Ecological Applications, 25(2), 2015, pp. 546–558

Model-data assimilation of multiple phenological observations to constrain and predict leaf area index

Toni Viskari,[1,3] Brady Hardiman,[1] Ankur R. Desai,[2] and Michael C. Dietze[1]

[1]*Department of Earth and Environment, Boston University, Boston, Massachusetts 02215 USA*
[2]*Department of Atmospheric and Oceanic Sciences, University of Wisconsin, Madison, Wisconsin 53706 USA*

Abstract. Our limited ability to accurately simulate leaf phenology is a leading source of uncertainty in models of ecosystem carbon cycling. We evaluate if continuously updating canopy state variables with observations is beneficial for predicting phenological events. We employed ensemble adjustment Kalman filter (EAKF) to update predictions of leaf area index (LAI) and leaf extension using tower-based photosynthetically active radiation (PAR) and moderate resolution imaging spectrometer (MODIS) data for 2002–2005 at Willow Creek, Wisconsin, USA, a mature, even-aged, northern hardwood, deciduous forest. The ecosystem demography model version 2 (ED2) was used as the prediction model, forced by offline climate data. EAKF successfully incorporated information from both the observations and model predictions weighted by their respective uncertainties. The resulting estimate reproduced the observed leaf phenological cycle in the spring and the fall better than a parametric model prediction. These results indicate that during spring the observations contribute most in determining the correct bud-burst date, after which the model performs well, but accurately modeling fall leaf senesce requires continuous model updating from observations. While the predicted net ecosystem exchange (NEE) of CO_2 precedes tower observations and unassimilated model predictions in the spring, overall the prediction follows observed NEE better than the model alone. Our results show state data assimilation successfully simulates the evolution of plant leaf phenology and improves model predictions of forest NEE.

Key words: data assimilation; ecosystem demography model; ED2; leaf-out; phenology; prediction; senescence; terrestrial ecosystem models.

Introduction

Leaf phenological cycles of budburst, leaf expansion, senescence, and leaf drop are a major control on growing season length and thus have a large impact on net ecosystem productivity, carbon allocation patterns, and the annual oscillation of atmospheric CO_2 (Richardson et al. 2013). The presence or absence of leaves drives local radiation budgets and micrometeorological variables by changing the turbulent, radiative, and consequently, thermal properties within the canopy (Moore et al. 1995). The annual cycle of spring emergence and fall senescence in deciduous forests is affected by plant traits, as well as ambient atmospheric and soil conditions, with temperature having a large impact on the timing of these transitions (White et al. 1997). Phenology also varies from tree to tree with a large interannual and spatial variation (Richardson et al. 2010). Currently, predicting how phenological cycles will react to climate change is a large source of uncertainty in ecosystem carbon flux projections (Richardson et al. 2012).

Manuscript received 12 March 2014; revised 23 June 2014; accepted 4 August 2014. Corresponding Editor: D. S. Schimel.
[3] E-mail: toni.viskari@fmi.fi

Climate and ecological processes are strongly coupled, and models predicting phenological cycles must account for this coupling (Levis and Bonan 2004). Approaches to modeling phenology range widely from determining parameters by fitting curves to observations (Bradley et al. 2007) to estimating a temperature period that triggers phenological changes (Zhang et al. 2004). However, a unifying feature of models based on these approaches, and included in most ecosystem models, is their tendency to have trouble predicting both the start and the length of the growing season (Kucharik et al. 2006, Richardson et al. 2012). Since phenological cycles have high interannual variability and a complex coupling with the climate that is not yet fully understood, phenological parameters determined by fitting data to models have had limited utility in predicting future phenological cycles (Richardson et al. 2012). Thus accurately predicting phenology, even if the relationship between phenology and temperature is improved, requires a method that updates model states or parameters with observations.

Data assimilation provides a potential means for constraining phenology in ecosystems models by synthesizing information from multiple different data sources. In doing so, data assimilation creates a statistically optimal estimate based on the relative

uncertainties of the different data sources and the model. Data assimilation is widely used in the geosciences (Bertino et al. 2003, Chen 2011, Viskari et al. 2012) and is an essential part of weather forecasting (Lorenc 1986, Rabier et al. 2000). The use of data assimilation in ecological modeling is relatively new but is increasing rapidly (McKane et al. 1997, Raupach et al. 2005, Williams et al. 2005, Zobitz et al. 2011, Dietze et al. 2013).

Data assimilation techniques can be divided between parameter and state data assimilation approaches. In parameter data assimilation, model parameters are estimated by comparing the differences between model output and observations (e.g., Braswell et al. 2005). The use of parameter data assimilation has to date been more common in ecological systems. For example, Jeong et al. (2012) and Medvigy et al. (2013) used parameter data assimilation to illustrate how the phenological cycle of photosynthetic capacity affects gross primary productivity (GPP). Parameter data assimilation provides insight into processes and especially to the coefficients describing these processes, but assumes that the equations are sufficiently accurate descriptions of the processes in question. In addition, predictions made with coefficients estimated from parameter data assimilation cannot eliminate the propagating uncertainty that stems from unexplained variability and uncertainty in the initial conditions. Therefore, for ecological processes where current models have large residual error, parameter data assimilation is unlikely to lead to noticeable improvements in predictive capacity.

While previous phenological modeling has focused on parameter estimation, the substantial interannual variability in phenological processes and the limited success of phenology models suggests that it may be more effective to focus on the estimation of the status of the canopy itself. In state data assimilation, the state variables (e.g., quantity of leaves present) are estimated by updating state predictions with observational data (Montzka et al. 2012). By continuously updating the state predictions with observation data, state data assimilation creates a more certain estimate of the current, which reduces the uncertainty in subsequent predictions. Thus, if a reliable descriptive model of the system is available, state data assimilation provides better predictions than parameter data assimilation.

While state data assimilation has been implemented for constraining the carbon cycle (Kaminski et al. 2013, Zhou et al. 2013) and the hydrological cycle (Seo et al. 2009), it is not yet as widely used as parameter data assimilation and to our knowledge has not previously been applied to phenological modeling. Additionally, the majority of state data assimilation examples above have focused on single variable state estimation instead of simultaneously estimating multiple variables using multiple data constraints. It is important to note, though, that while both approached are considered data assimilation, the challenges in implementing and using state data assimilation differ from those of parameter data assimilation.

We applied state data assimilation to predict the phenological cycle for a well-studied, mature, even-aged, northern hardwood, deciduous forest in northern Wisconsin using the ecosystem demography 2 model (Moorcroft et al. 2001, Medvigy et al. 2009). As variables associated with phenological events have large inherent variability and measurement uncertainty, our primary motion was to find out if state data assimilation is a more suitable approach for estimating of the phenological cycle than fitting individual parameters. In order to predict both spring and fall phenological transitions, two phenological state variables, leaf area index (LAI) and percent leaf elongation were estimated, along with associated parameters representing the spring and fall days of year when the canopy is at 50% peak LAI. The interaction of phenology with LAI is often ignored (Asner et al. 2003), so an additional objective of our study is to clarify this relationship and determine whether different LAI observations can be considered measurements of the same variable. Finally, we examined what challenges need to be addressed before implementing a functional operational state data assimilation system for phenological systems. To demonstrate an approach that could be applied in real time to forecast phenological changes, we only used observations current or preceding the state prediction instead of smoothing the prediction to encompass both past and future data. Observations from two different data sets were used to determine their impact on the phenological forecast and to evaluate the benefits from fusing multiple data sources.

Methods

Data assimilation

The Kalman filter (KF; Kalman 1960, du Plessis 1967) is a sequential state data assimilation method, where a state estimate is created by combining information from new observations and previous states. For a forward filtering problem, where observations ahead in time are not available, KF produces a statistically optimal solution. KF is essentially an iterative two-step process: (1) during the state propagation step (forecast step), the process model evolves both the state and the associated uncertainty to the next observation time; and (2) during the observation updating step (analysis step), a new state estimate and an associated uncertainty is determined from the available information sources weighted by their individual uncertainties

These two steps are applied iteratively as new observations become available. Thus, the state estimate is propagated to the next observation time during the first step, and then the state estimate for that time is updated based on new observations. The propagated forecast state, also referred to as the background state,

contains information from prior observations propagated to the current observation time. By using a dynamical forecast model instead of a simple diagnostic model, the forecast is constrained by the current understanding of the system dynamics and evolves over time. An ecological example of the Kalman filter is CarbonTracker, which constrains GPP and plant respiration using CO_2 mole fraction observations (Peters et al. 2007).

A general challenge for KF applications is estimating the uncertainties in the prediction, which can be both computationally and methodologically difficult to approximate. For this application we chose the ensemble adjustment Kalman filter (EAKF; Anderson 2010) from among a number of variants of KF as the data assimilation technique. The EAKF is similar to the more well-known ensemble Kalman filter (EnKF, Evensen 1994, Evensen 2003) in which an ensemble of model runs, each with a different state vector, is used to propagate uncertainty to the next observation time. Ensemble methods are an alternative to transforming the matrix of state estimate uncertainty as in extended nonlinear versions of the KF technique (Jazwinski 1970, Gelb 1974). Compared to analytical approaches to KF, which involve additional computational burdens, such as solving for the Jacobian matrix of the model, ensemble methods only require the nonlinear forward model, as the nonlinear model components affect the prediction uncertainty. In both the EnKF and EAKF, the forecast state is determined by the sample mean and covariance over the ensemble. However, the methods differ in how they update the new state in the analysis step based on observation. In EnKF, the new states are stochastically sampled from the estimated analysis state distribution, while in EAKF, the updated state vectors are instead each linearly adjusted so that the ensemble mean and variance matches the mean and variance of the estimated state distribution. In this way the information from the previous state vectors is not lost during the state update phase. The associated background state uncertainty is calculated from the spread in the state vector values.

The EAKF used in this analysis was implemented within a larger workflow structure provided by the data assimilation research testbed (DART; Anderson et al. 2009). DART is an open source EAKF framework developed at National Center for Atmospheric Research (NCAR) that allows users to insert a model in a few simple steps and explore a wide array of possible data assimilation approaches. DART is widely used, well documented, and freely *available online*.[4] All the interfaces required to couple the ecosystem demography model version 2 (ED2) model to DART are part of the predictive ecosystem analyzer (PEcAn) ecoinformatics workflow (Dietze et al. 2013, LeBauer et al. 2013; workflow *available online*).[5]

The state vector in EAKF, and in KF in general, can contain both observed and unobserved variables. Observed variables are adjusted by direct comparison with observations. Unobserved variables are adjusted based on their correlation with observed variables using a linear regression model fit across the ensemble of model runs. As all the chosen variables in this analysis are from the same location and are related, there was no need to limit the error covariances between different vector members (Hamill et al. 2001, Whitaker and Hamill 2002). In order to simplify the initial testing, the uncertainties were not inflated (Anderson and Anderson 1999, Anderson 2009).

A state data assimilation application requires three components: (1) observations of the state necessary for the observation update, (2) a forward model, which predicts the state for the next observation time, and (3) an observation operator, which translates the model state into an equivalent observation. We will describe each of these. The observation operators are discussed along with the respective observations.

Observations

Observations from Willow Creek (US-WCr) eddy covariance tower (45.8059° N, 90.0799° W) in the Chequamegon-Nicolet National Forest in Northern Wisconsin, USA (Cook et al. 2004) were used to test the method. This site is part of the larger Chequamegon Ecosystem Atmosphere study project (ChEAS; Chen et al. 2008) and is part of the AmeriFlux network (Baldocchi et al. 2001). Willow Creek is a mature, even-aged, northern hardwood, deciduous forest dominated by sugar maple, with smaller quantities of ash and basswood. The average stand age is approximately 80 years. Eddy covariance measurements have provided near-continuous observations of half-hourly CO_2 and H_2O exchange between the forest and atmosphere since 1999. We used observations from years 2002–2006, for which both phenological and meteorological measurements were available.

Changes in leaf area index (LAI, one-sided projected/silhouetted green leaf area per unit ground surface area sensu definition four in Asner et al. [2003]) are strongly dependent on phenology, and thus, LAI was the target variable observed in this study. Measurements of LAI, especially those employing optical methods, can be subject to numerous sources of observation error (Asner et al. 2003). Much of this uncertainty is mitigated by our selection of a forest with relatively uniform composition (temperate, broadleaf, deciduous). This uniformity contributes to a more sharply defined phenological transition period and limits the interacting sources of uncertainty present in LAI estimates in forests contain-

[4] http://www.image.ucar.edu/DAReS/DART/

[5] http://github.com/PecanProject

ing both deciduous/evergreen and broadleaf/needleleaf species (Asner et al. 2003). It should be noted that the maximum and minimum LAI depend on the size, density, and species of the vegetation rather than phenology. Thus state data assimilation will not affect those values. We used two methods of observing LAI: areal-averaged measurements from satellites and point measurements of intercepted radiation from a tower.

MODIS measurements.—Moderate resolution imaging spectrometer (MODIS; Barnes et al. 1998) measures reflected solar and emitted thermal radiation on 36 separate spectral bands. The measurements used here are from NASA's Terra and Aqua satellites, which have a swath of 2330 km and produce a global coverage every one to two days. We used the MODIS MOD15A2 LAI product (Yang et al. 2006*a*), which has a resolution of 1 km and is combined into 8-d averages. In generating the data product, the daily observations are screened based on cloud cover and other quality assurance/quality control (QAQC) criteria, so an 8-d average does not necessarily contain observations from all 8 d. The default MODIS LAI QAQC provided by the product was used as the quality control metric. The MODIS-derived LAI estimates from hereafter are referred to as MODIS LAI.

The MODIS observations used were for the single pixel corresponding to the tower location. Since the MODIS product output is the LAI value, no separate observation operator was necessary to simulate observations from model variables for the data assimilation process. However, this prevents the examination of different time windows than those provided and the independent estimation of associated uncertainties. Here, MODIS uncertainties were taken from the given standard deviations given by the MODIS product unless it was smaller than the threshold uncertainty 0.66 m^2/m^2 in which case the threshold uncertainty was used (Yang et al. 2006*b*; see statement *available online*).[6] This uncertainty, however, is the averaged uncertainty over all land cover types. It is not known how well this value applies to the deciduous forests of our study area as, to our knowledge, there have been no further studies assessing MODIS LAI uncertainty. Further, MODIS estimates that a fraction of the forest in this pixel is conifer, as indicated by the nonzero wintertime LAI, yet census data of plots surrounding the tower demonstrate a only a minor conifer component, adding to uncertainty in MODIS LAI values.

Flux tower radiation measurements.—Since LAI is not measured directly from the flux towers, we used paired above-canopy (30 m) and below-canopy (2 m) measurements of photosynthetically active radiation (PAR; 400–700 Nm solar radiation; LI-190S Quantum Sensor, LI-COR, Lincoln, Nebraska, USA). The LAI estimated from the PAR sensors is referred to as tower LAI.

PAR at the lowest level of the canopy (I_{2m}) can be modeled as (Monsi and Saeki 2005)

$$I_{2m} = I_{30m} e^{-k(m+LAI)} \tag{1}$$

where I_{30m} is above canopy PAR, m represents the area of branches and other nonphotosynthetic canopy surfaces, and k represents the extinction coefficient. From Eq. 1, it is straightforward to estimate flux tower LAI (LAI_F) as a function of the fraction of I_{2m} over I_{30m}

$$LAI_F = \frac{-\log \dfrac{I_{2m}}{I_{30m}}}{k} - m \tag{2}$$

where k was set to 0.52 based on calibrations done in the region (Cook et al. 2008), and m was approximated to be 1.45 ± 0.27 m^2/m^2 (mean ± SD) using the average beginning of the year (prebud-burst period; January, February, March) radiation measured by the flux tower in 2002–2005 and assuming LAI_F is zero over that period. This assumption is reasonable given the lack of evergreen species within the field of view of the below canopy PAR sensors. The standard deviation for k was approximated as 0.05 units based on Chen and Black (1991).

As a nonlinear function of two separate observations (I_{2m} and I_0), tower LAI cannot be linearly obtained from the model output and thus was precomputed for each measurement time rather than incorporated into the observation operator. The observation operator averages the LAI_F over a chosen time window and only includes measurements midday (12:00 local time). In addition, only I_{2m} and I_{30m} values that were above cut-off values of 10 and 1000 $\mu mol \cdot m^2 \cdot s$, respectively, were included in the average. The cut-off values were chosen to be high enough to eliminate cloudy days. The observation uncertainty for PAR was approximated as having a coefficient of variation of ~10% from the AmeriFlux quality assurance measurement reports (Boden et al. 2013). The compound uncertainty for a single tower LAI measurement is independently calculated for each time from the uncertainties of PAR, m, and k with the combined uncertainty formula

$$\sigma_{LAI} = \sqrt{\left[\frac{-\log \dfrac{I_{2m}}{I_0}}{k}\sqrt{\left(\frac{\sqrt{\left(\dfrac{\sigma_{I_{2m}}}{I_{2m}}\right)^2 + \left(\dfrac{\sigma_{I_{30m}}}{I_{30m}}\right)^2}}{-\log \dfrac{I_{2m}}{I_{30m}}}\right)^2 + \left(\frac{\sigma_k}{k}\right)^2}\right]^2 + \sigma_m^2} \tag{3}$$

where σ_{LAI}, σ_{I2m}, σ_{I30m}, σ_k, and σ_m are standard deviations associated with LAI_F, I_{2m}, I_{30m}, k, and m, respectively. The uncertainty for the averaged tower measurements (σ_{flux}) is

[6] http://landval.gsfc.nasa.gov/ProductStatus.php?ProductID=MOD15

$$\sigma_{\text{flux}} = \sqrt{\frac{\sum_{i=1}^{N}\sigma_{\text{LAI},i}^2 + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N} r_{i,j}\sigma_{\text{LAI},i}^2}{N}} \quad (4)$$

where N is the number of measurements included in the average, $r_{i,j}$ is the autocorrelation coefficient between two observations and i and j are the running observation indexes for the averaging window. By calculating the autocorrelation function over summer months, the autocorrelation was approximated as 0.15 for a lag of 1 d and 0 after that.

Forward model

The ecosystem demography model version 2 (ED2; Medvigy et al. 2009) is a terrestrial biosphere model and was used as the forward model to predict LAI, as well as fluxes of carbon, water, and energy. ED2 uses a size- and age-structured approximation to incorporate both stand-scale vertical canopy structure and landscape-scale heterogeneity while still representing regional-scale dynamical processes. Both the climate forcing and initial tree distribution are read in to the model off-line based on micrometerological measurements made from the flux tower and vegetation census data from the tower footprint, respectively (measurements *available online*).[7]

The prescribed phenology subroutine in ED2 (Medvigy et al. 2009), calculates a green leaf factor (GLF), which represents the elongation fraction of the leaves, from 0 to 1, according to

$$\text{GLF} = \frac{1}{1 + (a \times \text{doy})^b} \quad (5)$$

where doy is the day of year, a is the reciprocal of the day of year at which 50% canopy expansion occurs, and b is an exponent controlling the rate of change. The parameters a and b were empirically derived from MODIS phenology observations presented in Zhang et al. (2003) and depend both on the year and the season. The model switches between spring and fall phenological schemes on the first day of January and August, respectively. In ED2, LAI is a product of leaf biomass, specific leaf area, and stem density. Leaf biomass in ED2 is an empirical allometric function of dbh (1.4 m) scaled by leaf elongation, thus the LAI predicted by ED2 is directly proportional to GLF.

In Eq. 5, GLF is an emergent variable, as the state at one point in time is not directly affected by preceding values. For data assimilation purposes, the formulation was changed to generate a dynamical model that still retains the original prediction when run outside of data assimilation. To do this we adjust the leaf expansion rate parameter (b) as

$$b_{t+\Delta t} = \frac{1}{\log(a \times (\text{doy} - 1))} \log\left(\frac{1}{\text{GLF}(t)} - 1\right) \quad (6)$$

$$\text{GLF}(t + \Delta t) = \frac{1}{1 + (a \times \text{doy})^{b_{t+\Delta t}}} \quad (7)$$

where t represents time and Δt is the time step. This formulation produces the same GLF values as Eq. 5, but makes it a function of the GLF at the previous time. This is crucial for the data assimilation application, as now GLF(t) affects GLF($t + \Delta t$). Due to issues with numerical instability, Eqs. 6 and 7 were only used if GLF was larger than 0.0001 and smaller than 0.9999. Otherwise, GLF was determined from Eq. 5. Altering GLF in Eqs. 6 and 7 affects subsequent LAI and b_t. LAI predicted by the model without assimilation observations shall be referred to as model LAI. LAI produced with data assimilation will be referred to as estimated LAI.

Testing setup

The state vector (x) for DART includes three variables: LAI, GLF, and parameter a from Eqs. 6 and 7. It is important to note that DART uses ED2 to predict LAI at the observation time by propagating the previous state estimate. The initial spring ensemble values for a were drawn from a normal distribution with a mean of 0.007362885 d^{-1} (~16 May) and a standard deviation of 0.0003776807 d^{-1} (~7 d) calculated from offline ED2 run over the study period (Zhang et al. 2003). Both LAI and GLF are initially set to 0. By affecting the start of the leaf-out during spring and senescence in fall, parameter a generates an initial spread for GLF and LAI in the ensemble. Because a is season-specific, the mean and spread of this parameter is reinitialized to mean of 0.003637635 d^{-1} (~2 October) and a standard deviation of 0.00007665897 d^{-1} (~6 d) when switching between spring and fall phenologies on August 1st. The initial value for b for each year was set as −65 for spring and 58 for fall. These values were chosen averages of the offline ED2 run. Since b is constantly recalculated according to Eq. 6, the initial choice has only a very minor impact on results. This was confirmed by comparing results with varying initial values of b.

The spin-up run for the ED2 variables started from spring 1998. As the LAI observations for the winter months were missing and the state variables change only later in the spring, the ensemble runs were started on 11 April, which was before the beginning of the leaf-out for all values used for a. As the MODIS observation windows always start on 1 January, data were assimilated for each year separately in order to match the measurement windows and the DART background state.

Initially, tower measurements were averaged over the same 8-d time window as the MODIS measurements.

[7] http://flux.aos.wisc.edu/data/cheas/wcreek/flux/

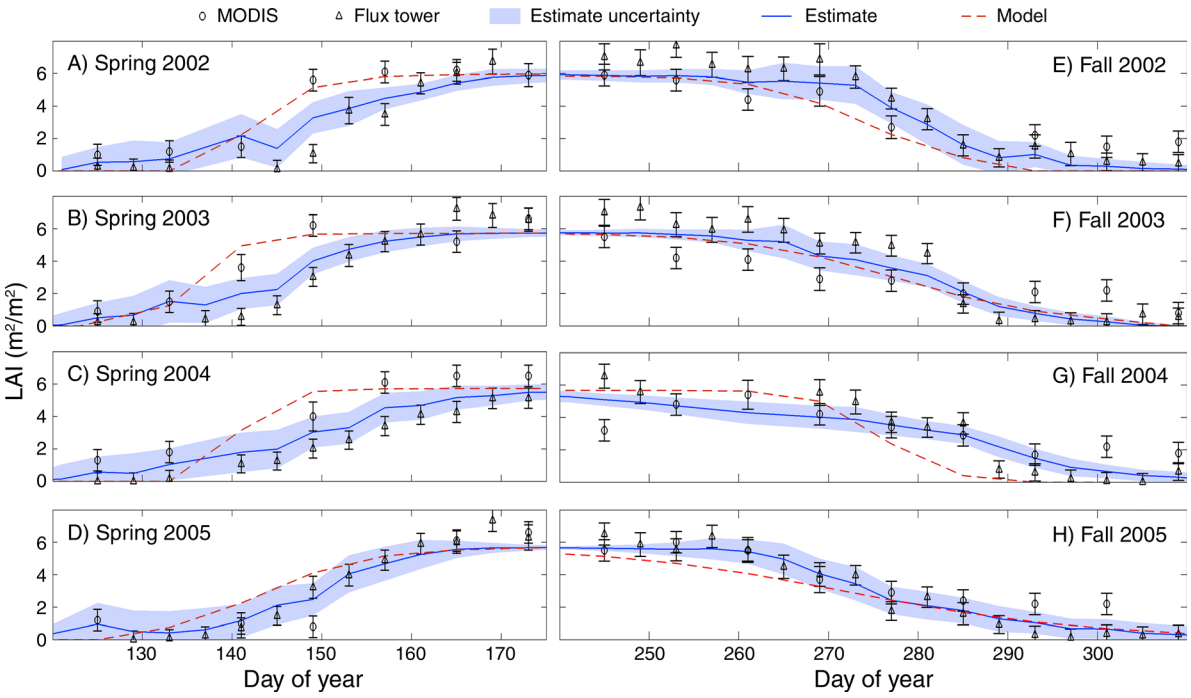

FIG. 1. Leaf area index (LAI) as observed by moderate resolution imaging spectrometer (MODIS; circles), as well as the flux tower (triangles) with an 8-d measurement window, as predicted by the model (dashed red line) and as estimated by data assimilation research testbed (DART; solid blue line) for years 2002–2005. Values are means ± SD. The state estimate uncertainty is presented by the blue shading. Panels (A–D) are for spring leaf-out and panels (E–H) for fall senescence.

The ED2 prediction is always for the day in the middle of the MODIS and tower averaging time window.

Based initial results, we observed that phenological events caused rapid changes in LAI values, especially during the spring leaf-out. To assess whether more frequent observations would improve our ability to capture these transitions, the state data assimilation was repeated with tower measurements averaged over 4-d intervals instead of 8-d intervals. As the MODIS observations used here come from a data product, similar tightening of the time window could not be done with them. The reduced measurement window length consequently increases the tower LAI uncertainty by approximately a factor of square root of 2 as the average was calculated over a smaller sample size, though the exact error estimate depends upon the number of observations meeting QAQC and their magnitude according to Eq. 4.

To determine how changes in the leaf-phenological state affected the carbon, we calculated daily average net ecosystem exchange (NEE). The flux tower NEE measurements (LI-6262, LI-COR, Lincoln, Nebraska, USA) are at half-hour intervals and were averaged from gap-filled measurements over the day if more than 25% of the raw observations for that day were available. Both the estimated and modeled NEE were calculated from the ED2 gross primary production (GPP) minus the sum of the autotrophic and heterotrophic respiration fluxes. The NEE estimate was averaged from the NEEs of individual ensemble members.

RESULTS

Observed tower and MODIS LAI values differed in magnitude and were not fully synchronized in their phenological dynamics (Fig. 1). During spring leaf-out, tower LAI tended to lag behind MODIS LAI by about 7 d. During fall leaf senescence, MODIS LAI declines sooner but slower than tower LAI, although in both cases the most rapid change in the value occurs around the same date. Across the summer there appears no systematic difference between their values. The modeled LAI, though, has a higher peak LAI than observations. During early spring and late fall, MODIS LAI is notably higher than any of the other LAIs.

Springtime modeled LAI increases around the same time as MODIS LAI, preceding both tower and estimated LAI. Modeled leaf-out occurs at a faster rate (greater slope) toward the end of spring, while both observations decelerate after the initial jump in value, which causes modeled LAI to peaks sooner than observed LAIs. After data assimilation, estimated LAI generally shows a small initial increase preceding tower observations of leaf-out, but successfully predicts the timing of the main leaf flush, splitting the difference between the observations. The rate of change for estimated LAI also slows towards the end of spring in accordance with the observations. Overall, the estimated

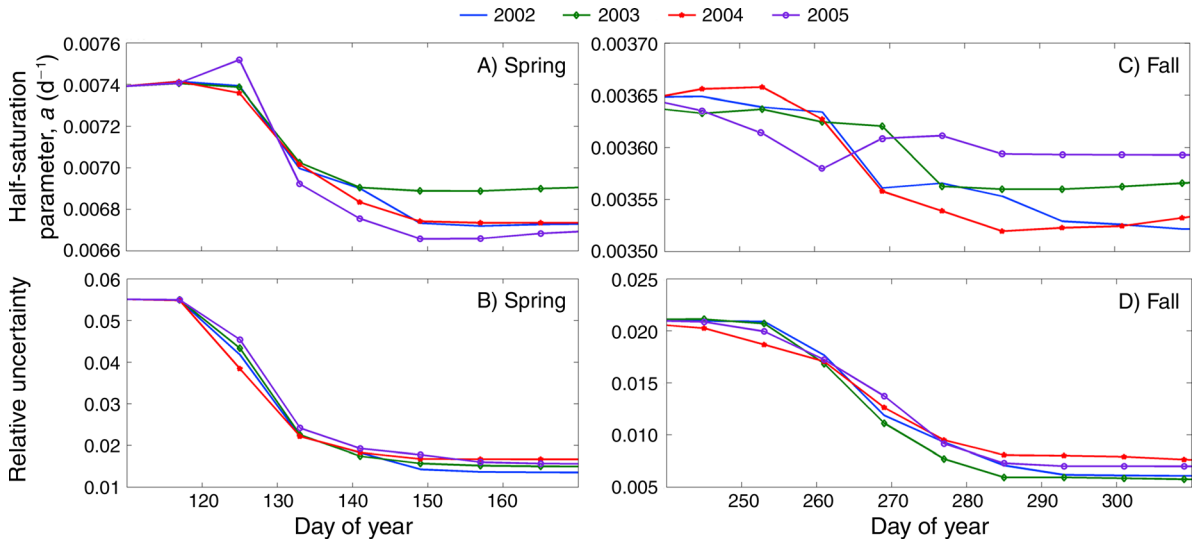

FIG. 2. The evolution of half-saturation parameter a as estimated by DART for years 2002–2005. The flux tower observations were averaged over an 8-d measurement window. Panels (A–C) are the ensemble means for each year. Panels (B and D) show relative uncertainty, calculated as the ensemble spreads (standard deviation of a) divided by the ensemble mean (mean a) for each year. Panels (A and B) are for spring leaf-out and panels (C and D) are for fall senescence.

LAI matches the observed end of leaf out better than the model alone.

During fall, modeled LAI generally begins to decrease abruptly, similar to tower LAI, although modeled LAI lacks the long tail of tower LAI and differs in timing of the main leaf senescence. In contrast, estimated LAI has a longer, slower decrease before a faster drop in value after approximately day of year 270 similar to MODIS LAI. Estimated LAI predicts the timing of the large decrease as observed by both MODIS and the tower better than modeled LAI. With the exception of 2005, the MODIS and the tower observations show notable differences in timing, but the estimated LAI is generally a compromise between them. In 2005, the MODIS and tower LAIs are close to each other both in value and phase.

In addition to estimating state variables LAI and GLF, the assimilation also constrained the parameter a, which approximates the timing of the midpoint for leaf on and for leaf off (Fig. 2). It is important to note that a has a different value in the spring than in the fall. During both spring and fall, the ensemble mean of a initially changes rapidly with the ensemble spread simultaneously decreasing. Eventually, though, all ensemble members for a year converge to around the same static mean a with the spread reducing from ~8 d to ~2 d. According to these results, the state estimate initially changes the leaf-out date considerably until most ensemble members settle around the same date. However, at this point the relative spread for GLF (not shown for clarity) has grown in value and maintains the LAI spread. Finally, since GLF cannot be larger than 1 or smaller than 0, the GLF ensemble values converge at the end of each phenological transition. As LAI is calculated from GLF, this results in consequent convergence of ensemble members.

Notably the pattern of evolution of a differs during the spring and the fall. During leaf-out, the ensemble mean for a always decreases from the initial mean. This suggests that the default parameter values used in the spring were biased toward too early leaf out, which is seen in Fig. 1. In contrast, during the fall senescence, the mean of a remains within the same, narrow range at the beginning and at the end, but varies greatly over the fall instead of plateauing quickly to a certain value. Thus the state estimate and the fitted parameters roughly agree on the period of time when the fall senescence occurs. Additionally, a generally becomes static after about 20 days in the spring compared to the approximately 40 days it takes in the fall.

Reducing the length of the tower averaging window increases both the temporal variability and measurement uncertainty in tower LAI, which is especially obvious at the beginning of the fall senescence (Fig. 3). LAI estimates from more frequent assimilation are more variable and have a larger ensemble spread than the less frequently estimated LAI. On the whole, when using the shorter averaging time window, the estimated LAI is generally closer to tower LAI than when using a longer averaging time window. Using more frequent tower observations, the relative uncertainty for a is generally ~1% smaller at the end of spring and ~0.4% smaller at the end of fall than with less frequent tower observations (Fig. 4).

To determine how the state assimilation affected ecosystem carbon fluxes, daily average NEE was compared both to the ED2 model simulation and to the flux tower observations (Fig. 5). Adjusting LAI, and

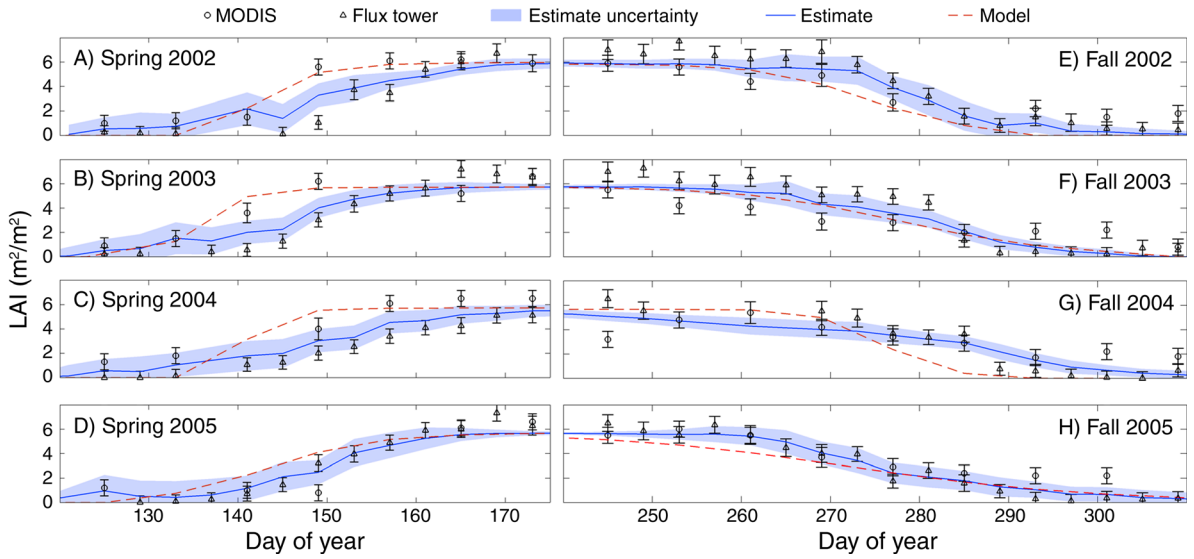

FIG. 3. LAI as observed by MODIS (circles), as well as the flux tower (triangles) with a 4-d measurement window, as predicted by the model (dashed red line) and as estimated by DART (solid blue line) for years 2002–2005. Values are means ± SD. The state estimate uncertainty is presented by the blue shading. Panels (A–D) are for spring leaf-out and panels (E–H) for fall senescence.

consequently GLF, influenced the estimated NEE, especially in the spring and fall. During spring, estimated NEE increased before modeled or observed NEE, even though estimated LAI increased later in the spring than the modelled LAI. The early increase in NEE was observed to be due to an increase in gross primary production (GPP) at low LAI rather than a decrease in ecosystem respiration (Re; figure not shown). However, later in the spring and at the beginning of summer, the estimated NEE is closer to the observed values than the modeled NEE. In the fall, there is no clear difference between how estimated and modeled NEE perform compared to the observed values.

DISCUSSION

This study had four objectives: (1) provide a proof-of-concept for the operational assimilation of phenological data, (2) identify and isolate issues that need to be addressed before full implementation of data assimilation in such a system is possible, (3) determine whether all observations of LAI can be treated as similar, and (4) find out if state data assimilation is preferable for phenological systems to methods that focus solely on

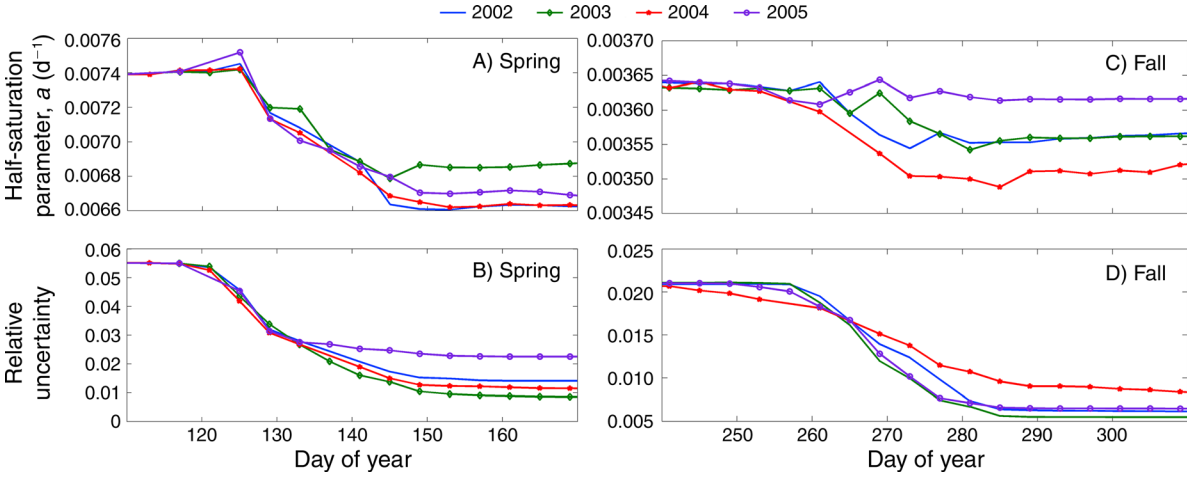

FIG. 4. The evolution of half-saturation parameter *a* as estimated by DART for years 2002–2005. The flux tower observations were averaged over 4-day measurement window. In the upper panels are the ensemble means for each year. Panels (B and D) show relative uncertainty, calculated as the ensemble spreads (standard deviation of *a*) divided by the ensemble mean (mean *a*) for each year. The left (right) panels are for spring leaf-out (fall senescence).

parameters. The first and last objectives were addressed by the analysis of results, as they demonstrate that data assimilation produced estimates of LAI closer to observed LAI than modeling alone, illustrating the utility of data assimilation in accounting for the numerous factors affecting leaf phenology. Especially in the fall, the noise in both estimated state variables and parameters indicates that fall phenology is difficult to predict without actively incorporating information from observations. By contrast, for spring leaf-out, the largest source of uncertainty is the initial bud-burst, but with a correct choice of parameters, the leaf expansion phase appears more predictable even without assimilating input from observations. This is understandable, as spring leaf-out is a continuous process of growth after the initial bud-burst, but fall leaf senescence is influenced by many environmental factors resulting in noncontinuous decrease in LAI.

During spring, estimated LAI reached maximum value at least 7 d later than the modeled LAI. In the fall, there is no constant timing or rate difference between modeled and estimated LAI, but there is a clear difference between modeled and estimated LAI evolution, and hence the phenological phase. Thus state data assimilation can be expected to improve accurate carbon cycle modeling, which is affected by the phenological changes (Richardson et al. 2010). However, when examining NEE, neither the modeled nor estimated NEE was clearly closer to the observed NEE values. This indicates that the midgrowing-season NEE in the model should be improved, as it appears to be insensitive to phenology. However, it should be stressed that even in this case, the estimated approach produced similar results to a model tuned specifically for those years and that location.

These results are encouraging for future data assimilation projects and show the benefit of assimilating multiple different observations when estimating the state. The results, however, also raise several points to consider when moving forward with using state data assimilation in this field.

Observations of the same variable do not necessarily represent the same process

One of our core questions was whether LAI measurements from different instruments can be compared directly. When comparing the spring and fall phenological events in different years, there are notable differences in timing and magnitude between the two data types. In the spring, tower measurements generally show leaf-out starting about a week later than the MODIS observations. The MODIS leaf-out preceding flux tower leaf-out in the spring has been noted before at different sites and is currently unexplained, although it is speculated to be due to the averaging window used by MODIS (Ahl et al. 2006). In contrast, during fall senescence, MODIS measurements show a slow decline preceding a sharp drop unlike the sudden decline in most tower measurements.

MODIS LAI is determined from differences in reflected radiation in different spectral bands, while tower LAI is calculated from intercepting biomass between the two PAR sensors. Due to these differences, when the leaves begin to change color, the MODIS product will cause a change in LAI, while the tower measurements will not show a large change in transmission since the leaves still intercept light. This does not, however, reduce the effectiveness of incorporating different observations, especially since all available measurements are derived observations. Rather, it indicates that the estimate would be further improved by using a model that would simultaneously provide separate predictions for leaf expansion and drop vs. changes in leaf color, corresponding to the different observations. Neither alone is the correct definition of LAI, since physiological processes drive photosynthesis and transpiration, while the physical presence of leaves drives canopy turbulence, soil temperature, and evapotranspiration.

Representativeness of the model and the observations

For the purposes of this study, we assumed that all model predictions and observations were of the same forest composition. This is not strictly true, as the MODIS pixel is much larger than the inventoried area of the flux tower footprint, which is in turn larger than the view field of the PAR sensor mounted on the tower itself. For example, the larger MODIS preleaf-out LAI indicates the presence of conifers in the MODIS pixel that are not observed in the flux tower footprint.

Additionally, the approach used by ED2 to calculate the LAI from the inventoried forest composition contains its own inherent sampling uncertainty, uncertainty in the allometric parameters, and process error in assuming a static allometry and one-dimensional radiative transfer rather than a dynamic response to the light environment (e.g., light-foraging and self-shading). Likewise, the tower measurements are made at a single point, which may differ in magnitude and timing of LAI compared to the rest of the forest. For these reasons, although the forest composition used for the model is based on measurements made around the flux tower, ED2 can still overestimate/underestimate the maximum LAI when compared to tower or MODIS measurements. This would result in a bias between predictions and observations and could cause unrealistic state and parameter estimations. For example, if the model were to overestimate LAI, then state data assimilation would prevent full leaf-out in order to keep predicted LAI at a lower value.

These results show the importance of correct spatial scale when applying data assimilation. Data assimilation readily lends itself to integrating data from widely varying scales that would otherwise be difficult to combine. However, the choice of model is an important

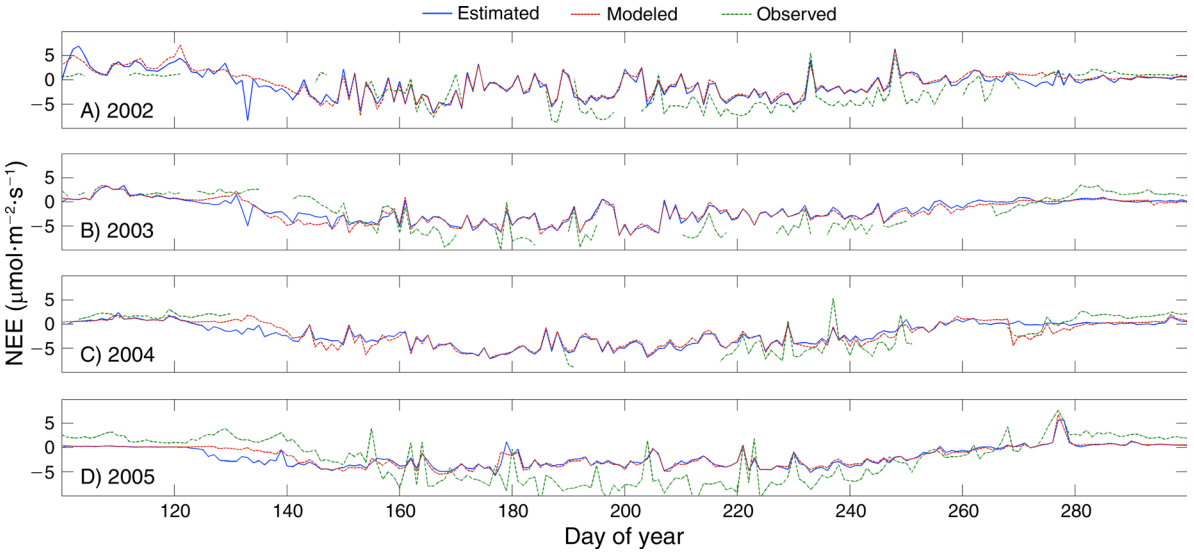

FIG. 5. Net ecosystem exchange (NEE) as estimated by DART (solid blue line), as predicted by the model (dashed green line) and as observed by flux tower (dotted red line) for years 2002–2005.

component of this process. For example, ED2 can simultaneously represent multiple resolutions in the model output. Future research will simultaneously determine the modeled LAI corresponding to the larger area observed by MODIS, as well as the smaller patch observed by the flux tower. This approach will need to deal with the additional complexities of resolving multiple representations of LAI in the state vector.

The impact of increased spatial scale on the LAI estimation

The model was used to estimate LAI at a single site using observations assumed to be representative of that site. Future efforts will simultaneously estimate the state over a larger spatial area (hundreds of kilometers). A regional approach should benefit the state estimate at a single site, as it will be further constrained by the information from nearby sites (Desai 2010). Before doing this, though, additional work will be required to determine the spatial covariance of phenological patterns, how to best represent the dependencies between different plant types at different sites, and how to accurately include the impact of the terrain heterogeneity. Still, regional coherence in phenological patterns is robust in most areas (Zhang et al. 2006) given the spatial patterns of climate variability, implying that spatial multisite assimilation would specifically benefit phenological assimilation.

Improving phenological modeling

The phenological model used here is a modified version of a very simple day-of-year model. This limits the state estimate, especially at the beginning and at the end of a phenological event. A number of other phenological models exist in the literature that are based on climate or growing degree days (Melaas et al. 2013), but to our knowledge there is an absence of true dynamical phenological models. We envision an approach that separates the initiation of phenological state change from the rate of that change. The former should be switched from a deterministic, threshold-based trigger to a probabilistic hazard model approach (Lin and Zhu 2012), enabling the day by day estimation of the probability of change conditional on not having yet switched states. Further improvements in modeling phenology will naturally also improve state estimation. Implementing state data assimilation concurrently with new model development would be beneficial by allowing testing of how different choices impact the predictions. Based on our results, it appears feasible that a more detailed phenological model of spring leaf-out could efficiently predict the phenological phase if the initial budburst is correctly timed from observations. However, the fall senescence is influenced by so many factors that it is unlikely that even a more developed phenological model would need continuous observations at least every 8 d to accurately predict the phenological state.

Impact of frequent/sparse observations

Phenological events tend to be rapid, especially spring leaf-out. Thus it is natural to question whether 8-d observation periods are too long, especially since LAI can change greatly within that time period (Elmore et al. 2012). More frequent observations might create a more continuous, and possibly realistic, state estimate.

While more observations have clear benefits for the state estimate, it is also important to understand that assimilating asynchronous and/or asymmetric observations can have unintended consequences. When comparing Figs. 2 and 4, it is clear that using more frequent

observations, and thus updating the state estimate more often, allowed the tower measurements have a larger impact on the state estimate despite the absence of any a priori reason to trust one data source over the other. However, as a consequence, less frequent satellite measurements have a reduced impact on state estimate. Future data assimilation efforts should consider keeping all data sources binned to the same frequency or risk artificially inflating reliance on a single data source. Additionally, increasing the frequency of the tower measurements also often increases the observation noise as can be seen when comparing Figs. 2 and 4. Thus when determining the frequency of the assimilated observations, it is important to weigh in these factors as well.

Another factor to consider when deciding frequency of included observations is how doing so affects simultaneous parameter data assimilation. In state data assimilation, during each observation update, state variable uncertainties are reduced. For dynamic variables, such as the leaf elongation, uncertainty will increase through time. For static parameters, though, each observation update will further reduce their uncertainty. Thus, they will converge to a certain value sooner, which can be harmful for the prediction if they converge too soon. This can be seen especially in the fall, where, with the exception of 2004, the relative uncertainties decrease faster with more frequent observations.

Our results demonstrate the utility of data assimilation for producing robust model estimates of LAI by incorporating multiple sources of observation and accounting for numerous environmental factors affecting leaf-phenology. These results can be readily expanded upon by including additional sources of data and observations at a number of intermediate scales. While several issues remain to be resolved in future research, our results demonstrate that state data assimilation is a valuable tool for modeling and predicting the leaf-phenological cycle and, consequently, the exchange of carbon between deciduous forests and the atmosphere. Shifting toward estimating net ecosystem carbon exchange will also improve confidence evaluation for data assimilation as with NEE it is easier to examine the cumulative effects of different ecosystem carbon fluxes. Beyond phenology, state data assimilation could also be applied to other ecological problems, such as satellite data algorithms like those used in MODIS, which produces time-series observations of the state. The use of data assimilation would allow us to include the effects of state evolution in this process and thus potentially improve predictive ecosystem models.

Acknowledgments

We thank the editor and the anonymous reviewers for their helpful comments and thoughts. This project was funded with the National Science Foundation grant DBI-1062547 to M. Dietze and DBI-1062204 to A. Desai. Flux tower observations were supported by the Department of Energy Ameriflux Network Management Project. The data assimilation workflow Data Assimilation Research Testbed (DART) was developed by the Data Assimilation Research Section (DAReS) group at the National Center for Atmospheric Research (NCAR). We thank Tristan Quaife and Shawn P. Serbin for their advice. We also thank Jaclyn Hatala Mathes, Afshin Pourmokhtarian, and Joshua Mantooth for their input.

Literature Cited

Ahl, D. E., S. T. Gower, S. N. Burrows, N. V. Shabanov, R. B. Myneni, and Y. Knyazikhin. 2006. Monitoring spring canopy phenology of a deciduous broadleaf forest using MODIS. Remote Sensing of Environment 104:88–95.

Anderson, J. 2009. Spatially and temporally varying adaptive covariance inflation for ensemble filters. Tellus 61A:72–83.

Anderson, J. 2010. A non-gaussian ensemble filter update for data assimilation. Monthly Weather Review 138:4186–4196.

Anderson, J., and S. L. Anderson. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Monthly Weather Review 127:2741–2758.

Anderson, J., T. Hoar, K. L. Raeder, H. Liu, N. Collins, R. Torn, and A. Arellano. 2009. The data assimilation research testbed: a community facility. Bulletin of the American Meteorological Society 90:1283–1296.

Asner, G. P., J. M. O. Scurlock, and J. A. Hicke. 2003. Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. Global Ecology Biogeography 12:191–205.

Baldocchi, D., et al. 2001. FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bulletin of the American Meteorological Society 82:2415–4234.

Barnes, W. L., T. S. Pagano, and V. V. Salomonson. 1998. Prelaunch characteristics of the moderate resolution imaging spectroradiometer (MODIS) on EOS-AM1. IEEE Transactions on Geoscience and Remote Sensing 36(4):1088–1100.

Bertino, L., B. Evensen, and H. Wackernagel. 2003. Sequential data assimilation techniques in oceanography. International Statistics Review 71(2):223–241.

Boden, T. A., Krassovski, and B. Yang. 2013. The AmeriFlux data activity and data system: an evolving collection of data management techniques, tools, products and services. Geoscientific Instrumentation, Methods and Data Systems 2:165–176.

Bradley, B. A., R. W. Jacob, J. F. Hermanence, and J. F. Mustard. 2007. A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. Remote Sensing of Environment 106(2):137–145.

Braswell, B. H., W. J. Sacks, E. Linder, and D. S. Schimel. 2005. Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations. Global Change Biology 11(2):335–355.

Chen, J. K., J. Davis, and T. P. Meyers. 2008. Ecosystem–atmosphere carbon and water cycling in the upper Great Lakes Region. Agriculture and Forest Meteorology 148:155–157.

Chen, J. M., and T. A. Black. 1991. Measuring leaf area index of plant canopies with branch architecture. Agriculture and Forest Meteorology 57:1–12.

Chen, P. 2011. Full-wave seismic data assimilation: theoretical background and recent advances. Pure and Applied Geophysics 168:1527–1552.

Cook, B. D., P. V. Bolstad, J. G. Martin, F. A. Heinsch, K. J. Davis, W. Wang, A. R. Desai, and R. M. Teclaw. 2008. Using light-use and production efficiency models to predict photosynthesis and net carbon exchange during forest canopy disturbance. Ecosystems 11:26–44.

Cook, B. D., et al. 2004. Carbon exchange and venting anomalies in an upland deciduous forest in northern Wisconsin, USA. Agriculture and Forest Meteorology 126(3–4):271–295.

Desai, A. R. 2010. Climatic and phenological controls on coherent regional interannual variability of carbon dioxide flux in a heterogeneous landscape. Journal of Geophysical Research—Biogeosciences 115:G00J02.

Dietze, M., D. S. Lebauer, and R. Kooper. 2013. On improving the communication between models and data. Plant, Cell and Environment 36:1575–1585.

du Plessis, R. M. 1967. Poor man's explanation of Kalman filters: or how I stopped worrying and learned to love matrix inversion. North American Rockwell Electronics Group, Anaheim, California, USA.

Elmore, A. J., S. M. Guinn, B. J. Minsley, and A. D. Richardson. 2012. Landscape controls on the timing of spring, autumn and growing season length in mid-Atlantic forests. Global Change Biology 18:656–674.

Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic models using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research—Oceans 99(C5):10143–10162.

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

Gelb, A. 1974. Applied optimal estimation. MIT Press, Cambridge, Massachusetts, USA.

Hamill, T., J. S. Whitaker, and C. Snyder. 2001. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. Monthly Weather Review 129:2776–2790.

Jazwinski, A. H. 1970. Stochastic processes and filtering theory. Mathematics in science and engineering. Volume 64. Academic Press, New York, New York, USA.

Jeong, S.-J., D. Medvigy, E. Shevliakova, and S. Malyshev. 2012. Uncertainties in terrestrial carbon budgets related to spring phenology. Journal of Geophysical Research—Biogeosciences 117:G01030.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. Transactions of the ASME—Journal of Basic Engineering 82D:35–45.

Kaminski, T., et al. 2013. The BETHY/JSBACH carbon cycle data assimilation system: experiences and challenges. Journal of Geophysical Research—Biogeosciences 118:1414–1426.

Kucharik, C. J., C. C. Barford, M. El Maayar, S. C. Wofsy, K. Monson, and D. D. Baldocchi. 2006. A multiyear evaluation of a dynamic global vegetation model at three AmeriFlux forest sites: vegetation structure, phenology, soil temperature, and CO_2 and H_2O vapor exchanges. Ecological Modelling 196:1–31.

LeBauer, D. S., D. Wang, K. T. Richter, C. C. Davidson, and M. C. Dietze. 2013. Facilitating feedbacks between field measurements and ecosystem models. Ecological Monographs 83(2):133–154.

Levis, S., and G. B. Bonan. 2004. Simulating springtime temperature patterns in the community atmosphere model coupled to the community land model using prognostic leaf area. Journal of Climate 17:4531–4540.

Lin, F.-C., and J. Zhu. 2012. Continuous-time proportional hazards regression for ecological monitoring data. Journal of Agricultural, Biological, and Environmental Statistics 17(2):163–175.

Lorenc, A. C. 1986. Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society 112:1177–1194.

McKane, R. B., E. B. Rastetter, G. R. Shaver, K. J. Nadelhoffer, A. E. Giblin, J. A. Laundre, and F. S. Chapin, III. 1997. Reconstruction and analysis of historical changes in carbon storage in Arctic tundra. Ecology 78(4):1188–1198.

Medvigy, D., S.-J. Jeong, K. L. Clark, N. S. Skowronski, and K. V. R. Schafer. 2013. Effects of seasonal variation of photosynthetic capacity on the carbon fluxes of a temperate deciduous forest. Journal of Geophysical Research—Biogeosciences 118(4):1703–1714.

Medvigy, D., S. C. Wofsy, J. W. Munger, D. Y. Hollinger, and P. R. Moorcroft. 2009. Mechanistic scaling of ecosystem function and dynamics in space and time: ecosystem demography model version 2. Journal of Geophysical Research—Biogeosciences 114:G01002. http://dx.doi.org/10.1029/2008JG000812

Melaas, E. K., A. D. Richardson, M. A. Friedl, D. Dragoni, C. M. Gough, M. Herbst, L. Montagnani, and E. Moors. 2013. Using FLUXNET data to improve models of springtime vegetation activity onset in forest ecosystems. Agriculture and Forest Meteorology 171–172:46–56.

Monsi M., and T. Saeki. 2005. On the factor light in plant communities and its importance for matter production. Annals of Botany 95:549–67.

Montzka, C., V. R. N. Pauwels, H.-J. H. Franssen, X. Han, and H. Vereecken. 2012. Multivariate and multiscale data assimilation in terrestrial systems: a review. Sensors 12(12):16291–16333.

Moorcroft, P. R., G. C. Hurtt, and S. W. Pacala. 2001. A method for scaling vegetation dynamics: the ecosystem demography model (ED). Ecological Monographs 71(4):557–586.

Moore, K. E., D. R. Fitzjarrald, R. K. Sakai, M. L. Goulden, J. W. Munger, and S. C. Wofsy. 1995. Seasonal variation in radiative and turbulent exchange at a deciduous forest in central Massachusetts. Journal of Applied Meteorology 35:122–134.

PEcAn. 2012. Predictive ecosystem analyzer. http://pecanproject.github.io/index.html

Peters, W., et al. 2007. An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. Proceedings of the National Academy of Sciences USA 104(48):18925–18930.

Rabier, F., H. Järvinen, E. Klinker, J. F. Mahfouf, and A. Simmons. 2000. The ECMWF operational implementation of four-dimensional variational assimilation. Part I: experimental results with simplified physics. Quarterly Journal of the Royal Meteorological Society 126:1143–1170.

Raupach, M. R., P. J. Rayner, D. J. Barrett, R. S. DeFries, M. Heimann, D. S. Ojima, S. Quegan, and C. C. Schmullius. 2005. Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. Global Change Biology 11(3):378–397.

Richardson, A. D., et al. 2012. Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis. Global Change Biology 18:566–584.

Richardson, A. D., et al. 2010. Influence of spring and autumn phenological transitions on forest ecosystem productivity. Philosophical Transactions of the Royal Society B 365:3227–3246.

Richardson, A. D., T. F. Keenan, M. Migliavacca, Y. Ryu, O. Sonnentag, and M. Toomey. 2013. Climate change, phenology, and phenological control of vegetation feed-backs to the climate system. Agriculture and Forest Meteorology 169:156–173.

Seo, D. J., L. Cajina, R. Corby, and T. Howieson. 2009. Automatic state updating for operational streamflow forecasting via variational data assimilation. Journal of Hydrology 367:255–275.

Viskari, T., E. Asmi, A. Virkkula, P. Kolmonen, T. Petäjä, and H. Järvinen. 2012. Estimation of aerosol particle number distribution with Kalman Filtering. Part 2: simultaneous use of DMPS, APS and nephelometer measurements. Atmospheric Chemistry and Physics 12:11781–11793.

Whitaker, J. S., and T. M. Hamill. 2002. Ensemble data assimilation without perturbed observations. Monthly Weather Review 130:1913–1924.

White, M. A., P. E. Thornton, and S. W. Running. 1997. A continental phenology model for monitoring vegetation

responses to interannual climatic variability. Global Biogeochemical Cycles 11:217–234.
Williams, M., P. A. 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.
Yang, W., D. Huang, B. Tan, J. C. Stroeve, N. V. Shabanov, Y. Knyazikhin, R. R. Nemani, and R. B. Myneni. 2006*a*. Analysis of leaf area index and fraction of par absorbed by vegetation products from the Terra MODIS sensor: 2000–2005. IEEE Transactions on Geoscience and Remote Sensing 44(7):1829–1842.
Yang, W., N. V. Shabanov, D. Huang, W. Wang, R. E. Dickinson, R. R. Nemani, Y. Knyazikhin, and R. B. Myneni. 2006*b*. Analysis of leaf area index products from combination of MODIS Terra and Aqua data. Remote Sensing of Environment 104:297–302.
Zhang, X., M. A. Friedl, and C. B. Schaaf. 2006. Global vegetation phenology from moderate resolution imaging spectroradiometer (MODIS): evaluation of global patterns and comparison with in situ measurements. Journal of Geophysical Research—Biogeosciences 111:G04017.
Zhang, X., M. A. Friedl, C. B. Schaaf, and A. H. Strahler. 2004. Climate controls on vegetation phenological patterns in northern mid- and high latitudes inferred from MODIS data. Global Change Biology 10(7):1133–1145.
Zhang, X., M. A. Friedl, C. B. Schaaf, A. H. Strahler, J. C. F. Hodges, F. Gao, B. C. Reed, and A. Hueto. 2003. Monitoring vegetation phenology using MODIS. Remote Sensing of Environment 84:471–475.
Zhou, T., P. Shi, G. Jia, and Y. Luo. 2013. Nonsteady state carbon sequestration in forest ecosystems of China estimated by data assimilation. Journal of Geophysical Research—Biogeosciences 118:1–16.
Zobitz, J., A. R. Desai, D. J. P. Moore, and M. A. Chadwick. 2011. A primer for data assimilation with ecological models using Markov chain Monte Carlo (MCMC). Oecologia 167(3):599–611.

Supplemental Material

Data Availability

Data associated with this paper have been deposited in Dryad: http://dx.doi.org/10.5061/dryad.2s71b