AGU100 ADVANCING EARTH AND SPACE SCIENCE

Check for updates

Key Points:
- The Simple Biosphere Model (SiB4) reproduces observed grassland behaviors, supporting its innovative prognostic phenology strategy
- Grassland mean annual leaf area index, gross primary production, and ecosystem respiration increase linearly with mean annual precipitation
- For a U.S. drought, the grassland LAI drought response was not as strong as the overlying climate response

Supporting Information:
- Supporting Information S1

Correspondence to:
K. D. Haynes,
Katherine.Haynes@colostate.edu

Representing Grasslands Using Dynamic Prognostic Phenology Based on Biological Growth Stages: Part 2. Carbon Cycling

Katherine D. Haynes[1], Ian T. Baker[1], A. Scott Denning[1], Sebastian Wolf[2], Georg Wohlfahrt[3], Gerard Kiely[4], Renee C. Minaya[5], and John M. Haynes[6]

[1]Department of Atmospheric Science, Colorado State University, Fort Collins, CO, USA, [2]Department of Environmental Systems Science, ETH Zurich, Zurich, Switzerland, [3]Institute for Ecology, University of Innsbruck, Innsbruck, Austria, [4]Environmental Research Institute, University College Cork, Cork, Ireland, [5]Center for Science Education, University Corporation for Atmospheric Research, Boulder, CO, USA, [6]Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, USA

Abstract Grasslands are one of the most widely distributed and abundant vegetation types globally, and land surface models struggle to accurately simulate grassland carbon dioxide, energy, and water fluxes. Here we hypothesize that this is due to land surface models having difficulties in reproducing grassland phenology, in particular in response to the seasonal and interannual variability of precipitation. Using leaf area index (LAI), net primary productivity, and flux data at 55 sites spanning climate zones, the aim of this study is to evaluate a novel prognostic phenology model (Simple Biosphere Model, SiB4) while simultaneously illustrating grassland relationships across precipitation gradients. Evaluating from 2000 to 2014, SiB4 predicts daily LAI, carbon, and energy fluxes with root-mean-square errors <15% and individual biases <10%; however, not including management likely reduces its performance. Grassland mean annual LAI increases linearly with mean annual precipitation, with both SiB4 and the Moderate Resolution Imaging Spectroradiometer (MODIS) showing a 0.13 increase in LAI per 100-mm increase in precipitation. Both gross primary production and ecosystem respiration increase with growing season length by ~8.5 g C m^{-2} per day, with SiB4 and Fluxnet estimates within 18%. Despite differences in mean annual precipitation and growing season length, all grassland sites shift to seasonal carbon sinks one month prior to peak uptake. During a U.S. drought, MODIS and SiB4 had nearly identical LAI responses, and the LAI change due to drought was less than the LAI change across the precipitation gradient, indicating that grassland drought response is not as strong as the overlying climate response.

Plain Language Summary Grasslands are very important ecosystems, making understanding and modeling them essential. Using data at 55 diverse grassland locations, we investigate their behavior while evaluating a model (SiB4). The annual amount of growth increases with the annual amount of rainfall, and the growing season gets longer as the amount of rainfall increases. The amount of growth drops due to drought, but this decrease is not as much as the decrease due to lower rainfall. The model is able to predict these patterns and closely matches when grasslands begin growing for the season. However, not including grass cutting at the end of the season likely reduces its performance, since many of the locations are heavily managed. The similarities between the model and the observations suggests that the model can be useful for carbon cycle predictions, ecosystem process studies, and climate change evaluations.

1. Introduction

Grasslands are one of the most widely distributed and abundant types of vegetation globally. These ecosystems cover 40% of the non-ice land area while existing in a wide variety of ecoregions on all continents except Antarctica (Suttie et al., 2005; White et al., 2000), and they account for slightly over one third of the terrestrial gross primary productivity (GPP) (Beer et al., 2010; Suttie et al., 2005; White et al., 2000). Grassland species grow across a broad range of climatic conditions, from desert regions that receive less than 250 mm of annual precipitation (e.g., Gremer et al., 2015) to regions that receive more than 2,000 mm (e.g., Wolf et al., 2011). Grassland species are very hardy, living in locations where the mean annual temperature (MAT) is

below freezing (e.g., Corradi et al., 2005) to over 20 °C (e.g., Ardö et al., 2008). In order to exist across such a wide range of growing conditions, grasslands vary dramatically in height and coverage, from sparse tussock with bunch grasses to short calcareous grasslands and steppes to tallgrass prairies (Blair et al., 2014; Suttie et al., 2005).

Since grasslands occur across environments with highly varying annual rainfall amounts, the response of grasslands to precipitation varies and continues to be widely studied (e.g., Cui et al., 2017; Konings et al., 2017). While changes in grassland productivity do not coincide with precipitation changes in humid environments (Jaksic et al., 2006), grasslands are highly responsive to precipitation in drier environments (Moran et al., 2014). For these ecosystems, variability in grassland growth and productivity varies spatially across sites, and the aboveground net primary production (ANPP) generally increases linearly with rainfall amount (Jia et al., 2015; Sala et al., 1988; Shi et al., 2014; Zhu & Southworth, 2013). In addition to this large-scale tendency across precipitation gradients, for any given location grasslands are highly responsive to temporal variations in precipitation (Hsu et al., 2012; Knapp et al., 2015). Even though individual sites have been studied for long-term responses of ANPP to annual precipitation, there is still uncertainty of the underlying mechanisms driving interannual variability in grassland production (Moran et al., 2014). Life history, geography, light availability, nutrient limitations, current year precipitation, previous year precipitation, and previous year ANPP all impact current year production, although the relative importance of these factors is still uncertain (Moran et al., 2014; Oesterheld et al., 2001; Sala et al., 2012).

Three model-data intercomparison studies all find that current land surface models (LSMs) have difficulties simulating grasslands skillfully. The first, a study by Schwalm et al. (2010) looked at 22 terrestrial biosphere models across 44 sites, finding that modeling performance of net ecosystem exchange (NEE) was worse at nonforested sites and hypothesizing that the loss of model skill is linked to phenology (i.e., soil moisture and temperature seasonality). A study by Schaefer et al. (2012) investigated GPP estimates from 26 models against 39 eddy covariance flux tower sites across the United States and Canada, similarly finding that the models performed the best at forested sites and the worst at crop, grass, and savanna sites. The reason for the difficulty in accurately simulating grassland sites was the overestimation of GPP under dry conditions. Finally, a study by Raczka et al. (2013) evaluated 17 LSMs against observations from 36 North American flux towers, several of which were grassland sites. The modeled fluxes for the grassland sites had the poorest correlations with observations, and none of the models had skill at reproducing interannual correlations and magnitudes of variability. On the seasonal scale, the LSMs significantly overestimated annual grassland GPP by ~40%, ecosystem respiration (RE) by ~60%, and NEE by ~60%. General causes of the large biases were an inability of models to properly simulate responses to soil moisture, drought, and humidity stress combined with an overextension of the growing season by 2 to 3 months.

Work has been ongoing to improve grassland predictions in LSMs. The ORCHIDEE-GM model (Chang et al., 2013, 2015) has incorporated a grassland land management module inspired by the Pasture Simulation Model (PaSim; Graux et al., 2011; Riedo et al., 1998; Vuichard et al., 2007). By integrating the cutting, grazing, fertilization, and animal modules from PaSim, ORCHIDEE-GM is able to more accurately represent the carbon cycle of managed grasslands. Comparing results to without management at 11 sites, ORCHIDEE-GM reduced the annual NEE bias and root-mean-square errors by 53% and 20%, respectively (Chang et al., 2013). ORCHIDEE-GM is able to simulate regional gradients of annual grassland productivity and ruminant livestock density across Europe; however, the model tends to systematically overestimate the absolute values of productivity in most regions (Chang et al., 2015).

Since a common pattern emerging from these studies is the difficulty LSMs have in capturing grassland phenology, the Simple Biosphere Model (SiB4) introduces an innovative prognostic phenology method focusing on grasslands. While predictive phenology modeling encompasses a rapidly growing body of literature, phenological methods have focused primarily on crops and forests and can be broadly categorized into two main approaches. First, seasonal transitions can be determined using a rule-based approach, typically using growing degree days (Richardson et al., 2012; White et al., 1997). This approach has been used for predicting crop growth (Levis et al., 2016; Lokupitiya et al., 2009), seasonal deciduous forests (Chen et al., 2016), and stress deciduous vegetation (Dahlin et al., 2015, 2017). Second, vegetation states can be continually characterized. This is typically done using a growing season index (GSI) (Jolly et al., 2005; Stöckli et al., 2008), and this approach is often used in combination with data assimilation to constrain biophysical properties (Lowman & Barros, 2018; Stöckli et al., 2011; Viskari et al., 2015). One method that continuously predicts the leaf

Table 1
Site Information

| Site	Lat	Lon	MAT	MAP	M	Years	Reference
AT-Neu	47.12	11.32	6.3	852	M	02-12	Wohlfahrt et al. (2008, DOI:10.18140/FLX/1440121)
AU-DaP	−14.06	131.32	27.3	1,250	M	08-13	Beringer (2013a), Beringer, Hutley, et al. (2016), Beringer, Mchugh, et al. (2017), and Isaac et al. (2017)
AU-Stp	−17.20	133.35	26.0	640	M	08-13	Beringer (2013b), Beringer, Hutley, et al. (2016), Beringer, Mchugh, et al. (2017), and Isaac et al. (2017)
AU-Yan	−34.99	146.29	17.3	472	M		
BR-Ji1	−10.80	−62.36	23.9	1,664	M	00-02	Saleska et al. (2013) and von Randow et al. (2004, DOI:10.3334/ORNLDAAC/1174)
BW-GhG	−21.15	21.74	20.5	463	M		
CA-Let	49.71	−112.94	5.4	398	M	00-07	Flanangan and Adkinson (2011) and Fluxnet Canada Team (2016, DOI:10.3334/ORNLDAAC/1335)
CH-Cha	47.21	8.41	9.5	1,136	M	06-12	Merbold et al. (2014)
CH-Fru	47.12	8.54	7.2	1,651	M*	06-12	Imer et al. (2013)
CH-Oe1	47.29	7.73	9	1,100	M*	02-08	Ammann et al. (2009)
CN-Cng	44.59	123.51	5	400		07-10	Dong et al. (2011)
CN-Du2	42.05	116.32	2.0	319	M	06-08	Chen et al. (2009)
CN-Hug	49.33	120.05	−2.0	367	M		
CN-Xi2	43.55	116.67	2	400	M		
CZ-BK2	49.49	18.54	6.7	1,316	M*	04-11	Marek et al. (2011)
DK-Eng	55.69	12.19	8.0	613		05-08	Soussana et al. (2007)
IE-Dri	51.99	−8.75	9.2	1,084		03-05	Peichi et al. (2011)
IT-MBo	46.01	11.05	5.1	1,214	M*	03-06	Marcolla et al. (2011)
IT-Tor	45.84	7.58	2.9	920		08-13	Galvagno et al. (2013)
ML-AgG	15.34	−1.48	29.7	350	M		
NG-Nga	11.52	11.1	28.4	800	M		
NL-Hor	52.24	5.07	10	800	M		
PA-SPs	9.32	−79.63	25.2	2,272		07-09	Wolf et al. (2011)
RU-EST	71.62	128.75	−13	190	M		
RU-Ha1	54.73	90.00	−0.1	592	M	02-04	Belelli Marchesini et al. (2007)
RU-Sam	72.37	126.50	−12.5	125	M	02-14	Boike et al. (2013), Runkle et al. (2013), Sachs et al. (2010), and Zhang et al. (2012, DOI:10.18140/FLX/1440185)
RU-Tuv	50.15	94.45	−4.7	393	M		
SD-Dem	13.28	30.48	26	320	M	05-09	Ardö et al. (2008)
SN-Dhr	15.40	−15.43	29	404	M		
UK-AMo	55.79	−3.24	7.7	984	M		
UK-Pol	55.29	−3.40	7.65	1,261	M		
UK-Tad	51.21	−2.82	9.86	823	M		
US-Arc	35.55	−98.04	14.8	843	M	05-06	Fischer et al. (2007) and Torn (2005, DOI:10.17190/AMF/1246026)
US-Atq	70.47	−157.40	−9.7	93	M		
US-Bkg	44.34	−96.84	6.0	586	M	04-10	Gilmanov et al. (2005) and Meyers (2004, DOI:10.17190/AMF/1246041)
US-Brw	71.32	−156.63	−12.6	85	M	00-07	Kwon et al. (2006) and Zona & Oechel (1998, DOI:10.17190/AMF/1246041)
US-CPE	40.82	−104.77	8.3	389	M	00-14	Lauenroth and Sala (1992), Milchunas and Lauenroth (2001), and Uresk et al. (2015, DOI:10.3334/ORNLDAAC/145)
US-CRP	40.73	−104.30	8.3	389	M		

Table 1 *(continued)*

Site	Lat	Lon	MAT	MAP	M	Years	Reference
US-Ctn	43.95	−101.85	8.2	432	M	07-08	Meyers (2007) and Schmidt et al. (2011, DOI:10.17190/AMF/1246117)
US-FPe	48.31	−105.1	5.5	335	M	00-08	Gilmanov et al. (2005) and Meyers (2000, DOI:10.17190/AMF/1246053)
US-FR1	29.93	−98.01	19.56	867	M		
US-IB2	41.84	−88.24	9.04	930		04-11	Matamala (2004) and Matamala et al. (2008, DOI:10.17190/AMF/1246066)
US-ICt	68.61	−149.30	−7.4	318	M	07-11	Euskirchen et al. (2007) and Euskirchen et al. (2016, DOI:10.17190/AMF/1246131)
US-Jrn	32.60	−106.85	16.9	250		00-14	Huenneke (2011, DOI:10.6073/pasta/5a17c5689767d95056ec63842b981998), and Peters et al. (2012)
US-Kon	39.08	−96.56	12.8	867	M	07-12	Brunsell (2006), Brunsell et al. (2014), and Logan & Brunsell (2015, DOI:10.17190/AMF/1246068)
						00-14	Blair & Nippert (2018, DOI:10.6073/pasta/eebbc157d1cf9be85a7327961c89859a)
US-Nun	40.68	−104.82	8.3	389	M		
US-NRM	39.99	−105.38	1.5	800		08-14	Humphries (2018, DOI:10.6073/pasta/cd03361ac0c1cbc686ad71da7e14f79d)
US-Ra1	39.14	−96.52	12.63	855	M		
US-Sdh	42.07	−101.41	8.8	510	M	04-09	Billesbach & Arkebauer (2004, DOI:10.17190/AMF/1246136)
US-Seg	34.36	−106.70	13.7	273	M	07-12	Anderson-Teixeira et al. (2011), Litvak (2007), and Petrie et al. (2015, 2016, DOI:10.17190/AMF/1246124)
						00-14	Muldavin & Moore (2016, DOI:10.6073/pasta/b913c188d07335b3e88fd585c1cd20b5) and Collins (2016, DOI:10.6073/pasta/ 1ec555306b53f8f506d48b0779b936ef)
US-SGS	40.81	−104.19	8.3	389	M		
US-Var	38.41	−120.95	15.8	559	M*	00-14	Baldocchi (2000), Ma et al. (2007), and Grant et al. (2012, DOI:10.17190/AMF/12465984)
US-Wkg	31.74	−109.94	15.6	407	M	04-14	Brunsell (2006), Scott (2004), and Scott et al. (2010, DOI:10.17190/AMF/1246112)
US-Wlr	37.52	−96.86	13.52	881	M	01-04	Cook and Coulter (2001) and L Coulter et al. (2006, DOI:10.17190/AMF/1246115)
ZA-Kru	−25.00	31.50	21.9	547	M	00-13	Scholes et al. (2001) and Archibald et al. (2009)

Note. The first three columns show the site name, latitude, and longitude. Site name is colored by grassland type: blue for C3 arctic, green for C3 nonarctic, orange for mixed C3/C4, and red for C4. The fourth and fifth columns show the mean annual temperature (°C) and the mean annual precipitation (mm). The sixth column shows an M if MODIS LAI data are available (* for adjusted sites). The seventh and eighth columns show FLUXNET (black) and LTER (purple) data information: the last two digits of the starting/ending years and references, respectively.

state without using a GSI is an approach by Caldararu et al. (2014), where vegetation states are determined using a mechanistic model with the premise that plant phenology is based on an optimal strategy for carbon gain. SiB4 combines these approaches to continually predict the vegetation state using growth stages (phenophases) determined from climate, weather, and environmental relationships to dictate carbon allocation, leaf photosynthetic capacity, and litterfall.

The focus of this study is twofold. First is to evaluate the performance of the grassland phenology in SiB4, which was described in detail in Haynes et al. (2019). SiB4 introduces an innovative approach to use commonalities that exist between grasslands growing in vastly different climates to model this vegetation type in an effective manner for site, regional and global studies. Work has been ongoing to generalize grassland behavior, in particular in response to rainfall, light, and fire (e.g., Gilmanov et al., 2010; Knapp et al., 2004; Koerner & Collins, 2014), and the SiB4 modeling approach expanded on these findings to develop a modeling strategy that takes advantage of consistencies in grassland behavior. This study assesses the performance of SiB4 in predicting grassland behavior across various climates and temporal scales, and the ability of SiB4
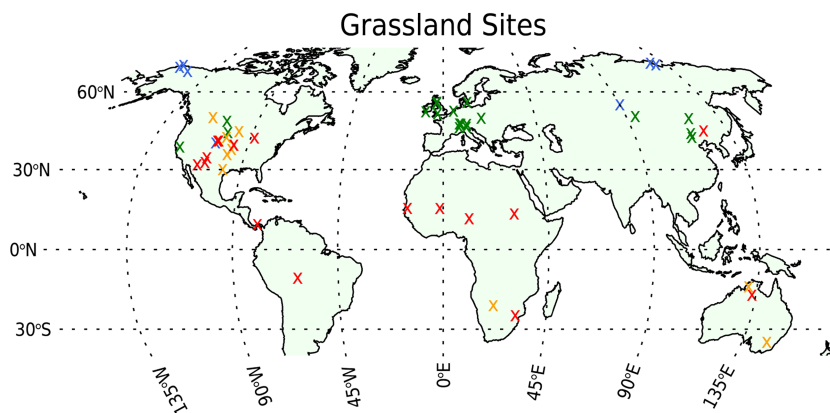
Figure 1. Grassland site map. Blue sites are C3 tundra grasslands (C3A, 8 sites), green sites are C3 temperate grasslands (C3G, 20 sites), orange sites are mixed grasslands (CXG, 10 sites), and red sites are C4 grasslands (C4G, 17 sites).

to reproduce the grassland relationships seen in this data set provides validation for this new modeling approach.

The second goal of this study is to investigate grassland behavior across precipitation gradients and to examine the impact of drought on LAI using a rich data set. Despite a substantial body of literature studying grasslands, there have been few worldwide grassland studies combining information from both flux and productivity data. Gilmanov et al. (2010) presented worldwide grassland productivity relationships with precipitation and radiation; however, they only used flux data. An additional study by Gilmanov et al. (2005) used both flux and satellite vegetation data to extrapolate grassland behavior; however, this was done only for the Great Plains region in the United States. Similarly, studies have used North American and South African grassland sites to learn about grassland consistencies between vastly different locations (Knapp et al., 2004, 2006; Koerner & Collins, 2014); however, these studies only used two sites. This study uses flux towers, *in situ* measurements, and satellite data, compiling three different data sets containing grassland carbon and energy metrics at 55 global sites spanning climate zones.

2. Methods

2.1. Sites

The grassland sites investigated in this study are listed in Table 1 and shown in Figure 1. We have been able to include sites from all continents where grasslands exist. The sites span climate zones, with site conditions ranging from MATs of −13 to 29.7 °C and mean annual precipitation (MAP) of 85 to 2,272 mm.

For fluxes, this study uses eddy covariance measurements acquired by the Fluxnet community at 34 sites (Table 2). We evaluated NEE, GPP, RE, latent heat flux (LE) and sensible heat flux (SH). When available, we compare SiB4 to Fluxnet-derived estimates of GPP and RE, even though this means comparing two modeled quantities. For sites providing multiple estimates of NEE, GPP, and RE, we selected the method closest to capturing the annual mean GPP across all calculations and years at each site (for details, see supporting information section S1). FLUXNET2015 provided filled data, and for these sites we used all the data with the selected method. Sites from the other networks were not filled, and the specific handling of the missing data is described in section 2.3.

For aboveground biomass, this study uses 8-day leaf area index (LAI) available from 2000–2014 at 48 grassland sites from the Moderate Resolution Imaging Spectroradiometer (MODIS) Land Product Subsets Project (ORNL DAAC, 2008a, 2008b). These MODIS LAI data are 8-day composites that have been specifically subset for predefined areas of $\sim 8 \times 8$ km centered on the towers. In doing this, the algorithm selects the optimal pixels from all acquisitions during the 8-day period (Knyazikhin et al., 1999). To capture phenology, seasonality, and interannual variability consistently across our broad range of sites, we use MODIS 8-day samples of LAI, rather than the more sparse site-level data.

Since the preprocessed MODIS subsets are designed to represent specific sites, we use LAI data for all of their available grassland sites and do not attempt to process global data at the tower sites that are not available. Even with using the customized site product, we found that a few of the sites include a significant fraction

Table 2
Fluxnet Site Networks

Network	Sites	Source
AmeriFlux	US-Arc, US-Bkg, US-Brw, US-Ctn, US-FPe, US-IB2, US-ICt, US-Kon, US-Sdh, US-Seg, US-Wlf	ORNL DAAC (2015) http://ameriflux.lbl.gov/
CarboEuropeIP/ GHG-Europe	IE-Dri	European Fluxes Database Cluster http://www.europe-fluxdata.eu/home/data
CarboItaly	IT-MBo	European Fluxes Database Cluster http://www.europe-fluxdata.eu/home/data
FLUXNET2015	AT-Neu, CH-Cha, CH-Fru, CH-Oe1, CN-Cng, CN-Du2, CZ-BK2, DK-Eng, IT-Tor, PA-SPs, RU-Ha1, RU-Sam, SD-Dem, US-Atq, US-Var, US-Wkg, ZA-Kru	FLUXNET2015 Data set http://fluxnet.fluxdata.org/data/ fluxnet2015-dataset
Fluxnet Canada	CA-Let	Canadian Carbon Program (CCP) http://fluxnet.ornl.gov/site_list/Network/3
LBA	BR-Ji1	Large Scale Biosphere-Atmosphere Experiment in Amazonia (LBA) https://daac.ornl.gov/LBA/lba.shtml
OzFlux	AU-DaP, AU-Stp	http://www.ozflux.org.au

of nongrass vegetation. To prevent a potential bias from the influence of evergreen vegetation, we examined the satellite images of the MODIS subsets. If a substantial fraction of the preprocessed MODIS subset area is covered by forested vegetation and if there appears to be a constant offset in LAI likely due to this coverage, we adjust the data by subtracting a constant, site-specific offset. After close inspection, we adjusted the MODIS LAI at 5 of the 48 sites. This offset is set to the minimum nongrowing season LAI, shifting the time series to minimal values when the grasslands are dormant, thus removing the influence of any interfering vegetation. Although evergreen vegetation grows during the growing season, since they retain their leaves all year, these changes occur slowly and are minimal compared to the large seasonality in the surrounding grasslands, justifying the use of a constant offset.

For annual productivity, this study uses annual aboveground and belowground net primary productivity (ANPP and BNPP) collected from five Long-Term Ecological Research (LTER) sites. All five sites are located in the United States, but they cover a diverse set of grasslands, from alpine (US-NRM) to desert (US-Seg and US-Jrn) to short-grass steppe (US-CPE) to tallgrass prairie (US-Kon). Information regarding the years available at each site along with site reference and description information is provided in purple in Table 1.

2.2. The Simple Biosphere Model (SiB4)

SiB4 is a mechanistic, self-consistent, prognostic land surface model that integrates heterogeneous land cover, environmentally responsive prognostic phenology, dynamic carbon allocation, and carbon pools that cascade from live biomass to surface litter to soil organic matter (see Haynes et al., 2019 for details). In SiB4, 10-min carbon, energy, water, and vegetation predictions arise from functional relationships that hold across climate zones and vegetation type. Phenology stage, determined daily, links carbon fluxes to carbon pools by specifying the allocation fractions and maximum photosynthetic rate. SiB4 uses five stages to represent grasslands as they develop and mature through the growing season: (1) leaf-out, (2) growth, (3) maturity, (4) stress/senescence, and (5) dormant. The timing and length of each stage is not prescribed, but rather is dynamic and diagnosed using three elements: (1) growth index, (2) environmental potential, and (3) daylength potential. Large values indicate high potential, with decreasing potential as the values become smaller. A phenological index (*PI*) is created by multiplying these together, and *PI* is high when it is beneficial to grow new leaves and when the environmental conditions support photosynthesis (leaf-out and growth stages). In contrast, *PI* is low when there is no benefit in growing new leaves or when conditions

are not suitable for photosynthesis (stress/senescence and dormant stages). To distinguish between the phenology stages, SiB4 applies prescribed threshold values to *PI*.

Once the phenology stage is determined, carbon is allocated to the pools and they are updated. Using the newly calculated vegetation state, SiB4 calculates radiation, hydrology, and land-atmosphere exchanges as in previous versions: Farquhar carbon assimilation rates (Farquhar et al., 1980); Ball-Berry stomatal conductance (Collatz et al., 1991, 1992); prognostic canopy air space temperature, moisture, and CO_2 (Baker et al., 2003); canopy, soil, and snow hydrology (Baker et al., 2008); two-stream approximation radiative transfer (Sellers et al., 1986, 1996); and interactive carbon fluxes and pools (Haynes et al., 2019).

A novel element to this phenology approach is the growth index, which determines the benefit of growing new leaves versus allocating carbon belowground. When vegetation coverage is low, the growth potential is high as each additional leaf has a significant impact on the photosynthesis rate; however, as the canopy closes adding more leaves becomes less beneficial and the growth potential decreases. The LAI at which the growth potential begins decreasing depends on the 10-year running average precipitation rate, which dictates how much vegetation the climate can support. Both the daylength and environmental potentials are similar to those used in the GSI approach; however, the environmental potential incorporates humidity, temperature, and soil moisture.

To predict litterfall, SiB4 uses three loss mechanisms: (1) daylength, (2) freezing temperatures, and (3) water deficiency. Both the freezing and water deficiency mechanisms use Q10 relationships with PFT-specific parameters. SiB4 also includes a grazing scheme, where 40% of annual NPP is grazed, corresponding to moderate grazing practices (Fetzel et al., 2017; Knapp et al., 1998).

This paper focuses on the three grassland plant functional types (PFTs): C3 tundra (C3A), C3 non-tundra (C3G), and C4 grassland (C4G). Mixed temperate C3 and C4 grasslands use the category CXG, where both C3 and C4 PFTs are simulated and combined using site-specific area-weighted contributions. At each site, the C3 and C4 contributions match the distribution growing at that location. All sites use PFT-specific parameters that are not tuned to any individual site. The parameter values were initially set to literature values and optimized in the course of this work to reach one set of values per PFT. By using common parameters, we hypothesize that this methodology is sufficiently applicable to all climates within each PFT to be useful for modeling land surface properties, ecological processes, and land-atmosphere exchanges around the globe.

All simulations in this study begin in January 1997 and run through December 2014. To provide consistency and compatibility across sites while testing the process interconnections controlling SiB4, the meteorological driver data and soil characteristics are from global data sets. For weather, this study uses hourly data (surface incident shortwave and longwave radiation; surface pressure, mixed-layer temperature, water vapor mixing ratio, and wind speed; and convective and large-scale precipitation) from the Modern-Era Retrospective Analysis for Research and Applications (MERRA), gridded at 0.5° by 0.625° (Rienecker et al., 2011). Following the procedure in Baker et al. (2010), to ensure realistic rainfall, each hourly convective and large-scale precipitation value is scaled equally such that the monthly total rainfall matches the monthly precipitation in the Global Precipitation Climatology Project (GPCP) Version 1.2 product (Huffman et al., 2001). The soil characteristics, including soil texture and reflectance, are specified from the International Geosphere-Biosphere Programme (IGBP) (Global Soil Data Task, 2000). The scale of the driver data is larger than the footprint of a flux tower; however, Sellers, Heiser, and Hall (1992) and Sellers, Berry, et al. (1992) showed theoretically and numerically that differences between sites are directly related to vegetation type and that model performance is effectively scale invariant with regard to the forcing meteorology, making comparisons of SiB4 predictions to tower observations robust.

Without a doubt, the grassland sites featured in this study have different land use histories, ecosystem states, and species-specific characteristics. If the goal of this study were to perform the best simulations at these sites, we would use site-specific driver data to capture local states and responses, improving model-observation agreement (and SiB4 is set up to incorporate such site-specific data easily). However, our larger goal is to test the extent to which SiB4 can simulate large-scale patterns from functional relationships. Using global data sets not only illustrates how well the processes included in SiB4 capture overarching grassland behaviors, but also provides estimates of likely errors for global simulations. Locations that perform poorly can be used to learn about important processes that are not included or whose modeling needs improvement.

To initialize the carbon pools, SiB4 calculates steady-state carbon pools using an accelerated equilibrium approach. Initializing with steady-state conditions implies mature ecosystems with no disturbances where growth balances decay (NEE ~ 0) and is a typical assumption for biogeochemical models (Richardson et al., 2012; Schaefer et al., 2008; Xia et al., 2012). The model repeatedly simulates 1997–2014, and after each iteration the steady-state pool sizes are determined analytically by setting the time derivatives in the pool equations to 0 (Schaefer et al., 2008). Each subsequent 18-year cycle begins from the previous steady-state carbon pool sizes. Spin-up continues until the beginning, ending, and steady-state pools are within 1% of each other. By satisfying this requirement, the model will no longer reflect the arbitrary pool sizes from which it started and instead will represent the effects of climate and other time-variant forcing.

During the initialization process, SiB4 continuously updates the soil moisture and running-mean variables. Since 10 years is typical for models to initialize soil moisture (Corbin et al., 2010; Wang et al., 2007), beginning each iteration with the previous soil moisture reservoirs ensures ample time for soil moisture to come to a state consistent with the driver data and model assumptions.

2.3. Evaluation and Analysis Procedures

For all results, SiB4 is sampled in the same manner as the observations. For LAI, we sampled SiB4 on the same days the MODIS data are reported (45 comparisons per year). Monthly and annual means are the means of the corresponding 8-day samples. At high-latitude sites with low solar zenith angles during the winter dormant months, MODIS has considerable missing data. To avoid biases in the annual means, we used a constant value to fill these missing values. Since non-missing values typically started and ended each growing season with an LAI of ~0.15, we used this to fill the winter missing values. For Fluxnet, we compared same-day hourly values (8,760 comparisons per site year adjusted for leap years less any missing data that are not evaluated) and daily mean values (365 comparisons per site year). Monthly and annual means are the average of all nonmissing daily data for each time period. To provide an overview of the data, scatter plots of SiB4 against LAI and flux observations are shown in supporting information section S2.

For site-specific evaluations, we calculate four comparative statistics: root-mean-square errors (RMSEs), mean bias errors (MBEs), correlation coefficients (COR), and normalized standard deviations for the model (NSD, normalized by the observed SD). Using daily data, we computed one value per site per statistic for both the entire time period and the growing season only. To determine the growing season, we applied thresholds to the model and observations and evaluated all times when either was above the threshold in order to include mismatches in the timing of green-up and senescence. For NEE we used days below zero; for SH we used days above 0; for LE we used days above 10% of the site maximum; and for LAI and carbon fluxes we used days above 20% of the maximum. The results are compiled into box plots showing the distribution of values across sites. Correlations are significant at the 95% level if they are above 0.3 as calculated from the two-tailed t test. To obtain this for LAI, we used 45 degrees of freedom calculated from 3 degrees of freedom per year (green-up, growing season, and browning) over 15 years. For fluxes, since these measurements are reported hourly, we used 48 degrees of freedom calculated from 6 degrees of freedom per year (dormancy, leaf-out, early growing season, peak growing season, late growing season, and senescence) and assumed the data exist for an average of 8 years.

For diurnal cycle evaluations, we calculate the monthly mean NEE diurnal cycle across all years at each site. To compare Northern Hemisphere and Southern Hemisphere sites, the months are shifted such that the month with the maximum NEE drawdown represents the peak growing season. The diurnal cycle analysis groups all sites together by grassland type, and the monthly NEE analysis shows the seasonal cycle for all sites ordered by increasing precipitation from top to bottom using the site MAP calculated from years with valid observations. Six sites have missing months, which are shown as 0 since they do not occur during the growing season.

To evaluate phenology, we used LAI to investigate four seasonal metrics: start of the growing season (SOS), day of maximum (DOM) to indicate the peak of the growing season, end of the growing season (EOS), and growing season length (GSL). We chose LAI for this analysis since it has the longest time record for the most sites, and we again applied a threshold to mean seasonal cycles for both MODIS and SiB4 LAI calculated using a spline fit to both data sets. SOS occurs on the day when LAI exceeds 20% of the maximum value, the DOM is the day of year when LAI is at its maximum value, EOS occurs once the LAI drops below 20% of the maximum site-specific value, and GSL is the number of days occurring between the SOS and the EOS. We chose the 20% threshold after performing sensitivity tests with thresholds ranging from 10–30%,

choosing the value that best captured the predominant features while avoiding spurious events. Since this study includes worldwide locations with the SOS day-of-year (DOY) occurring throughout the entire year, sites that have the EOS DOY before the SOS DOY have been shifted by 180 days to align with the Northern Hemisphere growing season.

To evaluate annual productivity, we compared mean annual ANPP at five LTER sites and BNPP at two LTER sites. For each site, the NPP measurements for every year available between 2000 and 2014 were averaged to obtain one mean NPP value per site. If necessary, site productivity was converted from dry weight to carbon using the standard 0.5 conversion factor (Verchot et al., 2006). SiB4 productivity was calculated in the same way: annual productivity above and below ground was calculated per calendar year, and the same years measured at each LTER site were averaged to create the SiB4 mean annual NPP per site. We show the results using a bar graph, where whiskers show the measured and modeled range in annual productivity.

Large-scale climatological relationships are investigated using monthly and annual mean LAI and NEE calculated over the entire 15-year time period. We use monthly mean LAI and NEE to show the mean seasonal cycle for all sites. For every site, the multiyear monthly means are calculated and the timing is shifted such that month 0 has the highest LAI or maximum uptake for that site. To compile them into a single figure, the monthly values at each site represent a single row in a contour diagram where the sites are compiled together and ordered by increasing precipitation. We also analyze multiyear annual means by calculating each annual mean from matching 8-day and daily values and then averaging these annual means across all years. We use annual means rather than growing season means to be able to compare sites across climates without introducing biases due to the length of the growing season. To generalize these relationships, we fit both linear and third-order polynomials to the data. We report the coefficient of determination (r^2) value and slope of the line and show the fit to the third-order polynomial.

To analyze interannual variability, we investigated two severe droughts in the central North American grasslands from 2010–2012 using a subset of nine sites across the U.S. Great Plains (Table S2). The region is characterized by a west-east gradient in precipitation ($\sim$200 to 1,000 mm year^{-1}) and a north-south temperature gradient ($\sim$6 to 21 °C). The large-scale environmental gradients across the grassland result in large-scale changes in vegetation, from deserts in the west to short grasslands in the central region, to tallgrass in the east. The north-south temperature gradient is correlated with a change in C3 and C4 coverage: C3 species dominate in the cool/dry portions and C4 species dominate in the warm/wet areas (Lauenroth et al., 1999).

The first drought occurred in the southern plains and began in late 2010 (Seager et al., 2014). The dry conditions, centered over Texas and extending into seven southern states, continued through the summer of 2011, breaking heat records and causing one of the most catastrophic short-term droughts in U.S. history (Nielsen-Gammon, 2011). The second drought occurred across the entire Great Plains region and was one of the most widespread and severe droughts in over one hundred years (Hoerlin et al., 2014). The precipitation deficits started to evolve in May across the Great Plains, eventually affecting more than half of the U.S., and by July 62% of the country experienced moderate to exceptional drought (National Climatic Data Center, 2013). During 2012, the annual mean temperature was 1.8^o C above average, with the warmest spring and second warmest summer in the period of 1895–2012 (Wolf et al., 2016). After peaking in July and August, the drought faded away gradually from September to October (Wang et al., 2016).

To evaluate grassland drought response, we investigated the 2010, 2011, and 2012 annual mean LAI and precipitation. To compare the response of LAI due to drought to the change in LAI across the climate gradient, for each site we regressed the mean annual total and drought LAI against the corresponding precipitation. For the sites affected by both the 2011 and 2012 droughts, the drought LAI and MAP was the mean between the two years. For the sites affected in 2012, the drought LAI and MAP were the 2012 values only. To single out the response in individual years, we regressed the standardized anomalies in LAI against the standardized anomalies in precipitation.

3. Results: SiB4 Evaluation
### 3.1. Site Evaluation
To evaluate the performance of daily SiB4 predictions across all sites, Figure 2 shows box plots of RMSE, MBE, correlation, and NSD for LAI, carbon, and energy fluxes. These metrics are calculated using all non-missing data and growing season only. For LAI, the results are similar regardless of including dormancy.
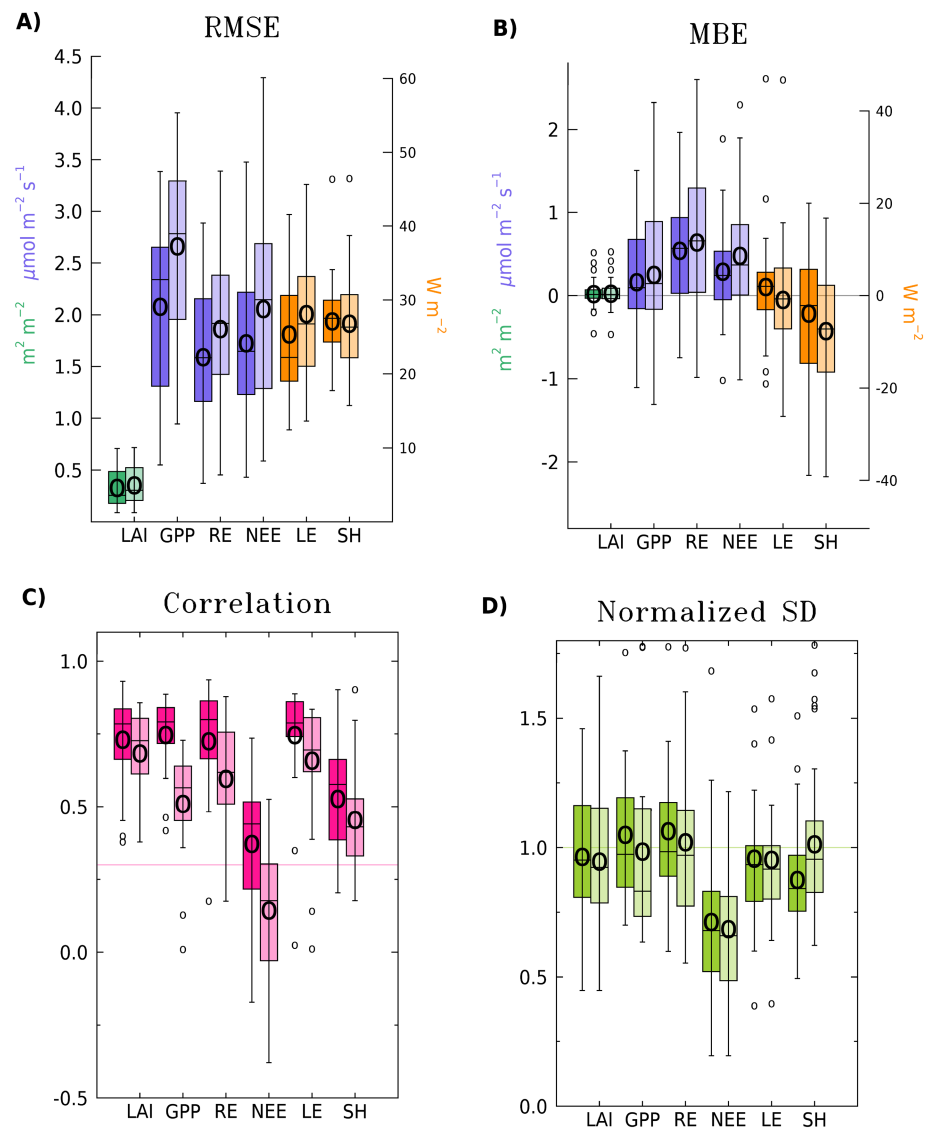
Figure 2. Box plots of evaluation statistics across sites, with results shown for LAI, GPP, RE, NEE, LE, and SH, respectively. For each variable, the left box (darker) shows the results using all non-missing data and the right box (lighter) shows the results for the growing season only. The boxes enclose the interquartile range (IQR), the whiskers extend out to 1.5 times the IQR (sites beyond this are shown with a small circle), the median is the horizontal line within the box, and the mean is the center of the dark circle in the box. The distributions shown are (A) root-mean-square errors (RMSE); (B) mean bias error (MBE); (C) correlation, with the 95% significance line highlighted; and (D) normalized standard deviation.

The median RMSE error between SiB4 and MODIS is 0.26 and all RMSE errors are < 0.72 (corresponding to normalized errors less than 30%). SiB4 does not have a prevalent LAI bias across grasslands, with a median MBE of 0.01. Both the upper and lower MBE quartiles lie within 0.1; however, two sites have negative biases > 0.2 and three sites repeatedly overestimate the LAI by more than 0.3, resulting in positive biases. All sites have statistically significant correlations (Figure 2C), with a median correlation coefficient of 0.78 between SiB4 and MODIS LAI, and 90% of the sites have correlation coefficients > 0.5. The correlations drop when only the growing season is evaluated; however, all sites remain significantly correlated. Comparing the observed and simulated standard deviations, SiB4 slightly underestimates the LAI variability such that the mean difference in spread is 0.04, and this underestimation increases to 0.08 for the growing season. All sites have a NSD between 0.45 and 1.6 (Figure 2D). Low RMSE errors, negligible MBE over all years, high correlations, and NSD's near one indicates SiB4 is able to capture the predominant seasonal and
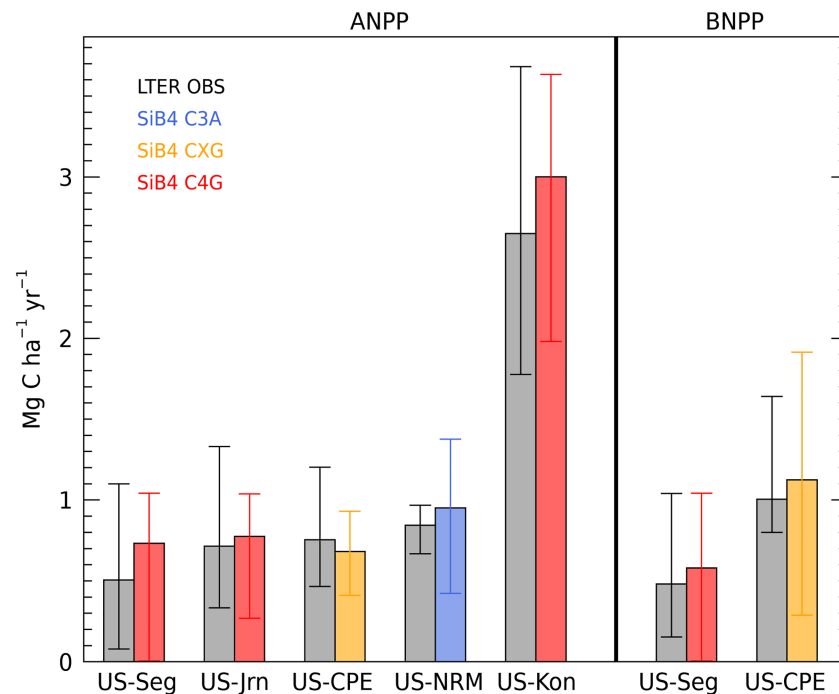
Figure 3. Mean annual NPP aboveground (left) and belowground (right). LTER observations are shown in black and SiB4 results are shown in colors by PFT. Whiskers show the minimum and maximum annual NPP across years.

interannual variability in LAI across sites, because large differences for any year would increase the RMSE to error values closer to observed values and consistent early or late phenology would lead to a larger MBE.

For GPP (second column in Figure 2), the mean RMSE between SiB4 and observed is 2.1 μmol C m^{-2} s^{-1}, and this increases to 2.8 when evaluating the growing season only. The mean and median MBE are 0.14 and 0.03 μmol C m^{-2} s^{-1}, respectively, indicating SiB4 does not have an overall bias in simulating GPP. All sites have significant correlations of GPP, with half of the sites having correlations > 0.77; however, the correlations drop by ∼0.18 for the growing season. The variability is similar between measured and simulated for both the entire time-series and for growing season only, as the mean NSD is within ∼0.05 of 1.

For RE, RMSE errors are ∼0.5 μmol C m^{-2} s^{-1} smaller than GPP errors; however, the bias is larger, indicating that SiB4 overestimates the respiration at 75% of the sites, with mean and median biases of ∼0.4 μmol C m^{-2} s^{-1} that are slightly larger for the growing season. This offset in RE is likely due to the initial pool values in SiB4 representing steady-state conditions while many of the sites are carbon sinks. All sites have correlation coefficients > 0.5, with half of the sites having RE correlations ≥ 0.74. SiB4 simulates the overall variability in RE well, with the NSD mean being 1.06 and the median being 0.98. It should be kept in mind that these exact differences between SiB4 and Fluxnet observations are subject to the choice of partitioning method at many of the sites; however, since we have selected methods that are near the average across all partitioning choices, it is expected that these results are not biased by using any particular methodology.

Looking at NEE, SiB4 underestimates the variability in NEE with a median NSD of 0.72 μmol C m^{-2} s^{-1}. The median daily RMSE between SiB4 and observed across sites is 1.8 μmol C m^{-2} s^{-1}, the median MBE is 0.24 μmol C m^{-2} s^{-1}, and the median NSD is 0.72 μmol C m^{-2} s^{-1}, which indicates that at nearly all sites SiB4 underestimates the summer drawdown of carbon dioxide from grasslands. The correlation coefficients for NEE are low, with 75% of the sites have non-significant correlations during the growing season. Because NEE is the difference between GPP and RE, the model error is at best difficult to interpret and at worst may actually be hidden in NEE evaluations due to compensating errors.

Energy flux evaluations for SiB4 compared to Fluxnet observations are shown in the last two columns of Figure 2. Compared to carbon fluxes, SiB4 does not capture the variability in energy fluxes as well, with both LE and SH having mean and median NSDs < 1. Evaluating throughout the year, daily RMSEs are all < 40 W m^{-2} for LE and < 32 W m^{-2} for SH (with one exception), and 75% of the sites have energy flux RMSEs
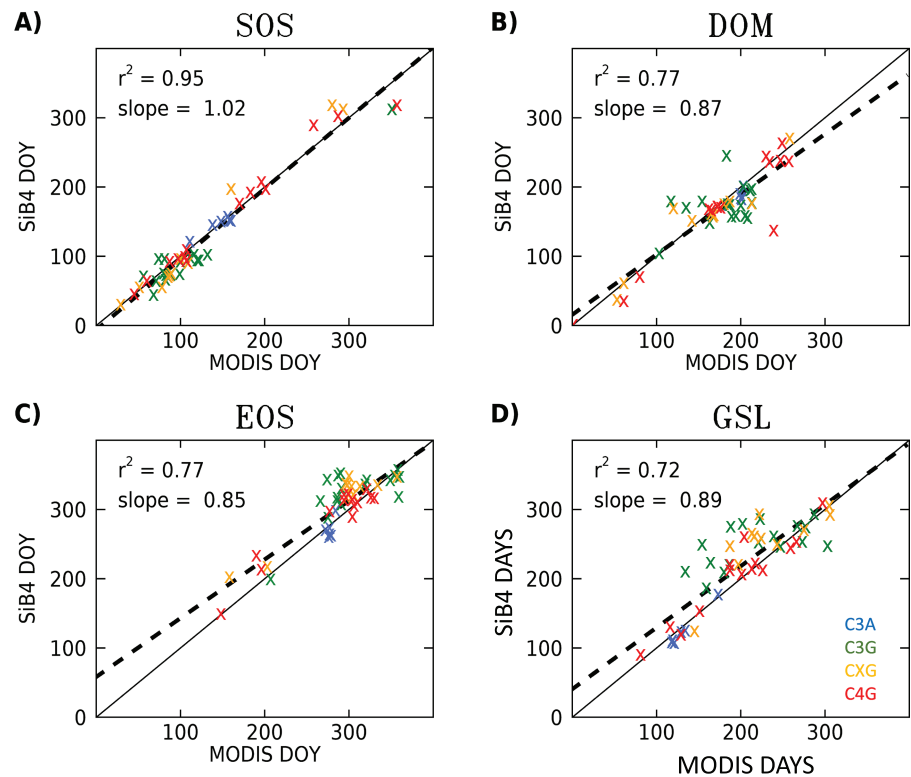
Figure 4. LAI phenology scatter plots, colored by grassland type. (A) Start of season (SOS), (B) day of maximum (DOM), (C) end of season (EOS), and (D) growing season length (GSL).

<30 W m^{-2}. The errors for LE increase at all sites when the growing season is evaluated; however, the errors for SH do not systematically increase, but instead the mean and median remain constant while the errors increase at a few sites but also decrease at a few sites. SiB4 does not have a bias in LE but has a mean negative bias of ~-8 W m^{-2} for SH. SiB4 has significant LE and SH correlations at all but three sites, with higher LE correlations.

To show how these errors arise, we assembled time series of modeled and observed 8-day LAI, daily energy and carbon component fluxes, and monthly NEE at five sites: an arctic grassland (RU-Ha1), a Mediterranean seasonal C3 grassland (US-Var), a temperate productive C3 grassland (AT-Neu), a C4 desert grassland (US-Wkg); and a C4 tallgrass prairie (US-Kon). We chose these sites because they sample the three different grassland PFTs; highlight different climates and grassland behaviors; and all have LAI data, multiyear Fluxnet records, and Fluxnet data that include the separation of NEE into the carbon components. Figures S8–S12 show these results, and section S3 provides a full discussion of these sites.

3.2. Annual Productivity

Mean annual NPP at five sites is shown in Figure 3. Four of the five sites have similar ANPP, despite different climates. Both US-Seg and US-Jrn are desert sites, while US-CPE is a temperate short-grass steppe and US-NRM is an alpine meadow. Compared to the low ANPP at these sites, the ANPP at US-Kon, a tallgrass prairie, is nearly 3 times more. SiB4 is able to simulate this difference, capturing both the mean ANPP and the range of ANPP within 0.3 Mg C ha^{-1} year^{-1}. BNPP is measured at two sites, and despite similar ANPP at these two sites, US-CPE has nearly twice as much BNPP as US-Seg. SiB4 again is able to predict this difference, capturing the mean and maximum annual NPP within 0.2 Mg C ha^{-1} year^{-1}.

3.3. Phenology

Seasonal metrics of LAI are shown in Figure 4. Comparing SiB4 against MODIS observations, correlations of all four phenology metrics (SOS, DOM, EOS, and GSL) are significant. SOS values are shown in Figure 4A. SiB4's correlation with MODIS is very tight, with an r^2 value of 0.95, and both MODIS and SiB4 show a general pattern of earlier SOS with increasing MAP. Most of the Northern Hemisphere temperate sites begin growing during their spring in April, and the Southern Hemisphere temperate sites follow this same pattern

and begin growing in October (180 days shifted). At arctic sites, the SOS is shifted later and typically occurs in late May once the temperatures have warmed enough to melt the snow and support photosynthesis. The sub-Saharan sites begin growing when the rainy season begins, typically in late July.

The days of year when the LAI peaks are shown in Figure 4B. The correlation between MODIS and SiB4 DOM is less than that for SOS, with a value of 0.77. DOM ranges from January to September, with most sites peaking in June or July. SiB4 tends to model slightly early DOMs for this cluster of sites, which is what reduces the correlation and causes an offset in the linear fit. For the sub-Saharan, Southern Hemisphere, and Mediterranean sites, the DOM in SiB4 and in MODIS are within a few days of each other. At three sites (CH-Cha, RU-Tuv, and US-Arc), there is more than a 50-day difference in DOM between MODIS and SiB4, and all these sites have at least two peaks of LAI during the growing season. While SiB4 captures the multiple peaks, their magnitude and timing does not exactly match MODIS; and averaged over multiple years, these relatively minor discrepancies aggregate to create large differences in the mean DOM. Some of these discrepancies may be due to the 8-day temporal average rather than daily values, while some may be due to management practices and other factors that are not included in SiB4.

Site EOS comparisons and values are shown in Figure 4C. SiB4 does not predict browning-down as well as greening-up, with large discrepancies in EOS at numerous sites. The r^2 between SiB4 and MODIS is 0.77, and SiB4 tends to end the season later than MODIS, particularly for sites with MAP > 1,000 mm. Considered by PFT, the EOS for the C3A sites occurs in late September, with SiB4 and MODIS agreeing on the EOS within 8 days. The majority of the non-arctic sites have their EOS clustered at the end of the year, and timing differences between SiB4 and MODIS are greatest for C3G.

Due to the large span in climatic conditions, grasslands have growing seasons ranging from 70 days to over 300 days (Figure 4D). Since GSL results from a combination of SOS and EOS timing, it is no surprise that the correlation of this phenology metric is lower. Additionally, due to less skill at predicting EOS timing, SiB4 has higher GSL variability. Arctic and sub-Saharan sites have the shortest growing seasons, being water and temperature limited, respectively. SiB4 and MODIS match GSL well at sites with minimal precipitation and short growing seasons; however, compared to MODIS, SiB4 overestimates GSL at sites with growing season lengths from 150 to 250 days. Despite this, both MODIS and SiB4 show that longer GSL coincides with higher MAP, and an analysis of GSL and temperature shows that GSL has a relationship with temperature using the site MAT difference from an optimal value (supporting information section S4).

3.4. Diurnal Variability

Carbon and energy fluxes have strong diurnal cycles over land, driven by incoming solar radiation (Betts, 2002). Plants take up carbon dioxide during the daytime yet respire both day and night. During the growing season, this causes net CO_2 uptake during the day and release overnight. As the growing season progresses, the amplitude of the diurnal cycle increases until the peak growing season, after which the diurnal cycle amplitude decreases to minimal values during dormancy.

Figure 5 shows monthly mean diurnal cycles of NEE for the four grassland PFTs. Three months prior to peak uptake (top left panel), the growing season is just beginning for temperate sites; thus, the amplitude of the seasonal cycle is only $\sim$5 μmol m^{-2} s^{-1} for C4 grasslands (red) and $\sim$2 μmol m^{-2} s^{-1} for C3 grasslands (green). The cycle is negligible for arctic grasslands (blue) because their short growing season has not yet begun. Early in the growing season, SiB4 matches the amplitude of the seasonal cycle with the Fluxnet observations for C3 and C4 grassland sites, yet SiB4 overestimates the annual mean amplitude of the diurnal cycle at mixed grassland sites by $\sim$3 μmol m^{-2} s^{-1} (yellow). The timing of the diurnal cycle for SiB4 closely matches the C3 grassland observations; however, SiB4 begins photosynthesis too late in the day at the C4 sites, as the uptake lags the observations by $\sim$2 hr throughout the morning.

As the growing season progresses, diurnal cycle amplitudes increase and differences in timing begin to emerge in both SiB4 and the observations. Two months prior to peak uptake, SiB4 overestimates the amplitudes for the mixed and C4 grasslands by $\sim$50% ($\sim$4 and $\sim$5 μmol m^{-2} s^{-1}, respectively). One month prior to peak uptake, SiB4 comes within 1.5 μmol m^{-2} s^{-1} of the observed amplitudes for the mixed and C3 grasslands, while continuing to overestimate the diurnal cycle amplitude for C4 grasslands by $\sim$3 μmol m^{-2} s^{-1}. With the amplitude increase throughout the growing season, these differences now represent errors of 12%, 15%, and 22% for C3, mixed, and C4 grasslands, respectively. Arctic C3 grasslands have a much longer duration of daytime uptake due to longer days, while C4 grasslands have less; and this pattern is seen in the SiB4 results and in the data.
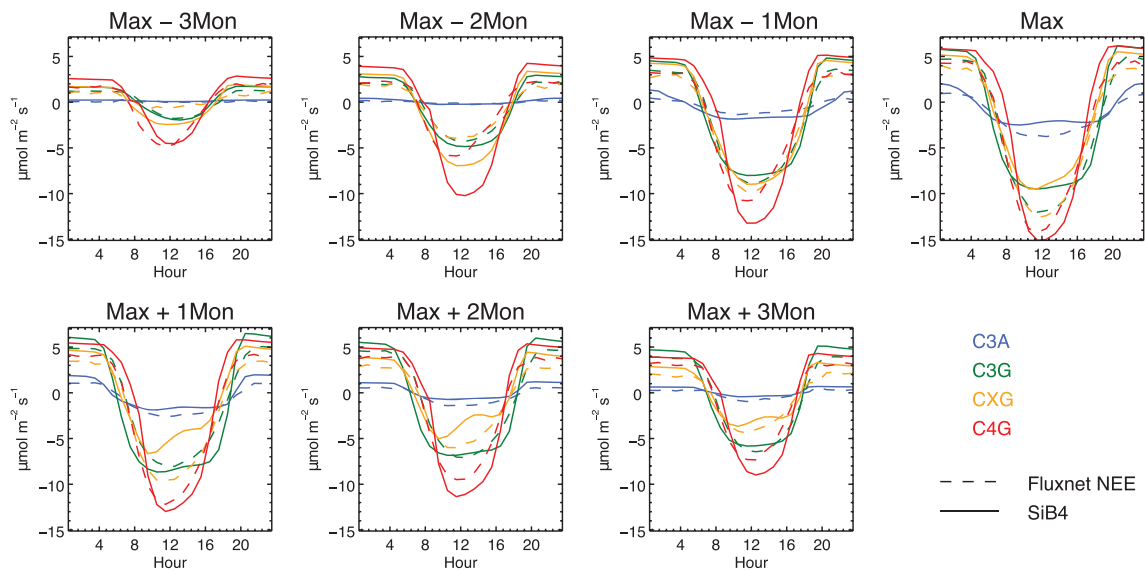
Figure 5. Monthly mean diurnal cycles per grassland PFT.

At the peak of the growing season (top right panel), the diurnal cycle amplitudes are the largest. Among the PFTs, C4 grasslands take up the most carbon during the day, with maximum net uptake rates of ~14 μmol m^{-2} s^{-1} at noon. Mixed and C3 grasslands take up less carbon during the day than C4 grasslands, and SiB4 underestimates the peak uptake rates for these PFTs by ~2 μmol m^{-2} s^{-1} (~20%). As expected, tundra grasslands take up the least carbon during the day, with maximum uptake rates of only ~3 μmol m^{-2} s^{-1} for SiB4 and ~4 μmol m^{-2} s^{-1} in the observations. At night SiB4 overestimates the respiration for all PFTs compared to the observations by up to 15% (2 μmol m^{-2} s^{-1}). This could be because SiB4 is representing a mature ecosystem in long-term balance or possibly because nighttime flux data may be biased low despite u* filtering.

SiB4 captures the basic pattern in having the fewest hours of uptake for C4 grasslands and the most hours of uptake in the arctic. SiB4 over-accentuates this pattern, underestimating the length of daytime uptake for C4 grasslands by ~3 hr and overestimating the length of daytime uptake for C3 grasslands by ~1 hr. For C3 grasslands, the underestimation of maximum daytime uptake results in a U-shape diurnal cycle in SiB4, compared to the observed V-shape. This midday uptake underestimation is a classic model problem arising from the timing of radiation data and the treatment of within-canopy light extinction (Baker et al., 2003). Since SiB4 adjusts the maximum rubisco velocity (V_{max}) with phenology, underestimating V_{max} during peak growing season could be contributing to the midday differences seen between the model and the observations.

Following the peak seasonal uptake (bottom row), SiB4 simulates the late summer decrease in uptake for C4 and C3 grasslands with errors throughout the day of < 20%. For mixed grasslands, SiB4 decreases the uptake throughout the afternoon more than seen in the observations, with differences up to 50% (~5 μmol m^{-2} s^{-1}) indicating that the model becomes stressed too early in the day at these sites.

Figure 5 illustrates that the diurnal cycle varies between grassland PFTs and evolves throughout the year. Although SiB4 has specific differences of up to 50% during some parts of the day compared to Fluxnet data, these differences occur in individual PFTs and months. Comparing across all PFTs, months, and times of day, SiB4 captures the general amplitude and timing patterns of the grassland diurnal cycles as they progress through the growing season.

4. Results: Large-Scale Grassland Behaviors
### 4.1. Seasonal Patterns
The length of the growing season for grasslands ranges from a few months to year-round, and LSMs have a difficult time accurately simulating monthly carbon fluxes for these complex ecosystems; thus, we analyzed the mean seasonal cycle across all grassland sites to see if any patterns emerge across climate zones
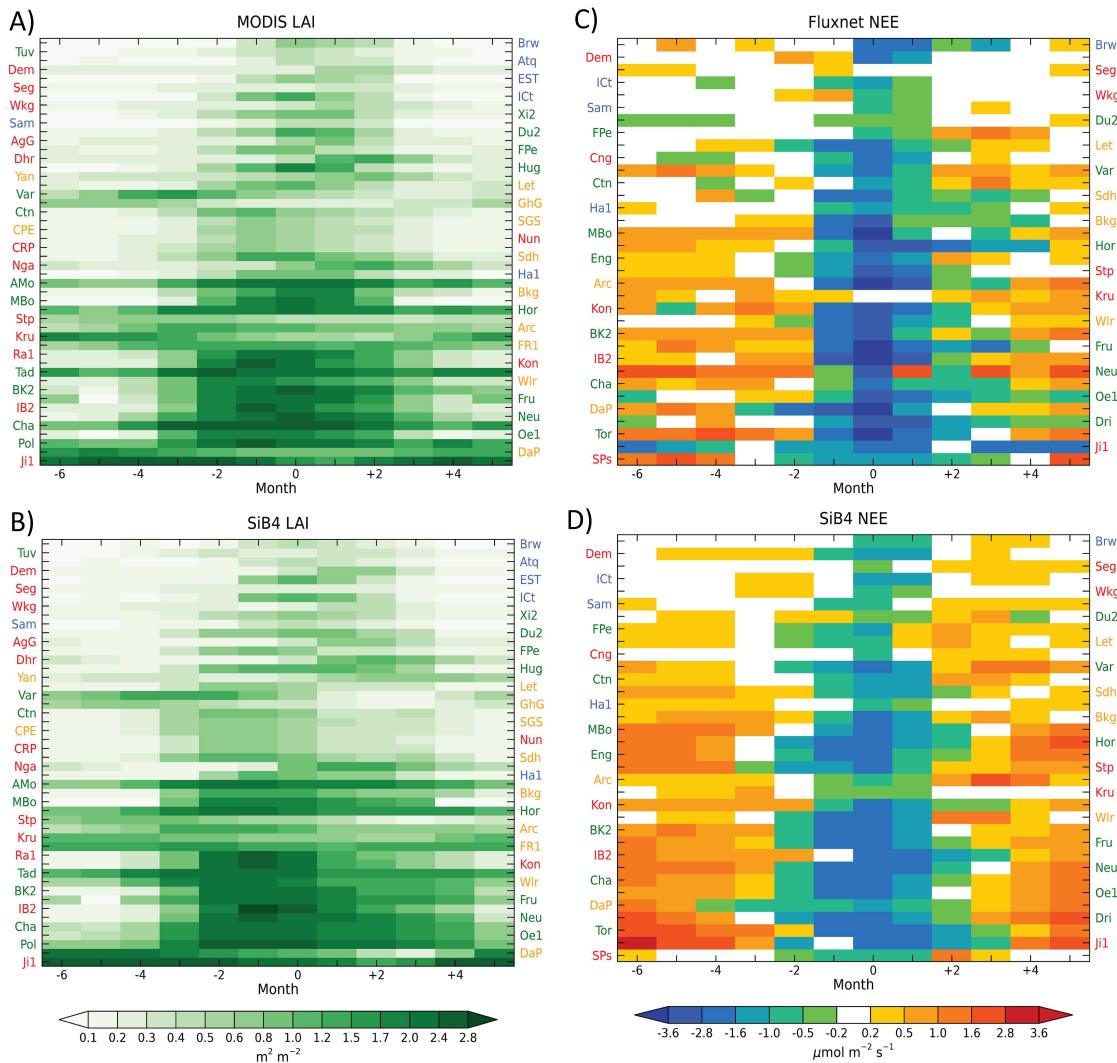
Figure 6. Mean seasonal cycles for all sites compiled together into a contour plot. Each row shows the 2000–2014 monthly mean values at a single site, and the sites are ordered by increasing MAP from top down. (A) MODIS LAI. (B) SiB4 LAI. For both of these, the months are shifted such that the month with the observed maximum observed LAI is at 0. (C) FLUXNET NEE. (D) SiB4 NEE. For both of these, the months are shifted such that the corresponding maximum drawdown (most negative NEE) is at 0.

(Figure 6). Looking at observed monthly mean LAI (Figure 6A), desert sites have small seasonal amplitude in LAI and short growing seasons (e.g., the row labeled Seg near the top of the figure). As the precipitation increases, both the amplitude and length of the seasonal cycle increase (e.g., the row labeled Pol near the bottom). Despite these increases, the shape of the seasonal cycle primarily remains symmetrical around the month of maximum LAI, indicating that the length of time for green-up and brown-down are similar. While numerous sites experience strong seasonality, several of the temperate C3 and tropical C4 grassland sites have non-negligible LAI year-round. SiB4 appears to capture these dominant features in LAI (Figure 6B), as the pattern appears similar to that seen in the observations.

Seasonal cycles of NEE for all sites are shown in Figures 6C and 6D. A striking pattern emerges: all sites take up carbon (negative NEE) for approximately 3 months, and this drawdown begins 1 month before the maximum drawdown. From 2 to 6 months prior to maximum uptake, all sites except one are sources of carbon. This pattern remains despite changes in productivity and increases in annual precipitation. In general, drawdown continues for one month after maximum uptake; however, the wetter, more productive sites remain sinks of carbon for longer periods of time. SiB4 captures this overall behavior as demonstrated in Figure 6D; however, the magnitude of the growing season uptake is underestimated and at some sites the model shifts to a weak carbon sink one month earlier than observed.
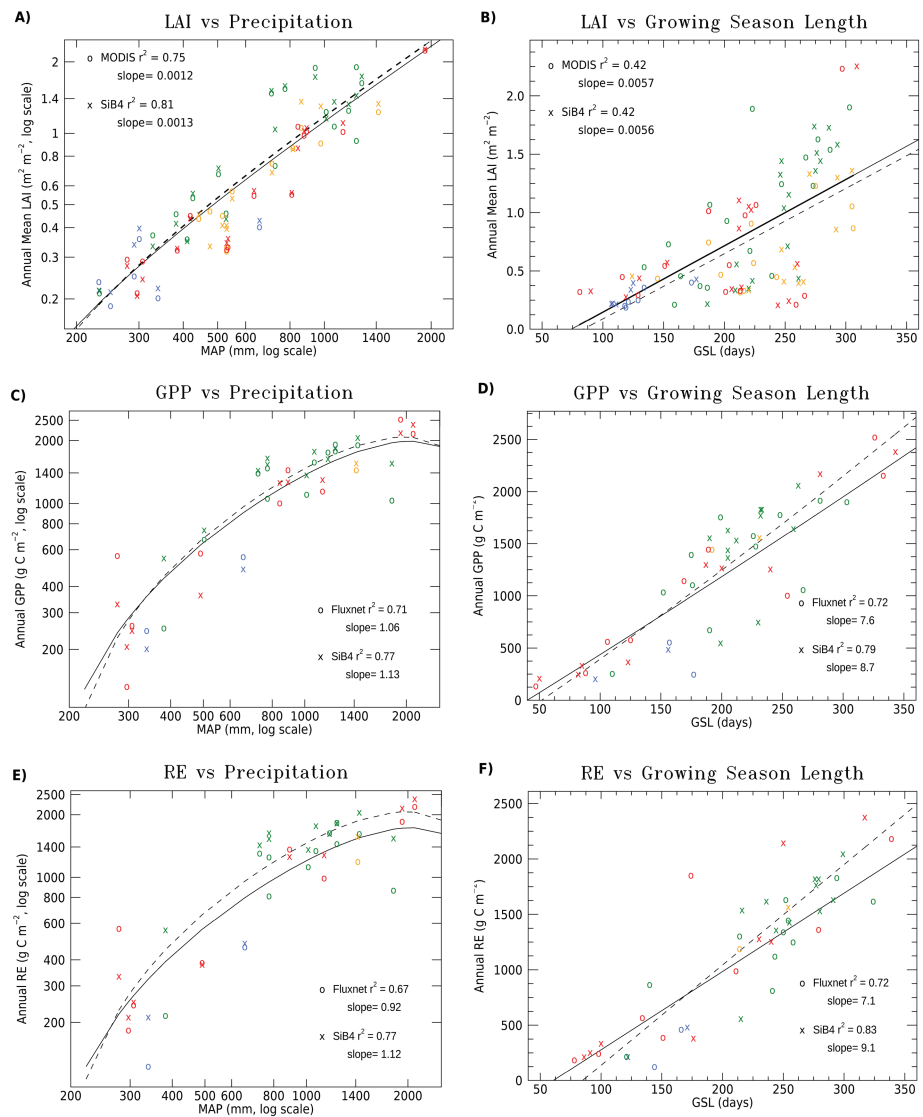
Figure 7. Large-scale grassland relationships, for both observations (o) and SiB4 (x). The colors show the PFT (C3A is blue, C3G is green, CXG is orange, and C4G is red). The r^2 and slope of the linear relationship is shown on each plot. Best fit lines from a third-order polynomial are shown in black (solid is observed and dashed is SiB4). (A) LAI versus MAP, (B) LAI versus GSL, (C) GPP versus MAP, (D) GPP versus GSL, (E) RE versus MAP, and (F) RE versus GSL.

4.2. Climatological Gradients

Grassland productivity is related to precipitation, and Figure 7A shows a strong linear relationship between annual mean LAI and MAP. Both observed and modeled correlations are tight and significant, with an observed r^2 of 0.75 and a simulated r^2 of 0.81. The observed and modeled slopes are nearly identical, indicating that for grasslands, a 100-mm increase in precipitation results in a 0.125 increase in LAI. Annual carbon fluxes are also related to annual precipitation; however, these relationships are weaker (Figures 7C and 7E). For carbon uptake, the observed r^2 is 0.71 and the slope is 1.06, indicating that for every 100 mm of precipitation, annual GPP will increase by 106 g C m^{-2}. Precipitation explains slightly more of the modeled GPP variability than observed, with an r^2 of 0.77. SiB4 also slightly overestimates the response of GPP to precipitation with a higher slope by 0.07. The relationship between RE and MAP is slightly weaker in the observations, with an r^2 of 0.67. The observed slope of the relationship between RE and precipitation is also slightly less than the GPP slope, at 0.92. As seen in GPP, the modeled relationship between RE and precipitation is stronger than observed, with an r^2 of 0.77 and a slope of 1.12.
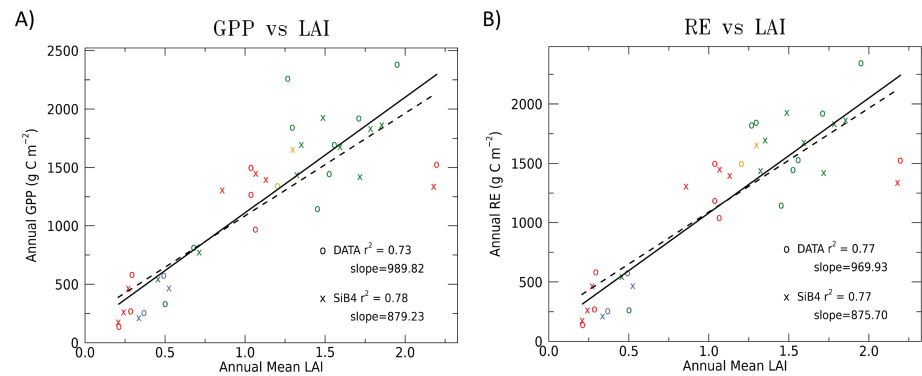
Figure 8. Large-scale annual carbon flux and LAI relationships, for both observations (o) and SiB4 (x). (a) Mean annual GPP versus LAI. (b) Mean annual RE versus LAI.

It was found that a third-order polynomial closely matches both the observed and simulated data, and using this relationship highlights regions of model-data mismatches. For GPP, the observed and modeled relationship is similar at sites with annual precipitation < 800 mm; however, for wetter sites, SiB4 overestimates the annual GPP, particularly at temperate C3 grasslands. For RE, SiB4 overestimates RE compared against the flux observations at all sites, with bigger differences as the site precipitation increases.

Over the past decades, studies have indicated that the Northern Hemisphere growing season duration has significantly increased due to both an earlier start and delayed ending (Jeong et al., 2011; Menzel et al., 2006; Myneni et al., 1997; Park et al., 2016). To investigate potential impacts this may have on grasslands and the carbon cycle, we investigated the relationships of LAI, GPP, and RE with GSL (Figures 7B, 7D, and 7F). Opposite of precipitation, the relationship between LAI and GSL is weak, with an observed and simulated r^2 values of 0.42. Both slopes were similar, indicating that annual mean LAI increases 0.05 per day of growing season added; however, considerable site differences are seen in both the observations and the model.

GPP and RE both have stronger relationships with GSL. Fluxnet r^2 values for GPP and RE against GSL are identical at 0.72, and SiB4 r^2 values are slightly higher at ~0.8. The slopes of the GPP and RE relationships against GSL is also similar, with values of ~7.4 and ~8.8 for Fluxnet and SiB4, respectively. The higher slopes for SiB4 results from the tendency of SiB4 to underestimate the carbon flux at sites with short growing seasons and overestimate the flux at sites with longer growing seasons. It is interesting to note that GSLs calculated with GPP differ substantially from GSLs calculated using RE. Using carbon fluxes to estimate growing season length dramatically shortens the growing season length at the desert site (US-Seg). Additionally, a shift toward longer GSLs in numerous C3G sites when using RE to determine the length can clearly be seen in both Fluxnet and SiB4.

To investigate the relationship between aboveground carbon biomass and carbon flux, Figure 8 shows GPP and RE plotted against LAI. Uptake is strongly related to LAI, both in the observations and in SiB4 (r^2 values of 0.73 and 0.78, respectively). Although the slope of the relationship is higher in the observations than in SiB4, both suggest that every 0.1 increase in annual mean LAI is associated with a ~92 g C m^{-2} increase in GPP per year. SiB4 underestimates the carbon uptake at temperate grassland sites with annual mean LAI ≥ 1.3. For RE, the observed r^2 is slightly higher than for GPP, bringing it equal to the modeled r^2. Both observations and SiB4 show that the slope of RE against LAI is very slightly lower than the slope of GPP, with SiB4 again underestimating the respiration for temperate C3 grassland sites.

4.3. Interannual Variability: Drought Response

Across the Great Plains of North America, there is a strong linear relationship between annual mean LAI and MAP (Figure 9A, dashed lines), with r^2 values of 0.86 and 0.90 for MODIS and SiB4, respectively. Both MODIS and SiB4 have identical slopes showing that for a 100-mm increase in precipitation the annual mean LAI will increase by 0.13. This coincides with the well-accepted linear relationship between MAP and ANPP across the Great Plains (Huxman et al., 2004; Knapp & Smith, 2001; Sala et al., 1988), and the slope is nearly identical to the slope obtained using all 48 sites (Figure 7A), indicating that this response is robust.

To compare the response of LAI due to drought to the change in LAI across the climate gradient, for each site we regressed the mean drought LAI against the corresponding drought MAP (Figure 9A, sites with
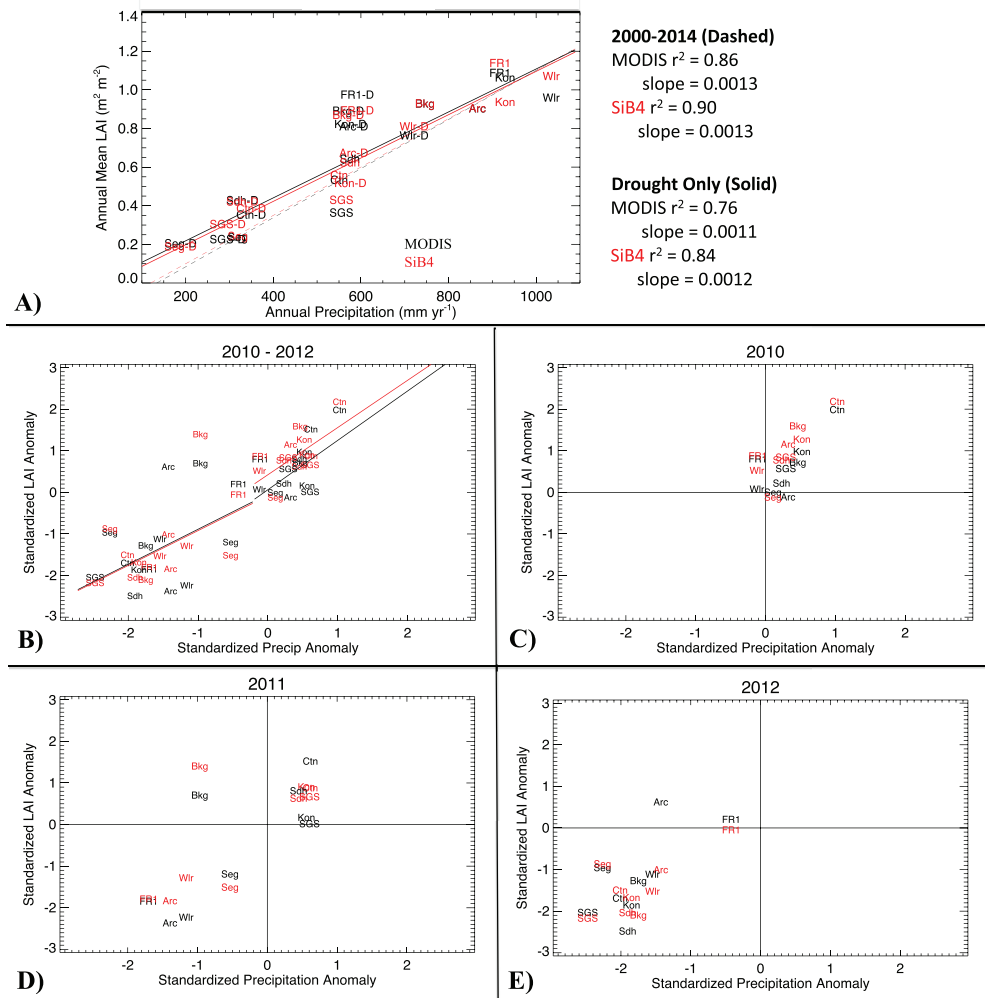
Figure 9. Relationships between annual precipitation (mm year^{-1}) and annual mean LAI, where each site is indicated by the site name, with black showing the MODIS values and red showing the SiB4 results. (A) The climatological gradient for the Great Plains sites using the long-term means (dashed lines), and the gradient for the sites using only the drought years (solid lines, site names ending in "-D"). (B) Standardized LAI anomalies against standardized precipitation anomalies for 2010–2012, with each year plotted separately. (C) As in (B), but for 2010 only. (D) As in (C), but for 2011 only. (E) As in (C), but for 2012 only.

"-D," solid lines). The drought lowered the MAP and LAI at all sites. SiB4 captures the mean drought and nondrought LAI with differences from MODIS of < 0.2 across all sites. Fitting a line to the values, the relationship if strongly linear: MODIS has a slope of 0.0011 and an r^2 of 0.76, while SiB4 has a slope of 0.0012 and an r^2 value of 0.84. The difference in slope is less than 10%.

The slope for the drought LAI is lower than the slope calculated using the long-term mean values. This indicates that across all sites, the response to drought is less than the change in LAI due to climatological differences in precipitation. Written another way, for a given precipitation value, if that precipitation is the MAP then the LAI will be lower than if that precipitation is only a single-year anomaly from a climatologically higher MAP. This is consistent with other research suggesting grasslands have developed a tolerance to drought and may have mechanisms to provide a buffer against precipitation deficits (Knapp & Smith, 2001; Lei et al., 2016).

To calculate the net change in LAI due to the droughts, we plotted all the anomalies for 2010–2012 together and calculated separate liner fits to the positive and negative anomalies to isolate the LAI changes due to decreased precipitation (Figure 9B). When MAP is more than average, SiB4 overestimates the LAI; however, during the drought years when the MAP is lower than average, MODIS and SiB4 have nearly identical best-fit

lines with a slope of 0.84. The change in LAI due to drought is only weakly linear however, as the r^2 values are only 0.22 and 0.27, respectively. However, with correlation coefficients of 0.47 and 0.52 and assuming 27 degrees of freedom (one per site per year), this relationship is still significant at the 95% level using the standard two-tailed t test.

Walking through the drought progression, in 2010 all sites have MAP and LAI near or above the long-term mean (Figure 9C). When the sites are hit by drought, the LAI is significantly lowered (Figures 9D and 9E). In 2011, the drought affected the southwestern and south-central sites (FR1, Arc, Wlr, and Seg), causing lower MAP and corresponding lower LAI. One exception to this is BKG, which had a lower MAP for the year but a higher mean LAI. This site is further north than the region impacted by the 2011 drought and has a strong seasonal cycle; thus, the reduced precipitation is not due to the drought and does not occur during the peak growing season to have an impact on LAI. The 2012 drought was widespread, and all of the sites had lower MAP than average. SiB4 predicted that all sites had lower LAI, while MODIS showed lower LAI for all sites except Arc and FR1. The 2012 drought was less intense at both these sites, and it could be the vegetation was actually recovering from the record moisture deficit and heat in 2011.

5. Discussion

Large-scale grassland behaviors emerge from combining LAI, flux, and NPP measurements. While each observation method alone is associated with shortcomings, combining this information leads to a broader and more robust understanding of grassland ecology. Additionally, predicting these different metrics all within reasonable error margins illustrates model skill, rather than tuning to optimize performance against one specific data set.

Annual mean grassland LAI has a significant linear relationship with MAP, supporting a large body of work showing that productivity in these ecosystems increases linearly with rainfall amount on the global scale (Jia et al., 2015; Shi et al., 2014; Zhu & Southworth, 2013). Both MODIS LAI data and SiB4 simulations show that every 100-mm increase in rainfall results in a 0.125 increase in mean annual LAI. Droughts lower both MAP and LAI, and the ability of SiB4 to differentiate mean annual LAI between drought and non-drought conditions with responses nearly identical to those seen by MODIS shows that the grassland phenology model is capable of responding to environmental conditions and predicting interannual variability.

GPP and RE also have significant relationships with MAP. Although the strength of these flux relationships is stronger in SiB4 than in the observations, these sites suggest that across large scales, grassland GPP increases by $\sim$110 g C m^{-2} annually for every 100 mm increase in precipitation. Since both carbon uptake and aboveground biomass have strong linear relationships with precipitation, not surprisingly they are significantly related to each other and these sites indicate that for every 0.1 increase in LAI the annual GPP will increase by $\sim$110 g C m^{-2}.

Grassland phenology also has relationships with MAP. The start of the growing season shifts earlier as MAP increases, and SiB4 and MODIS SOS are tightly correlated, predicting SOS values within 6 days of each other. For EOS, SiB4 has larger errors and tends to delay senescence. Two important reasons EOS is not well predicted emerge. First is the lack of management practices in SiB4. Many of the sites where SiB4 has a delayed EOS compared to the observations are intensively managed. The lower performance of SiB4 at these sites is likely because SiB4 does not include grassland cutting and harvesting. Second is a missing driving process for senescence. At tropical sites in particular, senescence begins prior to temperature or moisture stress, suggesting that some grasslands have evolved to have a nearly fixed growing season length based on the mean rainy season. Despite EOS differences, both SiB4 and MODIS show that EOS shifts later as MAP increases.

GSL also shows responses to MAP and MAT. Both MODIS and SiB4 show that the SOS and EOS relationships with MAP combine to lead to a lengthening of the growing season by $\sim$1 day per 8-mm MAP increase. For temperature, regressing GSL against MAT differences from an optimal value (283 K for these sites) yielded a statistically significant linear response, with a 6-day shortening of the growing season for a 1 K MAT temperature difference from the optimal. These relationships are based on an analysis of LAI GSL; however, statistically significant differences in specific GSL arose from the choice of metric used to determine the growing season. GSL is shortest when determined using GPP (150 days on average), $\sim$15 days longer using RE (175 days on average), and the longest using LAI (193 days on average). We hypothesize

that these changes in GSL arise from the ecological behavior of the metrics being investigated. GSL is shortest using GPP since this only represents the time period plants are photosynthetically active. Next comes GSL using RE, which has contributions from both autotrophic and heterotrophic respiration. Since there are non-photosynthetic inputs to respiration, this metric results in longer values. Finally, GSL determined using LAI has the longest values because LAI results from complex interactions between carbon fluxes and represents a composite of ecological processes. While fluxes can start and stop quickly with changing environmental conditions, LAI emerges from fluxes, allocation, and growth processes, thus resulting in longer GSL estimates. This finding highlights the importance of thoroughly understanding the GSL calculation to properly understand the results and attribute any conclusions.

Shifting phenology, particularly earlier spring onset in high northern latitudes, illustrates an important ecological response to climate change (Armano et al., 2010; Menzel et al., 2006; Post et al., 2018; Richardson et al., 2013). Changes in phenology and GSL will impose changes in carbon fluxes that will cascade into resultant carbon storage pools (Forkel et al., 2016; Keenan et al., 2014; Park et al., 2016). As vegetation responds dynamically to environmental changes, complex patterns in greening and browning have been observed (Goetz et al., 2011; Lara et al., 2018; Miles & Esau, 2016). Using this selection of global sites, we investigated how GSL is related to LAI, GPP, and RE. Both carbon flux components are significantly correlated with GSL, and site observations suggest that for every day added to the growing season, annual GPP and RE increase by ~8 g C m^{-2}. The LAI relationship with GSL was considerably weaker due to longer and more variable GSLs calculated when using LAI as the determinant. Despite site-specific differences in EOS and GSL, SiB4 captures the large-scale relationships of LAI, GPP and RE with GSL reasonably well. Although SiB4 overestimates the slope of both the GPP and RE responses by ~1.5 g C m^{-2}, compared against the large background flux from the existing growing season, this difference is relatively small ($< 1\%$).

Looking at monthly patterns in grassland carbon cycling, aboveground vegetation is relatively symmetrical about the month of peak LAI, and monthly patterns in LAI are closely matched between MODIS and SiB4. In contrast, grassland NEE is less symmetrical and exhibits model-data mismatches. From Fluxnet data, grasslands begin to be seasonal carbon sinks only one month prior to peak drawdown across all sites. This is a striking pattern that has both ecosystem and carbon cycling implications. Although the length of the seasonal sink increases with longer growing seasons, the temporal shape of the sink appears to be consistent across climatic conditions: grasslands rapidly switch from sources to sinks of carbon, reach their maximum seasonal sink magnitude after one month, and continue as decreasing seasonal carbon sinks for one to several months (depending on rainfall distribution) until the ecosystem returns to being a source of carbon through dormancy. Fluxnet data show that while both flux components increase through the growing season, RE lags GPP and has maximum efflux 1 to 2 months after peak drawdown (Figure S14).

SiB4 underestimated the magnitude of the peak drawdown and had a tendency to predict drawdown too early. Since NEE results from two counteracting components, causes for this behavior can be difficult to unravel; however, we have several ideas for the cause of this behavior. First off, site-specific land history and management practices that increase site productivity are not included in SiB4. Techniques such as burning, cutting, and fertilization have been shown to increase productivity (e.g., Buis et al., 2009; Imer et al., 2013; Koerner & Collins, 2014; Wohlfahrt et al., 2008). Not including these practices could lead to an underestimation of NEE compared to observed. Second, the rate of photosynthesis relative to aboveground biomass may not be correctly simulated. Although SiB4 seasonally changes the maximum rate of photosynthesis, it could be that this seasonality needs adjusting. Perhaps early and late in the season leaves are too productive, yet during peak drawdown the maximum rate of photosynthesis is underpredicted. Third, timing of respiration may be altering the NEE pattern. An investigation of monthly RE shows that SiB4 respiration one month prior to seasonal peak drawdown is higher than observed (Figure S14). Additionally, higher respiration extends longer following the peak drawdown in the observations that is not simulated by SiB4. These two patterns provide guidance for future improvements to SiB4 respiration; however, further study may be required to isolate the processes causing these changes since ecosystem respiration is comprised of both autotrophic and heterotrophic respiration. One hypothesis is that the autotrophic respiration is too high in SiB4 during green-up at the beginning of the season and too low during maturity and senescence. Another hypothesis is that the seasonality in heterotrophic respiration is shifted too early in the season, perhaps due to inaccurate temperature and/or moisture responses. It seems likely that all of these processes are contributing to the incorrect NEE seasonality.

SiB4 improves simulated grassland carbon and energy fluxes worldwide. Compared to model results presented in Raczka et al. (2013), SiB4 reduces simulated grassland annual GPP and RE biases from 2.4 and 3.0 Mg C ha^{-1} year^{-1} to 1.0 and 0.9 Mg C ha^{-1} year^{-1}. From another model-data comparison of GPP by Schaefer et al. (2012), SiB4 shows improvement in simulating carbon uptake. Comparing 26 models against estimated GPP at 4 grassland eddy covariance towers, Schaefer et al. (2012) show model-mean daily RMSE of 2.8 μmol m^{-2} s^{-1}, while SiB4 has a daily mean RMSE of 2.1 μmol m^{-2} s^{-1} across the 23 sites in this study with available GPP Fluxnet estimates. Looking at the two sites that are in common between the two studies, Schaefer et al. (2012) has RMSE values of 3.5 and 3.2 μmol m^{-2} s^{-1} for US-IB2 and US-Var, respectively. At these same sites, SiB4 has RMSE values of 2.3 and 1.7 μmol m^{-2} s^{-1}, showing a 50% reduction in predicted GPP compared to previous model results.

This study identified three key areas that will form the basis for upcoming incremental improvements to SiB4. First is to better represent management practices. Land history, management, and disturbance such as burning, cutting, fertilizing, and harvesting play an important role in grassland ecosystems. Not including these processes limits the skill of SiB4, particularly with respect to predicting site-specific productivity and growing season senescence. Further, expanding grazing to include spatial variations of grazing intensity and duration should also yield improved productivity predictions at heavily grazed locations. Second is to improve both autotrophic and heterotrophic respiration estimates to decrease efflux at the beginning of the growing season and increase efflux at the end of the growing season. Third is to investigate processes controlling senescence, including introducing a new process representing hypothesized biological adaptation to established growing season lengths. It appears that too-long growing seasons in SiB4 are compensating for underestimated maximum GPP, and improving the EOS predictions will help highlight errors in carbon assimilation rates that can be adjusted to yield further model improvements.

6. Conclusions

This study combined remotely sensed LAI data, eddy covariance flux measurements, and LTER NPP observations to investigate worldwide grasslands spanning climate variability. Grasslands have a large range of annual mean LAI that responds linearly with MAP (~0.125 LAI increase per 100-mm MAP). The LAI response due to drought was less than the mean LAI change across the precipitation gradient in both MODIS and SiB4, indicating that the drought response across sites is not as strong as the overlying climate response. As MAP increases, grassland growing season begins earlier in the year and extends later in the year, leading to a longer GSL; however, actual GSL values depend on the metric being used to identify the growing season, with the biggest differences at desert sites. Using GPP to predict GSL results in the shorter values (~150 days on average) while using LAI results in the longest GSL (~195 days on average). Annual GPP and RE have statistically significant linear relationships with MAP (~100 g C m^{-2} increase per 100-mm MAP) and GSL (~8.5 g C m^{-2} increase per growing season day). Looking at net carbon flux, monthly mean NEE patterns revealed that regardless of productivity or growing season length, all grasslands transition to seasonal sinks of carbon 1 month prior to peak uptake. Despite grasslands being one of the most diverse and difficult ecosystems to simulate, SiB4 is able to predict these observed site-level and large-scale grassland behaviors, illustrating that it is a useful predictive tool available for studies ranging from carbon cycle predictions to ecosystem process investigations to climate change evaluations.

References

Ammann, C., Spirig, C., Leifeld, J., & Neftel, A. (2009). Assessment of the nitrogen and carbon budget of two managed temperate grassland fields. *Agriculture, Ecosystems & Environment*, *133*(3-4), 150–162. https://doi.org/10.1016/j.agee.2009.05.006

Anderson-Teixeira, K., Delong, J. P., Fox, A. M., Brese, D. A., & Litvak, M. E. (2011). Differential responses of production and respiration to temperature 941 and moisture drive the carbon balance across a climatic gradient in New Mexico. *Global Change Biology*, *17*, 410–424. https://doi.org/10.1111/j.1365-2486.2010.02269.x

Archibald, S. A., Kirton, A., van der Merwe, M. R., Scholes, R. J., Williams, C. A., & Hanan, N. (2009). Drivers of inter-annual variability in net ecosystem exchange in a semi-arid savanna ecosystem, South Africa. *Biogeosciences*, *6*(2), 251–266. https://doi.org/10.5194/bg-6-251-2009

Ardö, J., Mölder, M., El-Tahir, B. A., & Elkhidir, H. A. M. (2008). Seasonal variation of carbon fluxes in a sparse savanna in semi arid Sudan. *Carbon Balance and Management*, *3*, 7. https://doi.org/10.1186/1750-0680-3-7

Armano, T., Smithers, R. J., Sparks, T. H., & Sutherland, W. J. (2010). A 250-year index of first flowering dates and its response to temperature changes. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *277*, 2451–2457.

Baker, I. T., Denning, A. S., Hanan, N., Prihodko, L., Uliasz, M., Vidale, P.-L., et al. (2003). Simulated and observed fluxes of sensible and latent heat and CO$_2$ at the WLEF-TV tower using SiB2.5. *Global Change Biology*, *9*, 1262–1277.

Baker, I., Denning, A., & Stöckli, R. (2010). North american gross primary productivity: Regional characterization and interannual variability. *Tellus*, *62B*, 533–549. https://doi.org/10.1111/j.1600-0889.2010.00492.x

Baker, I. T., Prihodko, L., Denning, A. S., Goulden, M., Miller, S., & da Rocha, H. R. (2008). Seasonal drought stress in the Amazon: Reconciling models and observations. *Journal of Geophysical Research*, *113*, G00B01. https://doi.org/10.1029/2007JG000644

Baldocchi, D. (2000). AmeriFlux US-Var Vaira Ranch-Ione. https://10.17190/AMF/1245984

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., et al. (2010). Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate. *Science*, *329*, 834–838. https://doi.org/10.1126/science.1184984

Belelli Marchesini, L., Papale, D., Reichstein, M., Vuichard, N., Tchebakova, N., & Valentini, R. (2007). Carbon balance assessment of a natural steppe of southern Siberia by multiple constraint approach. *Biogeosciences*, *4*, 581–595. https://doi.org/10.5194/bg-4-581-2007

Beringer, J. (2013a). Daly Pasture OzFlux Tower Site, OzFlux: Australian and New Zealand Flux Research and Monitoring Network. Available at http://www.ozflux.org.au/index.html. Retrieved from.

Beringer, J. (2013b). Sturt Plains OzFlux Tower Site, OzFlux: Australian and New Zealand Flux Research and Monitoring Network. Available at http://www.ozflux.org.au/index.html. Retrieved from https://doi.org/hdl:102.100.100/14230.

Beringer, J., Hutley, L., McHugh, I., Arndt, S., Campbell, D., Cleugh, H., & Wardlaw, T. (2016). An introduction to the Australian and New Zealand flux tower network OzFlux. *Biogeosciences*, *13*, 5895–5916. https://doi.org/10.5194/bg-13-5895-2016

Beringer, J., Mchugh, I., Hutley, L. B., Isaac, P., & Kljun, N. (2017). Technical Note: Dynamic INtegrated Gap-filling partitioning for OzFlux (DINGO). *Biogeosciences*, *14*, 1457–1460. https://doi.org/10.5194/bg-14-1457-2017

Betts, A. K. (2002). Diurnal cycle. In J. R. Holton, J. Pyle, & J. A. Curry (Eds.), *Encyclopedia of atmospheric sciences*. London: Academic Press.

Billesbach, D., & Arkebauer, T. (2004). AmeriFlux US-SdH Nebraska SandHills Dry Valley. https://10.17190/AMF/1246136

Blair, J., & Nippert, J. (2018). PAB01 aboveground net primary productivity of tallgrass prairie based on accumulated plant biomass on core LTER watersheds (001d, 004b, 020b). environmental data initiative. https://doi.org/10.6073/pasta/eebbc157d1cf9be85a7327961c89859a. Data accessed 4/14/2019

Blair, J., Nippert, J., & Briggs, J. (2014). Grassland ecology. In R. Monson (Ed.), *Ecology and the environment, the plant sciences* (pp. 389–423). Berlin: Springer Verlag. https://doi.org/10.1007/978-1-4614-7501-914

Boike, J., Kattenstroth, B., Abramova, K., Bornemann, N., Chetverova, A., Fedorova, I., et al. (2013). Baseline characteristics of climate, permafrost and land cover from a new permafrost observatory in the Lena River delta, Siberia (1998–2011). *Biogeosciences*, *10*(3), 2105–2128. https://doi.org/10.5194/bg-10-2105-2013

Brunsell, N. (2006). AmeriFlux US-Kon Konza Prairie LTER (KNZ). https://10.17190/AMF/1246068

Brunsell, N. A., Nippert, J. B., & Buck, T. L. (2014). Impacts of seasonality and surface heterogeneity on water-use efficiency in mesic grasslands. *Ecohydrology*, *7*(4), 1223–1233. https://doi.org/10.1002/eco.1455

Buis, G., Blair, J., Burkepile, D., Burns, C., Chamberlain, A., Chapman, P., et al. (2009). Controls of aboveground net primary production in mesic savanna grasslands: An inter-hemisphere comparison. *Ecosystems*, *12*, 982–995. https://doi.org/10.1007/s10021-009-9273-1

Caldararu, S., Purves, D. W., & Palmer, P. I. (2014). Phenology as a strategy for carbon optimality: A global model. *Biogeosciences*, *11*(3), 763–778. https://doi.org/10.5194/bg-11-763-2014

Chang, J., Viovy, N., Vuichard, N., Ciais, P., Campioli, M., Klumpp, K., et al. (2015). Modeled changes in potential grassland productivity and in grass-fed ruminant livestock density in Europe over 1961-2010. *PLOS ONE*, *10*(5), e0127554. https://doi.org/10.1371/journal.pone.0127554

Chang, J. F., Viovy, N., Vuichard, N., Ciais, P., Wang, T., Cozic, A., et al. (2013). Incorporating grassland management in ORCHIDEE: Model description and evaluation at 11 eddy-covariance sites in Europe. *Geoscientific Model Development*, *6*(6), 2165–2181. https://doi.org/10.5194/gmd-6-2165-2013

Chen, S., Lin, G., Huang, J., & Jenerette, G. (2009). Dependence of carbon sequestration on the differential responses of ecosystem photosynthesis and respiration to rain pulses in a semiarid steppe. *Global Change Biology*, *15*(10), 2450–2461. https://doi.org/10.1111/j.1365-2486.2009.01879.x

Chen, M., Melaas, E. K., Gray, J. M., Friedl, M. A., & Richardson, A. D. (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. *Global Change Biology*, *22*(11), 3675–3688. https://doi.org/10.1111/gcb.13326

Collatz, G. J., Ball, J. T., Grivet, C., & Berry, J. A. (1991). Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: A model that includes a laminar boundary layer. *Agricultural and Forest Meteorology*, *54*, 107–136.

Collatz, G. J., Ribas-Carbo, M., & Berry, J. A. (1992). Coupled photosynthesis-stomatal conductance model for leaves of c4 plants. *Australian Journal of Plant Physiology*, *19*, 519–538.

Collins, S. (2016). Below-Ground Net Primary Production (BNPP): Root ingrowth donuts in Chihuahuan Desert grassland and Creosote Shrubland at the Sevilleta National Wildlife Refuge, New Mexico (2005-present). Environmental Data Initiative https://doi.org/10.6073/pasta/1ec555306b53f8f506d48b0779b936ef

Cook, D., & Coulter, R. (2001). AmeriFlux US-Wlr Walnut River Watershed (Smileyburg). https://10.17190/AMF/1246115

Corbin, K. D., Denning, A. S., Lokupitiya, E. Y., Schuh, A. E., Miles, N. L., Davis, K. J., et al. (2010). Assessing the impact of crops on regional CO_2 fluxes and atmospheric concentrations. *Tellus*, *62B*, 521–532. https://doi:10.1111/j.1600-0889.2010.00485.x

Corradi, C., Kolle, O., Walter, K., Zimov, S. A., & Schulze, E.-D. (2005). Carbon dioxide and methane exchange of a north-east Siberian tussock tundra. *Global Change Biology*, *11*(11), 1910–1925. https://doi.org/10.1111/j.1365-2486.2005.01023.x

Cui, T., Martz, L., & Guo, X. (2017). Grassland phenology response to drought in the canadian prairies. *Remote Sensing*, *9*(12), 1258. https://doi.org/10.3390/rs9121258

Dahlin, K. M., Fisher, R. A., & Lawrence, P. J. (2015). Environmental drivers of drought deciduous phenology in the Community Land Model. *Biogeosciences*, *12*(16), 5061–5074. https://doi.org/10.5194/bg-12-5061-2015

Dahlin, K. M., Ponte, D. D., Setlock, E., & Nagelkirk, R. (2017). Global patterns of drought deciduous phenology in semi-arid and savanna-type ecosystems. *Ecography*, *40*(2), 314–323. https://doi.org/10.1111/ecog.02443

Dong, G., Guo, J., Chen, J., Sun, G., Gao, S., Hu, L., & Wang, Y. (2011). Effects of spring drought on carbon sequestration, evapotranspiration and water use efficiency in the Songnen meadow steppe in northeast China. *Ecohydrology*, *4*(2), 211–224. https://doi.org/10.1002/eco.200

Euskirchen, E., Bret-Harte, M., Shaver, G., Edgar, C., & Romanovsky, V. (2016). Long-term release of carbon dioxide from arctic tundra ecosystems in Alaska. *Ecosystems*, *20*, 960–974. https://doi.org/10.1007/s10021-016-0085-9

Euskirchen, E., Shaver, G., & Bret-Harte, S. (2007). AmeriFlux US-ICt Imnavait Creek Watershed Tussock Tundra. https://10.17190/AMF/1246131

Farquhar, G. D., von Caemmerer, S., & Berry, J. A. (1980). A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species. *Planta*, *149*, 78–90.

Fetzel, T., Havlik, P., Herrero, M., Kaplan, J. O., Kastner, T., Kroisleitner, C., et al. (2017). Quantification of uncertainties in global grazing systems assessment. *Global Biogeochemical Cycles*, *31*, 1089–1102. https://doi:10.1002/2016GB005601

Fischer, M., Billesbach, D., Berry, J., Riley, W., & Torn, M. (2007). Spatiotemporal variations in growing season exchanges of CO2, H2O, and sensible heat in agricultural fields of the southern Great Plains. *Earth Interactions*, *11*(17), 1–21.

Flanagan, L., & Adkinson, A. (2011). Interacting controls on productivity in a northern Great Plains grassland and implications for response to ENSO events. *Global Change Biology*, *17*(11), 3293–3311. https://doi.org/10.1111/j.1365-2486.2011.02461.x

Fluxnet Canada Team (2016). *FLUXNET Canada Research Network Canadian Carbon Program data collection, 1993-2014*. Oak Ridge, Tennessee, USA: ORNL DAAC. https://doi.org/10.3334/ORNLDAAC/1335

Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R., Heimann, M., Thonicke, K., et al. (2016). Enhanced seasonal CO_2 exchange caused by amplified plant productivity in northern ecosystems. *Science*, *351*(6274), 696–699. https://doi.org/10.1088/1748-9326/8/2/025008

Galvagno, M., Wohlfahrt, G., Cremonese, E., Rossini, M., Colombo, R., Filippa, G., et al. (2013). Phenology and carbon dioxide source/sink strength of a subalpine grassland in response to an exceptionally short snow season. *Environmental Research Letters*, *8*(2), 025008. https://doi.org/10.1088/1748-9326/8/2/025008

Gilmanov, T., Aires, L., Barcza, Z., Baron, V., Belelli, L., Beringer, J., et al. (2010). Productivity, respiration, and light-response parameters of world grassland and agroecosystems derived from flux-tower measurements. *Rangeland Ecology and Management*, *63*(1), 16–39. https://doi.org/10.2111/REM-D-09-00072.1

Gilmanov, T. G., Tieszen, L. L., Wylie, B. K., Flanagan, L. B., Frank, A. B., Haferkamp, M. R., et al. (2005). Integration of CO_2 flux and remotely sensed data for primary production and ecosystem respiration analyses in the Northern Great Plains: Potential for quantitative spatial extrapolation. *Global Ecology and Biogeography*, *14*(3), 271–292. https://doi.org/10.1111/j.1466-822X.2005.00151.x

Global Soil Data Task (2000). *Global Soil Data Products CD-ROM (IGBP-DIS). International Geosphere-Biosphere Programme, Data and Information System. Potsdam, Germany*. Oak Ridge, Tennessee, U.S.A: Oak Ridge National Laboratory Distributed Active Archive Center. http://www.daac.ornl.gov.

Goetz, S. J., Epstein, H. E., Bhatt, U. S., Jia, G. J., Kaplan, J. O., Lischke, H., et al. (2011). Recent changes in Arctic vegetation: Satellite observations and simulation model predictions. In G. Gutman, & A. Reissell (Eds.), *Eurasian arctic land cover and land use in a changing climate* (pp. 9–36). Dordrecht; Heidelberg: Springer Science+Business Media. https://doi.org/10.1007/978-90-481-9118-52

Grant, R. F., Baldocchi, D. D., & Ma, S. (2012). Ecological controls on net ecosystem productivity of a seasonally dry annual grassland under current and future climates: Modelling with ecosys. *Agricultural and Forest Meteorology*, *152*, 189–200. https://doi.org/10.1016/j.agrformet.2011.09.012

Graux, A.-I., Gaurut, M., Agabriel, J., Soussana, J.-F., & Baumont, R. (2011). Modelling the impacts of climate change on suckling grass-based systems with the Pasture Simulation Model. In D. Sauvant, J. Milgen, P. Faverdin, & N. Friggens (Eds.), *Modelling nutrient digestion and utilisation in farm animals* (pp. 364–374). Wageningen: Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-712-740

Gremer, J. R., Bradford, J. B., Munson, S. M., & Duniway, M. C. (2015). Desert grassland responses to climate and soil moisture suggest divergent vulnerabilities across the southwestern United States. *Global Change Biology*, *21*(11), 4049–4062. https://doi.org/10.1111/gcb.13043

Haynes, K. D., Baker, I. T., Denning, A. S., Stöckli, R., Schaefer, K., Lokupitiya, E. Y., & Haynes, J. M. (2019). Representing ecosystems using dynamic prognostic phenology based on biogical growth stages: Part 1. implementation in the simple biosphere model (sib4). *Journal of Advances in Modeling Earth Systems*, *11*. https://doi.org/10.1029/2018MS001540

Hoerlin, M., Eischeid, J., Kumar, A., Leung, R., Mariotti, A., Mo, K., et al. (2014). Causes and predictability of the 2012 great plains drought. *Bulletin of the American Meteorology Society*, *95*, 269–282. https://doi:10.1175/BAMS-D-13-00055.1.

Hsu, J. S., Powell, J., & Adler, P. B. (2012). Sensitivity of mean annual primary production to precipitation. *Global Change Biology*, *18*, 2246–2255.

Huenneke, L. (2011). NPP Study: Annual aboveground net primary production, summary by site. Environmental Data Initiative https://doi.org/10.6073/pasta/5a17c5689767d95056ec63842b981998

Huffman, G. J., Adler, R. F., Morrissey, M. M., Bolvin, D. T., Curtis, S., Joyce, R., & Susskind, J. (2001). Global precipitation at one-degree daily resolution from multisatellite observations. *Journal of Hydrometeorology*, *2*, 36–50.

Humphries, H. (2018). Aboveground net primary productivity data for Saddle grid, 1992 - ongoing, Environmental Data Initiative. https://doi.org/10.6073/pasta/cd03361ac0c1cbc686ad71da7e14f79d. Dataset accessed 5/03/2019

Huxman, T. E., Smith, M. D., Fay, P. A., Knapp, A. K., Shaw, R. M., Loik, M. E., et al. (2004). Convergence across biomes to a common rain-use efficiency. *Nature*, *429*, 651–654. https://doi.org/10.1038/nature02561

Imer, D., Merbold, L., Eugster, W., & Buchmann, N. (2013). Temporal and spatial variations of soil CO_2, CH_4 and N_2O fluxes at three differently managed grasslands. *Biogeosciences*, *10*(9), 5931–5945. https://doi.org/10.5194/bg-10-5931-2013

Isaac, P., Cleverly, J., McHugh, I., van Gorsel, E., Ewenz, C., & Beringer, J. (2017). Oz flux data: Network integration from collection to curation. *Biogeosciences*, *14*(12), 2903–2928. https://doi.org/10.5194/bg-14-2903-2017

Jaksic, V., Kiely, G., Albertson, J., Oren, R., Katul, G., Leahy, P., & Byrne, K. A. (2006). Net ecosystem exchange of grassland in contrasting wet and dry years. *Agricultural and Forest Meteorology*, *139*(3), 323–334. https://doi.org/10.1016/j.agrformet.2006.07.009

Jeong, S.-J., Ho, C.-H., Gim, H.-J., & Brown, M. E. (2011). Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982-2008. *Global Change Biology*, *17*(7), 2385–2399. https://doi.org/10.1111/j.1365-2486.2011.02397.x

Jia, X., Baoni, X., Shao, M., & Zhao, C. (2015). Primary productivity and precipitation-use efficiency in temperate grassland in the Loess Plateau of China. *PloS ONE*, *10*(8), e0135490. https://doi.org/10.1371/journal.pone.0135490

Jolly, W. M., Nemani, R., & Running, S. W. (2005). A generalized, bioclimatic index to predict foliar phenology in response to climate. *Global Change Biology*, *11*(4), 619–632. https://doi.org/10.1111/j.1365-2486.2005.00930.x

Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., & Richardson, A. D. (2014). Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Climate Change*, *4*, 598–604. https://doi.org/10.1038/nclimate2253

Knapp, A. K., Briggs, J. M., Hartnett, D. C., & Collins, S. L. (Eds.) (1998). *Grassland dynamics: Long-term ecological research in tallgrass prairie*. New York, NY, USA: Oxford University Press.

Knapp, A. K., Burns, C. E., Fynn, R. W. S., Kirkman, K. P., Morris, C. D., & Smith, M. D. (2006). Convergence and contingency in production-precipitation relationships in north american and south african c4 grasslands. *Oecologia*, *149*, 456–464. https://doi.org/10.1007/s00442-006-0468-2

Knapp, A. K., Carroll, C. J. W., Denton, E. M., La Pierre, K. J., Collins, S. L., & Smith, M. D. (2015). Differential sensitivity to regional-scale drought in six central u.s. grasslands. *Oecologia*, *177*(4), 949–957. https://doi.org/10.1007/s00442-015-3233-6

Knapp, A. K., & Smith, M. D. (2001). Variation among biomes in temporal dynamics of aboveground primary production. *Science*, *291*, 481–484. https://doi.org/10.1126/science.291.5503.481

Knapp, A. K., Smith, M. D., Collins, S. L., Zambatis, N., Peel, M., Emery, S., et al. (2004). Generality in ecology: Testing north american grassland rules in south african savannas. *Frontiers in Ecology and the Environment*, *2*(9), 483–491.

Knyazikhin, Y., Glassy, J., Privette, J. L., Tian, Y., Lotsch, A., Zhang, Y., et al. (1999). MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absored by Vegetation (FPAR) Product (MOD15) Algorithm Theoretical Basis Document, http://eospso.gsfc.nasa.gov/atbd/modistables.html

Koerner, S. E., & Collins, S. L. (2014). Interactive effects of grazing, drought, and fire on grassland plant communities in North America and South Africa. *Ecology*, *95*(1), 98–109.

Konings, A., Williams, A., & Gentine, P. (2017). Sensitivity of grassland productivity to aridity controlled by stomatal and xylem regulation. *Nature Geoscience*, *10*, 284–288. https://doi.org/10.1038/NGEO2903

Kwon, H., Oechel, W., Zulueta, R., & Hastings, S. (2006). Effects of climate variability on carbon sequestration among adjacent wet sedge tundra and moist tussock tundra ecosystems. *Journal of Geophysical Research*, *111*, G03014. https://doi.org/10.1029/2005JG000036

L Coulter, R., Pekour, M., Cook, D., Klazura, G., Martin, T., & Lucas, J. (2006). Surface energy and carbon dioxide fluxes above different vegetation types within ABLE. *Agricultural and Forest Meteorology*, *136*, 147–158. https://doi.org/10.1016/j.agrformet.2004.11.011

Lara, M. J., Nitze, I., Grosse, G., Martin, P., & McGuire, A. D. (2018). Reduced arctic tundra productivity linked with landform and climate change interactions. *Scientific Reports*, *8*, 2345. https://doi.org/10.1038/s41598-018-20692-8

Lauenroth, W. K., Burke, I. C., & Gutmann, M. P. (1999). The structure and function of ecosystems in the central North American grassland region. *Great PLains Research*, *9*, 223–259.

Lauenroth, W. K., & Sala, O. E. (1992). Long-term forage production of north american shortgrass steppe. *Ecological Applications*, *2*(4), 397–403. https://doi.org/10.2307/1941874

Lei, T., Pang, Z., Wang, X., Li, L., Fu, J., Kan, G., et al. (2016). Drought and carbon cycling of grassland ecosystems under global change: A review. *Water*, *8*, 460. https://doi.org/10.3390/w8100460

Levis, S., Badger, A., Drewniak, B., Nevison, C., & Ren, X. (2016). Clmcrop yields and water requirements: Avoided impacts by choosing rcp 4.5 over 8.5. *Climatic Change*, *146*, 501–515. https://doi.org/10.1007/s10584-10016-11654-10589

Litvak, M. (2007). AmeriFlux US-Seg Sevilleta grassland. https://10.17190/AMF/1246124

Logan, K., & Brunsell, N. (2015). Influence of drought on growing season carbon and water cycling with changing land cover. *Agricultural and Forest Meteorology*, *213*, 217–225.

Lokupitiya, E., Denning, S., Paustian, K., Baker, I., Schaefer, K., Verma, S., et al. (2009). Incorporation of crop phenology in Simple Biosphere Model (SiBcrop) to improve land-atmosphere carbon exchanges from cropland. *Biogeosciences*, *6*, 969–986. https://doi.org/10.5194/bg-6-969-2009

Lowman, L. E. L., & Barros, A. P. (2018). Predicting canopy biophysical properties and sensitivity of plant carbon uptake to water limitations with a coupled eco-hydrological framework. *Ecological Modelling*, *372*, 33–52. https://doi.org/10.1016/j.ecolmodel.2018.01.011

Ma, S., Baldocchi, D., Xu, L., & Hehn, T. (2007). Inter-annual variability in carbon dioxide exchange of an oak/grass savanna and open grassland in California. *Agricultural and Forest Meteorology*, *147*, 157–171.

Marcolla, B., Cescatti, A., Manca, G., Zorer, R., Cavagna, M., Fiora, A., et al. (2011). Climatic controls and ecosystem responses drive the inter-annual variability of the net ecosystem exchange of an alpine meadow. *Agricultural and Forest Meteorology*, *151*(9), 1233–1243. https://doi.org/10.1016/j.agrformet.2011.04.015

Marek, M. V., Janouš, D., Taufarová, K., Havránková, K., Pavelka, M., Kaplan, V., & Marková, I. (2011). Carbon exchange between ecosystems and atmosphere in the Czech Republic is affected by climate factors. *Environmental Pollution*, *159*(5), 1035–1039. https://doi.org/10.1016/j.envpol.2010.11.025

Matamala, R. (2004). AmeriFlux US-IB2 Fermi National Accelerator Laboratory-Batavia (Prairie site). https://10.17190/AMF/1246066

Matamala, R., Jastrow, J., Miller, R., & Garten, C. (2008). Temporal changes in C and N stocks of restored prairie: Implications for C sequestration strategies. *Ecological Applications*, *18*, 1470–1488.

Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., et al. (2006). European phenological response to climate change matches the warming pattern. *Global Change Biology*, *12*(10), 1969–1976. https://doi.org/10.1111/j.1365-2486.2006.01193.x

Merbold, L., Eugster, W., Stieger, J., Zahniser, M., Nelson, D., & Buchmann, N. (2014). Greenhouse gas budget (CO_2, CH_4 and N_2O) of intensively managed grassland following restoration. *Global Change Biology*, *20*(6), 1913–1928. https://doi.org/10.1111/gcb.12518

Meyers, T. (2000). AmeriFlux US-FPe Fort Peck. https://10.17190/AMF/1246053

Meyers, T. (2004). AmeriFlux US-Bkg Brookings. https://10.17190/AMF/1246040

Meyers, T. (2007). AmeriFlux US-Ctn Cottonwood. https://10.17190/AMF/1246117

Milchunas, D. G., & Lauenroth, W. K. (2001). Belowground primary production by carbon isotope decay and long-term root biomass dynamics. *Ecosystems*, *4*(2), 139–150. https://doi.org/10.1007/s100210000064

Miles, V. V, & Esau, I. (2016). Spatial heterogeneity of greening and browning between and within bioclimatic zones in northern west siberia. *Environmental Research Letters*, *11*(11), 115002. https://doi.org/10.1088

Moran, M. S., Ponce-Campos, G. E., Huerte, A., McClaran, M. P., Zhang, Y., Hamerlynck, E. P., et al. (2014). Functional response of U.S. grasslands to the early 21st-century drought. *Ecology*, *8*, 2121–2133.

Muldavin, E., & Moore, D. (2016). Core research site web seasonal biomass and seasonal and annual NPP data for the net primary production study at the Sevilleta National Wildlife Refuge, New Mexico (1999-present), Environmental Data Initiative. https://doi.org/10.6073/pasta/b913c188d07335b3e88fd585c1cd20b5

Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G., & Nemani, R. R. (1997). Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature*, *386*, 698–702.

National Climatic Data Center (2013). NOAA National Centers for Environmental Information, State of the Climate: National Overview for Annual 2012. http://www.ncdc.noaa.gov/sotc/national/201113

Nielsen-Gammon, J. W. (2011). The 2011 Texas Drought. Office of the State Climatologist Report, College Station, TX, http://climatexas.tamu.edu/files/2011_drought.pdf

ORNL DAAC (2008a). *MODIS collection 5 land products global subsetting and visualization tool*. Oak Ridge, Tennessee, USA: ORNL DAAC. Accessed August 2016. Subsets obtained for MOD15A2 product at various sites in spatial range: N=72.37N, S=35.00S, E=148.55E, W=157.40W, time period: 2000-02-18 to 2014-12-31, and subset size: 0.25x0.25 km, https://doi.org/10.334/ORNLDAAC/1241

ORNL DAAC (2008b). *MODIS Collection 5 land products global subsetting and visualization tool.* Oak Ridge, Tennessee, USA: ORNL DAAC. Accessed August 2016. Subsets obtained for MYD15A2 product at various sites in spatial range: N=72.37N, S=35.00S, E=148.55E, W=157.40W, time period: 2002-07-01 to 2014-12-31, and subset size: 0.25x0.25 km, https://doi.org/10.334/ORNLDAAC/1241

ORNL DAAC (2015). FLUXNET web page. Available online http://uxnet.ornl.gov from ORNL DAAC, Oak Ridge, Tennessee, USA. 1382 Accessed May 2016.

Oesterheld, M., Loreti, J., Semmartin, M., & Sala, O. E. (2001). Inter-annual variation in primary production of a semi-arid grassland related to previous-year production. *Journal of Vegetation Science*, *12*, 137–142.

Park, T., Ganguly, S., Tømmervik, H., Euskirchen, E. S., Høgda, K.-A., Karlsen, S. R., et al. (2016). Changes in growing season duration and productivity of northern vegetation inferred from long-term remote sensing data. *Environmental Research Letters*, *11*(8), 084001. https://doi.org/10.1088/1748-9326/11/8/084001

Peichi, M., Leahy, P., & Kiely, G. (2011). Six-year stable annual uptake of carbon dioxide in intensively managed humid temperate grassland. *Ecosystems*, *14*(1), 112–126.

Peters, D. P. C., Yao, J., Sala, O. E., & Anderson, J. P. (2012). Directional climate change and potential reversal of desertification in arid and semiarid ecosystems. *Global Change Biology*, *18*(1), 151–163. https://doi.org/10.1111/j.1365-2486.2011.02498.x

Petrie, M. D., Brunsell, N. A., Vargas, R., Collins, S. L., Flanagan, L. B., Hanan, N. P., et al. (2016). The sensitivity of carbon exchanges in Great Plains grasslands to precipitation variability. *Journal of Geophysical Research: Biogeosciences*, *121*, 280–294. https://doi.org/10.1002/2015JG003205

Petrie, M. D., Collins, S. L., Swann, A. M., Ford, P. L., & Litvak, M. (2015). Grassland to shrubland state transitions enhance carbon sequestration in the northern Chihuahuan Desert. *Global Change Biology*, *21*(3), 1226–1235. https://doi.org/10.1111/gcb.12743

Post, E., Steinman, B. A., & Mann, M. E. (2018). Acceleration of phenological advance and warming with latitude over the past century. *Scientific Reports*, *8*, 3927. https://doi.org/10.1038/s41598-018-22258-0

Raczka, B. M., Davis, K. J., Huntzinger, D., Neilson, R. P., Poulter, B., Richardson, A. D., et al. (2013). Evaluation of continental carbon cycle simulations with North American flux tower observations. *Ecological Monographs*, *83*(4), 531–556. https://doi.org/10.1890/12-0893.1

Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., 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*(2), 566–584. https://doi.org/10.1111/j.1365-2486.2011.02562.x

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., & Toomey, M. (2013). Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agricultural and Forest Meteorology*, *169*, 156–173. https://doi.org/10.1016/j.agrformet.2012.09.012

Riedo, M., Grub, A., Rosset, M., & Fuhrer, J. (1998). A pasture simulation model for dry matter production and fluxes of carbon, nitrogen, water and energy. *Ecological Modelling*, *105*, 141–183.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., et al. (2011). Merra: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, *24*(14), 3624–3648. https://doi.org/10.1175/JCLI-D-11-00015.1

Runkle, B. R. K., Sachs, T., Wille, C., Pfeiffer, E.-M., & Kutzbach, L. (2013). Bulk partitioning the growing season net ecosystem exchange of CO_2 in Siberian tundra reveals the seasonality of its carbon sequestration strength. *Biogeosciences*, *10*(3), 1337–1349. https://doi.org/10.5194/bg-10-1337-2013

Sachs, T., Giebels, M., Boike, J., & Kutzbach, L. (2010). Environmental controls on CH_4 emission from polygonal tundra on the microsite scale in the Lena river delta, Siberia. *Global Change Biology*, *16*(11), 3096–3110. https://doi.org/10.1111/j.1365-2486.2010.02232.x

Sala, O. E., Gherardi, L. A., Reichmann, L., Jobbágy, E., & Peters, Debra (2012). Legacies of precipitation fluctuations on primary production: Theory and data synthesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1606), 3135–3144. https://doi.org/10.1098/rstb.2011.0347

Sala, O. E., Parton, W. J., Joyce, L. A., & Lauenroth, W. K. (1988). Primary production of the central grassland region of the United States. *Ecology*, *69*(1), 40–45. https://doi.org/10.2307/1943158

Saleska, S., da Rocha, H., Huete, A., Nobre, A., Artaxo, P., & Shimabukuro, Y. (2013). LBA-ECO CD-32 Flux Tower Network Data Compilation, Brazilian Amazon: 1999-2006. Data set. Available on-line [http://daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, https://doi.org/10.3334/ORNLDAAC/1174

Schaefer, Kevin, Collatz, G. James, Tans, Pieter, Denning, A. Scott, Baker, Ian, Berry, Joe, et al. (2008). Combined simple biosphere/Carnegie-Ames-Stanford approach terrestrial carbon cycle model. *Journal of Geophysical Research*, *113*, G03034. https://doi.org/10.1029/2007JG000603

Schaefer, K., Schwalm, C. R., Williams, C., Arain, M. A., Barr, A., Chen, J. M., et al. (2012). A model-data comparison of gross primary productivity: Results from the North American Carbon Program site synthesis. *Journal of Geophysical Research*, *117*, G03010. https://doi.org/10.1029/2012JG001960

Schmidt, A., Hanson, C., Kathilankal, J., & Law, B. (2011). Classification and assessment of turbulent fluxes above ecosystems in North America with self-organizing feature map networks. *Agricultural and Forest Meteorology*, *151*, 508–520. https://doi.org/10.1016/j.agrformet.2010.12.009

Scholes, R., Gureja, N., Giannecchinni, M., Dovie, D., Wilson, B., Davidson, N., et al. (2001). The environment and vegetation of the flux measurement site near Skukuza, Kruger National Park. *Koedoe-African Protected Area Conservation and Science*, *44*(1), 73–83. https://doi.org/10.4102/koedoe.v44i1.187

Schwalm, C. R., Williams, C. A., Schaefer, K., Anderson, R., Arain, M. A., Baker, I., et al. (2010). A model-data intercomparison of CO_2 exchange across North America: Results from the North American Carbon Program site synthesis. *Journal of Geophysical Research*, *115*, G00H05. https://doi.org/10.1029/2009JG001229

Scott, R. (2004). AmeriFlux US-Wkg Walnut Gulch Kendall Grasslands. https://10.17190/AMF/1246112

Scott, R. L., Hamerlynck, E. P., Jenerette, G. D., Moran, M. S., & Barron-Gafford, G. A. (2010). Carbon dioxide exchange in a semidesert grassland through drought-induced vegetation change. *Journal of Geophysical Research*, *115*, G03026. https://doi.org/10.1029/2010JG001348

Seager, R., Goddard, L., Nakamura, J., Henderson, N., & Lee, D. E. (2014). Dynamical causes of the 2010/11 Texas-northern Mexico drought. *Journal of Hydrometeorology*, *15*, 39–68. https://doi.org/10.1175/JHM-D-13-024.l

Sellers, P. J., Berry, J. A., Collatz, G. J., Field, C. B., & Hall, F. G. (1992). Canopy Reflectance, photosynthesis, and transpiration. III. A reanalysis using improved leaf models and a new canopy integration scheme. *Remote Sensing of Environment*, *42*, 187–216.

Sellers, P. J., Heiser, M. D., & Hall, F. G. (1992). Relations between surface conductance and spectral vegetation indices at intermediate (100 m^2 to 15 km^2) length scales. *Journal of Geophysical Research*, *97*(D17), 19,033–19,059. https://doi.org/10.1029/92JD01096

Sellers, P. J., Mintz, Y., Sud, Y. C., & Dalcher, A. (1986). A Simple Biosphere Model (SiB) for use within general circulation models. *Journal Atmospheric Science*, *43*(6), 505–531.

Sellers, P. J., Randall, D. A., Collatz, G. J., Berry, J. A., Field, C. B., Dazlich, D. A., et al. (1996). A revised land surface parameterization (SiB2) for atmospheric GCMs. *Journal of Climate*, *9*, 676–705.

Shi, Z., Thomey, M. L., Mowll, W., Litvak, M., Brunsell, N. A., Collins, S. L., et al. (2014). Differential effects of extreme drought on production and respiration: Synthesis and modeling analysis. *Biogeosciences*, *11*(3), 621–633. https://doi.org/10.5194/bg-11-621-2014

Soussana, J., Allard, V., Pilegaard, K., Ambus, P., Amman, C., Campbell, C., et al. (2007). Full accounting of the greenhouse gas (CO_2, N_2O, CH_4) budget of nine Eurpean grassland sites. *Agriculture, Ecosystems & Environment*, *121*, 121–134. https://doi.org/10.1016/j.agee.2006.12.022

Stöckli, R., Rutishauser, T., Dragoni, D., O'Keefe, J., Thornton, P. E., Jolly, M., et al. (2008). Remote sensing data assimilation for a prognostic phenology model. *Journal of Geophysical Research*, *113*, G04021. https://doi.org/10.1029/2008JG000781

Stöckli, R., Rutishauser, T., Dragoni, D., O'Keefe, J., Thornton, P. E., Jolly, M., et al. (2011). A global reanalysis of vegetation phenology. *Journal of Geophysical Research*, *116*, G03020. https://doi.org/10.1029/2010JG001545

Suttie, J. M., Reynolds, S. G., & Batello, C. (2005). *Grasslands of the world*. Rome, Italy: Food and Agriculture Organization of the United Nations.

Torn, M. (2005). AmeriFlux US-Arc ARM Southern Great Plains control site Lamont. https://doi.org/10.17190/AMF/1246026

Uresk, D., Sims, P., & Dodd, J. (2015). NPP Grassland: Central Plains Experimental Range (SGS), USA, 1939-1990, R1. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/145

Verchot, L., Krug, T., Lasco, R. D., Ogle, S., Raison, J., Li, Y., et al. (2006). Chapter 6: Grassland. In S. Eggleston, L. Buendia, K. Miwa, T. Ngara, & K. Tanabe (Eds.), *2006 IPCC guidelines for national greenhouse gas inventories. volume 4: Agriculture, forestry, and other land use*. Hayama, Japan: Institute for Global Environmental Strategies (IGES). http://www.ipcc-nggip.iges.or.jp

Viskari, T., Hardiman, B., Desai, A. R., & Dietze, M. C. (2015). Model-data assimilation of multiple phenological observations to constrain and predict leaf area index. *Ecological Applications*, *25*(2), 546–558. https://doi.org/10.1890/14-0497.1

von Randow, C., Manzi, A. O., Kruijt, B., de Oliveira, P. J., Zanchi, F. B., Silva, R. L., et al. (2004). Comparative measurements and seasonal variations in energy and carbon exchange over forest and pasture in South West Amazonia. *Theoretical and Applied Climatology*, *78*(1), 5–26. https://doi.org/10.1007/s00704-004-0041-z

Vuichard, N., Soussana, J.-F., Ciais, P., Viovy, N., Ammann, C., Calanca, P., et al. (2007). Estimating the greenhouse gas fluxes of European grass1560 lands with a process-based model: 1. Model evaluation from in situ measurements. *Global Biogeochemical Cycles*, *21*, GB1004. https://doi.org/10.1029/2005GB002611

Wang, J.-W., Denning, A. S., Lu, L., Baker, I. T., Corbin, K. D., & Davis, K. J. (2007). Observations and simulations of synoptic, regional, and local variations in atmospheric co_2. *Journal of Geophysical Research*, *112*, D04108. https://doi.org/10.1029/2006JK007410

Wang, S., Huang, C., Zhang, L., Lin, Y., Cen, Y., & Wu, T. (2016). Monitoring and assessing the 2012 drought in the great plains: Analyzed satellite-retrieved solar-induced chlorophyll fluorescence, drought indices, and gross primary production. *Remote Sensing*, *8*(2), 61. https://doi.org/10.1007/978-1-4614-7501-9_14

White, R., Murray, S., & Rohweder, M. (2000). *Pilot analysis of global ecosystems: Grassland ecosystems*. Washington, DC, USA: World Resources Institute.

White, M. A., Thornton, P. E., & Running, S. W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles*, *11*(2), 217–234.

Wohlfahrt, G., Hammerle, A., Haslwanter, A., Bahn, M., Tappeiner, U., & Cernusca, A. (2008). Seasonal and inter-annual variability of the net ecosystem CO_2 exchange of a temperate mountain grassland: Effects of weather and management. *Journal of Geophysical Research*, *113*, D08110. https://doi.org/10.1029/2007JD009286

Wolf, S., Eugster, W., Potvin, C., & Buchmann, N. (2011). Strong seasonal variations in net ecosystem CO_2 exchange of a tropical pasture and afforestation in Panama. *Agricultural and Forest Meteorology*, *151*(8), 1139–1151. https://doi.org/10.1016/j.agrformet.2011.04.002

Wolf, S., Keenan, T. F., Fisher, J. B., Baldocchi, D. D., Desai, A. R., Richardson, A. D., et al. (2016). Warm spring reduced carbon cycle impact of the 2012 US summer drought. *Proceedings of the National Academy of Sciences*, *113*(21), 5880–5885. https://doi.org/10.1073/pnas.1519620113

Xia, J. Y, Luo, Y. Q., Wang, Y.-P., Weng, E. S., & Hararuk, O. (2012). A semi-analytical solution to accelerate spin-up of coupled carbon and nitrogen land model to steady state. *Geoscientific Model Development*, *5*, 1259–1271. https://doi.org/10.5194/gmd-5-1259-2012

Zhang, Y., Sachs, T., Li, C., & Boike, J. (2012). Upscaling methane fluxes from closed chambers to eddy covariance based on a permafrost biogeochemistry integrated model. *Global Change Biology*, *18*(4), 1428–1440. https://doi.org/10.1111/j.1365-2486.2011.02587.x

Zhu, L., & Southworth, J. (2013). Disentangling the relationships between net primary production and precipitation in Southern Africa savannas using satellite observations from 1982 to 2010. *Remote Sensing*, *5*(8), 3803–3825. https://doi.org/10.3390/rs5083803

Zona, D., & Oechel, W. (1998). AmeriFlux US-Brw Barrow. https://10.17190/AMF/1246041

References From the Supporting Information

Barr, A. G., Richardson, A. D., Hollinger, D. Y., Papale, D., Arain, M. A., Black, T. A., et al. (2013). Use of change-point defection for friction-velocity threshold evaluation in eddy-covariance studies. *Agricultural and Forest Meteorology*, *171-172*, 31–45. https://doi.org/10.1016/j.agrformet.2012.11.023

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., et al. (2006). Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: Algorithms and uncertainty estimation. *Biogeosciences*, *3*(4), 571–583. https://doi.org/10.5194/bg-3-571-2006

Xu, L., & Baldocchi, D. D. (2004). Seasonal variation in carbon dioxide exchange over a mediterranean annual grassland in California. *Agricultural and Forest Meteorology*, *1232*, 79–96. https://doi.org/10.1016/j.agrformet.2003.10.004