



Early spring onset increases carbon uptake more than late fall senescence: modeling future phenological change in a US northern deciduous forest

Aaron Teets^{1,2} · Amey S. Bailey³ · Koen Hufkens⁴ · Scott Ollinger⁵ · Christina Schädel¹ · Bijan Seyednasrollah⁶ · Andrew D. Richardson^{1,6}

Received: 5 August 2022 / Accepted: 27 November 2022 / Published online: 16 December 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

In deciduous forests, spring leaf development and fall leaf senescence regulate the timing and duration of photosynthesis and transpiration. Being able to model these dates is therefore critical to accurately representing ecosystem processes in biogeochemical models. Despite this, there has been relatively little effort to improve internal phenology predictions in widely used biogeochemical models. Here, we optimized the phenology algorithms in a regionally developed biogeochemical model (PnET-CN) using phenology data from eight mid-latitude PhenoCam sites in eastern North America. We then performed a sensitivity analysis to determine how the optimization affected future predictions of carbon, water, and nitrogen cycling at Bartlett Experimental Forest, New Hampshire. Compared to the original PnET-CN phenology models, our new spring and fall models resulted in shorter season lengths and more abrupt transitions that were more representative of observations. The new phenology models affected daily estimates and interannual variability of modeled carbon exchange, but they did not have a large influence on the magnitude or long-term trends of annual totals. Under future climate projections, our new phenology models predict larger shifts in season length in the fall (1.1–3.2 days decade⁻¹) compared to the spring (0.9–1.5 days decade⁻¹). However, for every day the season was longer, spring had twice the effect on annual carbon and water exchange totals compared to the fall. These findings highlight the importance of accurately modeling season length for future projections of carbon and water cycling.

Keywords Phenology · Biogeochemical modeling · AmeriFlux · Deciduous forests · Forest carbon cycling · Hubbard Brook · PhenoCam

Introduction

In deciduous forests, the timing of seasonal changes in leaf area is closely tied to ecosystem-level shifts in carbon, water, and nutrient cycling. When new photosynthetic tissues (leaves) emerge in the spring, deciduous forests transform from carbon sources to carbon sinks, and from evaporation- to transpiration-dominated systems. New leaves cause abrupt changes in canopy structure, affecting ecosystem-level sensible and latent heat exchange (Hogg et al. 2000), surface roughness (Young et al. 2021), and albedo (Ollinger et al. 2008; Hollinger et al. 2010). In the fall, as deciduous forests transition back to dormancy, they also transition to a more passive exchange of carbon, water, and nutrients. The timing of these transitions is critically important because they regulate how long deciduous forests remain active, directly affecting annual carbon uptake

Communicated by Rodrigo Vargas.

We improve phenology algorithms in ecosystem-level models using empirical observations from multiple data streams. The new phenology algorithms affected daily estimates and interannual variability of modeled carbon exchange, but did not have a large influence on the long-term trends of annual totals. Using the improved phenology algorithms, we assessed the relative influence of spring and fall on future ecosystem carbon exchange.

✉ Aaron Teets
aft49@nau.edu

Extended author information available on the last page of the article

(Goulden et al. 1996; Richardson et al. 2010) and water-use (Zha et al. 2010). Therefore, being able to predict the timing of seasonal transitions in deciduous forests is of great importance for biogeochemical modeling and future projection of land–atmosphere exchange (Richardson et al. 2012).

Growing seasons in deciduous forests are already lengthening in response to climate change worldwide (Piao et al. 2019). The timing of seasonal transitions in the spring and fall (i.e., the dates and duration of phenology changes) are particularly sensitive to changing temperatures. Increasing global temperatures are causing spring to arrive earlier (Jeong et al. 2011) and fall to arrive later (Liu et al. 2016) compared to just a few decades ago. These longer growing seasons have led to higher rates of annual gross primary productivity (GPP), and net ecosystem productivity (NEP) (Richardson et al. 2010; Dragoni et al. 2011; Keenan et al. 2014). However, increasing temperatures are also associated with higher rates of ecosystem respiration (Duveneck and Thompson 2017) and water use (Beamesderfer et al. 2020) that can counteract increases in forest carbon uptake (e.g. Sanders-Demott et al. 2020). Therefore, the future effects of longer growing seasons on ecosystem-level carbon and water cycling are not fully understood. Predictions range from slight increases in NEP (Richardson et al. 2009) to stable long-term trends due to increased rates of ecosystem respiration (Piao et al. 2007). Thus, future warming conditions will likely result in tradeoffs between increases in photosynthetic uptake and increases in ecosystem respiration.

Biogeochemical models are frequently used to better understand how ecosystems use carbon, water, and nutrients under current and future conditions. These models range in complexity from simple processes at individual sites to complex land–atmosphere feedbacks at the global scale. Although biogeochemical models are widely used to make inferences about ecosystem functioning, these models tend to have poor representations of phenology, oftentimes overestimating growing season length (Levis and Bonan 2004; Randerson et al. 2009; Richardson et al. 2012; Chen et al. 2016). Overestimating growing season length directly affects the amount of modeled carbon, water, or even nutrients exchanged in biological processes. Integrating empirical observations of phenology into existing biogeochemical models can overcome these weaknesses to represent growing season lengths more accurately (Reyer et al. 2013).

The phenology algorithms within biogeochemical models are used to simulate changes in leaf area index (LAI). Spring models conventionally trigger the onset of leaf development based on the accumulation of temperature over a set threshold (e.g., growing degree days, GDD). Simple spring models use accumulated warming temperatures (e.g., Lany et al. 2016; Menzel et al. 2006; Miller-Rushing and Primack 2008), but may also include photoperiod, water balance and/or chilling requirements (Morin et al. 2009; Körner and

Basler 2010). Like spring, fall senescence is also regulated by temperature (Estrella and Menzel 2006) and photoperiod (Estiarte and Peñuelas 2015). Fall phenology models use cold temperatures under a certain threshold, accumulating in a manner similar to GDD, called chilling degree days (CDD). Compared to spring, fall senescence is less well characterized in modeling efforts. This is in part because it has received less attention in the literature, and is less predictable because leaf senescence can be more suddenly affected by harsh weather such as frost, or wind (Gallinat et al. 2015). Despite many advances in modeling spring and fall phenology (e.g., Basler 2016; Hufkens et al. 2018), incorporating more sophisticated phenology models into biogeochemical models has lagged as a priority in model development.

Using empirical data to train phenology models is becoming more practical as the availability of phenology data has increased with automated methods of observation (Richardson et al. 2018). Near-surface remote sensing is a powerful way to monitor phenology in deciduous forests in a consistent and reproducible manner. Digital repeat photography (i.e., time lapse) is one example of near-surface remote sensing offering a cost-effective, automated tool for visually monitoring forest ecosystems from years to decades at an intermediate scale of observation between ground-based observations and satellite-based remote sensing (Richardson et al. 2018). The PhenoCam Network provides near-real time, open-access digital repeat photography and developed techniques to extract data that can quantify seasonal changes in vegetation phenology based on a greenness index (Sonnentag et al. 2012).

In this project, we use PhenoCam data to calibrate phenology models to regional long-term datasets, then use the new models to ask the following: (1) how does altering the spring and fall phenology algorithms affect future modeled carbon, water, and nutrient exchange in a northern deciduous forest? And (2) does phenological variation in spring or fall have a larger influence on modeled ecosystem-level carbon, water, and nutrient cycling? We predicted that altering the phenology algorithms would shorten the period of modeled photosynthetic activity and therefore reduce the amount of modeled carbon, water, and nutrients used annually. We hypothesized that longer springs would have a larger effect on carbon and water cycling compared to fall because day-length in the spring is longer and rates of incoming solar radiation are higher compared to the fall (Zhang et al. 2020). By addressing these questions, we improve our understanding on the effects of changing climate on phenology and how that regulates important ecosystem processes.

Methods

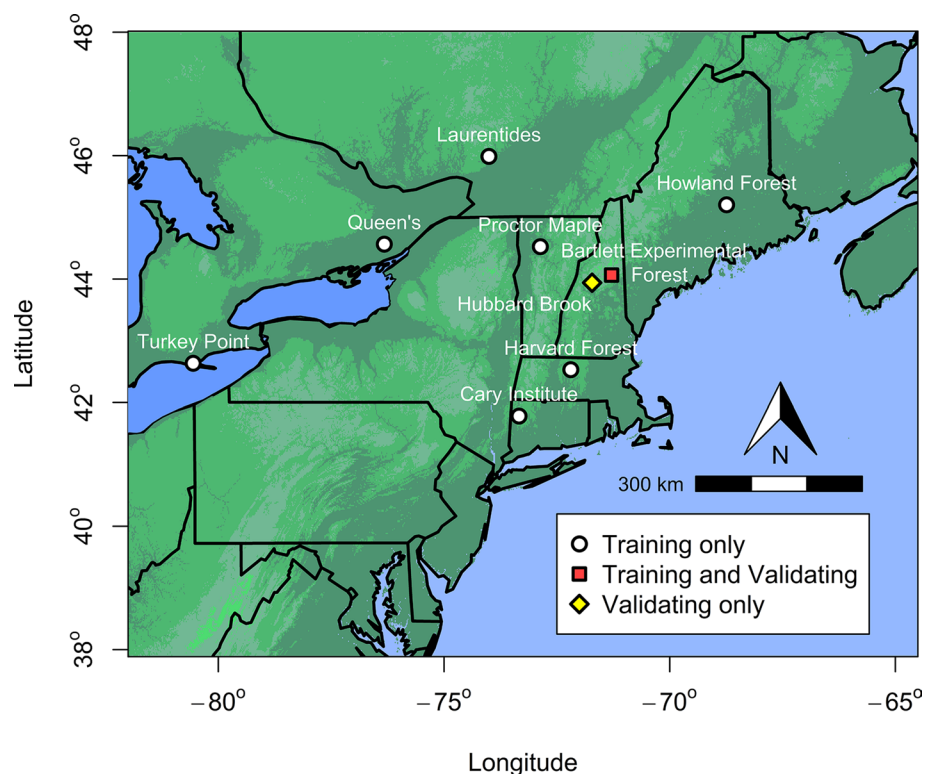
We used existing near-surface remote sensing data from eight different northern forests of North America to train spring and fall phenology models (Fig. 1). Using data from several different northern deciduous forests allows us to use more years of data to train models and makes models that are more generalizable across region. Each of the forests used for model training had some combination of maple (e.g., *Acer saccharum* and *Acer rubrum*), beech (*Fagus grandifolia*), and birch (e.g., *Betula alleghaniensis* and *Betula papyrifera*) species in the overstory. Each site had between 6 and 14 years of PhenoCam data available for training models, in total we used 96 site years of data to train the models.

Detecting transition dates using PhenoCam

We estimated transition dates in the spring and fall of each site year using long-term PhenoCam data (Richardson et al. 2018). Provisional data from the PhenoCam website (phenocam.nau.edu) were collated and used to estimate transitions at the beginning and end of each of the spring and fall transitions using phenocamR package (Hufkens et al. 2018) in the R Statistical Software (R Core Team 2020). We first extracted greenness values from a region of interest in PhenoCam images that was representative of the forest canopy, an approach that is well suited for

canopy-level, or “species ignorant” modeling. The relative greenness of a pixel (i.e., green chromatic coordinate or Gcc) in each region of interest is calculated by dividing the average green intensity value by the sum of the red, blue and green intensities across all pixels. To reduce the noise associated with illumination and suboptimal lighting, we extracted the 90th percentile of the Gcc across a 3-day moving window (Richardson et al. 2018). For more information on processing digital images to extract phenology dates, see Sonnentag et al. (2012). In the spring, we estimated the transition dates as the dates Gcc values reached the 10% and 80% of the maximum seasonal amplitude. In the fall, we found greenness was unable to predict transition dates late enough to capture the end of leaf senescence (based on ground observations at Hubbard Brook and measured leaf area index from Bartlett Experimental Forest [see Toda and Richardson 2018]). We found that using yellow chromatic coordinate (Ycc; equivalent to 1 minus the blue chromatic coordinate, Kosmala et al. 2016) was better suited for modeling the end of leaf senescence because it could predict fall transition dates later than Gcc (for more information, see Fig. S1 in the Supplementary Materials). For this reason, we used a combination of Gcc and Ycc to predict the start and end of fall senescence—the start of the fall transition was estimated as the date the Gcc reached 50% of the seasonal amplitude, and the end of the fall transition was the day Ycc reached 50% of the seasonal amplitude.

Fig. 1 Sites used to select and train phenology models based on PhenoCam data. Ground observations were used to validate model performance at Hubbard Brook. Future phenology projections and biogeochemical cycling (using PnET-CN) were modeled at Bartlett Experimental Forest



Model selection, calibration, and validation

We evaluated a variety of phenology transition date models that represent a range of environmental drivers and levels of complexity (Table 1). The spring models have been evaluated in many different forest types and represent some of the best-known phenology models used to predict the spring transition. The fall models have not been widely tested (but see Richardson et al. 2006; Delpierre et al. 2009) and the drivers of fall senescence in mesic deciduous forests are not well known. Our models require a combination of accumulating warming temperature, photoperiod, and accumulating chilling temperatures.

The spring and fall phenology models were calibrated using the phenoR package (Hufkens et al. 2018). We first parameterized models for the beginning and end of the season based on dates extracted from PhenoCam: 10% of the seasonal Gcc amplitude in the spring and 50% of the seasonal Ycc amplitude in the fall. Then, to model the duration of the transition period, we trained models to predict the end of the spring transition, and the beginning of the fall transition by holding all other parameter values constant except for the critical value of accumulating temperature. By doing so, we are effectively modeling four dates for each year—the initiation and completion of both spring onset and fall senescence for each model. Modeling the beginning and end of phenology transitions is important because the rate of leaf development or senescence can vary independently from year-to-year (Klosterman et al. 2018). We optimized models using a generalized simulated annealing algorithm (GenSA package; Xiang et al. 2013) in the phenoR package that identifies the global minima of different combinations of parameter values for each model. We ran 25 chains (i.e., individual

ensembles) of 40,000 iterations and identified the parameter sets with the lowest root mean square error (RMSE) as the parameters that best fit the data.

The calibrated models were then evaluated against the ground-based observations of spring and fall phenology transitions from Hubbard Brook Experimental Forest, in the White Mountains of New Hampshire, USA (Likens et al. 1970; Fahey et al. 2005). Spring leaf development and fall leaf senescence data were collected every 3–7 days during spring and fall since the early 1990s. Individual marked trees were given a score of 0–4 for both spring and fall transitions; 0 corresponding to no leaves and 4 corresponding to fully developed leaves (see Richardson et al. 2006). We compared estimates from our calibrated models to the linearly interpolated ground observations that ranged from 0–4. In the spring we compared our dates to the date the ground observations reached a 3 out of 4 (50% canopy closure), and in the fall we compared our dates to the date ground observations reached a 1 out of 4 (no more green in the canopy, half of the leaves have fallen) – see Richardson et al. (2006) for more information.

We selected the three best performing spring models for future projections based on the lowest Akaike information criterion (AIC) values. We included three spring models (each representing different drivers), and both fall models for future projections of phenology transitions and ecosystem processes. We included more models than just the top performing spring and fall models (i.e., also including models with $\Delta AIC > 2$) for quantifying uncertainty of our model outputs.

Future phenology projections

We modeled phenology transition dates into the future at Bartlett Experimental Forest to year 2100 with an ensemble

Table 1 Phenology models selected for parameterization and testing

Model name	Model code	Season	Drivers	# of param-eters	References
Linear regression	LIN	Spring	T	2	Hufkens et al. (2018)
Thermal time	TT	Spring	T	3	Cannell and Smith (1983)
Photothermal-time	PTT	Spring	T, P	3	Črepinšek et al. (2006)
M1	M1	Spring	T, P	4	Blümel and Chmielewski (2012)
Sequential	SQ	Spring	C, T	8	Hanninen (1990)
Parallel	PA	Spring	C, T	9	Hanninen (1990)
Sequential M1	SM1	Spring	C, T, P	9	Hufkens et al. (2018)
Parallel M1	PM1	Spring	C, T, P	10	Hufkens et al. (2018)
Cooling degree day	CDD	Fall	C	3	Richardson et al. (2006)
Cooling degree day and photoperiod	CDDP	Fall	C, P	3	Delpierre et al. (2009)

Models are listed by season and number of parameters. The driver codes refer to: *T* accumulating warming temperature, *P* photoperiod, *C* accumulating chilling temperature

of six different future climate scenarios from NA-CORDEX data archive (Mearns et al. 2017). The archive offers a selection of different downscaled future climate projections based on General Circulation Model (GCM) simulations in the Coupled Model Intercomparison Project, Phase 5 (CMIP5) multi-model ensemble (Maurer et al. 2007). We selected six climate models based on the criteria that the models (1) had daily data resolution, (2) had representative climate pathways (RCP) 4.5 and 8.5, and (3) were bias corrected based on Daymet gridded datasets. The Daymet corrected datasets were chosen because the phenology transition models described above were fit using Daymet data in the phenoR package—this minimized biases associated with transitioning from historic data to future climate projections. These two RCPs represent two projections of atmospheric greenhouse gas concentrations under high (RCP 8.5) and intermediate (RCP 4.5) human emission scenarios.

Biogeochemical modeling in PnET-CN

Finally, we performed a sensitivity analysis to better understand how phenology models affect rates of modeled carbon, water, and nitrogen cycling. We used the Photosynthesis and Evapotranspiration Model (PnET)—a simple biogeochemical model developed in New England forests (Aber and Federer 1992; Aber et al. 1996). The family of PnET models is built on the relationships that maximum photosynthetic rate is a function of leaf area and foliar nitrogen concentrations (see Ollinger et al. 2008) and that stomatal conductance is a function of the actual photosynthesis rate. We used a daily version of the PnET-CN model, a model that has been adapted to run at the daily timescale and includes both carbon (C) and nitrogen (N) cycling (Aber et al. 1997). PnET-CN has been tested in several biomes (Thorn et al. 2015) and has previously been validated at Bartlett Experimental Forest (Ollinger and Smith 2005). In the remainder of this manuscript, we refer to the daily version of PnET-CN simply as PnET.

Previous work has shown the default spring phenology model in PnET predicts spring onset too early, and the duration of leaf development lasts too long (Chiang and Brown 2007; Klosterman et al. 2018). Less is known about the performance of PnET's fall phenology models. We replaced the default phenology models with new spring and fall phenology models described above, and ran the model to year 2100 to assess the effects the phenology algorithm has on model output. The default spring phenology model for spring onset in PnET is a simple growing degree day (GDD) model (i.e., an ecodormancy model) that begins on day of year 61 and accumulates daily temperatures above 0 °C until a user-defined parameter threshold is reached. This is similar to the Thermal-time (TT) model used in model selection and validation section. Once leaves emerge, they continue to develop

until a second user-defined GDD threshold is reached. Fall senescence in PnET is indirectly affected by temperature and is triggered by the date that leaves in different canopy layers reach a negative carbon balance (due to cooling temperatures and shortening day lengths). We used a similar two-tiered approach of simulating the start and end of the spring and fall transitional periods.

We replaced the default phenology models in PnET with regionally calibrated models while keeping all other inputs and drivers constant. We ran the model at Bartlett Experimental Forest, hindcasting from 2004 to 2020 and forecasting to year 2100. This site was selected because an existing AmeriFlux tower established in 2004 (Ouimette et al. 2018; Richardson 2004) allows us to evaluate model performance compared to eddy covariance data. For the period overlapping with tower measurements (2004–2020), we used temperature, precipitation, and solar radiation from the DayMet interpolated dataset as PnET model drivers (Thornton et al. 2020). For future model runs, we used projections of temperature, precipitation, and solar radiation from one of the NA-CORDEX future climate projections. We selected the projection with the lowest bias compared to Daymet data over the 2004–2020 period. The projection (EC-Earth RCA4) had an average level of warming for RCP 8.5 over the next 80 years (~0.5 °C per decade) compared to other models. We held other model drivers constant (atmospheric CO₂, O₃, and deposition of NH₄ and NO₃) to remove variability not directly related to phenology. We used input parameters initially developed by Aber et al. (1996) from site-specific measurements from Harvard Forest and later refined by Ollinger et al. (2002) for implementation of PnET at Hubbard Brook. Vegetation-specific input parameters such as foliar nitrogen and specific leaf weight were developed directly from field measurements in the region.

We compared the daily modeled GPP, NEP, and ecosystem respiration to the corresponding daily tower based estimates from 2006 to 2020. We then compared how annual totals of modeled carbon exchange, as well as evapotranspiration (ET), water-use efficiency (WUE), and nitrogen mineralization (Nmin), were affected by altering the phenology subroutine. To ensure the model runs were close to equilibrium, we used long spin-up times (i.e., the length of time a model is run since initiation, by recycling climate driver data) and chose not to include past disturbance after spin-up was complete. Using long spin-up times is standard practice for these types of ecosystem-level biogeochemical models because they alter carbon and nutrient pools to balanced levels that more accurately represent natural systems. However, long spin-up times result in close to “steady-state” conditions. We decided to use the long spin-up times without past disturbance simulation because they resulted in more realistic carbon and nutrient pool sizes, albeit they also resulted in modeled NEP that was closer to steady state

than the eddy covariance data suggest [Bartlett Experimental Forest is a consistent carbon sink (Ouimette et al. 2018)]. We account for this when comparing model outputs relative to fixed phenology models (i.e., spring and fall transitions occurring at the same day of year), because these differences isolate the effects of the phenology models on carbon, water, and nutrient cycling outputs.

Results

Model selection and validation

By training spring and fall phenology models based on data from the eight regional PhenoCam sites, we found our spring models were better at predicting transition dates compared to fall models (Fig. 2) based on root mean square error (RMSE) and r^2 values. The spring models consistently fit the data better (with an average RMSE of 6.7 days) than the fall models (an average RMSE of 8.7). We determined the best fitting

spring models were the sequential M1 (SM1), the parallel (PA), and the photothermal-time (PTT) models based on AIC (Table 2). These three models required different drivers, only sharing warming temperature as the common driver across models. More complex spring models (including photoperiod and, or chilling requirements) tended to outperform the simplest models (i.e., those requiring only warming temperatures) but having more variables or parameters did not always result in improved model fits based on AIC. For example, the PTT model fit the spring validation dataset as well as any other model (based on RMSE and r^2) but required a third of the parameters of the other top models. For the fall models, the simpler model that only requires accumulating cold temperature (CDD) performed better than the other fall model which also includes photoperiod based on RMSE or r^2 .

Models performed slightly better for the validation dataset (i.e., 28 years of ground observations at Hubbard Brook) compared to the regional training dataset (i.e., 96 sites years of PhenoCam data from 8 sites). This

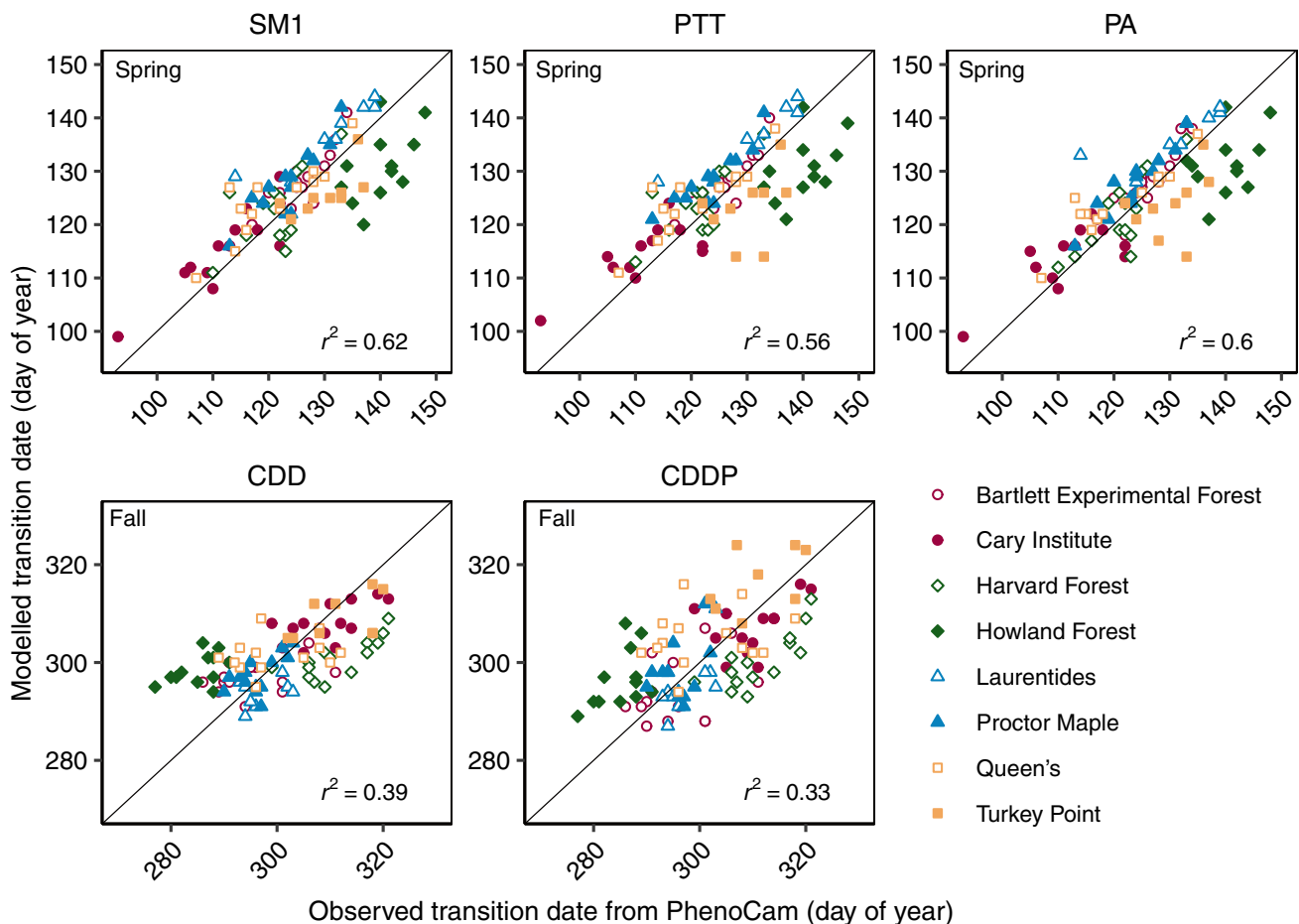


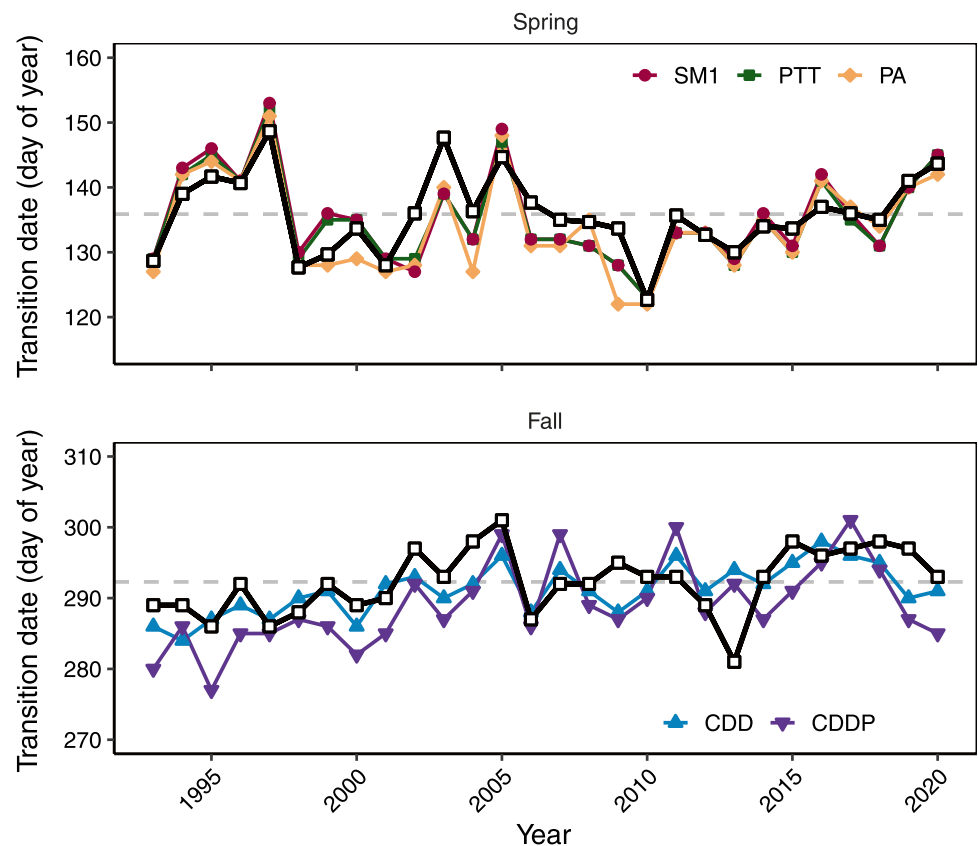
Fig. 2 Modeled vs. observed spring and fall transition dates using the three best fitting spring models and the two fall models for all site years and sites (represented by different shapes and colors) used for

model training. The spring models represent the initiation of spring onset, and the fall models represent the completion of fall senescence. The drawn line is a one-to-one line

Table 2 Model performance for start of spring and end of fall phenology models based on the training dataset (8 sites across the region) and the validation dataset (Hubbard Brook) based on ground observations

Model code	Season	AIC training	RMSE training	r^2 Training	RMSE validation	r^2 validation
SM1*	Spring	371.1	6.2	0.62	4.0	0.69
PTT*	Spring	373.9	6.7	0.56	3.6	0.73
PA*	Spring	374.0	6.3	0.60	4.3	0.73
M1	Spring	374.3	6.6	0.56	4.5	0.73
TT	Spring	376.4	6.7	0.55	3.5	0.71
PM1	Spring	384.4	6.7	0.55	3.7	0.72
SQ	Spring	386.5	6.8	0.55	3.6	0.70
LIN	Spring	398.6	7.6	0.40	5.5	0.45
PnET	Spring	–	19.7	0.27	10.3	0.46
CDD*	Fall	415.5	8.3	0.39	4.1	0.30
CDDP*	Fall	432.6	9.0	0.33	6.1	0.28
PnET	Fall	–	12.1	0.09	9.3	0.03

Models followed by an asterisk (*) indicate the models that were used for future projections. Model performance for spring and fall dates estimated from PnET are included for comparison, but we cannot report the AIC for because the PnET algorithms were not parameterized

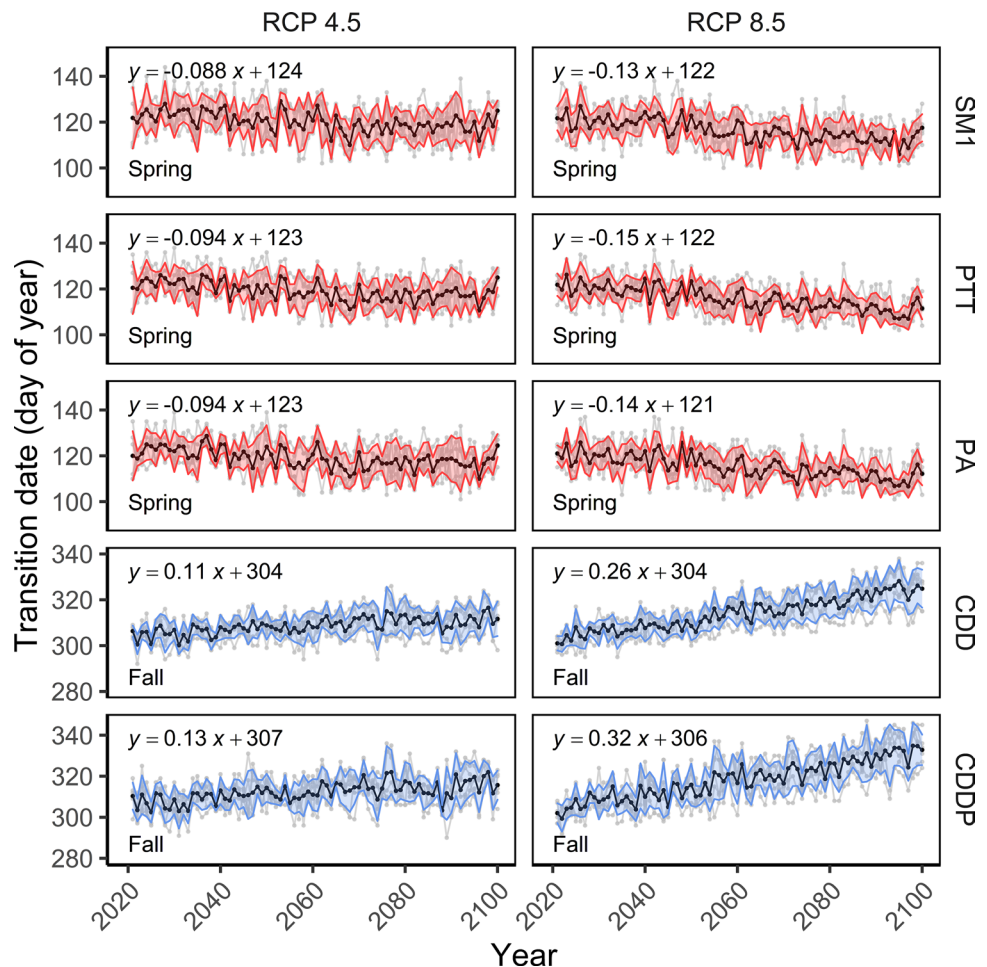
Fig. 3 Validation of best performed spring onset (top panel) and fall senescence (bottom panel) models at Hubbard Brook. The different colors and shapes represent the different phenology models that were parameterized and tested against the ground observations (black squares)

validation shows our models captured the magnitude and variability of an independent dataset reasonably well (Fig. 3), justifying the conclusion that the models are not overfit to the training dataset. This also suggests that models trained using near surface remote sensing data (in this case, PhenoCam data) can be validated using independently derived ground-based measurements that rely on entirely different methods (also see Fig. S2).

Future phenology projections

Each of the calibrated models predicted earlier spring and later fall transition dates under all future climate scenarios (Fig. 4) with larger shifts in growing season length during the fall compared to spring. Spring models predict an average advancement of 0.9 and 1.4 days decade⁻¹ for RCP 4.5 and RCP 8.5, respectively. In comparison, the

Fig. 4 Projected spring and fall transition dates from 2020 to 2099 using two different representative concentration pathways (RCP). Light gray lines represent individual projections from the CMIP5 future climate projections from the NA-CORDEX Program (Mearns et al. 2017). The ensemble mean (average of the six individual climate projections) is in black and standard deviation around the mean is represented with a colored ribbon. The spring transition dates are in red and fall transition dates in blue. The equation represents a least squares linear regression of the ensemble mean over time with the y-intercept representing the projected transition date at year 2020. All regression lines are significant at an alpha of 0.05



average fall senescence dates were 1.2 and 2.9 days later decade⁻¹ for RCP 4.5 and 8.5, respectively. Based on this ensemble of future climate scenarios, our models predict the growing season in northern deciduous forests of North America will lengthen by 2–5 days decade⁻¹ depending on the climate scenario, with more pronounced changes projected in the fall compared to the spring.

Implications for daily PnET model output

Incorporating new spring and fall models into PnET resulted in transition dates that matched the interannual variability of observed phenology dates better than the original PnET phenology algorithm. The unrepresentatively high interannual variability predicted by PnET suggested the original phenology algorithm was overly sensitive to temperature compared to the new models. The original PnET spring algorithm performed reasonably well at estimating the beginning of the spring transition dates compared to the validation dataset at Hubbard Brook ($r^2=0.44$; RMSE = 10.3), but not as well as the new spring models ($r^2=0.69$ –0.73; RMSE = 3.6–4.3). For the fall, the original PnET algorithm

did well at predicting the average end of fall date (October 21st compared to the average observed date of October 19th) at Hubbard Brook, but the high variance in PnET's estimates resulted in overall poor fits with observed transitions ($r^2=0.03$; RMSE = 9.3). Therefore, we found our new models performed much better at predicting independent ground observations compared to PnET.

Incorporating new spring and fall phenology models into PnET also resulted in shorter phenology transitions that had better agreement with transition duration estimated from PhenoCam data at Bartlett Experimental Forest. The original PnET spring phenology model tended to predict the onset of spring too early in the season (27 April) followed by a disproportionately slow leaf development ending much later than observations (25 Jun). This almost 2 months of leaf development was uncharacteristically long compared to the average 16-day spring transitions estimated from PhenoCam. The new spring phenology models resulted in the initiation of leaf development starting later (08 May) and end of leaf development ending much earlier (25 May) than the original PnET model, resulting in a leaf development period of 16–17 days on average. The duration of the

fall transition period using the PnET model was also longer than using the new models. The PnET fall phenology model started fall senescence approximately around the same time as the new models (26 Sep) but the end of fall senescence was on average 9 days longer (25 Oct) than the new models. As a result, the PnET fall transitions lasted nearly a month, while the observations and new model predictions lasted approximately 3 weeks. The early initiation in spring, and late fall senescence from PnET phenology models resulted in longer season lengths overall, but the long duration of transition periods resulted in a shorter period with maximum LAI (Fig. 5). The more rapid phenology transitions resulting from the new models showed clear improvements

in spring LAI, but for the years with LAI data available, the PnET fall models resulted in better fits. However, there were other problems associated with fall transitions produced by PnET. The beginning of the fall transition was a fixed date (i.e., no interannual variability), and the end was sometimes 2 weeks later than observations from PhenoCam. As a result, the duration and the end of the fall transitions were found to be unrealistically variable.

We found that altering the phenology algorithms in PnET resulted in relatively small improvements between daily modeled and measured ecosystem carbon exchange (Fig. 5). The most notable improvements were found between modeled and measurement NEP. The more rapid phenology

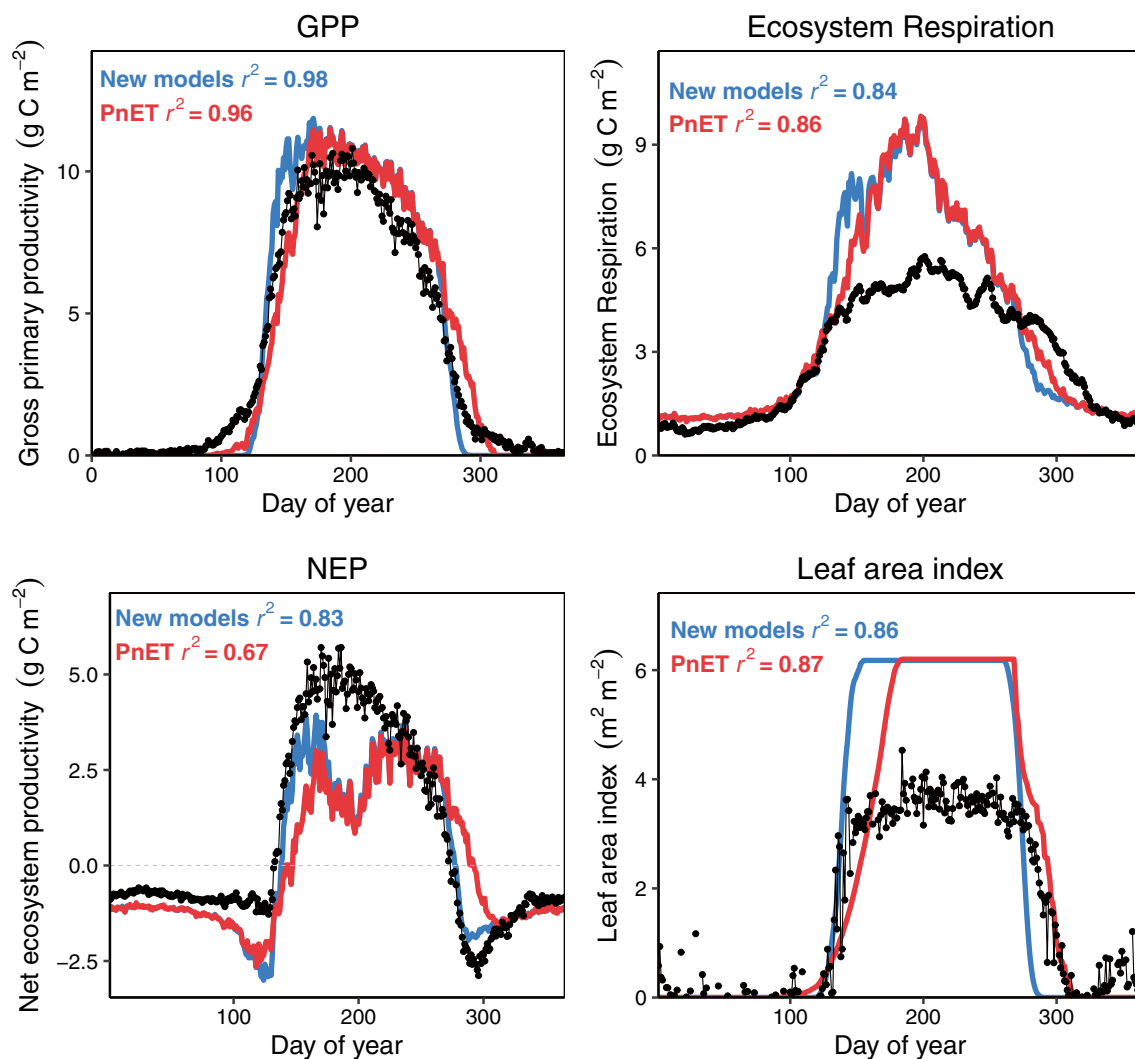


Fig. 5 Average daily observed (black) ecosystem carbon exchange and leaf area index (LAI) at Bartlett Experimental Forest versus daily modeled values using the original PnET phenology model (red) and the average of the new phenology models (blue). Ecosystem carbon exchange variables [gross primary productivity (GPP), ecosystem respiration, and net ecosystem productivity (NEP)] are the mean daily

estimates from 2004 to 2019. The r^2 values represent the relationship between average daily measured vs. modeled values. Modeled and measured LAI are mean daily values from 2013 to 2015. LAI was measured using the fraction of PAR (fPAR) technique described in Toda and Richardson (2018)

transitions from the new models resulted in better agreement with measured NEP. The slow transitions from the original PnET models resulted in unrepresentatively slow increases in NEP, and in the late fall senescence resulted unrealistically late net carbon uptake (Fig. 5). However, daily GPP and ecosystem respiration were less impacted by the phenology algorithms than we had expected; the original PnET phenology subroutine produced daily estimates of carbon exchange that performed as well as using the new phenology models. The slower transitions used in the original PnET model fit the gradual increases and decreases found in tower-based measurements of GPP and ecosystem respiration reasonably well (Fig. 5). As a result, we found the new spring and fall models resulted in minimal differences in the agreement between modeled and measured values of GPP and ecosystem respiration.

Our models demonstrated smaller advancements ($2.6\text{--}3.6$ days decade⁻¹) in spring compared to the PnET spring model (advancing at 3.8 days decade⁻¹) using the projection with the lowest bias compared to Daymet data (EC-Earth RCA4, RCP 8.5). In the fall, the PnET phenology model predicted the end of canopy senescence around the same time as the new models, but with much higher interannual variability than the new models. The original PnET fall model predicted smaller future changes in fall transitions than new models, with fall arriving 1.1 day decade⁻¹ later compared to $1.9\text{--}2.0$ days decade⁻¹ using the new models.

Implications for annual PnET model outputs

Despite the differences in phenology model structure, the new phenology models had relatively small effects on annual ecosystem carbon and water cycling. The effects of longer growing seasons using the original PnET phenology algorithm (i.e., earlier spring onset and later fall senescence compared to the new models) were compensated by longer and slower transitions to and from maximum LAI (i.e., the right answer was obtained for the wrong reason). As a result, there were only subtle differences in annual outputs of carbon and water cycling resulting from altering the models. Compared to other annual model outputs, the largest differences were seen in modeled GPP and NEP (Fig. 6). New models resulted in lower interannual variability of GPP and NEP compared to the original PnET phenology model. Using the original PnET phenology models resulted in modeled GPP increasing at a slightly faster rate (4.3 ± 0.5 g C m⁻² year⁻¹) compared to the new models (3.7 ± 0.1 g C m⁻² year⁻¹). This relatively small difference in modeled GPP over the sampling period resulted in a very slight increasing trend for modeled NEP using the PnET phenology model (0.1 ± 0.5 g C m⁻² year⁻¹), compared to the slight decreasing trend in

NEP using the new models (-0.3 ± 0.2 g C m⁻² year⁻¹). These differences demonstrate that incorporating new phenology models does indeed result in differences in annual modeled GPP and NEP over long timescales. Altering the phenology models had smaller effects on projections of water use (transpiration and WUE) and nitrogen cycling (Nmin). Similar to the carbon exchange variables, the timing of transpiration was altered with the optimized phenology models and concentrated over a shorter period, but over annual timescales the effects were fairly small. As a result, the future projections of annual transpiration and WUE show minor sensitivity to altering the PnET phenology model to the new models.

By comparing the output using one fixed transition (i.e., no interannual variability of spring or fall transitions) to output with both seasons fixed to constant transition dates, we can isolate the effect of either spring or fall transitions on model output. Using these fixed phenology dates, we found that longer seasons in the spring had larger effects on modeled carbon and water cycling compared to longer seasons in the fall (Fig. 7). For every day the season was extended in the spring, GPP increased by an average of 10.2 ± 0.2 g C m⁻² compared to 4.8 ± 0.1 g C m⁻² for the fall. Extended seasons in the spring or fall had smaller effects on ecosystem respiration than GPP, with slopes of 4.4 ± 0.1 g C m⁻² day⁻¹ and 3.3 ± 0.1 g C m⁻² day⁻¹ for spring and fall, respectively. As a result, NEP was also more sensitive to the spring transition date compared to fall, increasing by 5.7 ± 0.1 g C m⁻² day⁻¹ and 1.5 ± 0.1 g C m⁻² day⁻¹ for spring and fall, respectively. We found similar effects on water cycling; longer growing seasons in the spring had larger effects on transpiration (0.29 ± 0.005 cm m⁻² day⁻¹) compared to fall (0.14 ± 0.003 cm m⁻² day⁻¹), as with water use efficiency. The longer seasons had relatively little effect on nitrogen mineralization on annual scales. Based on these models' outputs, our analysis shows that for each day the season gets longer, spring can have as much as double the effect on the net carbon uptake and water use compared to fall.

Our calibrated phenology algorithms predicted springs will have smaller shifts in growing season length compared to the fall (Fig. 4, in days decade⁻¹), but our biogeochemical model predicts springs will have larger effects on carbon cycling for every day of change (Fig. 7, in g C m⁻² day⁻¹). By combining these projections, we estimate spring will have a larger overall effect on NEP in the future (increasing by $5.2\text{--}8.7$ g m⁻² C decade⁻¹ in the spring compared to $1.7\text{--}4.8$ g C m⁻² decade⁻¹ in the fall), but spring and fall will have similar influences on GPP (increasing by $9.3\text{--}15.3$ g C m⁻² decade⁻¹ in the spring compared to $5.3\text{--}15.4$ g C m⁻² decade⁻¹ in the fall). We found that increases in respiration in the fall were higher relative to GPP, therefore spring phenology had a larger overall effect on NEP.

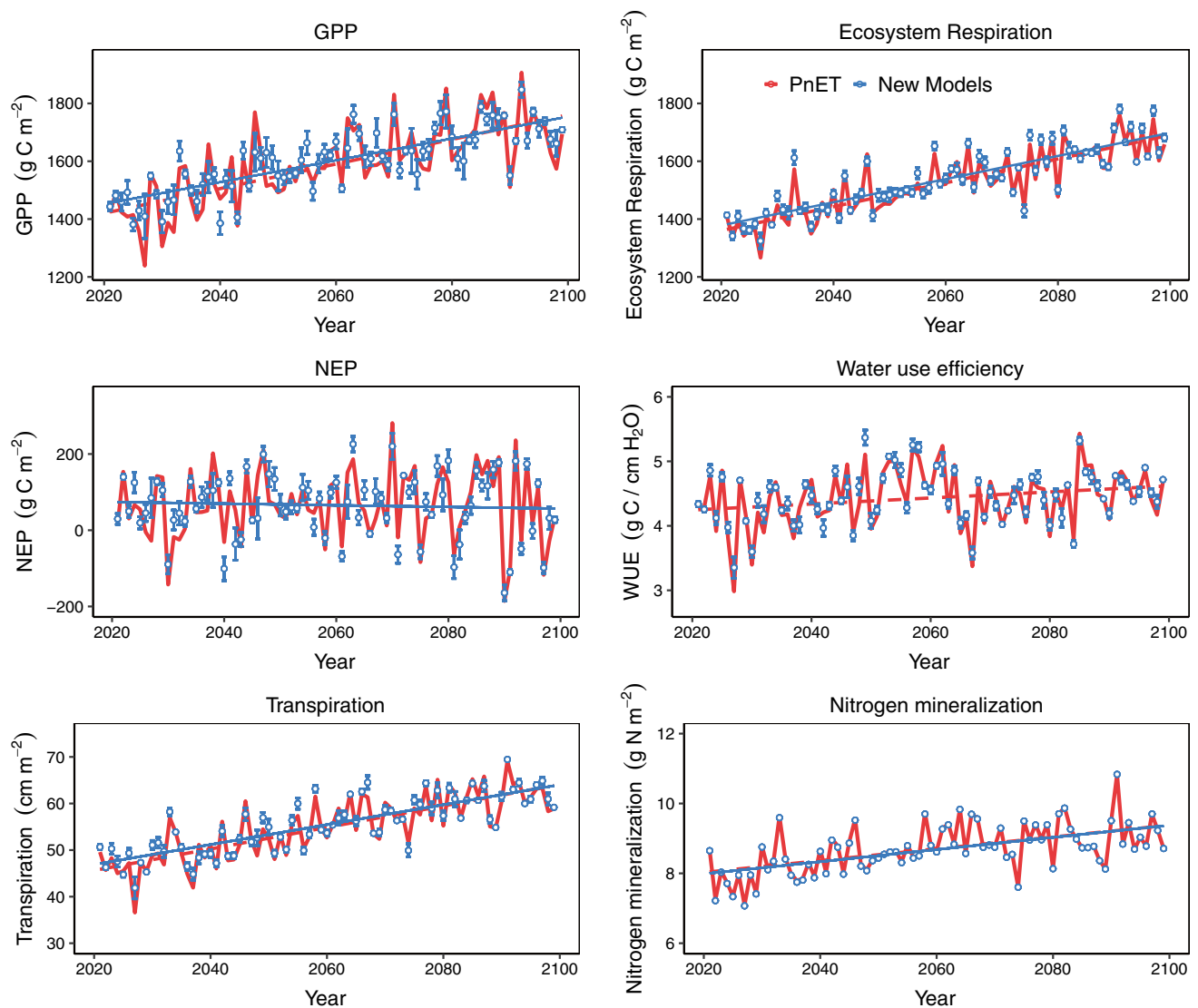


Fig. 6 Annual modeled carbon exchange [gross primary productivity (GPP), ecosystem respiration, and net ecosystem productivity (NEP)], water-use efficiency, transpiration, and nitrogen mineralization using PnET phenology subroutine (red line and dashed red trendline) modeled to year 2100 compared to output using the new models (blue cir-

cles and solid blue trendline). The error bars represent the standard deviation of the outputs using different combinations of the new models. Trend lines represent linear regressions of the timeseries—only regression lines that were significant at an alpha of 0.05 are included

Discussion

Spring and fall phenology are often overlooked aspects of biogeochemical modeling despite their importance in defining the duration of photosynthetic carbon uptake and plant water use each year. Here, we use regional phenology data to train leaf development and leaf senescence algorithms in an ecosystem-level biogeochemical model. Using empirical data to train phenology algorithms is becoming easier due to the growing amount of open-source phenology data (Crimmins et al. 2017; Richardson et al. 2018; Templ et al. 2018) and tools (Hufkens et al. 2018) available. We found

that incorporating the new phenology algorithms into PnET resulted in outputs that better fit measured NEP (as seen in Desai et al. 2010), but the model did not show any notable improvements for GPP and ecosystem respiration. On the daily scale, we show that the phenology algorithms exert a strong control on the timing of carbon and water cycling, but had surprisingly small effects on the magnitude and long-term trends of the future annual predictions. By improving the phenology subroutine in biogeochemical models, we strengthen a well-documented weakness of ecosystem modeling (Levis and Bonan 2004; Randerson et al. 2009; Richardson et al. 2012).

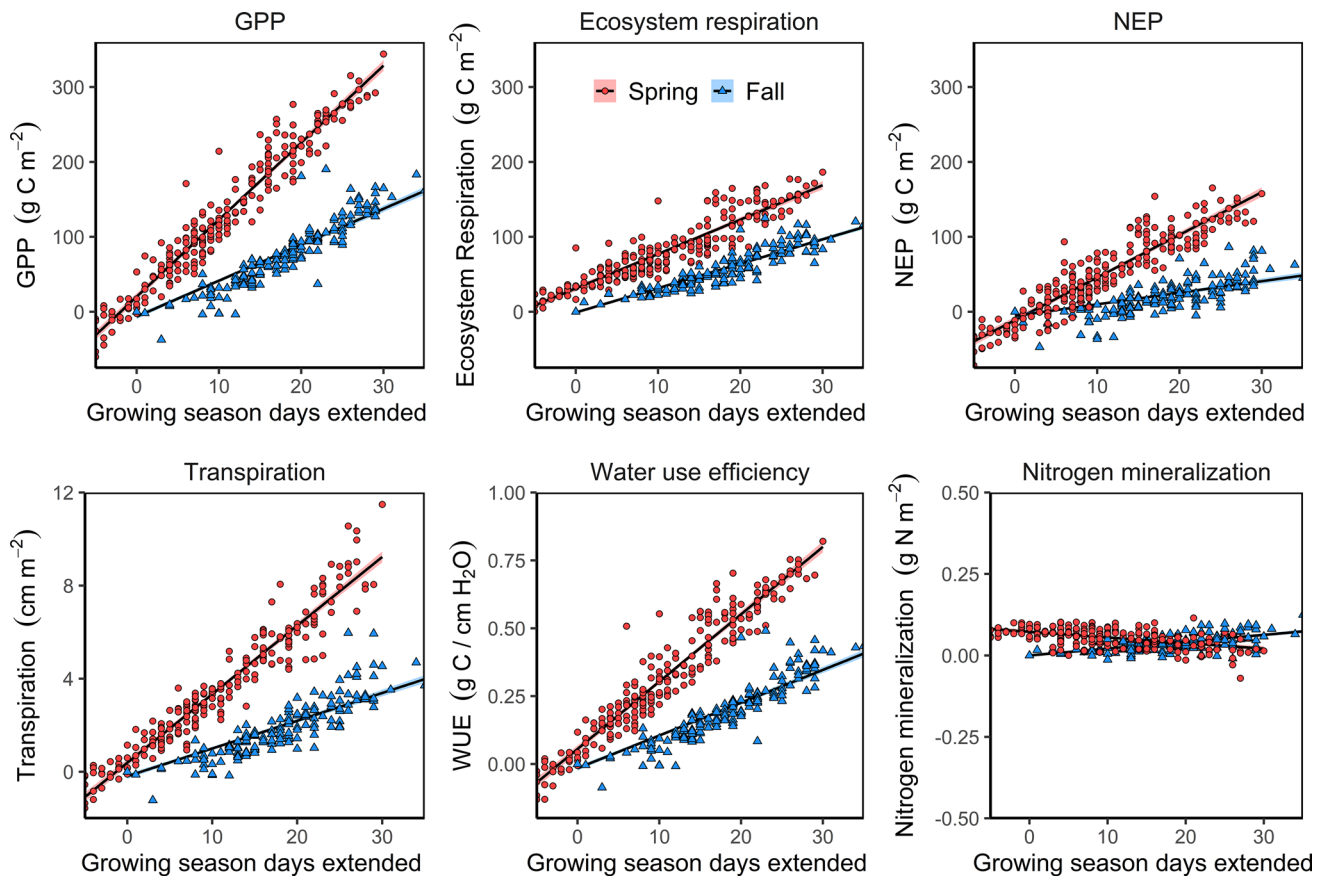


Fig. 7 Relative influence of spring and fall phenological transitions on annual carbon cycling [gross primary productivity (GPP), ecosystem respiration, and net ecosystem productivity (NEP)], transpiration, water use efficiency, and nitrogen mineralization. These values were derived by taking the difference between one-season-fixed models and models with both spring and fall fixed phenology transitions to

remove interannual variability unrelated to phenology transitions. The red circles represent models with new spring phenology algorithms and fixed fall transition dates (i.e., no interannual variability) and the blue triangles represent models with new fall phenology algorithms and fixed spring transition dates. Trend lines represent linear regressions—all regression lines were significant at an alpha of 0.05

We found that parameterizing phenology models using empirical data from multiple PhenoCam sites allowed us to produce phenology models that are broadly applicable to the northeastern deciduous forests of North America. We did not find that spring models requiring chilling requirements showed any improvements over models that used warming temperatures and photoperiod, similar to other North American (Melaas et al. 2016) and European (Basler 2016) tree species. This likely suggests the species in this region have sufficient winter chilling temperatures to release bud dormancy (Vitasse et al. 2011). Previous work has shown that including chilling temperatures can improve species specific models (Chiang and Brown 2007; Kaduk and Los 2011; Jeong et al. 2012; Flynn and Wolkovich 2018) but our canopy-level approach did not benefit from chilling temperatures enough to justify the additional model parameters. It is possible that in the future, warmer winters in North America may require more sophisticated models that include chilling requirements (Morin et al. 2009), but this is not currently

the case in these northern deciduous forests. This illustrates an inherent weakness in extrapolating beyond current conditions into a future where phenology may be affected by other factors.

Our results demonstrate that fall transition dates derived from PhenoCam are reproducible and robust enough to train fall models that can be validated against ground observations. We found that using the yellowness of a forest canopy from PhenoCam, in addition to greenness, can be a useful metric for predicting the end of fall senescence in these forest types (Kosmala et al. 2016). We also found that incorporating these fall models into PnET allowed us to improve predicted leaf senescence dates produced from the original phenology algorithms that were overly sensitive to climate drivers. These are promising findings for the future of modeling fall senescence in deciduous forests. However, we also find our fall models were unable to capture the year-to-year variability as well as the spring models—also found by Vitasse et al. (2011). Our more complex fall model (CDDP,

including photoperiod and temperature) also did not show any improvements over the simpler temperature only model. As previously documented, we found fall transitions were more difficult to characterize compared to spring transitions because they can be unexpectedly sudden in response to stressful environmental conditions including heavy rainfall events, heat stress (Xie et al. 2018), frost, or wind events (Gallinat et al. 2015). Compared to the literature, our models predict future changes in spring and fall transition dates that are either on track with, or less than, observations from other North American deciduous forests over the last few decades (de Beurs and Henebry 2005; Jeong and Medvigy 2014; Piao et al. 2019). Our models project growing season length will be more affected in the fall compared to the spring as temperatures are expected to increase over the next few decades [as documented by Jeong et al. (2011) and Zhu et al. (2012) from the early 1980s to mid 2000s]. Combined, our projections suggest growing season lengths will change by 2.0–4.7 days decade⁻¹ depending on future climate projections. These are similar to other recent projections for north-eastern North America (Piao et al. 2019) but are less than historical observations (Wang et al. 2015).

Using the original PnET spring algorithm, the beginning of spring was triggered too early, had higher variability, and lasted much longer than observed transitions from PhenoCam (as seen previously in Klosterman et al. 2018). We also found that original PnET algorithms predicted highly variable end of fall senescence dates, and because the start of senescence was fixed (i.e., no interannual variability), the duration of the fall season was poorly constrained. Despite the clear improvements our models had on season lengths generated by PnET, we were surprised that altering the phenology algorithms only marginally affected the magnitude of modeled carbon, water, and nutrient cycling on annual timescales (contrary to our hypothesis that shorter seasons would lead to decreased modeled carbon, water, and nutrient use). On a daily timescale, we found incorporating the new algorithms resulted in faster seasonal changes in carbon uptake and transpiration. As a result, the slower phenological transitions between active and dormant periods in the original PnET model compensated for uncharacteristically long growing season lengths, and annual totals were similar in magnitude to the new model outputs. We also note the longer seasons predicted by the original PnET algorithms are likely benefited by the 14% of the stand composed of evergreen conifers (Ouimette et al. 2018). Conifers tend to have longer periods of carbon uptake compared to their deciduous counterparts, and the tower-based GPP occurring before leaves are present and after leaves senesce is likely from uptake from conifer species, coincidentally resulting in better agreement with the original PnET outputs. Despite this, our new models resulted in notable improvements in

PnET's ability to model NEP compared to tower-based estimates.

All of our simulations suggested increasing trends in GPP and plant water use (i.e., transpiration) at Bartlett Experimental Forest over the next 80 years (see also Chen et al. 2016). Our projected GPP increases of 5–8 g C m⁻² day⁻¹ are on the upper end of the empirical observations in the literature (Falge et al. 2002; Piao et al. 2007; Baldocchi 2008). Projections of future NEP and WUE, however, demonstrated relatively stable long-term trends. Modeled NEP showed increased variability under future climate change, but no increasing trend due to counteracting increases in GPP and ecosystem respiration in response to longer growing seasons and elevated temperatures (as also described by Piao et al. 2007; Duveneck and Thompson 2017). The stable long-term trend in NEP in this study was largely due to the long spin-up times used in the model calibration process (Thornton and Rosenbloom 2005) and our decision not to simulate past disturbance. We accounted for the effects of long spin-up times (and incidentally, increasing temperature) by taking the difference of modeled outputs from a fixed spring and fall phenology model. Differences between dynamic and fixed phenology models demonstrated that phenological change (i.e., the longer growing seasons) did, in fact, have a positive effect on annual modeled NEP.

Our results indicate that, compared to the fall, longer growing seasons in the spring have a larger effect on carbon and water cycling for every day of change, as hypothesized. Compared to fall phenology, spring phenology had a larger effect on annual GPP relative to ecosystem respiration and therefore also had a larger effect on NEP (contrary to findings by Wu et al. 2013a, b). This may be partly due to higher temperatures in the spring also having higher rates of incoming solar radiation and soil moisture compared to the fall. Increasing temperatures in fall months may actually result in net carbon losses despite the later leaf senescence (Piao et al. 2008) because ecosystem respiration can respond to increasing fall temperatures more than GPP (Barichivich et al. 2013). These findings are, however, different from Keenan et al. (2014), where later fall phenology was associated with higher rates of both NEP and GPP compared to the spring. Presumably, the larger effects on modeled ecosystem carbon and water balance in our models are due to longer daylength during leaf development compared to leaf senescence allowing for higher levels of photosynthetic uptake (see Zhang et al. 2020). The seasonal effects on ecosystem respiration were less apparent because ecosystem respiration is less affected by daylength and more affected by temperature. While our projections suggest that both early springs and late falls will result in increased carbon uptake, spring was clearly more influential on annual model outputs due to the longer days and higher rates of GPP early in the season compared to the fall.

Conclusions

Training phenology models at four different seasonal transitions (start of onset, end of onset, beginning of senescence, and end of the senescence) allowed us to improve a biogeochemical model's ability to estimate a deciduous forest's growing season length. The optimized phenology subroutine resulted in season lengths that were shorter and more representative of phenology observations. Our work establishes a theoretically rigorous, but still straightforward, method to improve existing ecosystem-level biogeochemical models to be more representative of actual ecosystem processes. By highlighting direct effects of longer growing seasons on ecosystem carbon and water cycling, we quantitatively showed the importance of improving phenology models in modeling ecosystem-scale processes.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00442-022-05296-4>.

Acknowledgements We thank the many collaborators in their efforts to support PhenoCam data collection, including site PIs and technicians.

Author contribution statement AT and ADR designed the project, performed the analysis, and wrote the manuscript. ASB collected and curated ground-observation data. KH, CS, and BS contributed to programming and analysis. AT and SO performed the modeling. All authors reviewed and provided editorial comments.

Funding The development of the PhenoCam network has been funded by the Northeastern States Research Cooperative, NSF's Macrosystems Biology program (awards EF-1065029 and EF-1702697), and DOE's Regional and Global Climate Modeling program (award DE-SC0016011). Research at the Bartlett Experimental Forest is supported by the USDA Forest Service's Northern Research Station. We acknowledge funding support from the National Science Foundation for the Hubbard Brook Long Term Ecological Research (LTER) program, awards #DEB-1114804 and #DEB-1637685 and the Harvard Forest LTER award #DEB 1832210; and support from the Northeastern States Research Cooperative #12DG11242307065.

Data availability Data used to train phenology models is available through the PhenoCam Network (phenocam.nau.edu). Eddy covariance data used in the analysis are available through the AmeriFlux Network (<https://ameriflux.lbl.gov/sites/siteinfo/US-Bar>).

Code availability The R code used to download data, extract transition dates, then train models can be found at the repository: https://github.com/teetsaf/PhenoCam_data_download_and_modeling. The repository also contains instructions and C++ code for incorporating new phenology models into PnET.

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

- Aber JD, Federer CA (1992) A generalized, lumped-parameter model of photosynthesis, evapotranspiration and net primary production in temperate and boreal forest ecosystems. *Oecologia* 92:463–474. <https://doi.org/10.1007/BF00317837>
- Aber JD, Reich PB, Goulden ML (1996) Extrapolating leaf CO₂ exchange to the canopy: a generalized model of forest photosynthesis compared with measurements by eddy correlation. *Oecologia* 106:257–265. <https://doi.org/10.1007/BF00328606>
- Aber JD, Ollinger SV, Driscoll CT (1997) Modeling nitrogen saturation in forest ecosystems in response to land use and atmospheric deposition. *Ecol Modell* 101:61–78. [https://doi.org/10.1016/S0304-3800\(97\)01953-4](https://doi.org/10.1016/S0304-3800(97)01953-4)
- Baldocchi D (2008) Breathing of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. *Aust J Bot* 56:1–26. <https://doi.org/10.1071/BT07151>
- Barichivich J, Briffa KR, Myneni RB, Osborn TJ, Melvin TM, Ciais P, Piao S, Tucker C (2013) Large-scale variations in the vegetation growing season and annual cycle of atmospheric CO₂ at high northern latitudes from 1950 to 2011. *Glob Chang Biol* 19:3167–3183. <https://doi.org/10.1111/gcb.12283>
- Basler D (2016) Evaluating phenological models for the prediction of leaf-out dates in six temperate tree species across central Europe. *Agric for Meteorol* 217:10–21. <https://doi.org/10.1016/j.agrformet.2015.11.007>
- Beamesderfer ER, Altaf Arain M, Khomik M, Brodeur JJ, Burns BM (2020) Response of carbon and water fluxes to meteorological and phenological variability in two eastern North American forests of similar age but contrasting species composition: a multiyear comparison. *Biogeosciences* 17:3563–3587. <https://doi.org/10.5194/bg-17-3563-2020>
- Blümel K, Chmielewski FM (2012) Shortcomings of classical phenological forcing models and a way to overcome them. *Agric for Meteorol* 164:10–19. <https://doi.org/10.1016/j.agrformet.2012.05.001>
- Cannell M, Smith R (1983) Thermal time, chill days and prediction of budburst in *Picea sitchensis*. *J Appl Ecol* 20:951–963. <https://doi.org/10.2307/2403139>
- Chen M, Melaas EK, Gray JM, Friedl MA, Richardson AD (2016) A new seasonal-deciduous spring phenology submodel in the Community Land Model 4.5: impacts on carbon and water cycling under future climate scenarios. *Glob Chang Biol* 22:3675–3688. <https://doi.org/10.1111/gcb.13326>
- Chiang JM, Brown KJ (2007) Improving the budburst phenology subroutine in the forest carbon model PnET. *Ecol Modell* 205:515–526. <https://doi.org/10.1016/j.ecolmodel.2007.03.013>
- Črepinšek Z, Kajfež-Bogataj L, Bergant K (2006) Modelling of weather variability effect on fitophenology. *Ecol Modell* 194:256–265. <https://doi.org/10.1016/j.ecolmodel.2005.10.020>
- Crimmins TM, Crimmins MA, Gerst KL, Rosemartin AH, Weltzin JF (2017) USA National Phenology Network's volunteer-contributed observations yield predictive models of phenological transitions. *PLoS ONE* 12:1–17. <https://doi.org/10.1371/journal.pone.0182919>
- de Beurs KM, Henebry GM (2005) Land surface phenology and temperature variation in the International Geosphere-Biosphere Program high-latitude transects. *Glob Chang Biol* 11:779–790. <https://doi.org/10.1111/j.1365-2486.2005.00949.x>

- Delpierre N, Dufrêne E, Soudani K, Ulrich E, Cecchini S, Boé J, François C (2009) Modelling interannual and spatial variability of leaf senescence for three deciduous tree species in France. *Agric for Meteorol* 149:938–948. <https://doi.org/10.1016/j.agrfor.2008.11.014>
- Desai AR, Helliker BR, Moorcroft PR, Andrews AE, Berry JA (2010) Climatic controls of interannual variability in regional carbon fluxes from top-down and bottom-up perspectives. *J Geophys Res Biogeosciences* 115:1–15. <https://doi.org/10.1029/2009jg001122>
- Dragoni D, Schmid HP, Wayson CA, Potter H, Grimmer CSB, Randolph JC (2011) Evidence of increased net ecosystem productivity associated with a longer vegetated season in a deciduous forest in south-central Indiana, USA. *Glob Chang Biol* 17:886–897. <https://doi.org/10.1111/j.1365-2486.2010.02281.x>
- Duveneck MJ, Thompson JR (2017) Climate change imposes phenological trade-offs on forest net primary productivity. *J Geophys Res Biogeosciences* 122:2298–2313. <https://doi.org/10.1002/2017JG004025>
- Estiarte M, Peñuelas J (2015) Alteration of the phenology of leaf senescence and fall in winter deciduous species by climate change: effects on nutrient proficiency. *Glob Chang Biol* 21:1005–1017. <https://doi.org/10.1111/gcb.12804>
- Estrella N, Menzel A (2006) Responses of leaf colouring in four deciduous tree species to climate and weather in Germany. *Clim Res* 32:253–267. <https://doi.org/10.3354/cr032253>
- Fahey TJ, Siccama TG, Driscoll CT, Likens GE, Campbell J, Johnson CE, Battles JJ, Aber JD, Cole JJ, Fisk MC, Groffman PM, Hamburg SP, Holmes RT, Schwarz PA, Yanai RD (2005) The biogeochemistry of carbon at Hubbard Brook. *Biogeochemistry* 75:109–176. <https://doi.org/10.1007/s10533-004-6321-y>
- Falge E, Baldocchi D, Tenhunen J, Aubinet M, Bakwin P, Berbigier P, Bernhofer C, Burba G, Clement R, Davis KJ et al (2002) Seasonality of ecosystem respiration and gross primary production as derived from FLUXNET measurements. *Agric for Meteorol* 113:53–74. [https://doi.org/10.1016/S0168-1923\(02\)00102-8](https://doi.org/10.1016/S0168-1923(02)00102-8)
- Flynn DFB, Wolkovich EM (2018) Temperature and photoperiod drive spring phenology across all species in a temperate forest community. *New Phytol* 219:1353–1362. <https://doi.org/10.1111/nph.15232>
- Gallinat AS, Primack RB, Wagner DL (2015) Autumn, the neglected season in climate change research. *Trends Ecol Evol* 30:169–176. <https://doi.org/10.1016/j.tree.2015.01.004>
- Goulden ML, Munger JW, Fan SM, Daube BC, Wofsy SC (1996) Measurements of carbon sequestration by long-term eddy covariance: methods and a critical evaluation of accuracy. *Glob Chang Biol* 2:169–182. <https://doi.org/10.1111/j.1365-2486.1996.tb00070.x>
- Hanninen H (1990) Modelling bud dormancy release in trees from cool and temperate regions. *Acta for Fenn* 213:1–47
- Hogg EH, Price DT, Black TA (2000) Postulated feedbacks of deciduous forest phenology on seasonal climate patterns in the Western Canadian interior. *J Clim* 13:4229–4243. [https://doi.org/10.1175/1520-0442\(2000\)013%3c4229:PFODFP%3e2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013%3c4229:PFODFP%3e2.0.CO;2)
- Hollinger DY, Ollinger SV, Richardson AD, Meyers TP, Dail DB, Martin ME, Scott NA, Arkebauer TJ, Baldocchi DD, Clark KL, Curtis PS, Davis KJ, Desai AR, Dragonik D, Goulden ML, Gu L, Katulzz GG, Pallardy SG, Pawu KT, Schmid HP, Stoy PC, Suyker AE, Verma SB (2010) Albedo estimates for land surface models and support for a new paradigm based on foliage nitrogen concentration. *Glob Chang Biol* 16:696–710. <https://doi.org/10.1111/j.1365-2486.2009.02028.x>
- Hufkens K, Basler D, Milliman T, Melaas EK, Richardson AD (2018) An integrated phenology modelling framework in R. *Methods Ecol Evol* 9:1276–1285. <https://doi.org/10.1111/2041-210X.12970>
- Jeong SJ, Medvigy D (2014) Macroscale prediction of autumn leaf coloration throughout the continental United States. *Glob Ecol Biogeogr* 23:1245–1254. <https://doi.org/10.1111/geb.12206>
- Jeong SJ, Ho CH, Gim HJ, Brown ME (2011) Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008. *Glob Chang Biol* 17:2385–2399. <https://doi.org/10.1111/j.1365-2486.2011.02397.x>
- Jeong SJ, Medvigy D, Shevliakova E, Malyshev S (2012) Uncertainties in terrestrial carbon budgets related to spring phenology. *J Geophys Res Biogeosciences* 117:1–17. <https://doi.org/10.1029/2011JG001868>
- Kaduk JD, Los SO (2011) Predicting the time of green up in temperate and boreal biomes. *Clim Change* 107:277–304. <https://doi.org/10.1007/s10584-010-9951-1>
- Keenan TF, Gray J, Friedl MA et al (2014) Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nat Clim Chang* 4:598–604. <https://doi.org/10.1038/NCLIMATE2253>
- Klosterman S, Hufkens K, Richardson AD (2018) Later springs green-up faster: the relation between onset and completion of green-up in deciduous forests of North America. *Int J Biometeorol* 62:1645–1655. <https://doi.org/10.1007/s00484-018-1564-9>
- Korner C, Basler D (2010) Phenology Under Global Warming. *Science* (80-) 327:1461–1462
- Kosmala M, Crall A, Cheng R, Hufkens K, Henderson S, Richardson AD (2016) Season spotter: using citizen science to validate and scale plant phenology from near-surface remote sensing. *Remote Sens* 8:1–22. <https://doi.org/10.3390/rs8090726>
- Lany NK, Ayres MP, Stange EE, Sillett TS, Rodenhouse NL, Holmes RT (2016) Breeding timed to maximize reproductive success for a migratory songbird: the importance of phenological asynchrony. *Oikos* 125:656–666. <https://doi.org/10.1111/oik.02412>
- Likens GE, Bormann FH, Johnson NM, Fisher DW, Pierce RS (1970) Effects of forest cutting and herbicide treatment on nutrient budgets in the Hubbard Brook. *Ecol Monogr* 40:23–47
- Levis S, Bonan GB (2004) Simulating springtime temperature patterns in the community atmosphere model coupled to the community land model using prognostic leaf area. *J Clim* 17:4531–4540. <https://doi.org/10.1175/3218.1>
- Liu Q, Fu YH, Zhu Z, Liu Y, Liu Z, Huang M, Janssens IA, Piao S (2016) Delayed autumn phenology in the Northern Hemisphere is related to change in both climate and spring phenology. *Glob Chang Biol* 22:3702–3711. <https://doi.org/10.1111/gcb.13311>
- Maurer EP, Brekke L, Pruitt T, Duffy PB (2007) Fine-resolution climate projections enhance regional climate change impact studies. *Eos Trans Am Geophys Union* 88:504–504. <https://doi.org/10.1029/2007eo470006>
- Mearns L, McGinnis S, Korytina D, Arritt R, Biner S, Bukovsky M, Chang HI, Christensen O, Herzmann D, Jiao Y, et al (2017) The NA-CORDEX dataset, version 1.0. NCAR Climate Data Gateway, Boulder CO, accessed 2021–08–31, <https://doi.org/10.5065/D6SJ1JCH>
- Melaas EK, Friedl MA, Richardson AD (2016) Multiscale modeling of spring phenology across Deciduous Forests in the Eastern United States. *Glob Chang Biol* 22:792–805. <https://doi.org/10.1111/gcb.13122>
- Menzel A, Sparks TH, Estrella N, Koch E, Aaasa A, Ahas R, Alm-Kubler K, Bissolli P, Braslavská O, Briede A et al (2006) European phenological response to climate change matches the warming pattern. *Glob Chang Biol* 12:1969–1976. <https://doi.org/10.1111/j.1365-2486.2006.01193.x>
- Miller-Rushing AJ, Primack RB (2008) Global warming and flowering times in Thoreau's Concord: a community perspective. *Ecology* 89:332–341. <https://doi.org/10.1890/07-0068.1>
- Morin X, Lechowicz M, Augspurger C, O'Keefe J, Viner D, Chuine I (2009) Leaf phenology in 22 North American tree species during

- the 21st century. *Glob Chang Biol* 15:961–975. <https://doi.org/10.1111/j.1365-2486.2008.01735.x>
- Ollinger SV, Smith ML (2005) Net primary production and canopy nitrogen in a temperate forest landscape: an analysis using imaging spectroscopy, modeling and field data. *Ecosystems* 8:760–778. <https://doi.org/10.1007/s10021-005-0079-5>
- Ollinger SV, Aber JD, Reich PB, Freuder RJ (2002) Interactive effects of nitrogen deposition, tropospheric ozone, elevated CO₂ and land use history on the carbon dynamics of northern hardwood forests. *Glob Chang Biol* 8:545–562. <https://doi.org/10.1046/j.1365-2486.2002.00482.x>
- Ollinger SV, Richardson AD, Martin ME, Hollinger DY, Frolking SE, Reich PB, Plourde LC, Katul GG, Munger JW, Oren R et al (2008) Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. *Proc Natl Acad Sci U S A* 105:19336–19341. <https://doi.org/10.1073/pnas.0810021105>
- Ouimette AP, Ollinger SV, Richardson AD et al (2018) Carbon fluxes and interannual drivers in a temperate forest ecosystem assessed through comparison of top-down and bottom-up approaches. *Agric for Meteorol* 256–257:420–430. <https://doi.org/10.1016/j.agrformet.2018.03.017>
- Piao S, Friedlingstein P, Ciais P, Viovy N, Demarty J (2007) Growing season extension and its impact on terrestrial carbon cycle in the Northern Hemisphere over the past 2 decades. *Global Biogeochem Cycles* 21:1–11. <https://doi.org/10.1029/2006GB002888>
- Piao S, Ciais P, Friedlingstein P, Peylin P, Reichstein M, Luyssaert S, Margolis H, Fang J, Barr A, Chen A, Grelle A, Hollinger DY, Laurila T, Lindroth A, Richardson AD, Vesala T (2008) Net carbon dioxide losses of northern ecosystems in response to autumn warming. *Nature* 451:49–52. <https://doi.org/10.1038/nature06444>
- Piao S, Liu Q, Chen A, Janssens IA, Fu Y, Dai J, Liu L, Lian X, Shen M, Zhu X (2019) Plant phenology and global climate change: current progresses and challenges. *Glob Chang Biol* 25:1922–1940. <https://doi.org/10.1111/gcb.14619>
- R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>
- Randerson JT, Hoffman FM, Thornton PE, Mahowald NM, Lindsay K, Lee YH, Nevison CD, Doney SC, Bonan G, Stöckli R, Covey C, Running SW, Fung IY (2009) Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models. *Glob Chang Biol* 15:2462–2484. <https://doi.org/10.1111/j.1365-2486.2009.01912.x>
- Reyer CPO, Leuzinger S, Rammig A, Wolf A, Bartholomeus RP, Bonfante A, de Lorenzi F, Dury M, Gloning P, Abou Jaoudé R et al (2013) A plant's perspective of extremes: terrestrial plant responses to changing climatic variability. *Glob Chang Biol* 19:75–89. <https://doi.org/10.1111/gcb.12023>
- Richardson AD, Bailey AS, Denny EG, Martin CW, O'Keefe J (2006) Phenology of a northern hardwood forest canopy. *Glob Chang Biol* 12:1174–1188. <https://doi.org/10.1111/j.1365-2486.2006.01164.x>
- Richardson AD, Hollinger DY, Dail DB, Lee JT, Munger JW, O'keefe J, (2009) Influence of spring phenology on seasonal and annual carbon balance in two contrasting New England forests. *Tree Physiol* 29:321–331. <https://doi.org/10.1093/treephys/tpn040>
- Richardson AD, Black TA, Ciais P, Delbart N, Friedl MA, Gobron N, Hollinger DY, Kutsch WL, Longdoz B, Luyssaert S et al (2010) Influence of spring and autumn phenological transitions on forest ecosystem productivity. *Philos Trans R Soc B Biol Sci* 365:3227–3246. <https://doi.org/10.1098/rstb.2010.0102>
- Richardson AD, Anderson RS, Arain MA, Barr AG, Bohrer G, Chen G, Chen JM, Ciais P, Davis KJ, Desai AR et al (2012) Terrestrial biosphere models need better representation of vegetation phenology: results from the North American carbon program site synthesis. *Glob Chang Biol* 18:566–584. <https://doi.org/10.1111/j.1365-2486.2011.02562.x>
- Richardson AD, Hufkens K, Milliman T, Aubrecht DM, Chen M, Gray JM, Johnston MR, Keenan TF, Klosterman ST, Kosmala M et al (2018) Tracking vegetation phenology across diverse North American biomes using PhenoCam imagery. *Sci Data* 5:1–24. <https://doi.org/10.1038/sdata.2018.28>
- Richardson AD (2004). AmeriFlux US-Bar Bartlett Experimental Forest. Dataset <https://doi.org/10.17190/AMF/1246030>
- Sanders-DeMott R, Ouimette AP, Lepine LC, Fogarty SZ, Burakowski EA, Contosta AR, Ollinger SV (2020) Divergent carbon cycle response of forest and grass-dominated northern temperate ecosystems to record winter warming. *Glob Chang Biol* 26:1519–1531. <https://doi.org/10.1111/gcb.14850>
- Sonnentag O, Hufkens K, Teshera-Sterne C, Young AM, Friedl M, Braswell BH, Milliman T, O'Keefe J, Richardson AD (2012) Digital repeat photography for phenological research in forest ecosystems. *Agric for Meteorol* 152:159–177. <https://doi.org/10.1016/j.agrformet.2011.09.009>
- Templ B, Koch E, Bolmgren K, Ungersböck M, Paul A, Scheifinger H, Rutishauser T, Busto M, Chmielewski FM, Hájková L et al (2018) Pan European Phenological database (PEP725): a single point of access for European data. *Int J Biometeorol* 62:1109–1113. <https://doi.org/10.1007/s00484-018-1512-8>
- Thorn AM, Xiao J, Ollinger SV (2015) Generalization and evaluation of the process-based forest ecosystem model PnET-CN for other biomes. *Ecosphere* 6:1–27. <https://doi.org/10.1890/ES14-00542.1>
- Thornton PE, Rosenbloom NA (2005) Ecosystem model spin-up: estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model. *Ecol Modell* 189:25–48. <https://doi.org/10.1016/j.ecolmodel.2005.04.008>
- Thornton MM, Wei Y, Thornton PE, Shrestha R, Kao S, Wilson BE (2020) Daymet: station-Level Inputs and Cross-Validation Result for North America, Version 4. ORNL DAAC. <https://doi.org/10.3334/ORNLDAAAC/1850>
- Thornton PE, Shrestha R, Thornton M, Kao SC, Wei Y, Wilson BE (2021) Gridded daily weather data for North America with comprehensive uncertainty quantification. Scientific Data. <https://doi.org/10.1038/s41597-021-00973-0>
- Toda M, Richardson AD (2018) Estimation of plant area index and phenological transition dates from digital repeat photography and radiometric approaches in a hardwood forest in the Northeastern United States. *Agric for Meteorol* 249:457–466. <https://doi.org/10.1016/j.agrformet.2017.09.004>
- Vitasse Y, François C, Delpierre N, Dufrêne E, Kremer A, Chuine I, Delzon S (2011) Assessing the effects of climate change on the phenology of European temperate trees. *Agric for Meteorol* 151:969–980. <https://doi.org/10.1016/j.agrformet.2011.03.003>
- Wang X, Piao S, Xu X et al (2015) Has the advancing onset of spring vegetation green-up slowed down or changed abruptly over the last three decades? *Glob Ecol Biogeogr* 24:621–631. <https://doi.org/10.1111/geb.12289>
- Wu C, Chen JM, Black TA, Price DT, Kurz WA, Desai AR, Gonsamo A, Jassal RS, Gough CM, Bohrer G et al (2013a) Interannual variability of net ecosystem productivity in forests is explained by carbon flux phenology in autumn. *Glob Ecol Biogeogr* 22:994–1006. <https://doi.org/10.1111/geb.12044>
- Wu C, Gough CM, Chen JM, Gonsamo A (2013b) Evidence of autumn phenology control on annual net ecosystem productivity in two temperate deciduous forests. *Ecol Eng* 60:88–95. <https://doi.org/10.1016/j.ecoleng.2013.07.019>
- Xiang Y, Gubian S, Suomela B, Hoeng J (2013) Generalized simulated annealing for global optimization: the GenSA package. *R J* 5:13–28
- Xie Y, Wang X, Wilson AM, Silander JA (2018) Predicting autumn phenology: how deciduous tree species respond to weather

- stressors. *Agric for Meteorol* 250–251:127–137. <https://doi.org/10.1016/j.agrformet.2017.12.259>
- Young AM, Friedl MA, Seyednasrollah B, Beamesderfer E, Carrillo CM, Li X, Moon M, Arain MA, Baldocchi DD, Blanken PD et al (2021) Seasonality in aerodynamic resistance across a range of North American ecosystems. *Agric for Meteorol* 310:1–19. <https://doi.org/10.1016/j.agrformet.2021.108613>
- Zha T, Barr AG, van der Kamp G, Black TA, McCaughey JH, Flanagan LB (2010) Interannual variation of evapotranspiration from forest and grassland ecosystems in western Canada in relation to drought. *Agric for Meteorol* 150:1476–1484. <https://doi.org/10.1016/j.agrformet.2010.08.003>
- Zhang Y, Commans R, Zhou S, Williams AP, Gentile P (2020) Light limitation regulates the response of autumn terrestrial carbon uptake to warming. *Nat Clim Chang* 10:739–743. <https://doi.org/10.1038/s41558-020-0806-0>
- Zhu W, Tian H, Xu X, Pan Y, Chen G, Lin W (2012) Extension of the growing season due to delayed autumn over mid and high latitudes in North America during 1982–2006. *Glob Ecol Biogeogr* 21:260–271. <https://doi.org/10.1111/j.1466-8238.2011.00675.x>

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Aaron Teets^{1,2}  · Amey S. Bailey³ · Koen Hufkens⁴ · Scott Ollinger⁵ · Christina Schädel¹ · Bijan Seyednasrollah⁶ · Andrew D. Richardson^{1,6}

¹ Center for Ecosystem Science and Society (ECOSS), Northern Arizona University, Flagstaff, AZ, USA

² Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, USA

³ USDA Forest Service, Northern Research Station, Durham, NH, USA

⁴ BlueGreen Labs, Melsele, Belgium

⁵ Earth Systems Research Center, University of New Hampshire, Durham, NH, USA

⁶ School of Informatics, Computing, and Cyber Systems (SICCS), Northern Arizona University, Flagstaff, AZ, USA