>0.23(89.7)</td>
<td>0.46(86.8)</td>
<td><strong>0.73</strong> (66.5)</td>
<td>0.70(61.6)</td>
<td>0.87(102.1)</td>
<td>0.86(93.3)</td>
<td>0.86(78.4)</td>
<td><strong>0.89</strong> (68.8)</td>
</tr>
<tr>
<td>4</td>
<td>Collelongo</td>
<td>0.57(70.7)</td>
<td>0.76(87.2)</td>
<td><strong>0.81</strong> (67.7)</td>
<td><strong>0.81</strong> (55.3)</td>
<td>0.72(91.2)</td>
<td>0.66(109.0)</td>
<td>0.73(97.7)</td>
<td><strong>0.81</strong> (83.0)</td>
</tr>
<tr>
<td>5</td>
<td>Kaamanen</td>
<td>0.71(40.6)</td>
<td>0.81(40.2)</td>
<td>0.86(27.2)</td>
<td><strong>0.88</strong> (25.2)</td>
<td>0.76(37.9)</td>
<td>0.70(39.0)</td>
<td>0.72(40.2)</td>
<td><strong>0.74</strong> (39.3)</td>
</tr>
<tr>
<td>6</td>
<td>Hyytiälä</td>
<td>0.65(39.6)</td>
<td>0.80(32.5)</td>
<td>0.83(28.7)</td>
<td><strong>0.84</strong> (28.3)</td>
<td>0.81(59.1)</td>
<td>0.82(50.9)</td>
<td>0.82(52.5)</td>
<td><strong>0.87</strong> (47.6)</td>
</tr>
<tr>
<td>7</td>
<td>El Saler</td>
<td>0.42(68.4)</td>
<td>0.53(65.4)</td>
<td>0.66(50.2)</td>
<td><strong>0.67</strong> (49.5)</td>
<td>0.89(95.7)</td>
<td>0.88(87.7)</td>
<td>0.89(81.1)</td>
<td><strong>0.91</strong> (75.4)</td>
</tr>
<tr>
<td>8</td>
<td>Santarejos KM83</td>
<td>0.52(157.8)</td>
<td>0.74(135.2)</td>
<td><strong>0.77</strong> (127.1)</td>
<td><strong>0.76</strong> (111.9)</td>
<td>0.59(166.1)</td>
<td>0.41(125.1)</td>
<td>0.39(118.0)</td>
<td><strong>0.68</strong> (92.0)</td>
</tr>
<tr>
<td>9</td>
<td>Tapajos KM67</td>
<td>0.47(147.2)</td>
<td>0.78(132.6)</td>
<td>0.78(131.0)</td>
<td><strong>0.81</strong> (100.2)</td>
<td><strong>0.45</strong> (146.7)</td>
<td>0.03(120.7)</td>
<td>-0.02(120.5)</td>
<td><strong>0.45</strong> (83.1)</td>
</tr>
<tr>
<td>10</td>
<td>Morgan Monroe</td>
<td>0.55(89.8)</td>
<td>0.66(102.5)</td>
<td>0.78(76.5)</td>
<td><strong>0.85</strong> (58.5)</td>
<td>0.56(111.7)</td>
<td>0.53(98.7)</td>
<td>0.57(88.7)</td>
<td><strong>0.73</strong> (75.1)</td>
</tr>
<tr>
<td>11</td>
<td>Boreas OBS</td>
<td>0.41(49.6)</td>
<td>0.53(45.7)</td>
<td>0.69(35.6)</td>
<td><strong>0.79</strong> (41.2)</td>
<td>0.84(68.4)</td>
<td>0.86(65.1)</td>
<td>0.85(63.7)</td>
<td><strong>0.83</strong> (70.8)</td>
</tr>
<tr>
<td>12</td>
<td>Lethbridge</td>
<td>0.48(53.7)</td>
<td>0.50(56.4)</td>
<td>0.72(40.1)</td>
<td><strong>0.81</strong> (32.5)</td>
<td>0.74(82.1)</td>
<td>0.75(81.5)</td>
<td>0.75(78.7)</td>
<td><strong>0.76</strong> (74.3)</td>
</tr>
<tr>
<td>13</td>
<td>Fort Peck</td>
<td>0.67(60.6)</td>
<td>0.74(58.9)</td>
<td>0.80(48.0)</td>
<td><strong>0.79</strong> (47.5)</td>
<td><strong>0.71</strong> (63.3)</td>
<td>0.66(69.1)</td>
<td>0.60(72.9)</td>
<td><strong>0.71</strong> (63.2)</td>
</tr>
<tr>
<td>14</td>
<td>Harvard Forest</td>
<td>0.67(62.8)</td>
<td>0.78(68.8)</td>
<td>0.86(46.6)</td>
<td><strong>0.89</strong> (37.5)</td>
<td>0.59(96.7)</td>
<td>0.51(104.5)</td>
<td>0.67(88.0)</td>
<td><strong>0.79</strong> (73.8)</td>
</tr>
<tr>
<td>15</td>
<td>Niwot Ridge</td>
<td>0.49(66.1)</td>
<td>0.61(63.0)</td>
<td><strong>0.73</strong> (46.4)</td>
<td>0.71(47.6)</td>
<td>0.84(96.7)</td>
<td>0.85(85.9)</td>
<td>0.86(78.9)</td>
<td><strong>0.88</strong> (72.5)</td>
</tr>
</tbody>
</table>

*Bold numbers show the best of the four model versions for each diagnostic and site.*

near-infrared soil albedos were set to arbitrary values of 0.18/0.36 for a dry top soil and 0.09/0.18 for a saturated top soil due to a lack of in-situ information at most sites. Stem Area Index was set to 0.08. Vegetation top/bottom heights were 35 m/1 m for tropical forests, 20 m/10 m for other forests, and 1 m/0.1 m for short vegetation. A climatological monthly Leaf Area Index for each site came from the 1982–2001 EFAI-NDVI data set [Stöckli and Vidale, 2004]. Since our intent was to perform a process-based analysis of a global model, PFT-dependent model parameters were not tuned to site-specific and species-specific conditions.

### 2.3.2. Initial Conditions and Spin-Up

[25] The model was initialized from its standard arbitrary initial conditions of 283 K vegetation, ground and soil temperatures, 30% (CLM3.0) 40% (CLMgw, CLMgw_rsoil, CLM3.5) volumetric soil water content and with empty ground snow and canopy interception water stores. Spin-up was achieved by repeating the full range of available years five times for each site (five spin-up cycles). Mean yearly latent and sensible heat fluxes were within 0.1 W m$^{-2}$ of those from the previous spin-up cycle after a single spin-up cycle (similar to PILPS 2a spin-up criteria [Chen et al., 1997]). More arid climates would need longer spin-up times since the water table there takes longer to adjust (see, e.g., the global simulations by Oleson et al. [2008]). Nevertheless, surface fluxes are not affected by variations of a very deep water table in such areas.

#### 2.3.3. Analysis

[30] Hourly model output from the last spin-up cycle was used for the analysis. At sites where 30 min measurements were available they were averaged to 60 min values.

### 3. Results

[31] Comparisons between modeled and observed LE and H in Figure 1 (Taylor Diagram) and Table 3 (R and RMSE) provide a quick overview of performance changes across model versions: the original CLM3.0; modifications using a groundwater scheme (CLMgw), addition of a bare soil evaporation resistance (CLMgw_rsoil) and further addition of PFT-dependent nitrogen limitation factor in the final model version (CLM3.5). (In the Taylor diagram [Taylor, 2001], four statistical quantities are geometrically connected: the correlation coefficient $R$, standard deviation of observations $\sigma_o$, standard deviation of the model $\sigma_m$, and the centered pattern root-mean-square error $E'$. The polar axis displays $R$ and the radial axes display the standard deviation of the modeled variable divided by the standard deviation of the observed variable $\sigma_o/\sigma$. The geometric relationship of this diagram is such that the distance between the 1.0 value of the X-axis and the plotted value show $E'$ and thus is a measure for the absolute model error. Root-mean-square error $E$ is given by $E = E' + E''$, where $E''$ is the mean bias. The four statistical moments are connected...
by: \( E^2 = \sigma^2_R + \sigma^2_e - 2\sigma_R \sigma_e R \). General changes in LE and H on the hourly and monthly timescale covering all sites are discussed first, followed by a close inspection of results at individual sites encompassing temperate, north boreal, mediterranean and the tropical ecosystems.

3.1 Latent Heat Flux

[32] The R of LE shown in Figure 1a increases with the new groundwater scheme (CLMgw, green crosses) compared to the original CLM3.0 (red asterisks). However, the variability of LE is now exaggerated compared to observations. The inclusion of the bare soil evaporation resistance (CLMgw_rsoil, cyan diamonds) generates more realistic LE variability, resulting in a higher R than with the groundwater formulation alone. A further improvement in both correlation and variability is achieved with the nitrogen limitation (CLM3.5, violet triangles). R for hourly LE at most sites increases from around 0.4–0.7 (CLM3.0) and 0.5–0.8 (CLMgw) to 0.7–0.9 (CLMgw_rsoil and CLM3.5). For all 15 sites hourly and monthly LE has a higher R and a lower RMSE for CLM3.5 compared to CLM3.0. Some sites display substantial improvements: e.g., the mediterranean site Castelporziano (R increases from 0.23 to 0.70 for hourly LE and from 0.21 to 0.81 for monthly LE) and the tropical site KM67 (R for hourly LE increases from 0.40 to 0.81) and for monthly LE from 0.03 to 0.68). Similarly, temperate ecosystems show a steady improvement (e.g., Velsalim or Morgan Monroe). High latitude ecosystems (e.g., Kaamanen and Hyttilä) are already well simulated by CLM3.0, but they also slightly improve. RMSE decreases at all sites (except for Velsalim at the monthly timescale) from CLM3.0 to CLM3.5 on both hourly and monthly timescale. The grassland sites Lethbridge and Fort Peck improve on both hourly and monthly timescale, but to a lesser extent than forest sites.

3.2 Sensible Heat Flux

[33] The changes in H, shown in Figure 1b, are not as easily generalized as LE, although the model changes appear to result in an overall improvement. Even though changes in LE are almost fully compensated by opposite changes in H, the new hydrology formulations do not affect R of H as much. This is due to a number of factors. First of all, H is smaller than LE for most sites. Secondly, while LE can completely be shut down by soil moisture, H is strongly coupled to net shortwave radiation through skin temperature, largely independent of the state of subsurface hydrology [Betts, 2004]. R is a good indicator for phase but not for magnitude in this case. For sites like Castelporziano, where the R of LE increased substantially, R of H remains constant (R increases from 0.87 to 0.89; hourly timescale). But RMSE of H decreases from 102.1 W m\(^{-2}\) to 68.8 W m\(^{-2}\).

Remaining high RMSE values should also be viewed with respect to uncertainties in observed fluxes (Table 2). This result suggests that the mean error and variability of H was improved with the new hydrology, and not the timing and phase of H. Indeed, in Figure 1b R values remain roughly the same for all four model versions. But the spread in the radial direction decreases and successively moves symbols closer towards observed variability at the 1.0 arc by use of the new formulations. Several sites actually show a slightly worse R with the new hydrology, but they still have a decreased RMSE compared to CLM3.0. For the two tropical sites hourly R for H becomes worse in CLMgw and CLMgw_rsoil and increases again with CLM3.5. Hourly and monthly RMSE for those sites significantly decreases by around 35–45%. Only small changes in R and RMSE on both hourly and monthly timescale can be seen for the two grasslands Fort Peck and Lethbridge. They cannot make use of the groundwater if the water table falls below their shallow rooting depth, which is most likely the case at those two sites.

3.3 Temperate

[34] Morgan Monroe State Forest is a deciduous temperate broadleaf forest in Indiana (USA). Monthly mean measured LE (Figure 2c; black) displays a clear seasonal cycle with a growing season between May and October. H (Figure 2d; black) peaks before leaf emergence in March and April [Schmirit al., 2003].

[35] CLM3.0 shows excess LE in winter and too low LE in summer. This results in a too low modeled variability of hourly LE and exaggerated variability of hourly H (Figure 2b; red). The modeled root zone soil moisture is low throughout the year compared to observed soil moisture (Figure 2a; black and red). The simulated soil moisture profile (Figure 3a; CLM3.0) provides insight into the processes responsible for these results: an impermeable and dry soil layer is formed after a few years of spin-up and inhibits further infiltration and water storage at deeper soil moisture levels. The main reason for this effect is found in the delicate interplay between soil physics and the numerical solution of the vertical soil water transfer. As discussed in Stöckli et al. [2007], the exponential relationship between soil hydraulic conductivity and soil water content in a finite difference numerical solution of Darcy’s equation can create a feedback below certain soil water content which successively decouples upper from lower soil layers through further inhibition of infiltration. We can see in Figure 3a that this “vicious loop” cannot be broken even by long precipitation events during spring.

[36] It was chosen to improve the physical and biophysical processes in order to support a stable numerical solution of soil water dynamics as documented in Oleseon et al. [2007]: a Topmodel-based surface and subsurface runoff scheme [Niu et al., 2005] coupled to a prognostic groundwater scheme [Niu et al., 2007] are mechanistic formulations of soil water dynamics which were not in the original CLM3.0. The new groundwater scheme (Figure 3b; CLMgw) increases soil moisture to a range where the numerical solution provides a more stable interaction between infiltration and seasonal water storage: there are no more dry impermeable soil layers. Seasonal LE and H fluxes in CLMgw have a more realistic variability (Figure 2b; green crosses) and R for hourly LE rises from 0.55 in CLM3.0 to 0.66 in CLMgw. It stays nearly constant for H (0.56 to 0.53). Although soil moisture in the upper 30 cm does not significantly increase (Figure 2a; green line), the new scheme has increased summer LE (Figure 2f; red and green lines) due to a higher soil moisture availability in lower depths (Figure 3b; CLMgw). But it also has increased off-season LE for the same reason (Figure 2e; red and green lines).
Figure 2. Model diagnostics at a temperate deciduous forest (Morgan Monroe State Forest, USA) during 2003: (a) soil moisture relative to saturation at 30 cm depth; (b) Taylor diagram with hourly statistics of LE and H fluxes; (c) monthly LE fluxes; (d) monthly H fluxes; (e) diurnal cycle of LE fluxes in February; (f) diurnal cycle of LE fluxes in August. Error bars show estimated uncertainties of observed turbulent fluxes. Legend: observations: black plus signs; CLM3.0: red asterisks; CLMgw: green crosses; CLMgw_rsoil: cyan diamonds; CLM3.5: violet triangles.
The implementation of a more realistic soil water treatment reveals a deficiency of the model: given enough soil water and low leaf-coverage during off-season periods, CLMgw simulates excessive bare soil evaporation compared to observations. The same problem was present in CLM3.0 but it was mostly hidden by the generally dry soil conditions. Addition of the empirically derived bare soil resistance [Sellers et al., 1992] offers a constraint for bare soil evaporation fluxes during periods of low leaf coverage in deciduous forests.

It leads to a significant improvement of the simulated terrestrial water cycle in CLMgw_rsoil (Figure 3c): the water


Table 1. Subsurface flow components and parameters used in the model. The time step of 30 min is used to calculate the subsurface flow components. The subscript *A* denotes the area-weighted average of the components, while *I*, *O* and *L* denote the components of the input, output and loss terms, respectively. The units are m s⁻¹, mm s⁻¹, mm and m³ s⁻¹.

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
<th>CLMgw_rsoil</th>
<th>CLM3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input terms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>I_Surf</em></td>
<td>Surface runoff</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><em>I_Drain</em></td>
<td>Drainage runoff</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><em>I_Soil</em></td>
<td>Soil water drainage</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Output terms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O_Surf</em></td>
<td>Surface runoff</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><em>O_Drain</em></td>
<td>Drainage runoff</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><em>O_Soil</em></td>
<td>Soil water drainage</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Loss terms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>L_Surf</em></td>
<td>Surface runoff</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><em>L_Drain</em></td>
<td>Drainage runoff</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><em>L_Soil</em></td>
<td>Soil water drainage</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**Note:** The values of 0.00 indicate that the corresponding components are not active in the model.
Figure 4. Model diagnostics at a north boreal tundra ecosystem (Kaamanen, Finland) during 2004: (a) monthly LE fluxes; (b) monthly H fluxes; (c) soil temperature at 30 cm; (d) terrestrial water storage; (e) accumulated surface runoff; (f) accumulated total runoff. Error bars show estimated uncertainties of observed turbulent fluxes. Legend: observations: black plus signs; CLM3.0: red asterisks; CLMgw: green crosses; CLMgw_rsoil: cyan diamonds; CLM3.5: violet triangles.
Figure 5. Model diagnostics at a Mediterranean hardwood forest (Castel Porziano, Italy) for 2003: (a) soil moisture at 30 cm depth; (b) Taylor diagram showing statistics from hourly LE and H fluxes; (c) monthly LE fluxes; (d) monthly H fluxes; (e) terrestrial water storage; (f) modeled versus NEE-derived GPP. Error bars show estimated uncertainties of observed turbulent fluxes. Legend: observations: black plus signs; CLM3.0: red asterisks; CLMgw: green crosses; CLMgw_soil: cyan diamonds; CLM3.5: violet triangles.
Castelporziano is dominated by H. In comparison to a more humid ecosystem, improving LE at a dry ecosystem only has a small effect on the diurnal course of H. However, as can be seen in Figure 5b, a better simulation of LE can shift the absolute magnitude and therefore seasonal variability of H towards observed values.

[45] CLM3.0 simulates a low magnitude and damped seasonal course of soil moisture compared to observed soil moisture at 30 cm depth (Figure 5a). It coincides with a low simulated TWS magnitude (the range between the minimum and the maximum TWS during a year) of around 60 mm (Figure 5e). Concurrently LE is almost completely absent in the summer months from May to August (Figure 5c), resulting in a too high H during this time period (Figure 5d). This result explains the exaggerated variability and low correlation of CLM3.0 and CLMgw displayed in Figure 5b. While the addition of groundwater storage (CLMgw) rises TWS magnitude to 120 mm, it cannot overcome the unrealistic drought stress during summer months. The bare soil resistance constrains off-season evaporation losses (Figure 5c; CLMgw_rsoil) and augments TWS magnitude to over 300 mm. The soil water storage capability of the groundwater scheme becomes effective when the soil model’s numerics and physics show a more stable interaction.

[46] The new hydrology of CLMgw_rsoil is able to supply the extensive water demand at this ecosystem during the dry summer 2003. A storage deficit of around 100 mm persists into the next year (Figure 5e). Although the off-season observed soil moisture levels correspond well to those modeled in CLMgw_rsoil, the model’s soil at 30 cm still dries out too much during summer. Deeper soil levels act as the large TWS buffer in this case. Reichstein et al. [2003] notes that the site’s vegetation has access to topographically induced groundwater (lateral groundwater recharge), which was not simulated here.

[47] Similarly to LE, modeled GPP (Figure 5f) becomes more realistic from CLM3.0 to CLMgw_rsoil during summer. But GPP and LE are now overestimated during other parts of the year. The new sun-shade canopy scheme implemented by Thornton and Zimmermann [2007] has a more realistic light interception parameterization for canopy-integrated photosynthesis but depends on the quantification of nitrogen as a controlling factor for this process. The standard model does not include nitrogen controls on photosynthesis. After soil hydrology is fixed in CLMgw_rsoil we now find that GPP is overestimated. PFT-dependent $V_{\text{max}}$ scaling factors $f(N)$ simulating nitrogen limitation are presented in Oleson et al. [2007] and applied in CLM3.5. As a result of the decreased light response (Figure 5f, violet triangles), GPP and LE slightly decrease during spring and autumn. However, this newly introduced formulation alone cannot account for the exaggerated fluxes. GPP (and to a lesser extent also LE) is still highly overestimated during the wet season.

3.6. Tropical

[48] The evergreen tropical broadleaf forest site KM83 south of Santarem (Brazil) represents a constant hot and humid climate [da Rocha et al., 2004]. 70% of the annual precipitation occur within the seven month long wet season from January to July.

[49] Figures 6c and 6d show that accumulated LE and H fluxes are simulated accurately during the wet season with CLM3.0, but the observed continuous increase in accumulated water flux throughout the dry season from August to December cannot be sustained, resulting in a very high Bowen ratio during this latter period. CLMgw, CLMgw_rsoil and CLM3.5 provide remedy for this deficiency: R for hourly LE steadily increases from 0.52 to 0.76 (Table 3). R for hourly H decreases from 0.59 to 0.41 for CLMgw and increases again to 0.68 for CLM3.5. The generally low H at this site (within the uncertainty range of observations) renders the correlation coefficient as an unsuitable measure for performance comparisons (this is even more evident at the other tropical site KM67). RMSE is a more robust measure. On the hourly timescale it decreases significantly from 166.1 W m$^{-2}$ to 92.0 W m$^{-2}$.

[50] Little difference is found between the aquifer water storage formulation only (CLMgw) and the use of an additional bare soil evaporation resistance formulation (CLMgw_rsoil). For instance, R for LE rises from 0.74 to 0.77 on the hourly timescale. Constant and high leaf coverage at this site provides a radiation-driven process for the control of excessive bare soil evaporation, so the addition of the missing resistance term is not critical for this evergreen tropical ecosystem.

[51] A comparison between modeled and measured soil moisture at 20 cm depth in Figure 6a does not provide much evidence for why dry season LE is enhanced in CLMgw_rsoil compared to CLM3.0; most of the model enhancements seem to influence lower soil depths. There were no soil moisture measurements reported for depths below 1 m. The modeled TWS cycle in Figure 6b provides insight into the relevant hydrological processes: while CLM3.0 has a very low TWS magnitude of less than 100 mm, CLMgw and CLMgw_rsoil push TWS magnitude to 400 mm. Seasonal soil water storage with such a high capacity is important for a tropical ecosystem since plant biophysical functioning in a seasonally dry climate depends on long-term soil moisture dynamics. This is supported by observational evidence: da Rocha et al. [2004] show that the Amazonian rainforest at KM83 can sustain transpiration throughout the dry season since it has access to deep soil water.

[52] As already shown for the mediterranean site, CLMgw_rsoil with the more realistic soil water cycle leads to overestimated LE and GPP. Including the parametric nitrogen limitation $f(N)$ in CLM3.5 results in a more realistic LE and H balance for both wet and dry season (Figures 6c and 6d). R for hourly LE remains roughly constant (0.76, compared to 0.77 for CLMgw_rsoil), but RMSE is reduced by around 15 W m$^{-2}$. On the other hand, R for hourly H significantly increases from 0.39 to 0.68 and RMSE is reduced by 26 W m$^{-2}$. GPP is still overestimated during the wet season. However, a slightly more realistic light response of GPP (Figures 6e and 6f) is achieved. In a high light environment such as the Amazon, stomatal conductance during daylight is mostly constrained by the maximum rate of carboxylation. The factor $f(N)$ has the largest absolute effects on GPP for these ecosystems. LE (and thus the surface energy partitioning) is influenced to a lesser extent, as LE is also controlled by the boundary layer.
Figure 6. Model diagnostics at a tropical evergreen forest (Santarem KM83, Brazil) during 2002: (a) soil moisture at 20 cm depth; (b) terrestrial water storage; (c) accumulated LE fluxes; (d) accumulated H fluxes; (e) modeled versus NEE-derived GPP; (f) mean light response curves for modeled and NEE-derived GPP (binned by incoming solar radiation). Error bars show estimated uncertainties of observed turbulent fluxes. Legend: observations: black plus signs; CLM3.0: red asterisks; CLMgw: green crosses; CLMgw_rsoil: cyan diamonds; CLM3.5: violet triangles.
vapor pressure gradient, bare soil evaporation and aerodynamic properties.

4. Discussion

[53] Turbulent heat and water fluxes of the original CLM3.0 show significant biases in tropical, mediterranean and temperate climatic environments. These biases result from a poor representation of soil moisture storage and its interaction with seasonal variations of the surface climate. Modeled plant transpiration generally shuts down during either summer or dry seasons due to a lack of soil moisture supply. Observations from the 15 flux tower sites, however, indicate that plants can sustain their physiological function during seasonal-scale and longer term drought periods. Subsurface hydrological processes on which these plants largely depend therefore need to be properly represented in land surface models in order to simulate the terrestrial carbon and water cycle [Reichstein et al., 2002]. This requirement gains further importance in view of the predicted temperature and precipitation changes in future climatic scenarios, which could severely affect ecosystem function during hotter and drier summer periods [Seneviratne et al., 2006].

4.1. Terrestrial Water Storage

[54] To achieve a higher water storage capacity in a land surface model, the total soil depth and other soil parameters are often modified as a first guess. The above findings, however, suggest that soil water storage capacity is a dynamic quantity. It does not primarily depend on soil physical parameters. It rather results from a consistent interplay between the soil and vegetation biophysical parameterizations on one side and the soil numerical scheme on the other side: they both depend on each other in order to provide a realistic simulation of the terrestrial water cycle.

[55] At the mediterranean and temperate sites only small improvements in surface fluxes result from the implementation of larger soil water storage capacity by use of a prognostic aquifer scheme. Soil water infiltration and storage are both still largely inhibited by excessive bare soil evaporation during off-season periods in those ecosystems. Further addition of a bare soil evaporation resistance finally results in a realistic TWS magnitude and concurrently in a substantial decrease of RMSE. Figures 3a–3c illustrate the underlying soil hydrological processes:

[56] 1. A dry soil can continuously inhibit vertical soil moisture fluxes and thus decrease seasonal water storage by hydrologically decoupling upper from lower soil layers.

[57] 2. Extending the storage pool by implementing a prognostic aquifer breaks the infiltration barrier by providing ample soil moisture to the root zone but TWS remains at a low seasonal magnitude (e.g., Figure 5e).

[58] 3. Bare soil evaporation during off-season periods was identified as the main process which dampens TWS magnitude for deciduous vegetation in temperate and mediterranean climate zones. With a more realistic off-season bare soil evaporation TWS becomes positive during the winter or wet season when moisture is stored in the soil. As a consequence transpiration fluxes during months of low rainfall (dry season) or large atmospheric demands (summer season) substantially improve.

[59] While a prognostic aquifer model [Niu et al., 2005, 2007] provides the physical framework for simulating large seasonal TWS fluctuations, the size of TWS magnitude depends on a dynamically varying set of involved soil and vegetation processes. The new hydrological formulations enhance TWS by 200–300 mm compared to the original CLM3.0, with quite beneficial effects for the simulated surface energy and water balances in seasonally dry climates. This result is highly consistent with comparisons between modeled and GRACE estimates of TWS at catchment scale presented by Oleson et al. [2008]. They show that CLM3.5 enhances TWS magnitude by 50–300 mm compared to CLM3.0, with improved correlations and substantial decreases in RMSE.

[60] In northern boreal regions like Kaamanen, however, TWS magnitude decreases by around 100 mm when groundwater storage is added (Figure 4e). This behavior is opposite to what one would expect. As in warm climates, soil water storage function of cold climates not only depends on storage capacity, but closely interacts with the dominant hydrological processes through time-delayed feedbacks: the analysis shows that snowmelt water can be stored in spring after soil thaw and should not completely run off into rivers like in the original formulation. Soil moisture storage seems to dampen the seasonal course of TWS at Kaamanen. While adding groundwater does not much affect turbulent surface fluxes in cold climates (Figures 4a and 4b) it could lead to improvements in high latitude runoff timing and magnitude (Figures 4e and 4f). This is documented in Oleson et al. [2008] by comparison of global simulated versus observed river discharge and runoff.

4.2. Nitrogen Limitation

[61] Results from the mediterranean and tropical sites suggest that the enhanced and more realistic water storage processes in the model can lead to excessive transpiration. The addition of a parameterized nitrogen control for photosynthesis decreases light sensitivity of stomatal opening as expected (Figures 5f, 6e, and 6f). The need for this parameterization only became evident after the new soil hydrology and the new canopy integration scheme was implemented: maximum photosynthesis rates in CLM3.0 were fixed, based on observed values. Low soil moisture levels furthermore limited the plant physiological activity in most climates. Parameterized nitrogen control became a necessity with the new hydrological modifications. While nitrogen is an important controlling factor for most terrestrial ecosystems ($J_N$ ranging from 0.60–0.84 in Oleson et al. [2007]), our results suggest that it mostly affects the surface energy and water balance in environments with high GPP. Tropical broadleaf forests have the lowest diagnosed nitrogen limitations among the 16 PFTs (highest $J_N = 0.84$). However, they mostly operate at high light levels, resulting in the largest nitrogen-controlled decreases in GPP in absolute terms. In comparison to GPP, LE is less sensitive to changes in stomatal conductance through nitrogen control because LE is a composite of transpiration and bare soil evaporation. The latter is independent of nitrogen availability. LE is further controlled by boundary layer aerodynamical
resistances, which only indirectly and weakly influence GPP. Nevertheless, results show a positive effect of the newly introduced \( f(N) \) on \( R \) and RMSE of hourly and monthly LE and \( H \) fluxes. The two tropical sites KM67 and KM83 show the largest decrease in RMSE (Table 3) by including \( f(N) \) compared to simulations with changes in soil hydrology alone. Boundary layer processes for these ecosystems are expected to benefit from this model enhancement in coupled simulations.

4.3. Open Questions

[62] Figures 5f and 6f show that yet another process might be missing. LE and GPP are still overestimated during the wet season at the Mediterranean and the tropical site. LE is less of a problem than GPP since LE is also driven by the atmospheric vapor pressure gradient and surface layer aerodynamics. It furthermore is composited from plant transpiration and bare soil evaporation, the latter being unrelated to stomatal functioning. Suggestions for missing processes are, e.g., a prognostic dry season phenomenology (which can also vary in tropical ecosystems [Myneni et al., 2007]) and dynamic allocation of leaf structure and photosynthates [Dickinson et al., 2002], which are not simulated in the standard CLM. Simulations for Mediterranean and tropical FLUXNET sites employing CLM3.5 with its full biogeochemistry scheme [Thornton et al., 2007] could shed some light into these open questions. Figures 2c and 2d show that CLM3.5 still cannot represent the \( H \) peak during March and April just before leaf emergence in temperate forests. This problem is common to many land surface models and might be related to model deficiencies in either phenology (too early leaf emergence) or surface litter cover (too much bare soil evaporation) and should be addressed in future studies.

5. Conclusion

[63] The Community Land Model version 3 includes mechanistic representations of terrestrial radiation, heat, water and carbon exchange processes, which have been developed from laboratory experiments and field studies. Deficiencies in the CLM3.0 soil hydrology have been revealed from long-term climate simulations, with sometimes negative effects on surface climate and plant biogeography. In this study new algorithms for removing these deficiencies were tested in off-line simulations at 15 FLUXNET tower sites.

[64] 1. The prognostic aquifer scheme [Niu et al., 2007] extends the soil storage pool of CLM3.0, but this enhancement only becomes effective when bare soil evaporation is curtailed by the application of an empirical bare soil resistance term [Sellers et al., 1992]. Soil water storage in models like CLM strongly depends on the interplay between soil numerics (nonlinear state-parameter dependence) and terrestrial biophysics. In this case excessive off-season bare soil evaporation in deciduous ecosystems inhibited groundwater storage by successively reducing long term soil moisture levels below a threshold at which hydraulic conductivity allows for vertical water transfer in the finite difference soil water scheme.

[65] 2. As a consequence of these two enhancements, CLM3.5 now includes a more dynamic soil water storage capacity: TWS magnitude increases in tropical, Mediterranean and temperate climates and decreases in cold climates. This result was mainly achieved by introduction of mechanistic hydrological processes and neither by extending the soil depth nor by modifying soil hydraulic parameters. In support of this conclusion [Gulden et al., 2007] find that a model with a prognostic aquifer is less sensitive to the largely unknown and spatially variable set of soil hydraulic parameters compared to a model with a deep soil alone. The uncertainty in the prescription of soil physical parameters in land surface models should therefore be mitigated by use of more mechanistic formulations for soil water storage. Furthermore this result justifies and facilitates comparisons between tower sites with similar vegetation but different soils.

[66] 3. Nitrogen control of photosynthesis (and therefore stomatal opening and transpiration) is needed in order to correctly partition energy into turbulent heat and water fluxes in environments with high GPP. This missing process was only uncovered after soil hydrological modifications led to a better simulated subsurface water balance and the new canopy integration scheme created a more realistic light response of photosynthesis. The original CLM3.0 was providing the right results for the wrong reasons: stomates in tropical and Mediterranean ecosystems were seasonally closing due to missing water supply while observations indicate that photosynthesis in these ecosystems is not so sensitive to drought effects.

[67] 4. Despite above improvements CLM3.5 still overestimates GPP during the wet season in Mediterranean and tropical ecosystems. Although the surface energy partitioning is less sensitive to stomatal response than GPP, we hypothesize that drought phenology or biogeochemical feedbacks involving the full terrestrial carbon-nitrogen cycle could be responsible for these differences. Local-scale and species-specific soil and vegetation properties and furthermore the general underestimation of eddy covariance fluxes might explain some differences between observed and modeled turbulent fluxes [Wilson et al., 2002; Foken et al., 2006]. The steady reduction of RMSE into the range of observation uncertainty (or below; e.g., for monthly fluxes at Kaamanen: RMSE = 10–20 \( \text{W m}^{-2} \)) in boreal, northern boreal and temperate climates as a result of the new mechanistic formulations is a strong indicator for the success of CLM’s new hydrology. In seasonally dry and tropical climates most uncertainty may still be on the model’s side, since monthly RMSE ranges between 30–50 \( \text{W m}^{-2} \), which is larger than estimated errors in observations.

[68] 5. A land surface model should as a first step include a realistic set of mechanistic formulations, which was the focus of this study. This leads to a better understanding of the role of ecophysiological drivers such as water, light and nitrogen in controlling photosynthesis at a range of ecosystems, and it thus helps to either support or invalidate some of our above hypotheses. It further makes the model suitable for global predictive applications across a range of spatial and timescales. As noted by Abramowitz [2005], there are, however, still considerable opportunities for improvements in such models. In a second step, the many empirical model parameters should be constrained in order to further reduce model uncertainty. The currently developed standardized,
gap-filled and bias-corrected FLUXNET Synthesis data set involving more than 200 tower sites (D. Papale and M. Reichstein, personal communication, 2007) provides a global set of observations suitable for a data assimilation exercise aimed at reducing parameter uncertainty in land surface models.

While at first the small number of 15 FLUXNET towers seems to be inappropriate for testing a globally applicable land surface model, we demonstrate that focusing on only four sites already effectively helps to identify and correct for major missing soil hydrological and vegetation biophysical processes in the model. As already shown by Stöckli and Vidale [2005], such a modeling framework with offline simulations allows for computationally inexpensive research and development of land surface models. FLUXNET provides valuable observations of quantities at timescales which are relevant in climate simulations. Despite lacking global coverage, FLUXNET statistically inherits the whole global set of ecosystems and climate zones. Although individual sites differ in absolute magnitude and timing of heat, water and carbon fluxes, they show similar patterns for sites within certain ecosystem and climate zones. Similarly, model deficiencies become visible as consistent patterns of time and phase shifts on diurnal and seasonal timescales across a number of sites, which was demonstrated here. While this study explored hourly-seasonal terrestrial processes, there is an increased number of FLUXNET sites with 10 years or longer coverage which allow a similar analysis for the interannual timescale.

References

Chen, T. H., et al. (1997), Cabauw experimental results from the project for intercomparison of land-surface parameterization schemes, J. Clim., 10, 1194–1215.
Friedlingstein, P., et al. (2006), Climate-carbon cycle feedback analysis: Results from the (CIMP-M interacts model intercomparison, J. Clim., 19), 3337–3353.


Sellers, P. J., M. D. Heiser, and F. G. Hall (1992), Relations between surface conductance and spectral vegetation indices at intermediate (100 m² to 15 km²) length scales, J. Geophys. Res., 97(D17), 19.033 – 19.059.


