Uncertainty in predictions of forest carbon dynamics

Ecological Applications, 21(5), 2011, pp. 1506–1522

Uncertainty in predictions of forest carbon dynamics: separating driver error from model error

L. Spadavecchia,[1] M. Williams,[1,3] and B. E. Law[2]

[1]*School of GeoSciences and NERC Centre for Terrestrial Carbon Dynamics, University of Edinburgh, Edinburgh EH9 3JN United Kingdom*
[2]*Oregon State University, Corvallis, Oregon 97331 USA*

Abstract. We present an analysis of the relative magnitude and contribution of parameter and driver uncertainty to the confidence intervals on estimates of net carbon fluxes. Model parameters may be difficult or impractical to measure, while driver fields are rarely complete, with data gaps due to sensor failure and sparse observational networks. Parameters are generally derived through some optimization method, while driver fields may be interpolated from available data sources. For this study, we used data from a young ponderosa pine stand at Metolius, Central Oregon, and a simple daily model of coupled carbon and water fluxes (DALEC). An ensemble of acceptable parameterizations was generated using an ensemble Kalman filter and eddy covariance measurements of net C exchange. Geostatistical simulations generated an ensemble of meteorological driving variables for the site, consistent with the spatiotemporal autocorrelations inherent in the observational data from 13 local weather stations. Simulated meteorological data were propagated through the model to derive the uncertainty on the CO_2 flux resultant from driver uncertainty typical of spatially extensive modeling studies. Furthermore, the model uncertainty was partitioned between temperature and precipitation. With at least one meteorological station within 25 km of the study site, driver uncertainty was relatively small (~10% of the total net flux), while parameterization uncertainty was larger, ~50% of the total net flux. The largest source of driver uncertainty was due to temperature (8% of the total flux). The combined effect of parameter and driver uncertainty was 57% of the total net flux. However, when the nearest meteorological station was >100 km from the study site, uncertainty in net ecosystem exchange (NEE) predictions introduced by meteorological drivers increased by 88%. Precipitation estimates were a larger source of bias in NEE estimates than were temperature estimates, although the biases partly compensated for each other. The time scales on which precipitation errors occurred in the simulations were shorter than the temporal scales over which drought developed in the model, so drought events were reasonably simulated. The approach outlined here provides a means to assess the uncertainty and bias introduced by meteorological drivers in regional-scale ecological forecasting.

Key words: carbon dynamics; data assimilation; ensemble Kalman filter; geostatistics; product–sum covariance model; process-based modeling.

Introduction

Modern catchment scale studies of environmental phenomena commonly use models for extrapolation and prediction (Running 1994, Runyon et al. 1994, Law et al. 2001*a*, Williams et al. 2001*b*, 2005). A key problem is upscaling detailed observations made at a small number of sites to a wider area, due to the expense and technical difficulties associated with direct observation (Thornton et al. 1997, Williams et al. 2005). Processed based models formalize knowledge of ecological processes, and allow observations at various scales to be incorporated into regional analyses (Heuvelink and Webster 2001, Canham et al. 2003, Williams et al. 2005). Such models typically require estimates of rate parameters and initial surface characteristics, along with a set of meteorological driving variables, from which estimates of the state vector are derived.

Regional analyses are complicated by the difficulty in measuring and setting parameters, and finding adequate data to drive the model. On one hand, parameters may be difficult or impossible to measure in practice, particularly if the rates of the processes they represent are slow relative to the observational period, with time constants greater than a few months (Williams et al. 2009). On the other hand, sourcing adequate data to drive the model over the required spatiotemporal extent

Manuscript received 3 July 2009; revised 12 March 2010; accepted 30 March 2010. Corresponding editor: D. S. Schimel. For reprints of this Invited Feature, see footnote 1, p. 1427.
[3] Corresponding author. E-mail: mat.williams@ed.ac.uk

may be difficult due to sparse sensor networks and missing observations resultant from sensor failure (Thornton et al. 1997). In general, optimization procedures are used to infer appropriate parameter sets (e.g., Williams et al. 2005), and interpolation schemes are used to gap-fill meteorological drivers (e.g., Thiessen 1911, Running et al. 1987, Hungerford et al. 1989, Daly et al. 1994, Hudson and Wackernagel 1994, Thornton et al. 1997, Goovaerts 2000). The errors resultant from these activities are difficult to quantify, and in the case of driver interpolation rarely explored (Fuentes et al. 2006).

Parameter errors can be quantified through a variety of techniques, usually based on Monte Carlo analyses. Parameters may be perturbed by a series of fixed percentages to probe the effect on the state vector (e.g., Van Oijen et al. 2005, Williams et al. 2005). More formally, the model can be parameterized using a Bayesian framework with parameter error determined from its posterior distribution (Verbeeck et al. 2006, Kennedy et al. 2008, Klemedtsson et al. 2008). Here we explore an alternative Bayesian technique, whereby an a priori parameter set is updated by comparing the model trajectory with observations via data assimilation using an ensemble Kalman filter. This technique has been popular amongst meteorologists and oceanographers (Evensen 1994, Eknes and Evensen 2002), and confers the advantage of balancing the observation and model error in an optimal sense (Maybeck 1979).

Quantification of error resultant from meteorological driver uncertainty may be assessed through geostatistical simulation techniques (Goovaerts 2001, Fuentes et al. 2006). Based on this uncertainty that these techniques quantify, a moderately large ($n = \sim 1000$) ensemble of equi-probable meteorological fields can be constructed from the available data, honoring the spatiotemporal autocorrelation structure of the data. The error magnitude of the state vector is quantified after propagating the ensemble through the model using Monte Carlo analysis of the n model estimates (e.g., Fuentes et al. 2006).

Data scarcity can increase driver error when using geostatistical upscaling of meteorological drivers over a region (Spadavecchia and Williams 2009). However, it is not clear how errors in the meteorological fields affect the state vector, particularly because errors are reduced with increasing temporal aggregation (Spadavecchia and Williams 2009). Processes which respond instantaneously to the driver fields (e.g., temperature controlled) are likely to have larger error magnitudes than those which integrate driving variables over time (e.g., soil moisture controlled). As a result, driver errors, which in some cases are appreciable (Spadavecchia and Williams 2009), may in fact cancel out over the model run.

We present an analysis of the sources and magnitude of model errors using the Data Assimilation Linked Ecosystem Carbon model (DALEC); a simple process-based ecosystem model of carbon dynamics, here modified to include water fluxes and carbon–water interactions. The model is multi-output, supplying estimates of C stocks, soil moisture and fluxes of carbon and water on a daily time step. The model is parameterized for a well-sampled ponderosa pine forest at Metolius, in central Oregon, USA using the ensemble Kalman filter (EnKF; Evensen 2003). Parameter uncertainty is propagated into model outputs to determine the associated uncertainty. The observed meteorology is then replaced with an ensemble of geostatistical simulations conditioned on observations surrounding the study site. The parameterized model is run again multiple times to sample the resultant uncertainty in net ecosystem exchange (NEE) due to driver uncertainty. Finally, a full uncertainty analysis is undertaken using Monte Carlo sampling of both parameter and driver sets, to examine the cumulative uncertainty of the NEE.

The objectives of this paper are to examine and compare the magnitude of model error resultant from parameter uncertainty and driver uncertainty at a daily timescale. Furthermore, the error magnitude resultant from uncertainty in a variety of daily driver fields is characterized to diagnose which fields are critical to constrain model predictions. In doing so this study addresses the following hypotheses:

*H*1) Driver error will be larger than the parameter error, since the likely range of parameters are well constrained locally by eddy flux data, while meteorological simulations are conditioned on patchy, spatially dispersed data.

*H*2a) Precipitation will contribute most to model uncertainty. Precipitation has the largest interpolation error, and ecosystem production is drought limited in the study region (Law et al. 2001*a*, Van Tuyl et al. 2005).

*H*2b) Temperature will contribute most to model uncertainty. Errors associated with precipitation will average out over time, as plant response to precipitation is resultant from drought. Drought integrates precipitation uncertainty over time through soil moisture content, so instantaneous temperature effects on heterotrophic processes will dominate the NEE error signal.

Methods

Study site

The Metolius young ponderosa pine site is located on a private forestry concession near the Metolius Research Natural Area (44°26′ N, 121°34′ W, elevation ~1165 m), about 10 miles (16.7 km) west of Sisters, Oregon, USA (Fig. 1). The site was clear-cut in 1978, and since then has naturally regenerated, with some thinning in 2002. The average age of trees in 2000, before thinning, was 16 years. The canopy layer is exclusively composed of *Pinus ponderosa*, with an understory of *Purshia tridentata* and

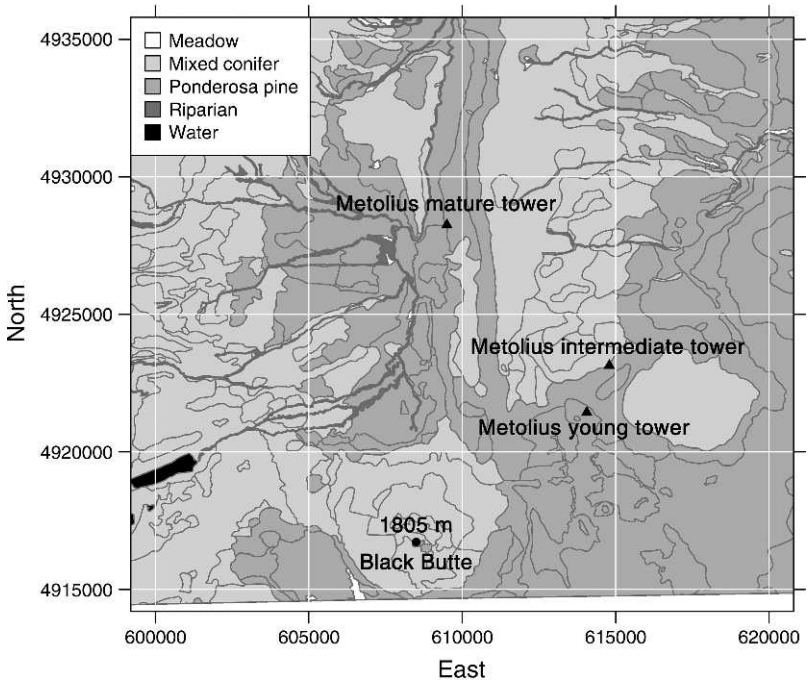

FIG. 1. Metolius young ponderosa pine site and surrounding area, Deschutes County, Oregon, USA. Two other AmeriFlux towers are situated to the north. The area is extensively forested with ponderosa pine and mixed coniferous vegetation (vegetation data courtesy of USDA Forest Service, Sisters Ranger Station, Sisters, Oregon).

Pteridium aquilinum, and an herb layer of *Fragaria vesca*. From 2000 to 2002, the site had a continuously functioning eddy covariance system, forming part of the AmeriFlux observational network. Fluxes were measured at ~9 m above the canopy. The site is characterized by warm dry summers and wet cool winters. Diurnal temperature variation can be high (1.5–18.6°C), and the site is prone to drought (mean annual precipitation = 402 mm, mean number of dry days = 224).

Observations from a nearby tower, at a pine stand of intermediate age (USMe2) were used in the meteorological component of this study. At USMe2, precipitation was measured using a rain gauge (model TE525WS; Texas Electronics, Dallas, Texas, USA) on the tower at 32 m above ground level and in a small natural clearing next to the main tower using the same type of rain gauge equipped with a snowfall adapter during the winter and spring months. Sensors of snow depth and air temperature close to the ground (at 1.6 m, $T_{a,1.6m}$) were mounted on a 2-m tower in the same location.

In the two decades encompassing the observation period, the three driest years were 2002, 2000, and 2003 summed by calendar year, and 2001, 1994, and 2003 summed by water year (Oct–Sep). Because there were several unusually dry years during the study period, cumulative effects on carbohydrate reserves could influence apparent responses to climate variables, particularly in the relatively shallow-rooted young forest.

Modeling daily exchanges of C and water

Canopy processes.—The model consists of a "big leaf" photosynthesis (GPP) and evapotranspiration (ET) model (aggregated canopy model, ACM; Williams et al. 1997) coupled to a module that tracks the allocation and mineralization of carbon, and a module that tracks the dynamics of soil moisture. This coupled model is henceforth referred to as the Data Assimilation Linked Ecosystem Carbon Model, or DALEC model (Fig. 2; Fox et al. 2009).

The ACM calculates GPP and ET as a function of vegetation properties (leaf area index, and foliar N for GPP), meteorology (maximum daily temperature, daily temperature range, maximum daily vapor pressure deficit, total daily irradiance) and soil properties (soil hydraulic resistance and soil water potential). The ACM model was parameterized from locally calibrated soil–plant–atmosphere model predictions of GPP and ET (Williams et al. 1996, Schwarz et al. 2004), using the approach laid out in Williams et al. (1997).

C cycling.—The carbon module apportions the predicted gross primary production (GPP) into autotrophic respiration and the growth of plant C pools (DALEC; Williams et al. 2005) and then tracks additions to and mineralization of litter and soil organic matter (SOM). DALEC requires the specification of ten carbon parameters to control the fate of C in the ecosystem. These parameters relate to the rate of decomposition, fraction of GPP respired, fraction of NPP allocated to

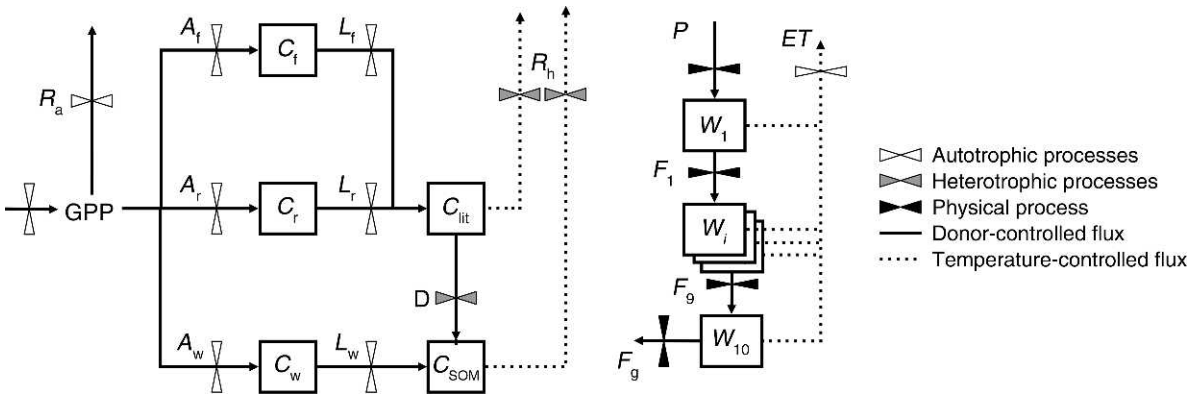

FIG. 2. Data assimilation linked ecosystem carbon (DALEC) C and water dynamics model. Pools are shown as boxes, while fluxes are represented as arrows. The left-hand plot illustrates the C module: GPP (gross primary production) is allocated to foliage (f), roots (r), or woody (w) material. Allocation fluxes are marked A, while losses are marked L. C loss is through respiration fluxes (R), split between autotrophic (a) and heterotrophic (h) sources. The right panel details the flow of water through the model: Precipitation (P) is allocated between 10 soil water layers (W_1–W_{10}). Vertical drainage flows (F_1–F_9) occur when soil layers are saturated. Water may be lost through gravitational drainage (F_g) to groundwater or evapotranspiration (ET).

foliage, fraction of remaining NPP allocated to fine roots, turnover rates of foliage, wood, fine roots, litter, and SOM, and the temperature sensitivity of litter and SOM mineralization. DALEC also requires an initial estimate of the C stock present in five pools: foliage, fine roots, woody stems, litter, and SOM (see Williams et al. 2005).

The model takes daily inputs of minimum temperature (T_{min}), maximum temperature (T_{max}), and precipitation (P). Temperature observations are converted to daily average temperature (T_a), maximum daily vapor pressure deficit (VPD), and solar radiation (RAD) using well-tested relationships (Running et al. 1987, Thornton et al. 1997). VPD is estimated using Murray's formula (Murray 1967), while RAD is predicted using the Allen model (Allen 1997). Details of these models are provided in the Appendix.

Modeling soil water dynamics and drought stress.—A simple daily model of soil water dynamics was constructed based on intensive hourly modeling studies at the site (Williams et al. 2001*a*, Schwarz et al. 2004). The model tracks water inputs and outputs in a 10 layer "bucket" model extending to 3 m in depth. Moisture drains from soil layers when water content exceeds field capacity. We used relationships from Saxton et al. (1986) and local measurements of soil texture to determine porosity and field capacity. Soil hydraulic resistance was determined based on soil texture, root biomass, and water fraction in each soil layer (Williams et al. 2001*a*). Soil water potential (Ψ_s) was generated from a locally determined empirical relationship ($\Psi_s = -1.74 + 3.997\theta$) on soil water fraction (θ). Rooting depth was determined as a function of root biomass using data from nearby ponderosa pine stands (Schwarz et al. 2004). More details of this modeling approach are provided in Fisher et al. (2008).

Data

Meteorological observations.—The 13 closest meteorological monitoring stations with data available for the period 2000–2002 were used to generate meteorological simulations at the study site location (Fig. 3). These stations were selected so that there would be a minimum of eight stations providing T_{min}, T_{max}, and P observations each day, on which to condition meteorological simulations. Observations were filtered such that the values would not exceed the state extremes for Oregon (National Oceanic and Atmospheric Administration (NOAA), Silver Spring, Maryland, USA). Meteorological records for stations >25 km away were also sourced to examine the effect of data scarcity on NEE uncertainty (Fig. 3). Data from a total of 112 stations were available.

Flux observations.—Three years of data from the Metolius young ponderosa pine plot (Law et al. 2001*c*, *d*) were used to parameterize DALEC. The data consisted of records of net ecosystem exchange (NEE), total ecosystem respiration (R_e), evapotranspiration (ET), and a set of meteorological observations, sampled at the daily time step. Direct observation of T_{min}, T_{max}, T_a, P, VPD, and RAD were made simultaneously with the flux data. Gaps in the data resulted from sensor failure and filtering to remove observations with low friction velocity (u^*), or physically implausible magnitudes (NEE > 25 $\mu mol \cdot m^{-2} \cdot s^{-1}$). Short gaps in daytime NEE were filled using the lookup table method (Falge et al. 2001). Gaps in latent heat exchange were filled using an empirical relation to measured net radiation. We generated daily net ecosystem exchange of CO_2 (NEE) data for days in which <25% of the 48 possible half hour measurements were gap filled; for the three year period of this study, this amounted to 684 daily NEE values. Soil respiration was measured using six automated chambers installed in 1999 (Irvine and Law 2002); total

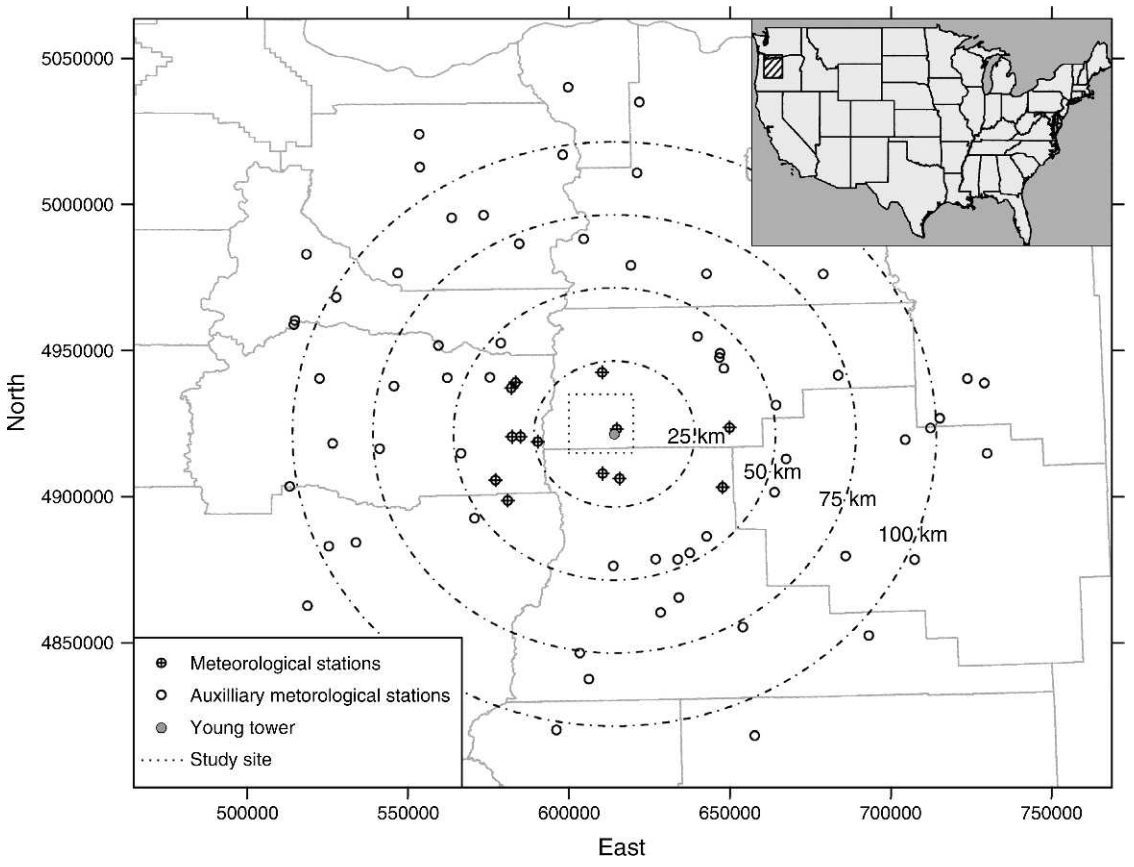

FIG. 3. Meteorological monitoring stations surrounding the Metolius young ponderosa pine site, indicated as a gray point. Crosshairs represent the 13 core stations used for the simulations, while auxiliary stations at increasing distance from the study site are indicated as open circles. Light gray lines indicate county boundaries. The extent of the area within the conterminous United States is indicated as a hatched black rectangle in the inset map.

daily effluxes were recorded on 401 days during 2000–2002.

Canopy density observations.—Observations of the leaf area index (LAI) of the forest canopy were used to constrain the parameterization of DALEC. Data were collected at four times during the three-year period, using an LAI-2000 plant canopy analyzer (LI-COR, Lincoln, Nebraska, USA; Table 1). Observations were collected on a 10-m^2 grid over 1 ha, and were corrected for clumping at the needle, shoot, and stand levels (Law et al. 2001*b*, *c*). These observations were related to the model foliar carbon estimate via direct measurements of the specific leaf mass from foliage samples, see Williams et al. (2005) for further details.

TABLE 1. Summary of total leaf area index (LAI) estimates for the young ponderosa pine site for four survey dates.

Date	LAI
19 Jul 2000	1.0
19 Aug 2001	1.46
22 May 2002	0.85
17 Sep 2002	1.67

Note: LAI estimates were derived by combining both tree canopy measurements (ponderosa pine) and understory surveys.

DALEC parameterization

Many of the parameters associated with the processes of photosynthesis, evapotranspiration and soil water physics have been derived from the literature or from previous research at the study site (Williams et al. 2001*a*, 2005, Schwarz et al. 2004). The most uncertain parameters are the 10 associated with respiration, turnover, and allocation of C among plant and soil pools. We added an eleventh parameter to these, the parameter from the ACM GPP model that relates foliar N content to photosynthetic capability, to include an estimate of uncertainty in the GPP calculations.

We used an ensemble Kalman filter (EnKF, Williams et al. 2005) to estimate the likely distributions for these uncertain parameters. The EnKF combines a model of a system (i.e., DALEC) with observations of that system over time (i.e., NEE and LAI observations). The model generates predictions of the state vector (C pools and fluxes, soil moisture and water fluxes) for each time step. NEE and LAI predictions are then compared with

independent observations. Based on an assessment of model forecast and observational uncertainty, the predicted NEE and LAI are adjusted. The model error covariance matrix, as determined in the EnKF, is then used to adjust the full state vector accordingly.

We adjusted the EnKF approach used in Williams et al. (2005) from a state estimation problem to a parameter estimation problem. We added the 11 model parameters to the state vector supplied to the EnKF. We set the model error on the fluxes and pools of C and water to relatively low values (0.01%) compared to the uncertainty on the 11 parameters (0.2%). In the EnKF the model error is added to each member of the state vector at each time step, causing a divergence in the parameter distribution around the mean for each member of the state vector. Assimilated observations adjust the mean state and the distribution around the mean, generally reducing its spread. The parameter uncertainty was selected so that its divergence was more than one order of magnitude larger than the state uncertainty, so that over time the filter sampled a greater proportion of parameter space than state space. Thereby the analyses adjusted parameter values, rather than states. The absolute magnitude of the parameter uncertainty was selected to allow shifts in parameters values of ~75% over an annual cycle, which we expected to encompass the likely uncertainty in initial estimates of parameters, while avoiding sharp shifts in parameter values at daily timescales, that disrupt the mass balance of the modeling.

To assess uncertainty in the estimate of short-term flux due to random sampling errors, we used the relative random flux error defined as the ratio of the standard error to the absolute value of the record mean flux. Random flux errors averaged 20% (Vickers et al. 2009), and the majority of the uncertainty is due to random sampling errors, not nonstationarity. Systematic errors are generally expected to sum to 12% (Falge et al. 2001). However, the nature of NEE (it can be positive or negative) means that defining errors by a coefficient of variation is unsuitable in data assimilation, so instead errors are set at 0.7 g $C{\cdot}m^{-2}{\cdot}d^{-1}$, approximately 20% of the typical summer NEE values. The coefficient of variation on LAI observations at the site was 10%, and so the error was set at this value (Law et al. 2001*d*). In earlier analyses we found an ensemble of 200 was adequate; here, with an enhanced state vector, we increased the number to 400.

The initial EnKF analysis used parameter estimates from an earlier study as prior estimates (Williams et al. 2005). After the initial analysis, the posterior parameter estimates were used to reinitialize the parameters, and the EnkF was run again. This process was repeated once more, at which point the parameter ensembles had stabilized. The model was then run in forward mode, with each of the final posterior parameter ensembles used in turn to evaluate the effect of parameter uncertainty on the NEE estimate (experiment 1).

Meteorological simulation

Sequential Gaussian simulation (SGS; Goovaerts 1997) was used to quantify the uncertainty of interpolated driving variables at the Metolius site. SGS may be regarded as an extension of the commonly used kriging technique (e.g., Hudson and Wackernagel 1994, Ashraf et al. 1997, Goovaerts 2000, Spadavecchia and Williams 2009). Kriging estimates represent the most likely value of the estimate given the surrounding observations, based on a probabilistic model. Kriging estimates have attached variances which are a valid measure of uncertainty when taken in isolation; however they are less useful for assessing the uncertainty of the regionalization as a whole (Goovaerts 1997). SGS expands on kriging by drawing equally possible realizations of the whole field from the probabilistic model, preserving the surface roughness of the estimated field and avoiding the characteristic smoothing effect of kriging (Goovaerts 1997,1999, 2001, Deutsch and Journel 1998). For complete details on the approaches, see the Appendix and Spadavecchia and Williams (2009). The approach outlined does not account for covariance between driving variables. Methods for co-kriging to account for these covariances have not to our knowledge been implemented in the space–time domain. Complications of achieving stable covariance matrices for inversion in the kriging calculations are a major challenge, and further research is required to implement such an approach.

Precipitation simulations

Due to the stochastic nature of rainfall events it was necessary to simulate precipitation in a two-stage process. First the probability of a rainfall event was simulated from a binary recoding of the observations data (precipitation indicator P_i) representing the probability of an event (i.e., $P_i = 0$ if $P = 0$, $P_i = 1$ if $P > 0$). Precipitation events were dispersed across the simulation grid by comparing the simulated probability (P_i^*) of rainfall with a draw from a random number generator (r). Grid nodes were coded for an event if $P_i^* > r$. Having established the grid nodes where precipitation takes place, we then simulated the amount of precipitation at these points using the methods outlined in the Appendix. See Spadavecchia and Williams (2009) for more details.

Partitioning driver uncertainty

1000 simulations of T_{min}, T_{max}, and P were generated at the Metolius site for the full three years of the study, conditioned on data from the eight closest spatial neighbors over a temporal window of ±10 days (88 observations) via SGS. Previous analyses have shown that temporal information from such a window improves the regionalization of precipitation data, it has no significant impact on temperature regionalization's (Spadavecchia and Williams 2009). Meteorological observations at the study site were excluded so as to

explore the uncertainty resultant from modeling C dynamics over sparsely sampled regions. The parameterized model was run with each of the 1000 simulations in turn, to inspect the variability in the predicted NEE ensemble (experiment 2). Two subsidiary experiments were run, calculating NEE using (experiment 2.i) locally observed temperatures, VPD, and RAD with simulated precipitation and (experiment 2.ii) locally observed precipitation with simulated temperatures, VPD, and RAD. Finally, having generated a parameter ensemble and 1000 equi-probable meteorologies, a sample of 1000 parameter and meteorology permutations was generated to test the combined effect of parameter and driver uncertainty on the model (experiment 3).

To test *H*2, the precipitation regime of the data was compared with the simulated rainfall trajectories. The number of days since a precipitation event ($n_{P=0}$) was calculated for the 1000 simulations generated in experiment 2.i. The number of days since a precipitation event in the local observations was substracted from $n_{P=0}$ to generate a metric of drought (Δ_P). Data where Δ_P was positive (i.e., simulations with longer dry spells than observed in the data) were used to examine the effect of drought on the uncertainty of the NEE trajectory.

Sparsity of meteorological conditioning data

Interpolation uncertainty is related to the distance to the nearest neighbors (Spadavecchia and Williams 2009). The effect of increasing data sparseness was investigated by conditioning simulations on data from increasing search radii (Fig. 3), ignoring weather stations closer than the threshold distances of 25, 50, 75, and 100 km. In each case, the closest four stations beyond the threshold distance were used to condition the simulations. The model was run with each of these meteorological ensembles to test the robustness of the comparison of meteorological and parameterization uncertainties on the uncertainty of the final NEE analysis (experiment 4).

Results

DALEC parameterization

Parameter optimization and confidence intervals.—The EnKF assimilated observations of NEE and LAI into an ensemble of 400 state vector predictions from DALEC, thereby generating estimates of the 11 parameters included in the state vector. The ensemble was subject to a chi-squared goodness of fit test, by comparing forward predictions of NEE against the NEE observations, as a check on the parameterizations. The chi-squared test was applied on the predictions and observations of daily NEE to test their similarity. Of the 400 parameter sets, 375 passed this test ($\chi^2 <= 717$, df = 656, $P > 0.95$) and were used in further analyses.

The posterior parameter means were similar in most cases to the parameter priors (Table 4), themselves outputs of an earlier optimization study. There were slight increases in magnitude of the mean estimate for some parameters (e.g., fraction of photosynthate respired, and fraction of NPP allocation to foliage), decreases for others (turnover rate of foliage) and no clear changes for some (e.g., heterotrophic respiration sensitivity). Visual inspection of the temporal evolution of the parameter ensembles over time (data not shown) revealed strong growth in the confidence intervals on the parameters during the first year of assimilation, indicating that prior ensemble variances were too narrow. Some parameters saw stabilization in confidence intervals by years 2 and 3 (e.g., turnover rate of foliage, heterotrophic respirations sensitivity) while others saw slower but continued growth (e.g., turnover rate of wood and mineralization rate of litter). The final distributions of the parameter ensembles at the end of the assimilation period, and the resultant total NEE distribution, were generally normal (Fig. 4), largely as a result of using relatively tight prior parameter estimates.

State estimation with optimized parameters.—The mean ensemble trajectory for four major components of the flux reasonably matched the observations of NEE and LAI, as expected due to optimization. The trajectories of total ecosystem respiration (R_e) and evapotranspiration (ET), data not used in the assimilation, were reasonably replicated also (Fig. 5). Quantitative tests revealed the model to be performing well with respect to all data streams (Table 2). We have previously noted that phenology module introduces mismatches with the data (Williams et al. 2005). This phenology problem results in lower amplitude of leaf area variation over each year than was observed, and causes an underestimate of maximum rates of C sequestration and evapotranspiration each summer.

The optimized model estimated a total carbon uptake of 422 g C/m^2 over three years, with a 95% confidence interval of ±211.2 g C. The NEE uncertainty was determined as the 95% ensemble CI, excluding the high and low 2.5% tails from the acceptable ensemble (n = 375). The uncertainty resultant from parameterization was therefore substantial, representing 51% of the total net flux. A detailed analysis of the NEE error indicated relatively unbiased estimates, with ~60% of model errors <0.5 g C·m^{-2}·d^{-1} (Fig. 6). A simple linear regression between the estimates and observations indicated a slope of 0.302 and an intercept of −0.299 (Fig. 6), suggesting that the model tended to smooth the NEE trajectory, underestimating the distribution at the extremes, with a small bias towards underestimation of the daily C uptake.

Meteorological simulation

One thousand simulations were drawn from the meteorological data using the specified covariance models (see Appendix for details). Estimates of VPD and RAD were generated from T_{min} and T_{max} via

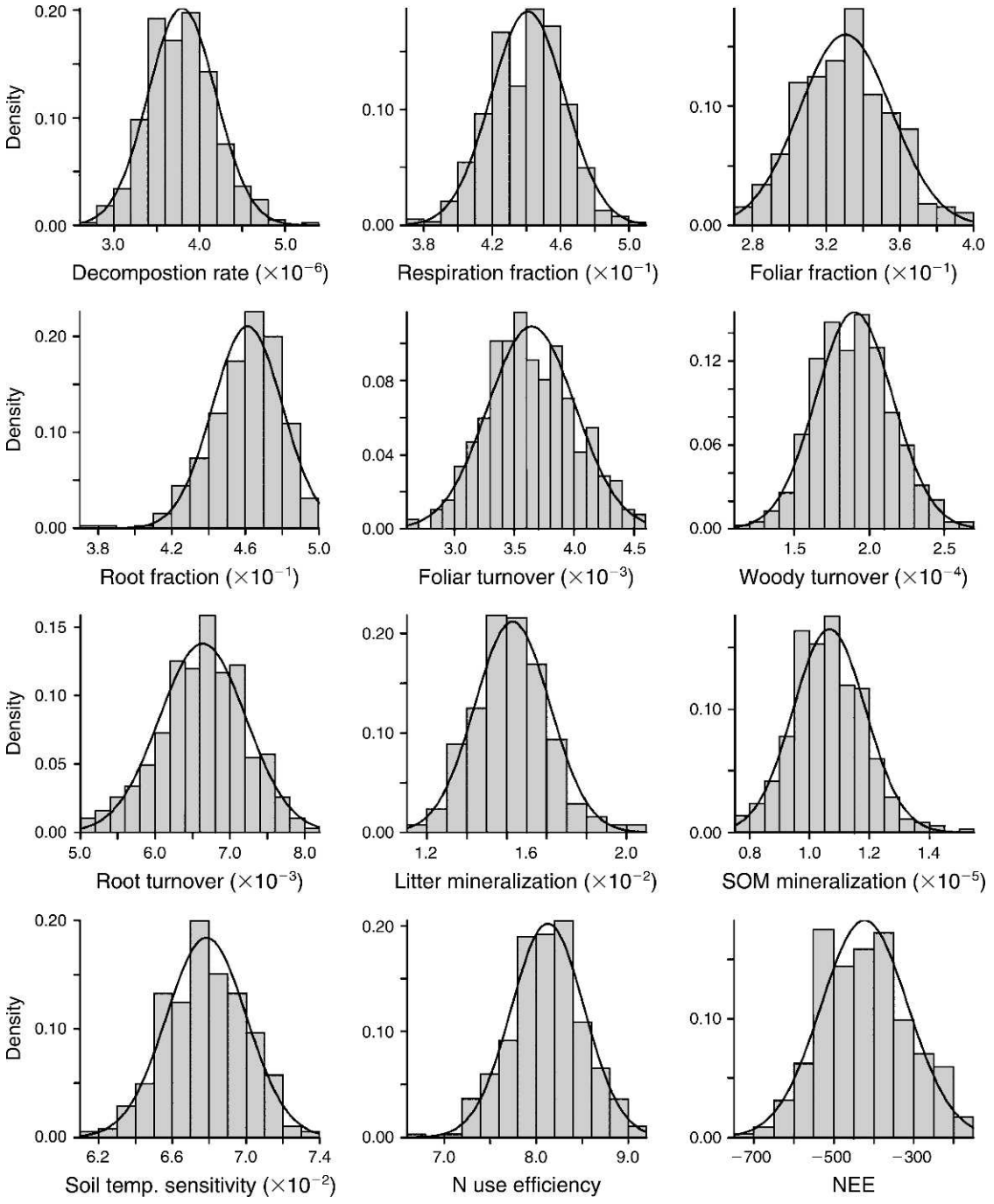

FIG. 4. Marginal parameter distributions retrieved from an ensemble of 375 elements derived from the ensemble Kalman filter and passing a goodness-of-fit test against the observed net ecosystem exchange (NEE) time series 2000–2002. The resultant total net ecosystem exchange (g C/m^2) over three years is also indicated. Numbers on the x-axis should be multiplied by the factor in parentheses. Abbreviations are: SOM, soil organic matter; temp., temperature.

Murray's formula ($A = 0.978$, $B = 22.23$, $C = 243.95$, where A, B, and C are empirical constants; see Appendix) and the Allen model ($K_r = 0.17$; where K_r is an empirical constant, see Appendix), which were calibrated locally. The observed meteorology was reproduced successfully for all variables (Fig. 7), with r^2 values > 0.8 for all variables except P, which had an r^2 of 0.63 and a considerable positive bias (Table 3).

The 1000 meteorological realizations were propagated through DALEC to sample the NEE uncertainty resultant from driver uncertainty. All experiments were run using the mean parameter set retrieved from the

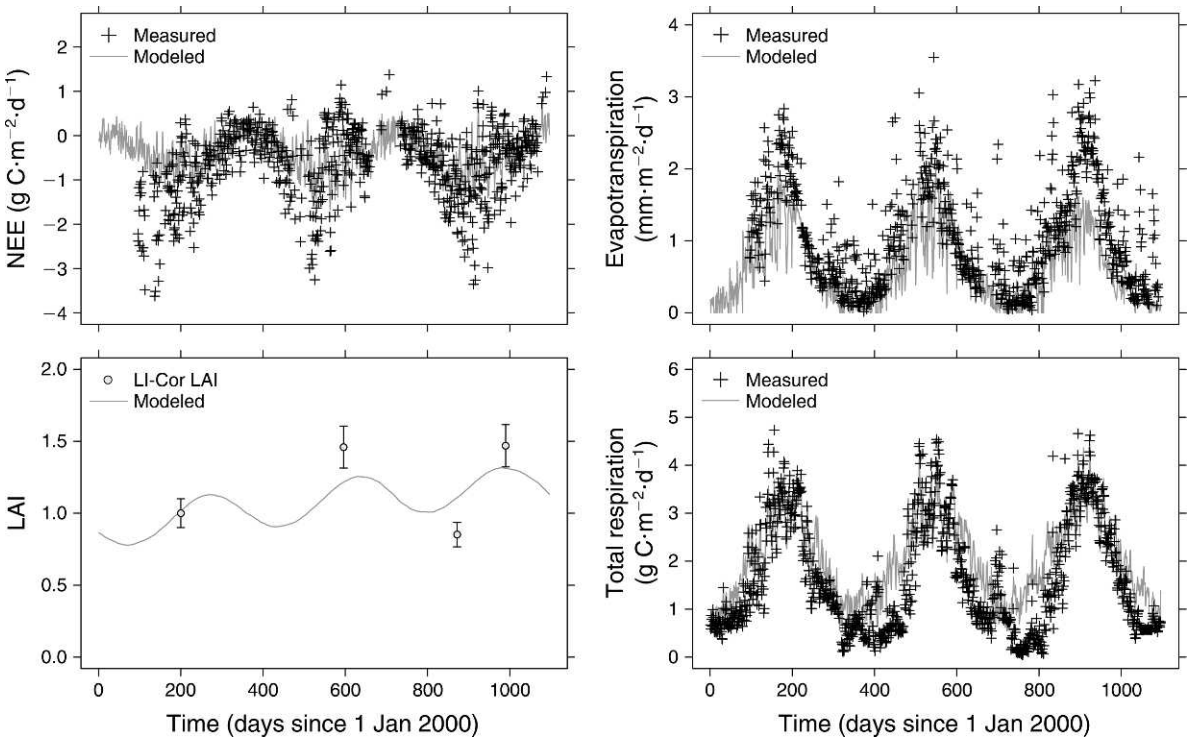

FIG. 5. Model data comparison for four fluxes. In all cases, the dark gray line indicates the mean trajectory of the 375-element ensemble trajectory retrieved from the ensemble Kalman filter. Observations are indicated as black crosses. The modeled LAI is compared with ground based Li-Cor LAI2000 (Li-Cor, Lincoln, Nebraska, USA) observations and MODIS satellite retrievals. Leaf area index (LAI) error bars are included to show the high variability of the satellite retrievals.

EnKF (Table 4). With all meteorological observations replaced with simulated values, the model predicted a total NEE of −379 g C/m^2 with a 95% confidence interval of ±31.36 g C/m^2. Replacing only P with simulated values (experiment 2.i) resulted in a total flux of −479 ± 10 g C/m^2. Replacing all temperature (T_{min}, T_{max}, T_a) and temperature-derived variables (VPD, RAD) with simulated values (experiment 2.ii) resulted in a total NEE of −339 ± 14 g C/m^2 (Table 5). Although NEE uncertainty attributable to the drivers was relatively small (typically <10%), larger differences in the total flux were observed. The directions of bias for P and temperature were opposite (33.3 and 13.3 g C·m^{-2}·yr^{-1}, respectively), and seemed to counter each other to some extent when the full meteorological uncertainty was propagated through the model (Table 5): The total NEE estimated in experiments 1 and 2 were well within one standard deviation of each other, but were found to be significantly different ($t = 7.76$, $P < 0.0001$).

The importance of drought stress was assessed by comparing the number of days since the last predicted rainfall event for each simulation against the observed rainfall regime (Δ_P). Positive values of Δ_P indicated that the model was going into drought while the observed P was >0 (misspecification). The maximum value of Δ_P was 87 days, while the mean Δ_P was 1.2, indicating an overall negative bias in the number of simulated rainy days: Thus, despite an overall overestimation of P (Table 3) there is an underestimation of rainfall frequency, with a mean simulated P frequency = 102 ± 31 vs. mean observed P frequency of 141 days per annum. Misspecified droughts had a mean length of 6.5 days with a standard deviation of 8.2 days. The mean

TABLE 2. Summary of model fits for various model outputs

Data source	r^2	RMSE	Kendalls τ	Bias	Gain
NEE (g C·m^{-2}·d^{-1})	0.39	8.70	0.44	−0.299 ***	0.302 ***
ET (mm)	0.55	9.90	0.55	0.126 ***	0.503 ***
LAI	0.70	0.12	0.67	0.75 (NS)	0.34 (NS)
R_e (g C·m^{-2}·d^{-1})	0.78	13.00	0.67	0.967 ***	0.654 ***

Note: NEE stands for net ecosystem exchange; LAI stands for leaf area index (unitless); R_e is total ecosystem respiration.
** $P < 0.01$; *** $P < 0.001$.

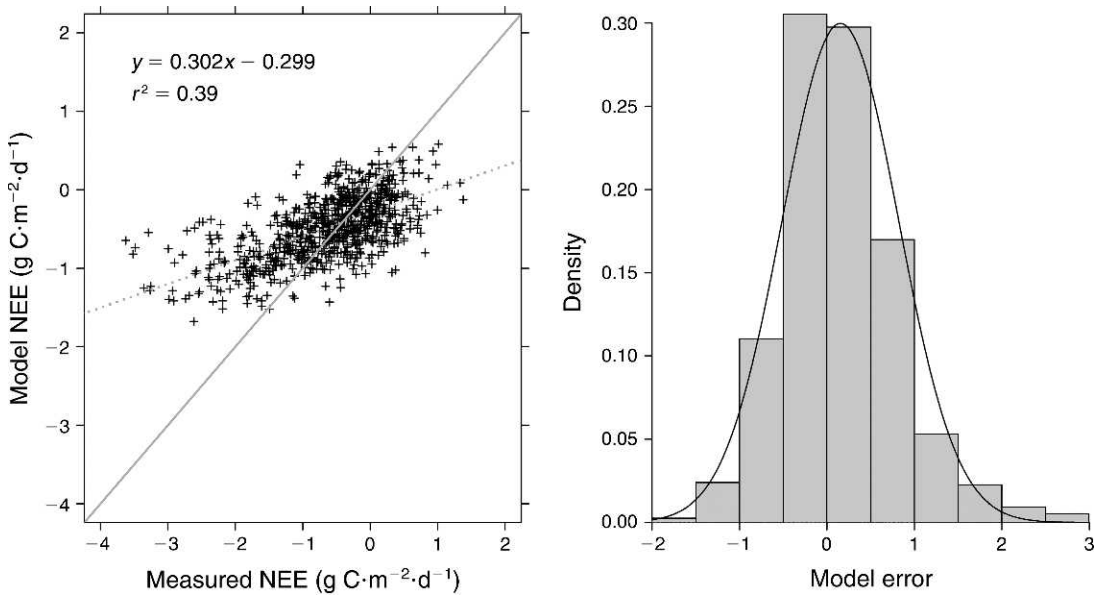

FIG. 6. Net ecosystem exchange (NEE) model data comparison. Model values are the 375-element ensemble means retrieved from the ensemble Kalman filter. The model error distribution is indicated in the right panel. The heavy right hand tail indicates an underestimate of the summer C uptake.

number of days between rainfall events for the observations was 4.1, with a standard deviation of 6.9.

Increasing Δ_P was linearly related to a decrease in modeled soil water content ($r^2 = 0.81$, $P < 0.0001$). The RMSE of modeled vs. observed NEE decreased with increasing drought stress ($r^2 = 0.31$, $P < 0.0001$). Drought was initiated after approximately 30 dry days, as indicated by the step change in Fig. 8. The background RMSE attributable to precipitation uncertainty was 0.7. When $\Delta_P < 30$ the RMSE was approximately equal to background levels (0.71). However, as $\Delta_P > 30$ the RMSE dropped to 0.56.

Monte Carlo sampling of NEE uncertainty

One thousand permutations of parameter and driver combinations were generated at random from the pool of 375 parameter sets and 1000 driver sets (sampling with replacement) and used to generate forward model runs. These runs predicted a NEE for 2000–2002 of -363 ± 105 g C/m^2 (mean ± SD; Table 5). A comparison of the daily flux estimates for the main experiments (1, 2, and 3) revealed broadly similar ensemble trajectories with expected variation in parameters, drivers, and in both (Fig. 9). However, the range of experiment 2 (driver variation only) was asymmetrical about the mean, with a greater deviation in the positive (weaker uptake) direction. As such, the summer extremes in uptake appear to be less well replicated in experiment 2 than experiment 1 (parameter variation only). Furthermore, winter uptake appeared weaker in experiment 2 in comparison with experiment 1.

NEE variability with expected parameter variation (experiment 1) exceeded that with expected meteorological variation (experiment 2). The robustness of this result was tested by increasing the variability of the meteorological ensemble in experiment 4. The amount of conditioning data was reduced to four spatial neighbors (i.e., met stations), while sequentially increasing the minimum distance to an observation (Table 6). In all cases, the NEE uncertainty attributable to meteorological uncertainty was less than the uncertainty attributable to parameter uncertainty (9–17% and 50%, respectively). The results of experiment 4 indicate a general increase in NEE uncertainty with increasing distance to conditioning data, although results from the 25-km threshold distance were more uncertain than the results from the >50-km threshold (Table 6).

A comparison of the cumulative NEE allowed an examination of the growth in uncertainty over time for the three main experiments (Table 5, Fig. 9). Again, the mean ensemble trajectories were broadly similar, with little difference in total uptake for inter-annual comparisons, which were approximately within one standard deviation of each other. However, the greater variability in NEE imposed by parameter uncertainty compounds to a much larger annual uncertainty than for driver uncertainty. The total cumulative uncertainty (experiment 3, Fig. 9c) was not very different from experiment 1, except for exaggerated extremes and a more pronounced end of season die-back, also exhibited in experiment 2 (Fig. 9b).

DISCUSSION

We were able to parameterize DALEC using the EnKF, resulting in a total net C flux estimate consistent with the data and previous literature for the site (Law et

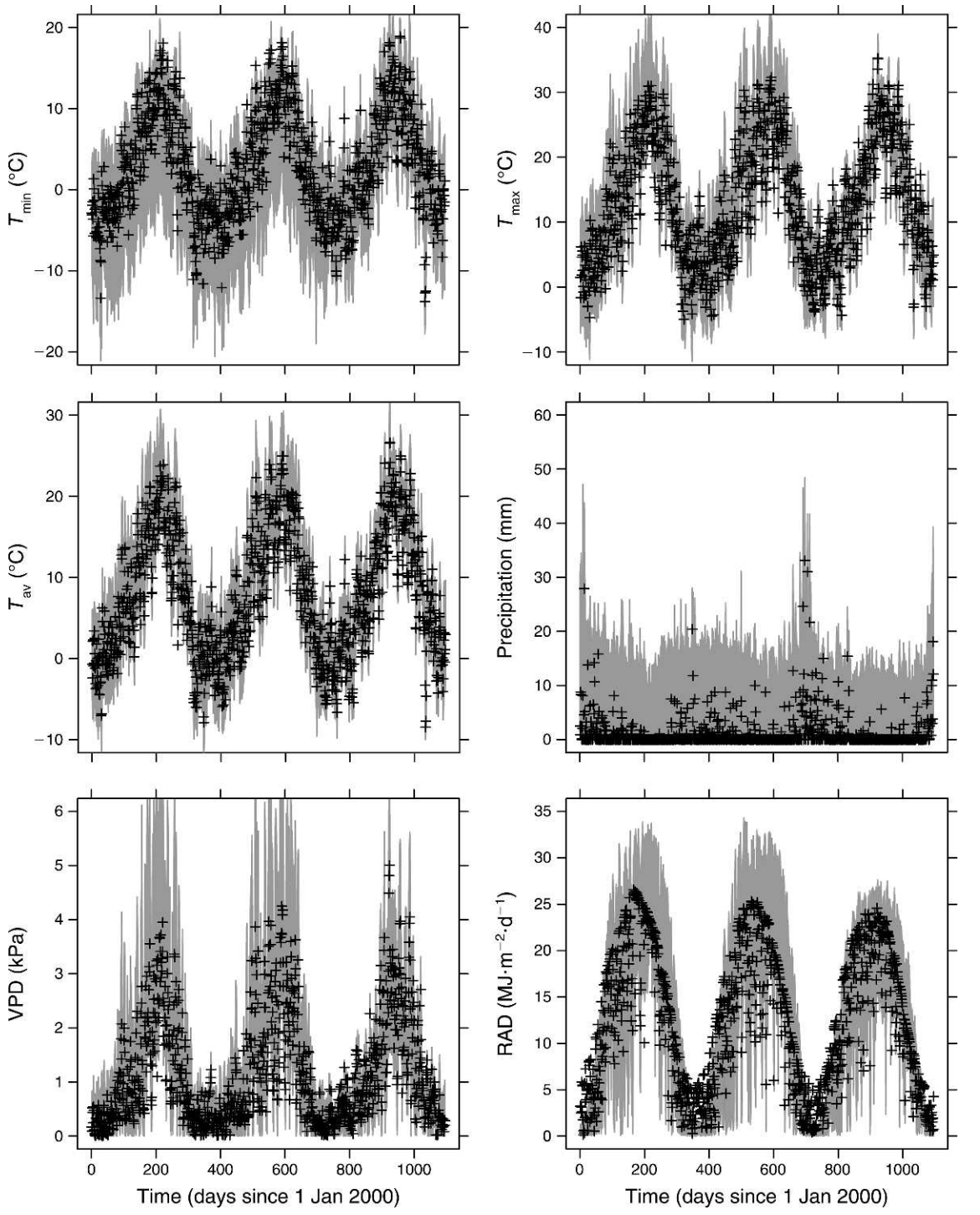

FIG. 7. One thousand meteorological time series derived from geostatistical simulation. Each element of the meteorological ensemble is indicated as a gray line, while observations are indicated as black crosses. Variables are: daily minimum temperature, T_{min}; daily maximum temperature, T_{max}; average daily temperature, T_{av}; maximum daily vapor pressure deficit, VPD; and solar radiation, RAD.

al. 2003, Williams et al. 2005). There was a large range of permissible parameter sets, resulting in cumulative NEE uncertainties over the three years of the study corresponding to ~50% of the total net flux (95% confidence interval of NEE expressed as a percentage of the total flux). The cumulative NEE over three years using EnKF for parameter estimation was 423 ± 109 g C/m^2 (mean ± SD of ensemble). This mean analysis is

TABLE 3. Simulation vs. observation comparison for the daily meteorology at the Metolius young ponderosa pine site.

Driver	Simulated	Observed	Bias	Gain	r^2
T_{min} (°C)	2.8 (6.6)	1.3 (5.6)	0.95	0.79	0.88
T_{max} (°C)	12.8 (9.2)	13.6 (9.8)	−0.34	1.04	0.96
T_{av} (°C)	7.78 (7.7)	8.7 (7.9)	−0.88	1.01	0.97
Precipitation (mm)	639.6 (164.0)	402.4 (60.9)	1.05	0.63	0.63
RAD ($MJ \cdot m^{-2} \cdot d^{-1}$)	12.4 (7.5)	13.6 (8.5)	−0.78	1.03	0.82
VPD (kPa)	1.2 (1.0)	1.3 (1.2)	−0.04	1.13	0.91

Notes: Simulated and observed values are daily means with standard deviations in parentheses. Precipitation is mean annual precipitation. Variables are: daily minimum temperature, T_{min}; daily maximum temperature, T_{max}; average daily temperature, T_{av}; maximum daily vapor pressure deficit, VPD; and solar radiation, RAD.

very close to that reported in an earlier study using the EnKF at the same site for state (rather than parameter) estimation, 419 ± 29 g C/m^2 (Williams et al. 2005). The smaller uncertainty associated with the state estimation approach was largely due to assimilation of a larger and more varied set of data; NEE data were supplemented with ecosystem respiration estimates from chambers, sap flow data to constrain GPP, and biometric data to constrain LAI.

There are still outstanding issues in model-data fusion studies concerning the estimation of model confidence intervals (Fox et al. 2009, Williams et al. 2009). The algorithm chosen here, EnKF, has been shown to produce narrower CI than algorithms that randomly sample parameter space, such as the Metropolis approach. In our EnKF applications we have set Gaussian priors on parameters in the state vector, whereas Metropolis approaches often set uniform priors. The advantage of Gaussian priors is that the algorithm avoids areas of unrealistic parameter space. The disadvantage is that a poor prior can prevent important areas of parameter space from being explored. In our case, we had considerable information for setting priors, based on previous analyses (Williams et al. 2005). We used the EnKF to determine to what degree the priors were reasonable, given data and their associated uncertainties. We found in all cases that parameter uncertainties grew during the assimilation process. Some parameter uncertainties stabilized but others continued to grow, for instance that on wood turnover. Thus, the parameter estimates from the EnKF should be used with care in any prognostic analyses, i.e., beyond 3 years. The information content of the eddy flux observations is not enough to constrain processes with longer time constants, like wood turnover. The relatively large parameter uncertainty could be reduced by, for instance, assimilation of wood increment and fine root turnover data, i.e., those data largely orthogonal to flux data.

Meteorological simulations for the three-year period had a high degree of variability, which decreased in the final year (Fig. 7). This decrease in uncertainty was due to observations at the nearby Metolius "intermediate" tower starting on project day 732 (1 January 2002). Geostatistical simulation techniques are able to reproduce the roughness of the driver fields, preserving data extremes, which may be particularly important for regionalization of precipitation. The precipitation signal is comprised of a background fluctuation ~0 mm, with rare but sizable events which may be on the order of 100 mm/d. Thus, reproduction of extreme events over the average behavior may be critical, and it is in this respect that SGS confers an advantage over kriging techniques. In general, the meteorology was satisfactorily replicated for the site, but issues of bias arose, particularly for precipitation. This finding may be associated with the location of the site, which is in the rain shadow of the Cascade Mountains where there is a steep gradient in precipitation from west to east over about 25 km (2200 mm/yr at the Cascade crest to 350 mm/yr near Sisters, Oregon).

TABLE 4. Summary of prior parameter estimates and the posterior mean and standard deviation generated by the ensemble retrieved from EnKF (ensemble Kalman filter) fitting.

Name	Parameter	Prior mean	Posterior mean	Posterior SD	Scale
t_1	decomposition rate	3.75	3.80	0.40	$\times 10^{-6}$
t_2	autotrophic respiration fraction	4.27	4.54	0.22	$\times 10^{-1}$
t_3	foliar allocation fraction	3.0	3.50	0.25	$\times 10^{-1}$
t_4	root allocation fraction	4.48	4.76	0.19	$\times 10^{-1}$
t_5	foliar turnover rate	4.14	3.64	0.37	$\times 10^{-3}$
t_6	woody turnover rate	1.54	1.96	0.26	$\times 10^{-4}$
t_7	root turnover rate	6.41	6.74	0.58	$\times 10^{-3}$
t_8	litter mineralization rate	1.59	1.81	0.19	$\times 10^{-2}$
t_9	SOM mineralization rate	0.97	1.14	0.12	$\times 10^{-5}$
t_{10}	heterotrophic process temperature sensitivity	6.8	6.70	0.22	$\times 10^{-2}$
t_{11}	photosynthetic scalar	8.0	8.2	0.39	

Note: Values for priors and posteriors should be scaled by the values in the Scale column.

TABLE 5. Total NEE estimates from various uncertainty sources.

Experiment	Source of variation	NEE (g C/m^2) 2000	2001	2002	Total	95% CI†
1	parameters	−120 (28)	−148 (44)	−155 (41)	−422 (107)	50
2	meteorology	−95 (8)	−129 (9)	−155 (8)	−379 (16)	9
2.i	precipitation	−128 (3)	−166 (5)	−185 (6)	−479 (10)	4
2.ii	temperature VPD and RAD	−84 (6)	−114 (8)	−140 (6)	−339 (14)	8
3	total	−95 (25)	−125 (36)	−144 (49)	−363 (105)	57

Note: Values reported are means with standard deviations in parentheses.
† The 95% confidence interval of NEE, expressed as a percentage of the total flux.

Positive bias in precipitation simulations resulted in a positive bias in NEE estimates when all other meteorological drivers were held at their observed values. While precipitation variability was comparatively large, its effects were temporally buffered by the effect of soil capacitance (Fig. 8). A reduction in RMSE was observed with increasing drought stress, reflecting a decrease in the positive bias imposed by the simulated precipitation. Drought stress manifested itself after ~30 days without rain. The mean length of misspecified drought events was 1.2 days, while on average the simulations had an decreased frequency of precipitation events with respect to the observations. On average, simulated dry spells were 6.5 days longer than those measured at the site. This difference indicates that the time scales on which precipitation errors occur in the simulations are much shorter than the temporal scales over which drought operates in the model. Thus the uncertainty in the meteorological ensembles is small enough to avoid misspecification of drought events with a significant impact on primary production.

Despite considerable uncertainty in the simulated driver sets, the resultant NEE uncertainty was 9% of the total flux, contributing only ~7% to the total combined NEE uncertainty, and well within the uncertainty attributable to parameterization. This result was robust under significant degradation of the meteorological data set, with a maximum driver uncertainty of <20% when conditioning simulations on four neighbors separated by distances greater than 100 km from the study site. We therefore reject $H1$, that the dominant source of NEE uncertainty is due to driver uncertainty.

Experiments on the effect of data scarcity indicated a general increase in NEE uncertainty with increasing distance to conditioning data (Table 6). When the nearest station was >100 km from the study site, rather than <25 km, uncertainty in NEE predictions introduced by meteorological drivers increased by 88%. Examining the error contribution of each driver to the NEE trajectory revealed interesting bias effects. The positive bias in simulations of precipitation elevated the estimated C uptake by ~30 g C·m^{-2}·yr^{-1}, while

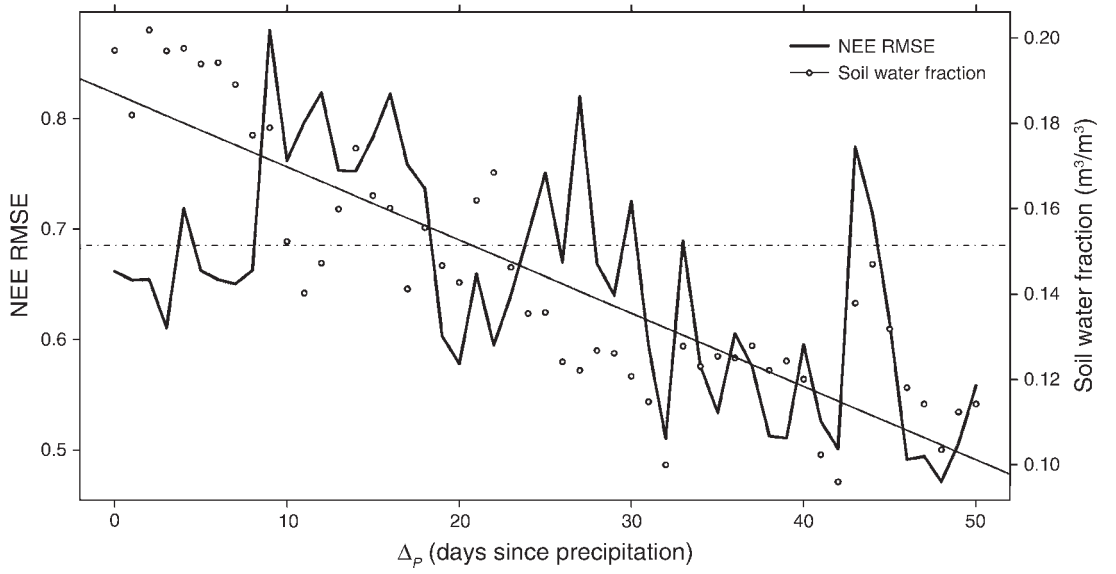

FIG. 8. Effect of drought on NEE error. Δ_P is the number of days simulated as dry on which precipitation events were measured: Δ_P records the number of days that $P_{\text{simulated}} = 0$ while $P_{\text{observed}} > 0$. As Δ_P increases the model goes into misspecified drought, as indicated by the modeled soil water fraction (right axis). The deviation in modeled and observed NEE trajectories attributable to misspecified drought is plotted on the left axis as the root mean squared error (RMSE). The background RMSE of the model resultant from precipitation uncertainty is ~0.7, indicated as a dot-dash line.

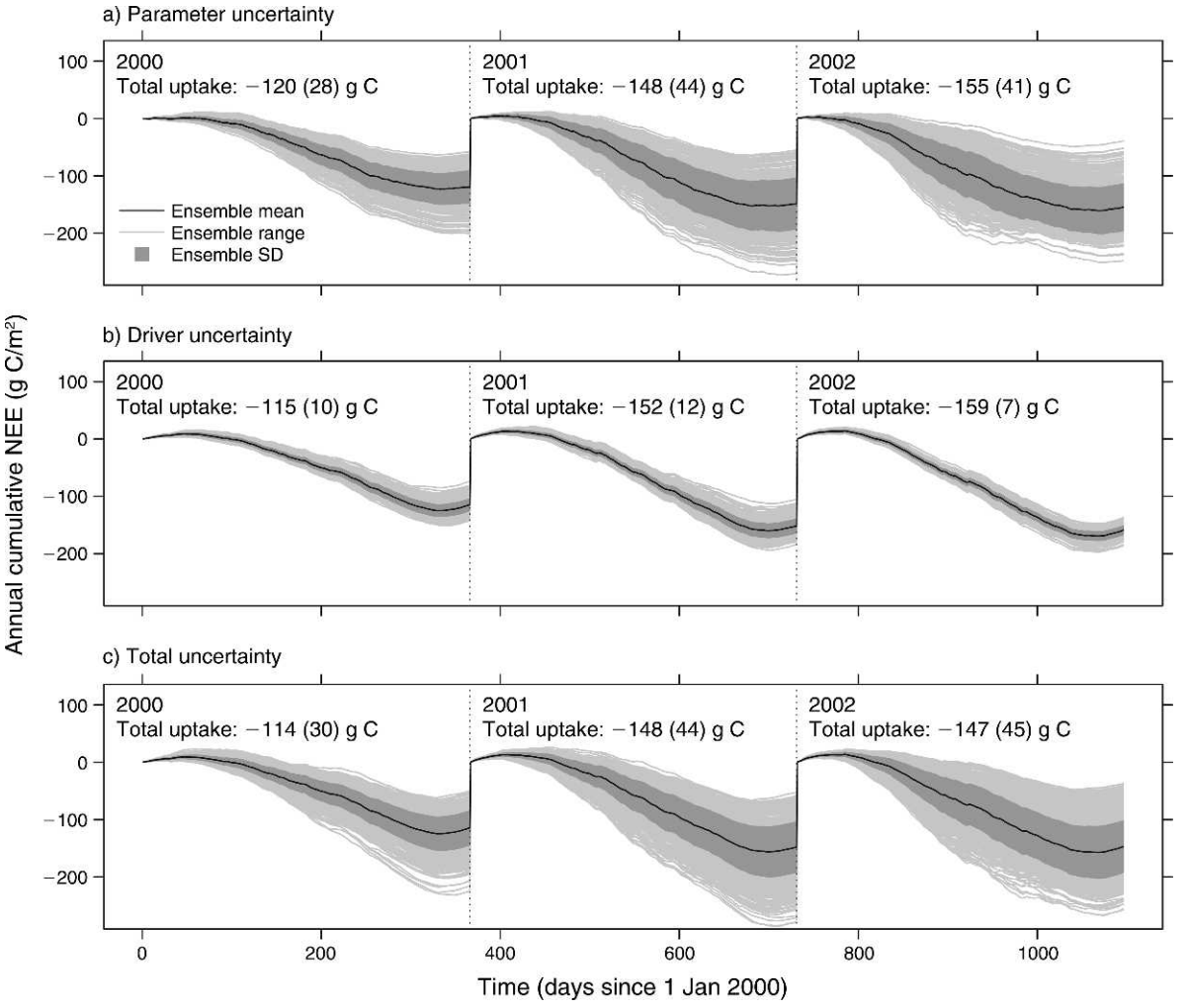

FIG. 9. Cumulative NEE estimates over three years (2000–2002) under different sources of uncertainty. The ensemble mean is indicated in black, while its uncertainty is represented as a dark gray region. The individual ensemble members are indicated in light gray. Ensemble uncertainty is resultant from (a) 375 parameter sets, (b) 1000 geostatistical simulations of meteorology, and (c) 1000 combinations of panels (a) and (b). Mean total uptake for each year is indicated at the bottom of the plots in g C/yr, with standard deviations indicated in parenthesis.

smoothing of the temperature signal (overestimation of mean T_{max}, underestimate of mean T_{max}) resulted in underestimation of C uptake by $\sim$10 g C·m^{-2}·yr^{-1}. These opposing signals act to cancel out when considering the total meteorological uncertainty, resulting in a less biased estimate of total NEE with respect to experiment 1 (observed meteorology), with a small uncertainty (Table 5). Bias issues in the meteorological simulations are a concern, and while in this study the opposing directions of precipitation and temperature bias reduce overall bias, it is not clear whether this was by chance alone. It is likely that the bias cancellation was fortuitous for our study site, and there may be significant bias problems for other locations and ecosystems. A broader study of these bias issues for regional meteorological drivers is thus vital.

Of the meteorological drivers considered, temperature appeared to have the largest impact on NEE uncertainty, as opposed to bias, with approximately twice the influence of precipitation on the signal (Table 5). As such we reject *H*2a, and accept the alternative proposition that instantaneous temperature variability dominates the flux uncertainty. This is likely due to the sensitivity of both GPP and heterotrophic respiration (via a Q_{10} relationship) to daily air temperature in DALEC. Interestingly, decoupling the effect of deriving VPD and daily insolation from temperature drivers in DALEC indicates that indirect estimation of these drivers have a minimal impact on the total NEE.

CONCLUSIONS

We were able to retrieve statistically permissible parameter sets at a data rich location, but still faced

TABLE 6. Effect of increasing data scarcity on NEE uncertainty.

Search threshold (km)	2000	2001	2002	Total	95% CI†
<25	−96 (9)	−129 (10)	−158 (8)	−384 (18)	9
>25	−91 (9)	−128 (10)	−140 (11)	−359 (20)	11
>50	−111 (11)	−156 (11)	−158 (14)	−425 (21)	10
>75	−108 (11)	−153 (12)	−149 (14)	−410 (24)	11
>100	−104 (13)	−162 (17)	−176 (16)	−442 (37)	17

Notes: The search threshold is the distance to the closest station. Values shown are mean NEE/m^2, with standard deviations in parentheses. Simulations were run using the mean parameter values from Table 4. Meteorological data scarcity was increased via a reduction in the proximity of conditioning data through exclusion of data below the search threshold. Simulations were conditioned on the four closest stations for each search threshold.

† The 95% confidence interval of NEE, expressed as a percentage of the total flux.

appreciable uncertainties in flux estimates resultant from parameter uncertainty. As such, spatially explicit modeling exercises may struggle to characterize the regional flux without considerable fieldwork, or investment in remote sensing methodologies to retrieve well-constrained parameter sets for the region of interest. Modeling the young ponderosa pine site at Metolius is challenging, because the system is aggrading rapidly. Observed annual increases in LAI result in increasing rates of C cycling. So the model parameterization must be able to allocate C to grow the plant tissues realistically. Further, parameters change with stand age and its associations with root access to soil water (Irvine et al. 2004, Schwarz et al. 2004).

We found considerable variability in simulated driver trajectories resulted in a small contribution to the net uncertainty. Issues of bias in meteorological upscaling are of much greater concern, but seemed to cancel out over time when propagated through the model. It is likely that the cancellation of bias due to temperature and precipitation is by chance alone, and further research into issues of bias in driver fields is warranted. In areas with very sparse meteorological stations (>100 km separation), then uncertainty in meteorological drivers becomes a more significant problem. The time scales on which precipitation errors accumulated in the simulations were shorter than the temporal scales over which drought operated in the model, and so there was little likelihood of misspecifying drought events.

We have presented a robust analysis of the relative magnitude of parameterization and driver errors using novel techniques. Quantification of the uncertainty associated with regionalized meteorological fields at relevant resolutions for catchment scale studies has been presented for the first time, and represents a key step in the application of data assimilation approaches on the catchment scale. Improved model parameterizations and calculations of bias in meteorological fields are a research priority for spatially explicit regional modeling exercises, especially where data may be sparse. Climate forecasts produce mean meteorological values for discrete grid-cells across a landscape. These mean values can be used to drive ecological forecasts at a similar spatial resolution. Our results here, using local meteorological data, emphasize the problems associated with ecological modeling using simple characterizations of landscape meteorology (i.e., sparse data), particularly in areas with complex terrain. Climate forecast data, used as drivers in ecological forecasts, need to be carefully assessed for non-linear effects. For instance, the mean grid cell meteorology may not produce the same ecosystem response as the mean output of ecosystem responses based on a statistical downscaling of the mean climate to the finer resolution of stand scales (Williams et al. 2001*b*) and an ensemble of model simulations. The geostatistical approach outlined here provides a means to assess such sub-grid scale variations in meteorology and to quantify their ecological effects. Ecological forecasts using gridded climate data will, of course, include an additional bias due to forecast error.

ACKNOWLEDGMENTS

This research was supported by an NERC studentship for L. Spadavecchia. The Metolius AmeriFlux research was supported by the U.S. Department of Energy Terrestrial Carbon Program (DOE grant #DE-FG02-06ER64318). Special thanks to James Irvine and Meredith Kurpius for supplying flux and respiration data, and all at the Oregon Climate Service for useful advice on meteorological data sources. Thanks also to Martin De Kauwe, who provided useful remarks on the manuscript. J. P. Gosling also provided many useful remarks on the methodology. We also thank all at the Sisters Forest Rangers Station, especially Pam Fahey, who provided the vegetation classification data, and two anonymous reviewers for their constructive comments.

LITERATURE CITED

Allen, R. 1997. Self-calibrating method for estimating solar radiation from air temperature. Journal of Hydrological Engineering 2:56–67.

Ashraf, M., J. C. Loftis, and K. G. Hubbard. 1997. Application of geostatistics to evaluate partial weather station networks. Agricultural and Forest Meteorology 84:255–271.

Canham, C. D. W., J. Cole, and W. K. Lauenroth. 2003. Models in ecosystem science. Princeton University Press, Princeton, New Jersey, USA.

Daly, C., R. P. Neilson, and D. L. Phillips. 1994. A statistical topographic model for mapping climatological precipitation over mountainous terrain. Journal of Applied Meteorology 33:140–158.

Deutsch, C. V., and A. G. Journel. 1998. GSLIB: geostatistical software library and users guide. Second edition. Oxford University Press, New York, New York, USA.

Eknes, M., and G. Evensen. 2002. An ensemble Kalman filter with a 1-D marine ecosystem model. Journal of Marine Science 36:75–100.

Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to

forecast error statistics. Journal of Geophysical Research 99:10143–10162.

Evensen, G. 2003. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dynamics 53:343–367.

Falge, E., D. Baldocchi, and R. J. Olson. 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. Agricultural and Forest Meteorology 107:43–69.

Fisher, R. A., M. Williams, M. D. Ruivo, A. L. de Costa, and P. Meira. 2008. Evaluating climatic and soil water controls on evapotranspiration at two Amazonian rainforest sites. Agricultural and Forest Meteorology 148:850–861.

Fox, A. M., M. Williams, A. D. Richardson, D. Cameron, J. Gove, M. Reichstein, T. Quaife, D. Ricciuto, E. Tomelleri, C. M. Trudinger, and M. T. Van Wijk. 2009. The REFLEX project: comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data. Agricultural and Forest Meteorology 149:1597–1615.

Fuentes, M., T. G. F. Kittel, and D. Nychka. 2006. Sensitivity of ecological models to their climate drivers: statistical ensembles for forcing. Ecological Applications 16:99–116.

Goovaerts, P. 1997. Geostatistics for natural resources evaluation. Oxford University Press, Oxford, UK.

Goovaerts, P. 1999. Geostatistics in soil science: state-of-the-art and perspectives. Geoderma 89:1–45.

Goovaerts, P. 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology 228:113–129.

Goovaerts, P. 2001. Geostatistical modelling of uncertainty in soil science. Geoderma 103:3–26.

Heuvelink, G. B. M., and R. Webster. 2001. Modelling soil variation: past, present, and future. Geoderma 100:269–301.

Hudson, G., and H. Wackernagel. 1994. Mapping temperature using kriging with external drift—theory and an example from Scotland. International Journal of Climatology 14:77–91.

Hungerford, R. D., R. R. Nemani, S. W. Running, and J. C. Coughlan. 1989. MTCLIM: a mountain microclimate simulation model. U.S. Forest Service Research Paper INT-414. U.S. Forest Service, Fort Collins, Colorado, USA.

Irvine, J., and B. E. Law. 2002. Contrasting soil respiration in young and old-growth ponderosa pine forests. Global Change Biology 8:1183–1194.

Irvine, J., B. E. Law, M. R. Kurpius, P. M. Anthoni, D. Moore, and P. A. Schwarz. 2004. Age-related changes in ecosystem structure and function and effects on water and carbon exchange in ponderosa pine. Tree Physiology 24:753–763.

Kennedy, M., C. Anderson, A. O'Hagan, M. Lomas, I. Woodward, J. P. Gosling, and A. Heinemeyer. 2008. Quantifying uncertainty in the biospheric carbon flux for England and Wales. Journal of the Royal Statistical Society Series A-Statistics in Society 171:109–135.

Klemedtsson, L., P. -E. Jansson, D. Gustafsson, L. Karlberg, P. Weslien, K. von Arnold, M. Ernfors, O. Langvall, and A. Lindroth. 2008. Bayesian calibration method used to elucidate carbon turnover in forest on drained organic soil. Biogeochemistry 89:61–79.

Law, B. E., A. H. Goldstein, P. M. Anthoni, M. H. Unsworth, J. A. Panek, M. R. Bauer, J. M. Fracheboud, and N. Hultman. 2001*a*. Carbon dioxide and water vapor exchange by young and old ponderosa pine ecosystems during a dry summer. Tree Physiology 21:299–308.

Law, B. E., F. M. Kelliher, D. D. Baldocchi, P. M. Anthoni, J. Irvine, D. Moore, and S. Van Tuyl. 2001*b*. Spatial and temporal variation in respiration in a young ponderosa pine forests during a summer drought. Agricultural and Forest Meteorology 110:27–43.

Law, B. E., O. J. Sun, J. Campbell, S. Van Tuyl, and P. E. Thornton. 2003. Changes in carbon storage and fluxes in a chronosequence of ponderosa pine. Global Change Biology 9:510–524.

Law, B. E., P. Thornton, J. Irvine, S. Van Tuyl, and P. Anthoni. 2001*c*. Carbon storage and fluxes in ponderosa pine forests at different developmental stages. Global Change Biology 7:755–777.

Law, B. E., S. Van Tuyl, A. Cescatti, and D. D. Baldocchi. 2001*d*. Estimation of leaf area index in open-canopy ponderosa pine forests at different successional stages and management regimes in Oregon. Agricultural and Forest Meteorology 108:1–14.

Maybeck, P. S. 1979. Stochastic models, estimation and control. Volume 1. Academic Press, New York, New York, USA.

Murray, F. W. 1967. On the computation of saturation vapor pressure. Journal of Applied Meteorology 6:203–204.

Running, S. W. 1994. Testing forest–BGC ecosystem process simulations across a climatic gradient in Oregon. Ecological Applications 4:238–247.

Running, S. W., R. R. Nemani, and R. D. Hungerford. 1987. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating evapotranspiration and photosynthesis. Canadian Journal of Forest Research 17:472–483.

Runyon, J., R. H. Waring, S. N. Goward, and J. M. Welles. 1994. Environmental limits on net primary production and light-use efficiency across the Oregon transect. Ecological Applications 4:226–237.

Saxton, K. E., W. J. Rawls, J. S. Romberger, and R. I. Papendick. 1986. Estimating generalized soil-water characteristics from texture. Soil Science Society of America Journal 90:1031–1036.

Schwarz, P. A., B. E. Law, M. Williams, J. Irvine, M. Kurpius, and D. Moore. 2004. Climatic versus biotic constraints on carbon and water fluxes in seasonally drought-affected ponderosa pine ecosystems. Global Biogeochemical Cycles 18:GB4007.

Spadavecchia, L., and M. Williams. 2009. Can spatio-temporal geostatistical methods improve high resolution regionalisation of meteorological variables? Agricultural and Forest Meteorology 149:1105–1117.

Thiessen, A. H. 1911. Precipitation averages for large areas. Monthly Weather Review 39:1082–1089.

Thornton, P. E., S. W. Running, and M. A. White. 1997. Generating surfaces of daily meterological variables over large regions of complex terrain. Journal of Hydrology 190:214–251.

Van Oijen, M., J. Rougier, and R. Smith. 2005. Bayesian calibration of process-based forest models: bridging the gap between models and data. Tree Physiology 25:915–927.

Van Tuyl, S., B. E. Law, D. P. Turner, and A. I. Gitelman. 2005. Variability in net primary production and carbon storage in biomass across Oregon forests—an assessment integrating data from forest inventories, intensive sites, and remote sensing. Forest Ecology and Management 209:273–291.

Verbeeck, H., R. Samson, F. Verdonck, and R. Lemeur. 2006. Parameter sensitivity and uncertainty of the forest carbon flux model FORUG: a Monte Carlo analysis. Tree Physiology 26:807–817.

Vickers, D., C. Thomas, and B. E. Law. 2009. Random and systematic CO_2 flux sampling errors for tower measurements over forests in the convective boundary layer. Agricultural and Forest Meteorology 149:73–83.

Williams, M., B. E. Law, P. M. Anthoni, and M. Unsworth. 2001*a*. Use of a simulation model and ecosystem flux data to examine carbon–water interactions in ponderosa pine. Tree Physiology 21:287–298.

Williams, M., E. B. Rastetter, D. N. Fernandes, M. L. Goulden, G. R. Shaver, and L. C. Johnson. 1997.

Predicting gross primary productivity in terrestrial ecosystems. Ecological Applications 7:882–894.

Williams, M., E. B. Rastetter, D. N. Fernandes, M. L. Goulden, S. C. Wofsy, G. R. Shaver, J. M. Melillo, J. W. Munger, S. -M. Fan, and K. J. Nadelhoffer. 1996. Modelling the soil–plant–atmosphere continuum in a *Quercus–Acer* stand at Harvard Forest: the regulation of stomatal conductance by light, nitrogen and soil/plant hydraulic properties. Plant, Cell and Environment 19:911–927.

Williams, M., E. B. Rastetter, G. R. Shaver, J. E. Hobbie, E. Carpino, and B. L. Kwiatkowski. 2001*b*. Primary production in an arctic watershed: an uncertainty analysis. Ecological Applications 11:1800–1816.

Williams, M., et al. 2009. Improving land surface models with FLUXNET data. Biogeosciences 6:1341–1359.

Williams, M., P. Schwarz, B. E. Law, J. Irvine, and M. R. Kurpius. 2005. An improved analysis of forest carbon dynamics using data assimilation. Global Change Biology 11:89–105.

APPENDIX

Meteorological calculations and data sets (*Ecological Archives* A021-070-A1).