Addressing Biases in Arctic-Boreal Carbon Cycling in the Community Land Model Version 5

Leah Birch[1], Christopher R. Schwalm[1], Sue Natali[1], Danica Lombardozzi [2], Gretchen Keppel-Aleks[3], Jennifer Watts[1], Xin Lin[3], Donatella Zona[4], Walter Oechel[4], Torsten Sachs[5], Thomas Andrew Black[6], and Brendan M. Rogers[1]

[1]Woodwell Climate Research Center
[2]National Center for Atmospheric Research
[3]University of Michigan
[4]San Diego State University
[5]GFZ German Research Centre for Geosciences
[6]University of BC

Correspondence: Leah Birch (lbirch@woodwellclimate.org) and Brendan Rogers (brogers@woodwellclimate.org)

Abstract. The Arctic-boreal zone (ABZ) is experiencing amplified warming, actively changing biogeochemical cycling of vegetation and soils. The land-to-atmosphere fluxes of CO_2 in the ABZ have the potential to increase in magnitude and feedback to the climate causing additional large scale warming. The ability to model and predict this vulnerability is critical to preparation for a warming world, but Earth system models have biases that may hinder understanding the rapidly changing ABZ carbon fluxes. Here we investigate circumpolar carbon cycling represented by the Community Land Model 5 (CLM5.0) with a focus on seasonal gross primary productivity (GPP) in plant functional types (PFTs). We benchmark model results using data from satellite remote sensing products and eddy covariance towers. We find consistent biases in CLM5.0 relative to observational constraints: (1) the onset of deciduous plant productivity to be late, (2) the offset of productivity to lag and remain abnormally high for all PFTs in fall, (3) a high bias of grass, shrub, and needleleaf evergreen tree productivity, and (4) an underestimation of productivity of deciduous trees. Based on these biases, we focus model development of alternate phenology, photosynthesis schemes, and carbon allocation parameters at eddy covariance tower sites. Although our improvements are focused on productivity, our final Model Recommendation results in other component CO_2 fluxes, e.g. Net Ecosystem Exchange (NEE) and Terrestrial Ecosystem Respiration (TER), that are more consistent with observations. Results suggest that algorithms developed for lower latitudes and more temperate environments can be inaccurate when extrapolated to the ABZ, and that many land surface models may not accurately represent carbon cycling and its recent rapid changes in high latitude ecosystems, especially when analyzed by individual PFTs.

1 Introduction

As the atmospheric concentration of CO_2 continues to rise, the Arctic-boreal Zone (ABZ) is expected to continue to warm more rapidly than the rest of the globe (Serreze and Francis, 2006; Serreze and Barry, 2011). The impacts of this accelerated warming are manifest across all major components of the ABZ — the cryosphere, hydrosphere, and biosphere (Duncan et al.,

1

2020). The multifaceted ABZ response to warming includes accelerated carbon cycling (Jeong et al., 2018), permafrost thaw, intensification of disturbance regimes (Alexander and Mack, 2016), changes in snow cover and ecosystem water availability (Callaghan et al., 2011; Biancamaria et al., 2011), and shifts in vegetation structure and composition (Beck et al., 2011; Forkel et al., 2016; Searle and Chen, 2017). These changes in the whole ABZ terrestrial ecosystem structure and function have important implications for global climate, given the region's strong biophysical coupling (Bonan et al., 1992; Bala et al., 2007; Rogers et al., 2013, 2015) and large, and potentially vulnerable, reservoirs of below and aboveground carbon, especially in the permafrost zone (Shaver et al., 1992; McGuire et al., 2009, 2010; Koven et al., 2015; Parazoo et al., 2018b; Natali et al., 2019; McGuire et al., 2018).

The responses of carbon cycling in the ABZ to changes in global climate are complex, interconnected, and may have compensating effects (Welp et al., 2016). For example, air and soil warming, in conjunction with a lengthening of the annual non-frozen period across the ABZ (Kim et al., 2012), stimulate plant productivity directly and indirectly through increased nutrient and water availability (Natali et al., 2014; Salmon et al., 2016). Warming and CO_2 fertilization have contributed to widespread "greening" across the ABZ, including shrubification (Myers-Smith et al., 2011, 2015) and northward treeline expansion (Lloyd and Fastie, 2003; Chapin et al., 2005), i.e. the encroachment of trees and shrubs into tundra regions. However, rapid warming across much of the ABZ is also accelerating decomposition, causing drought stress in warmer and drier landscapes (Carroll et al., 2011; Walker and Johnstone, 2014; Walker et al., 2015; Carroll and Loboda, 2017), and intensifying disturbance regimes such as wildfire and insect outbreaks (Turetsky et al., 2011; Kasischke et al., 2010; Rogers et al., 2018; Hanes et al., 2019); all of which contribute to the increasingly observed patterns of "browning" in the ABZ (Verbyla, 2011; Elmendorf et al., 2012; Phoenix and Bjerke, 2016).

As an emergent property of global change drivers in the ABZ, the seasonal cycle of CO_2 exchange across the ABZ has been experiencing changes in timing and magnitude of fluxes. Most critically regarding the magnitude of carbon fluxes, the atmospheric CO_2 concentration in the ABZ has been measured to be increasing between 30-60% during the last 60 years (Keeling et al., 1996; Randerson et al., 1999; Graven et al., 2013; Liptak et al., 2017; Jeong et al., 2018). Our current knowledge of the ABZ seasonal cycle of CO_2 suggests that much of the observed change in seasonal amplitude is due to increased vegetation productivity during the growing season, a result of CO_2 fertilization and warming (Forkel et al., 2016; Ito et al., 2016; Zhao et al., 2016). At the same time, fall and winter respiration constitute a large portion of the annual CO_2 budget (Euskirchen et al., 2014; Natali et al., 2019) and have been increasing with climate change (Belshe et al., 2012; Piao et al., 2008), making the implications for net sink-source dynamics uncertain (Ciais et al., 1995; McGuire et al., 2018). Hence, the current and anticipated state of carbon source-sink dynamics remains an open question in part due to the uncertainty in the dominant mechanisms and differential responses governing carbon fluxes across the ABZ.

Ground observations, satellite products, and process-based climate models are all used to understand interactions and feedbacks between changing environmental conditions and carbon cycling in the ABZ. *In situ* observations of carbon fluxes are required for mechanistic understanding but are often limited across time and space, especially in large and remote regions with extreme temperatures, like the ABZ (Virkkala et al., 2018, 2019). For example, respiration during the winter has long been assumed to be effectively zero, but better technology has slowly allowed the seasonal cycle story to grow (Natali et al., 2019).

Satellite observations provide near complete coverage in space and time, but are indirect observations of ecosystem properties, are ~~challenged~~ challenging in the ABZ due to low insolation in the winter months and extreme snow storms, and contain a variety of uncertainties related to sensor properties, atmospheric contamination, and processing (Duncan et al., 2020). The brevity of the growing season and lack of light in the ABZ throughout the year also contributes to biases in satellite measurements (Randerson et al., 1997). Process-based models, or terrestrial biosphere models (TBMs), are a particularly invaluable resource for examining mechanisms across spatial and temporal scales, even projecting carbon cycle feedbacks in the future under varying socioeconomic scenarios (Taylor et al., 2012; Eyring et al., 2016, CMIP5 and 6). However, due to different formulations, assumptions, mechanisms, model inputs, and parameterizations, TBMs display a wide range of CO_2 source-sink dynamics in the ABZ (Fisher et al., 2014; Huntzinger et al., 2013) and biases compared to observations (Schwalm et al., 2010; Schaefer et al., 2012). Given the criticality of the ABZ to future global carbon balance and the heterogeneity of landscape responses to warming, it is a high priority to understand and address the current biases in TBM carbon cycling.

The Community Land Model version 5.0 (CLM5.0) is the land component of the Community Earth System Mode version 2.0 (CESM2.0). CLM is one of the most widely-used land surface models and contributes to many global intercomparisons (Zhao and Zeng, 2014; Peng et al., 2015; Ito et al., 2016) and future climate projections relevant for scientists and policymakers (Piao et al., 2013, eg, IPCC). The current state-of-the-art release of the Community Land Model (Lawrence et al., 2019, CLM) incorporates several improvements to climatic fluxes and biogeochemistry relevant for the ABZ. A general improvement was observed globally for CLM5.0 compared to past versions of the model (i.e. CLM 4.0 and 4.5). However, a high bias in photosynthesis or gross primary productivity (GPP) at high latitudes remains a well-documented issue (Wieder et al., 2019) in CLM5.0. Thus, we explore the simulation of GPP along with the net ecosystem exchange (NEE) and terrestrial ecosystem respiration (TER) in order to identify biases in simulation of the seasonal carbon balance.

This study assesses the ability of CLM5.0 to accurately represent CO_2 fluxes with gridded model simulations, identifies deficiencies in the simulation of ABZ carbon fluxes, and provides a Model Recommendation for application in the ABZ. We provide a step-by-step diagnosis of the major factors contributing to biases in the simulation of the seasonal cycle of CO_2 fluxes in CLM5.0. We use FLUXCOM (Jung et al., 2017, 2019, a gridded product based on machine learning), the International Land Model Benchmarking Project (Collier et al., 2018, ILAMB) to assess model results, and in situ data from FluxNet (https://fluxnet.fluxdata.org) and Ameriflux (https://ameriflux.lbl.gov/). We focus our development on the simulation of CO_2 fluxes for each ABZ vegetation type in CLM5.0, representing the tundra and boreal forest. We use point-based simulations at eddy covariance (i.e., EC or flux) tower sites to inform the failure or success of each model development test of the phenology and photosynthesis modules in CLM5.0. We validate model development using gridded products and additional flux towers (withheld from the initial model development) before making our final Model Recommendation. As a result, we identify and resolve many of the known biases in the representation of phenology (Richardson et al., 2012), photosynthesis (Lawrence et al., 2019), and carbon allocation in CLM5.0, allowing a more realistic representation of carbon cycling in this rapidly changing ecosystem.

2 Methods

We investigate the seasonal cycle of ABZ CO_2 fluxes with CLM5.0 due to its widespread use and significant model improvements from the previous version. These updated processes include snow physics related to snow age and density, canopy snow interactions, active layer depth, groundwater movement, soil hydrology and biogeochemistry, and river transport (Li et al., 2013). Moving away from globally-constant values of plant traits that are challenging to measure, carbon and nitrogen cycle representations now use prognostic leaf photosynthesis traits (Ali et al., 2016, the maximum rate of electron transport or J_{max} and the maximum rate of carboxylation or V_{cmax}), carbon costs for nitrogen uptake, leaf nitrogen optimization, and flexible leaf stoichiometry. Stomatal physiology were updated with the Medlyn conductance model, replacing the Ball-Berry (Medlyn et al., 2011). Additionally plant hydraulics have undergone recent improvement in more realistic stress representation (Kennedy et al., 2019). One primary goal with these improvements was to allow for more physically-based parameters that could be informed by observational ecological data, ultimately allowing for better fidelity with hydrological and ecological processes. Land cover inputs to CLM5.0 were updated to capture transient land use changes from the satellite record.

2.1 Pan-Arctic CLM5.0 Simulation

We run CLM5.0 at 0.5° by 0.5° grid resolution with meteorology inputs (rainfall, snowfall, 2m air temperature, 2m specific humidity, surface pressure, downward shortwave radiation, downward longwave radiation, 10m wind speed, and cloud cover fraction) from the Global Soil Wetness Project (GSWP3v1,http://hydro.iis.u-tokyo.ac.jp/GSWP3/), which is a standard forcing dataset in the Land Surface, Snow and Soil Moisture Model Intercomparison Project (Van den Hurk et al., 2016, LS3MIP). GSWP3v1 has been shown to be appropriate and ~~less biased as a~~ the least biased forcing dataset for CLM5.0 simulations in the ABZ (Lawrence et al., 2019, CLM). We begin a simulation of the CLM5.0 release in 1850 and run through 2014 including default time series inputs of CO_2, aerosol deposition, nitrogen deposition, and land use change (Lamarque et al., 2010; Lawrence et al., 2016), which are available on NCAR's Cheyenne system (Computational and Laboratory, 2017). We implement a regional simulation of CLM5.0 north of 40° N across both hemispheres, allowing us to focus exclusively on ABZ processes. We confirm the improvements made to the newly updated CLM version 5.0 (Lawrence et al., 2019) through a comparison of CLM version 4.5 with the same input datasets.

For our control CLM5.0 simulation, we use an available equilibrated 1850 intialization on NCAR's Cheyenne system (Computational and Laboratory, 2017) with spun-up carbon pools. After model development, we again spin the model using this initial dataset, and we find GPP to equilibrate quickly, within 20 years (Supplement Fig. S1). To be conservative, we spin-up the model with our recommended model development versions for 100 years to ensure carbon fluxes have come to equilibrium. Then we use this equilibrated state as initial conditions in a production run simulation beginning in 1850 with all the same configuration and climatology as the CLM5.0 release control simulation.

Typical of land surface models, CLM5.0 represents vegetation through broad plant functional types (PFTs). CLM5.0 represents ABZ vegetation using five PFTs: needleleaf evergreen boreal trees (NETs), needleleaf deciduous boreal trees (NDTs), broadleaf deciduous boreal trees (BDTs), deciduous boreal shrubs (hereafter "shrubs"), and arctic C3 grasses (hereafter

"grasses"). We focus model development on PFT-specific comparisons, which allows a direct comparison with observational data. Any improvements to PFT specific carbon flux simulations have implications for changing vegetation distributions in the ABZ.

2.2 Model Benchmarking and Validation with FLUXCOM and ILAMB

Benchmarking is the process of quantifying model performance based on observational data considered to be the expected value or truth. We use FLUXCOM (Tramontana et al., 2016; Jung et al., 2017, 2019) to benchmark gridded CO_2 fluxes (i.e., gross primary productivity, terrestrial ecosystem respiration, and net ecosystem exchange, or GPP, TER, and NEE) in CLM5.0. FLUXCOM is an upscaled machine learning product based on FLUXNET eddy covariance towers. As a global product, FLUX-COM is particularly useful for filling spatial gaps in tower observations, especially in the relatively data-sparse ABZ. Derived from This product uses multiple reanalysis datasets and machine learning methods to train multiple predictors at flux tower sites. The resulting product is the mean of those ensembles, which also allows standard error to be calculated. We use the standard deviation around the mean to identify successful model development. Machine learning is a useful tool for this type of gap filling, as it does not care about geoghraphic locations, just the predictor space, which are the fluxes and environmental conditions. FLUXCOM is unable to simulate fluxes from fires and CO_2 fertilization accurately, which is why we focus model development on averages over the past couple decades, rather than specific years with forest fires and the increasing CO_2 amplitude trend. Any systemic problems with FLUXNET data would also exist within FLUXCOM, but validations of regions with sun-induced chlorophyll fluorescence (Köhler et al., 2015, SIF) add confidence in the FLUXCOM product. Additionally, derived from global MODIS-based vegetation layers Sulla-Menashe and Friedl (2019), FLUXCOM also generates has the ability to generate PFT-specific output. Although there are inconsistencies between the PFT classifications, such as the representation of "mixed forests" in FLUXCOM, this allows direct comparisons of the PFT-specific fluxes represented by CLM5.0.

For an independent set of comparisons that includes additional environmental variables, we also use the International Land Model Benchmarking System (Collier et al., 2018, ILAMB). ILAMB is an open-source land model evaluation system that provides a uniform approach to benchmarking and scoring model fidelity. It is a powerful tool to quickly and thoroughly investigate biases, seasonality, spatial distribution, and interannual variability in climate model output. We use ILAMB to benchmark fluxes of CO_2, moisture, and heat, in addition to several land surface properties essential for climate responses and feedbacks such as albedo and leaf area index (LAI). Although the focus of our model development is on GPP, TER and NEE tend to respond strongly to changes in productivity (Chapin et al., 2006; Schaefer et al., 2012; Chen et al., 2015). We benchmark these additional interdependent properties in ILAMB to ensure our development generates systematic improvements.

2.3 Point Simulation Protocol

Although FLUXCOM is an invaluable tool to fill spatial and temporal gaps in tower observations across the ABZ, it is by definition not as accurate as direct *in situ* observations of CO_2 fluxes, for instance measurements from EC towers, which also include helpful ancillary information such as detailed vegetation composition. After benchmarking the aggregated grid cell fluxes, we assess model performance at specific EC towers that measure year-round seasonal CO_2 fluxes, which is a standard

155 model development procedure (Stöckli et al., 2008a). We aggregate fluxes of CO_2 to monthly means from flux towers in the ABZ that are part of the FluxNet (https://fluxnet.fluxdata.org) and Ameriflux (https://ameriflux.lbl.gov/) networks. We screen the tower records to determine if the PFT type in CLM5.0 corresponds to the vegetation described by tower metadata. We choose towers and grid cells with at least three years of sample data before 2014, as that is the end data of GSWP forcing data for CLM5.0. Collectively, the chosen towers that conform to our data requirements span all PFT types over the ABZ

160 (Supplement Table S1). ~~To mediate the sparsity of data, we~~ We divide our observational data ~~further~~ into model development sites vs. ~~validation sites~~ evaluation sites to prevent overfitting of parameters. Our chosen model development sites are US-EML (Belshe et al., 2012), CA-QC2 (Margolis, 2018), CA-OAS (Black, 2016, BDT), and RU-SKP (Maximov, 2016), which encompass all of the CLM5.0 PFTs. We verify our work using additional flux tower measurements from FI-SOD (Aurela et al., 2016), RU-Tks (Aurela, 2016), CA-Sf1 (Amiro, 2016), US-Atq (Oechel et al., 2014), RU-Sam (Kutzbach et al., 2002-2014;

165 Holl et al., 2019; Runkle et al., 2013), and CA-Gro (McCaughey, 2016).

During our model development process, we examine phenology and photosynthesis schemes in CLM5.0 by running point simulations starting in 1901 using the same inputs as our gridded simulations. For each model issue listed in Section 2.4, we iteratively test hypotheses and ranges of parameters values that may improve the simulation at the representative towers. Points simulations allow for rapid deployment of model tests, while also conserving compute resources. This speed of computation

170 is invaluable for our multiple model development trajectories. We find that for our focus on phenology and photosynthesis, carbon fluxes equilibrate rapidly and a 20-year spin up is sufficient for point simulations (Supplement Fig. S1). We run the point simulations through 2014 and compared the years measured by flux towers with the same years simulated by CLM5.0. We acknowledge that the climatology experienced by a given flux tower and the reanalysis data used as a model input the model are different. Thus, we focus on the mean seasonal behavior of the flux towers and CLM5.0 to guide model development ~~and~~

175 ~~allow the yearly variance to serve~~ . The yearly variance serves as an uncertainty range for our characterization of flux tower behavior. Additionally, using the mean monthly CO_2 fluxes as calibration data can prevent over-fitting of CLM5.0 parameters. Each model development simulation for a specific PFT is also run for the other PFTs at the development sites (CA-QC2, CA-OAS, US-EML, and RU-SKP). After finalizing a given model development scheme, we implement the updates at the withheld EC sites (Supplement Table S1) and then in a gridded fashion across our ABZ regional domain.

180 ## 2.4 Model Development

We identify several issues in the phenology and photosynthesis schemes in CLM5.0 for the ABZ, which are detailed in Section 3.1. These can be categorized by (i) extrapolation of schemes and parameterizations designed for temperate vegetation, (ii) biases in the prediction of leaf photosynthetic traits and (iii) mis-specified carbon allocation parameters. We also incorporate a bug fix as detailed in the Supplement Section 3.

185 ### 2.4.1 Phenology Onset

The representation of spring and autumn phenology for deciduous trees, shrubs and grasses in CLM5.0 is based on a study of the conterminous United States (White et al., 1997) and extended to the ABZ. In the extratropics, plants initiate their photosynthetic

growing season in response to various climatic factors in spring, reach peak productivity in summer, and enter dormancy in autumn. This is parameterized in CLM5.0 by allowing spring onset to begin once a threshold for cumulative growing degree days is met, as determined by White et al. (1997) using relationships between temperatures and the satellite based Normalized Difference Vegetation Index (NDVI). Thus, in CLM5.0, onset is based on relationships derived from temperate latitudes and extrapolated to the ABZ. We find that this parameterization requires relatively warm temperatures for the ABZ before onset can begin, which causes a delay in the beginning of the growing season for deciduous plants in the ABZ.

To implement a more mechanistic approach to onset in the ABZ, we identify environmental thresholds that correspond to physiological changes during spring onset in high latitudes. Field observations consistently demonstrate that productivity initiation in the ABZ is governed by the cessation of freezing temperatures (Ueyama et al., 2013; Stöckli et al., 2008b) and the availability of soil water(Goulden et al., 1998). Additionally snow cover has been shown to influence the start of the growing season (Høye et al., 2007; Semenchuk et al., 2016). We use daily output from FLUXCOM and flux towers to identify the initiation of GPP in spring for different PFTs. We then compare the timing of productivity to a variety of CLM5.0 environmental variables known to correspond strongly with GPP onset (Chapin III and Shaver, 1996; Starr and Oberbauer, 2003; Borner et al., 2008), including soil temperature, soil moisture, soil ice content, air temperature, liquid and ice precipitation, snow depth, and latent and sensible heat fluxes (Supplement Fig. S2). We find that soil temperature (and thus soil ice content in the third soil layer with ∼10 cm depth), minimum 2m temperature, and snow cover undergo notable state transitions around the timing of GPP onset, enabling their use as a phenology threshold in CLM5.0.

For our ABZ deciduous phenology algorithm, we therefore allow photosynthesis to begin when the following environmental criteria occur:

1. the 10-day average soil temperature in the 3rd soil layer is above $0°$ C

2. the 5-day average minimum daily 2m temperature average is above $0°$ C

3. when only a thin layer of snow remains on the ground (< 10 cm).

Together, these metrics approximate when plants begin photosynthesis in spring, allowing for roots to uptake moisture in unfrozen soil, for air temperatures to be consistently above freezing, and for plants to no longer be covered in snow.

2.4.2 Phenology Offset

Fall deciduous phenology in CLM5.0 is based on the same study focused on the temperate latitudes (White et al., 1997). As with phenology onset, biases arise from the extrapolation of temperate zone relationships to the high latitudes. Using NDVI, senescence was identified to occur in autumn when daylight decreases ∼11 hours. This daylight threshold is then set to be a global constant in CLM5.0. Complete dormancy is reached after 30 days after this photoperiod threshold. In the ABZ, this threshold of 11 hours of total daylight generally causes plants to decrease productivity in October and to begin dormancy in November. In reality, vegetation should be reaching dormancy at the end of September in the high Arctic (Zhang et al., 2004), with senescence beginning in August (Corradi et al., 2005). Based on existing physiological studies of ABZ vegetation,

7

it is unclear if temperature or photoperiod are the driving factor that triggers fall senescence (Marchand et al., 2004; Eitel et al., 2019), or if a combination of both is necessary for ABZ senescence (Oberbauer et al., 2013). Therefore, we focus on photoperiod, which is seasonally more consistent across the ABZ and clearly crucial for photosynthesis. Based on observations at high latitudes, 15 hours is a more accurate timing for senescence above 65° N (Corradi et al., 2005; Eitel et al., 2019). We scale the photoperiod threshold linearly along a latitudinal gradient from 65° N until ~11 hours at 45° N such that the temperate latitudes retain the offset timing determined by White et al. (1997).

2.4.3 Day Length Scaling for Photosynthetic Parameters

The Farquhar model of photosynthesis for C3 plants uses two main parameters to represent photosynthetic capacity, J_{max} (the maximum rate of photosynthetic electron transport) and V_{cmax} (the maximum rate of Rubisco carboxylase activtiy). In the current release of CLM5.0, J_{max} and V_{cmax} are predicted by a mechanistic model of Leaf Utilization of Nitrogen for Assimilation (Ali et al., 2016, LUNA). Unlike previous versions of CLM, both J_{max} and V_{cmax} are prognostic in CLM 5.0, which allows for the vegetation to adjust to nutrients and environmental conditions. In our comparison of productivity in CLM5.0, we find that the prediction of J_{max} and V_{cmax} may be biased high in the ABZ (Rogers et al., 2017) when using algorithm values and schemes more appropriate for the tropics and temperate regions, which contributes to the overestimation of GPP by CLM5.0 across the ABZ.

Currently, J_{max} is scaled in LUNA using day length:

$$f(daylength) = \left(\frac{daylength}{12}\right)^2 \tag{1}$$

The function, $f(daylegth)$, is a scaling factor that is based on the formulation in Bauerle et al. (2012), which quantifies the relationship between day length and J_{max}. However, the denominator in this equation in CLM5.0 is set to 12 hours, when it should be the maximum day length possible at a particular latitude (Bauerle et al., 2012). While 12 hours is fairly representative for lower latitudes, this scale factor does not work for the ABZ where some regions experience up to 24 hour day light in summer, which allows f(daylength) > 1 in Equation 1, particularly around the summer solstice in June. To address this, we replace the default denominator of 12 hours with the geographically specific annual maximum hours of daylight that occur for a given grid cell.

2.4.4 Temperature Acclimation

Within both the Photosynthesis and LUNA schemes in CLM5.0, J_{max} and V_{cmax} are scaled from their values on the environmental leaf temperature and to 25° C using a modified Arrhenius temperature response function from (Kattge and Knorr, 2007, Supplement Fig. S3). Currently, J_{max} and V_{cmax} are allowed to acclimate to the plant's growth temperature, defined as the 10-day average 2-m temperature. However, the temperature acclimation function is limited to temperatures between between 11° C and 35° C and tuned to mostly temperate species (Kattge and Knorr, 2007). At temperatures outside of the acclimation range, the temperature acclimation function scales J_{max} and V_{cmax} to unusually high values (Supplement Fig. S3), likely due to the use of temperate species for parameterization tuning. The mean daily summer temperature in the ABZ above 60° N is below

11° C (Kalnay et al., 1996, NCEP/NCAR reanalysi)s), which implies vegetation at this latitude may never enter the range for temperature acclimation designated by Kattge and Knorr (2007). The temperature scaling done below 11° C is not based on any ABZ studies, nor does it match the previous scaling used in CLM5.0 parameterizations from Leuning (2002), which do contain

255 some field sites in the ABZ. At more southern locations in the ABZ, vegetation may fluctuate around this minimum threshold value of 11° C, allowing discontinuities to appear in the temperature scaling of J_{max} and V_{cmax} and influencing biases in the seasonality of CO_2 fluxes. Due the lack of observational data across the ABZ incorporated in the Arrhenius function for acclimation in Kattge and Knorr (2007) we choose to implement temperature scaling functions from Leuning (2002), which does not create a discontinuity in J_{max} and V_{cmax} at such a critical temperature for the ABZ. This is a standard implementation of

260 the Arrhenius temperature response function, which has been shown to work well at lower temperatures under present climate conditions in previous versions of CLM. The shift from Leuning (2002) to Kattge and Knorr (2007) was originally made due to improved process understanding of acclimation in photosynthesis, but Kattge and Knorr (2007) would only be suitable if ABZ sites had been included in the parameterization. Without those sites, it is an extrapolation of the southern scheme to the ABZ, which we find introduces significant biases.

265 ### 2.4.5 J_{max} and V_{cmax} Winter Default

The LUNA module calculates J_{max} and V_{cmax} dynamically during the growing season only. When plants are dormant in winter (non-growing season), CLM5.0 uses constant values (See equations in the Supplement Section 4).

$$J_{max,t} = \begin{cases} 50, & \text{during winter} \\ J_{max,t-1} + \min[mxcon, J_{max,opt} - J_{max,t-1}] \times J_{max,opt}, & \text{in growing season, if } J_{max,opt} \geq J_{max,t-1} \\ J_{max,t-1} + \max[-mxcon, J_{cmax,opt} - J_{max,t-1}] \times J_{max,opt}, & \text{in growing season, if } J_{max,opt} < J_{max,t-1} \end{cases}$$

270 $$V_{cmax,t} = \begin{cases} 85, & \text{during winter} \\ V_{cmax,t-1} + \min[mxcon, V_{cmax,opt} - V_{cmax,t-1}] \times V_{cmax,opt}, & \text{in growing season, if } V_{cmax,opt} \geq V_{cmax,t-1} \\ V_{cmax,t-1} + \max[-mxcon, V_{cmax,opt} - V_{cmax,t-1}] \times V_{cmax,opt}, & \text{in growing season, if } V_{cmax,opt} < V_{cmax,t-1} \end{cases}$$

$J_{max,opt}$ ($V_{cmax,opt}$) is predicted by LUNA as the optimal J_{max}($V_{cmax,opt}$)for the plant, which as the name states is optimal and does not account for limitations on enzyme resources. The maximum change constraint ($mxcon$) limits the amount of change for J_{max} (and V_{cmax}) based on the resources available to that plant, which can change every time step. This scheme allows J_{max} and V_{cmax} and thus photosynthesis to be co-limited by resources. We find the change constraint to be a reasonable

275 one to place on J_{max} and V_{cmax}, as it allows for the climatic history on the grid cell to influence the future prediction of leaf photosynthetic traits. However, this scheme for J_{max} and V_{cmax} is only active during the growing season. In winter when LAI=0 and plants are dormant, J_{max} and V_{cmax} are not predicted by LUNA and instead are given a default global place holder value. Thus, at the start of the growing season (or first day of spring), $J_{max,t}$ and $V_{cmax,t}$ are directly calculated from the last

~~day of winter~~ a winter constant value:

$$J_{max,\text{last day of winter}} = J_{max,t-1} = 50 \tag{2}$$

$$V_{cmax,\text{last day of winter}} = V_{cmax,t-1} = 85 \tag{3}$$

We find that this global default winter value strongly influences the prediction of J_{max} and V_{cmax} throughout the entire growing season (Supplement Fig. S4). In all of the ABZ PFTs, raising these default values increases mean growing season GPP, whereas decreasing them lowers GPP (Supplement Fig. S4). Furthermore, the constant winter values in Equation ?? represent a high bias globally in V_{cmax} (Lawrence et al., 2019), contributing additional bias. Due to the sensitivity of this choice and in an effort to leverage the physiological history of a given location, we choose to save the average predictions of J_{max} and V_{cmax} from the previous growing season for all PFTs ($J_{max,prevyr}$ and $V_{cmax,prevyr}$). We ~~use these pft specific values to initialize J_{max} and V_{cmax}, such that on the first day of the growing season,~~ preserve the LUNA equations, except these constant values.

$$J_{max,pft,\text{last day of winter}} = J_{max,pft,prevyr} \tag{4}$$

$$V_{cmax,pft,\text{last day of winter}} = V_{cmax,pft,prevyr} \tag{5}$$

~~Then Equations 2 and 3 continue the calculation J_{max} and V_{cmax} for the rest of the growing season.~~

2.4.6 Carbon Allocation

Finally, we investigate the sensitivities of parameters related to carbon allocation, which are relatively uncertain and strongly influence CO_2 fluxes in the ABZ, particularly the stem-to-leaf ratio and the root-to-leaf ratio. In CLM5.0, the parameter defining the root-to-leaf allocation ratio is set at a constant value of 1.5 for all PFTs. This is not an ideal configuration as boreal trees and tundra vegetation are structurally different than other plant types due to the need to cope with colder temperatures, which should be reflected in allocation to their roots, leaves, and other plant components. Even within a PFT, different species have been measured to have drastically different ratios of allocation to roots and leaves (Iversen et al., 2015), but for the purpose of circumpolar simulations, we limit our allocation parameters to the PFT level. Root to leaf ratios have been measured as consistently high for grasses and shrubs, meaning more allocation to roots than leaves (Chapin III, 1980; Iversen et al., 2015), as the large root systems are key to survival of these Arctic species (Archer and Tieszen, 1983). We, therefore, tested a higher root-to-leaf allocation of 2 for shrubs and grasses, which agrees relatively well with observations of tundra vegetation (Buchwal et al., 2013). For boreal trees, below ground allocation in evergreen conifers has been found to be higher than in deciduous trees (Gower et al., 2001; Kajimoto et al., 1999). Lowering the root-to-leaf ratio of DBT better represents the typically shallow root systems of deciduous boreal trees (Kobak et al., 1996), while being consistent with values implemented in other deciduous tree modeling studies (Arora and Boer, 2005, e.g., 0.75). Observations suggest that boreal NETs in general have more extensive root systems than the deciduous trees (Gower et al., 1997), thereby requiring more belowground resources, and that tundra shrubs and grasses allocate even more photosynthate belowground. Thus, observations provide support for NET root-to-leaf allocation to be larger than DBT, and we choose to allow the NET root-to-leaf allocation to remain at the CLM5.0 default value.

Regarding stem allocation, CLM5.0 includes an option for dynamic stem-to-leaf allocation. The ratio is based on NPP and can be used for woody trees and shrubs (Friedlingstein et al., 1999), generally acting to increases woody growth in favorable conditions (Vanninen and Mäkelä, 2005). Though this allocation was previously used in CLM4.5 and turned off for CLM5.0, it is an uncertain choice as noted explicitly in Lawrence et al. (2019). A problem with this allocation scheme was noted in the tropics (Negrón-Juárez et al., 2015), but not in temperate or high latitude climates. We test both options, comparing the static stem-to-leaf ratios used in CLM5.0 to the dynamic allocation option.

3 Results

We investigate the simulation of CO_2 fluxes in the arctic-boreal zone by CLM5.0 using gridded and point simulations. We identify biases in the carbon cycle in Section 3.1, present our mechanistic and additive improvements to phenology and photosynthesis in Section 3.2, and make our model recommendation in Section 3.3. This ABZ analysis focuses primarily on GPP fluxes in CLM5.0, but as expected, our assessment extends to TER and NEE due to the interdependence of these carbon fluxes on productivity (Chen et al., 2015). We assess simulation biases based on the comparison of CLM5.0 output against FLUX-COM and EC towers, as detailed in Section 2.2. We address the biases that arise from using the default CLM5.0 parameters and schemes described in Section 2.4, which can be classified as: (1) phenology onset, (2) phenology offset, (3) daylight scaling, (4) Leuning temperature scaling, (5) initial spring value of J_{max}/V_{cmax}, (6) dynamic stem-to-leaf carbon allocation, (7) realist root-to-leaf carbon allocation.

3.1 Biases in CLM5.0

The latest release of CLM5.0 substantially overestimates summer GPP in the ABZ by $\sim$3 gC m^{-2} day^{-1} or 40% (red line in Figures 1b and 2a). The magnitude of this bias is such that CLM5.0 estimates of GPP for high latitudes tundra vegetation are comparable with the more southern boreal forests (Fig. 1a). This lack of a latitudinal gradient in CLM5.0 is not supported by FLUXCOM and ILAMB benchmarking (Fig. 1f). Though most of the ABZ in CLM5.0 is over productive, we note that there is a large area with $GPP = 0$ in Siberia, indicating that this region is not photosynthesizing in CLM5.0 though it should be. Thus, the simulation of carbon fluxes in CLM5.0 is very heterogeneous with areas that are highly productive and areas that are not-functioning, or "dead-zones".

We next investigate the seasonal cycle of CO_2 fluxes in the ABZ. Across the ABZ, average productivity in the CLM5.0 simulation is high throughout the year compared to FLUXCOM, which is shown throughout the year in Figure 2a. From a seasonal perspective, CLM5.0 vegetation enters dormancy later than observations, as can be seen by the high biases in GPP in fall. The timing of photosynthesis in spring ~~does, however, appear~~ appears accurate when we look at the PFT-aggregated average of CO_2 fluxes. We see a similar high bias in TER in Figure 2b, as respiration is tightly coupled to the high biased GPP. The magnitude of peak summer NEE in CLM5.0 matches observational data better than GPP and TER. However, its seasonal cycle exhibits biases and timing issues related to spring drawdown, summer minimum, and fall peak NEE (Fir. 2c).
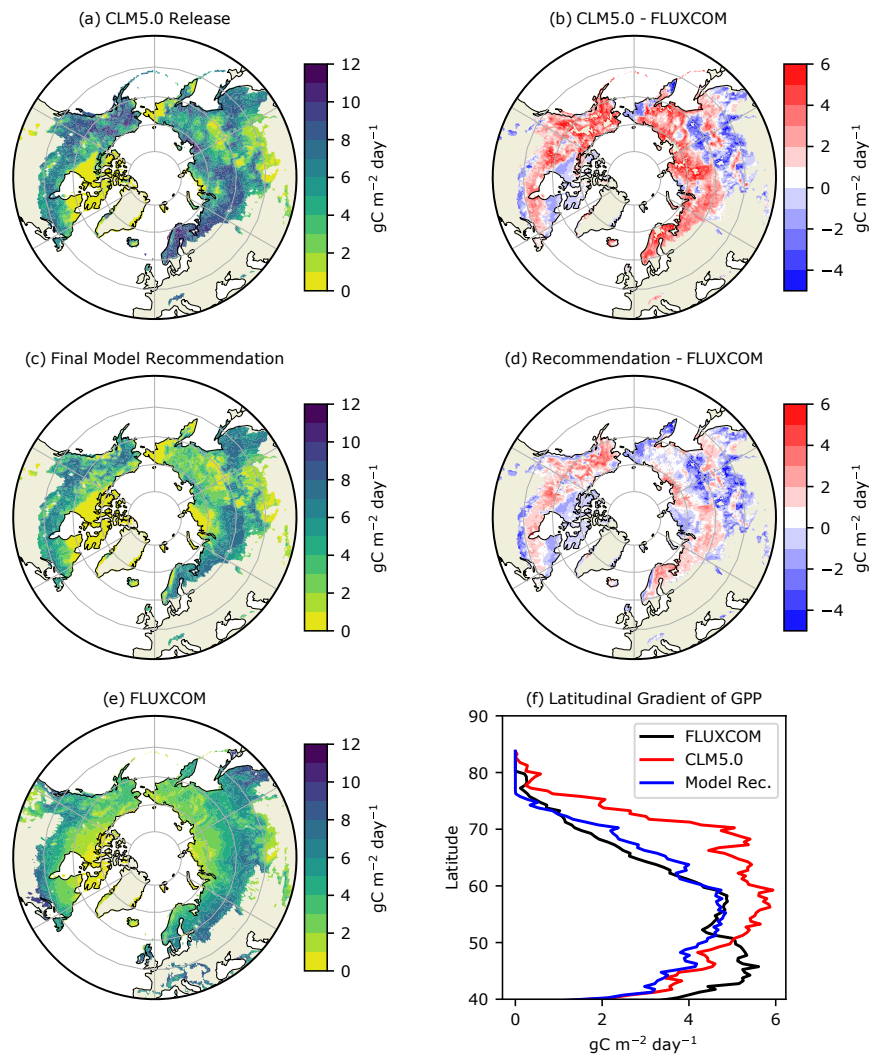
Figure 1. Average Summer (JJA) GPP (gC m^{-2} day $^{-1}$) for (a) CLM5.0 Release, (b) our Model Recommendation, and (c) FLUXCOM, with the difference between model simulation and FLUXCOM in (b) and (d). The latitudinal gradient of summer GPP is depicted in (f).
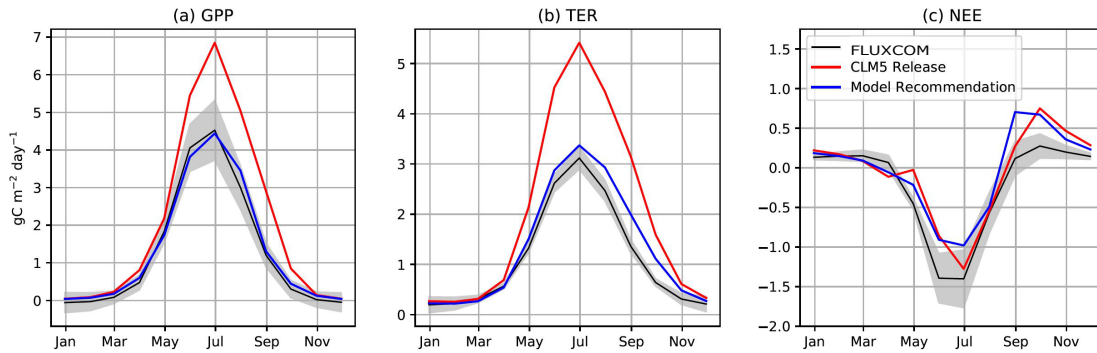
Figure 2. The annual cycle of (a) GPP, (b) TER, and (c) NEE from CLM5.0 and our Model Recommendation compared to FLUXCOM in the ABZ. Non-productive grid cells in the ABZ are removed from the average, meaning where LAI=0, which is standard procedure in CLM analysis (Lawrence et al., 2019).

Assessing the seasonality of CO_2 fluxes in the ABZ using the PFT-specific output of CLM5.0 reveals biases in phenology that are hidden when PFTs are aggregated together in a grid cell. In terms of phenology, we find that NET begins significant photosynthesis in February in CLM5.0 when air temperatures are well below freezing. This onset of NET photosynthesis is considerably early according to both FLUXCOM (Figure 3b) and our understanding of available liquid water for photosynthesis (Goulden et al., 1998). The peak productivity in NET occurs in June, instead of July as seen in observational data. In contrast, the onset of photosynthesis for deciduous trees and shrubs is consistently late (Fig. 3). The gridded CLM5.0 GPP output hides these biases, showing that on average onset in CLM5.0 matches observations (Fig. 1a). In contrast, the high bias of CLM5.0 productivity during late autumn was easily seen in the gridded CLM5.0 output, and we confirm that this bias is due to the shifted seasonal cycle of deciduous PFTs (Fig. 3).

Regarding the magnitude of photosynthesis, the CLM5.0 PFT-specific output indicates that both shrubs and grasses have a large high bias of GPP compared to observational data by a factor of 2-3 (Fig. 3a). We confirm that the tundra grass and shrub PFT specific output is often greater than or as productive as the boreal trees in CLM5.0 (Fig. 3 e,f vs. b,c,d). The deciduous boreal trees (NDT and BDT) have a low growing season GPP bias, while NET have a high productivity bias. By examining the spatial pattern of the average summer GPP (Fig. 1), one can see that there are many areas where $GPP = 0$ for many consecutive years, indicating that the PFT did not survive. In Siberia, a prominent "dead zone" occurs in what should be highly productive deciduous needleleaf larch forests. Smaller "dead zone" areas are present for all other PFTs across the ABZ (Supplement Fig. S5).

As with the aggregated CLM5.0 output, the PFT specific biases in TER are similar to those noted for GPP (Supplement Fig. S6). PFT-specific patterns in NEE also tend to follow the biases in GPP and TER, with the notable exception of spring in deciduous vegetation. For these PFTs, there is a large spike of CO_2 release to the atmosphere, up to 0.5 gC m^{-2} day^{-1} between April and May that does not match observations (Supplement Fig. S7). This is primarily due to the late onset photosynthesis
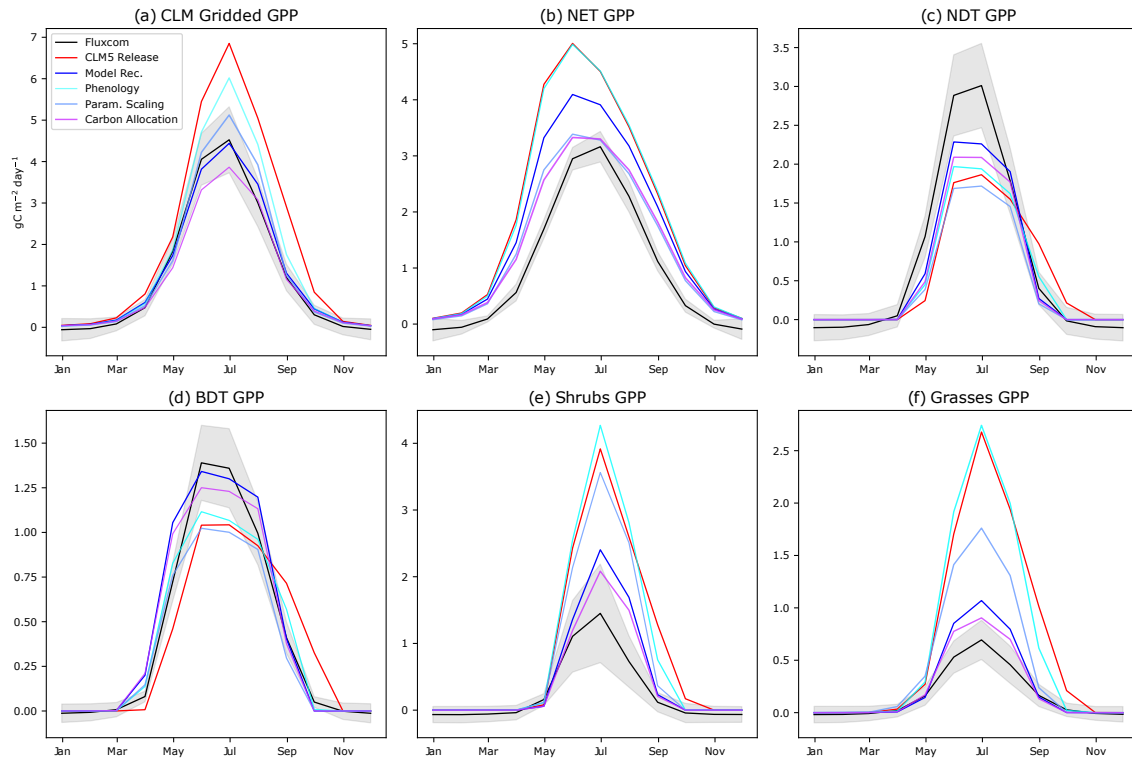
Figure 3. The annual cycle of GPP for CLM5.0 release with intermediate model development steps compared to FLUXCOM for: (a) aggregated gridded CLM5.0 output, (b) NET, (c) NDT, (d) BDT, (e) shrubs, (f) grasses. Phenology incorporates changes to onset/offset. ~~Temp~~Param. Scaling adds our scaling of J_{max} and V_{cmax} by daylight and Leuning. Carbon allocation adds changes to the root-to-leaf and stem-to-leaf parameters. The final model recommendation incorporates the bug fix (Supplement Section 3) and spring initialization of J_{max} and V_{cmax}.

at a time when TER is increasing due to warming air and soil temperatures. We note that the net balance of BDTs is near 0, rather than a sink, which agrees with our conclusions that deciduous trees are not productive enough relative to FLUXCOM and flux towers. Ultimately, the timing and magnitude of biases in TER and NEE confirm our need to focus on GPP at a first step in better representing seasonal CO2 fluxes in CLM5.0 for the ABZ.

Benchmarking CLM5.0, yields the following primary issues for the simulation of GPP across the ABZ:

1. The onset of GPP in deciduous PFTs in spring is consistently late across all PFTs.

2. The fall senescence of GPP is consistently late.

3. There is no latitudinal gradient in summer GPP, due in part to the high GPP bias in tundra grasses and shrubs.

4. NET begin photosynthesis early in winter and reach peak in productivity in June, instead of July.

5. NET have a high GPP bias throughout the growing season.

**14**

6. Deciduous trees (BDT and NDT) have a low GPP bias.

7. There are large areas of PFTs that have no productivity at all or are effectively dead in CLM5.0, particularly the NDT in
   Siberiak, representing larch forests (larix spp.).

3.2 Model Development at Flux Towers

We confirm these biases in the ABZ by comparing the CLM5.0 PFT-specific output to representative flux towers across the ABZ. For example, the southern boreal mixed forest site at CA-QC2 (Figure 4a) contains NET, BDT, and shrub PFTs. The BDT here is a dead zone, where $GPP = 0$, making it a useful site to understand what model thresholds may be influencing one PFT to die out in a simulation, while others remain productive. We also include a BDT site at CA-OAS (Figure 4b) to further investigate the simulation of deciduous trees at an alive grid cell in CLM5.0. The grasses and shrubs at Eight Mile Lake, Alaska are approximately five times more productive than observations (Figure 4c). In contrast, at the larch forest at RU-SKP (Figure 4d), we find that onset is late and GPP is roughly half the observed value, consistent with the low bias in NDT gridded output. We perform model development by examining each issue described in Section 2.4 sequentially on our model development sites: CA-QC2, CA-OAS, US-EML, and RU-SKP. Each flux tower measurement is carefully chosen to cover all CLM5.0 PFTs to understand their complex impacts on GPP across all PFTs, while also including a "dead-zone" to tease out a different kind of bias. Model development follows the procedure and biases described in Section 2; we begin with phenology, move on to the photosynthesis schemes, and end with adjusting the carbon allocation parameters. The use of flux towers and point simulations allow us to test a range of hypotheses for each model development objective. Successful model development is achieved when GPP is within standard deviations limits. We choose to make the model development additive because the improvements are generally small and justified observationally, but all together make for a substantially improved simulation of ABZ carbon fluxes (Fig. 3).

Regarding phenology, ~~high-latitude deciduous vegetation in CLM5.0 initiates photosynthesis too late in the spring. The parameterization from White et al. (1997) did not allow for ABZ vegetation to pass a growing degree day threshold until roughly May or later, as it was developed for temperate ecosystems and extrapolated to the ABZ. We find that~~ using new thermal and moisture thresholds for onset~~causes~~, the deciduous plants to begin photosynthesis earlier in the season, which more closely matches observations (Fig.3). ~~We also generate a more accurate fall simulation by replacing the 11 hour global offset photoperiod with a latitudinal gradient threshold, ranging from 15 hours to 11 hours. The timing of photosynthesis is critical to understanding changes in the carbon cycle in the ABZ, and we find these mechanistic schemes offer a better ABZ simulation of phenology (Fig.3).~~ However, the bias in the magnitude of photosynthesis is not improved by our phenology changes; In fact, the productivity of grasses and shrubs increases further with a growing season that begins earlier in Spring (Fig. 3e,f, comparing the red "CLM5.0 Release" line to the cyan "Phenology" line).

~~We test a variety of changes to CLM5.0 photosynthetic schemes and parameters as detailed in Section 2.4. Regarding~~ Next, regarding J_{max}, we find that modifying the function that scales J_{max} to accurately use the maximum number of daylight hours on each grid cell (Bauerle et al., 2012) decreases productivity across all PFTs (Figure 4 and Supplement Fig. S8). In particular,
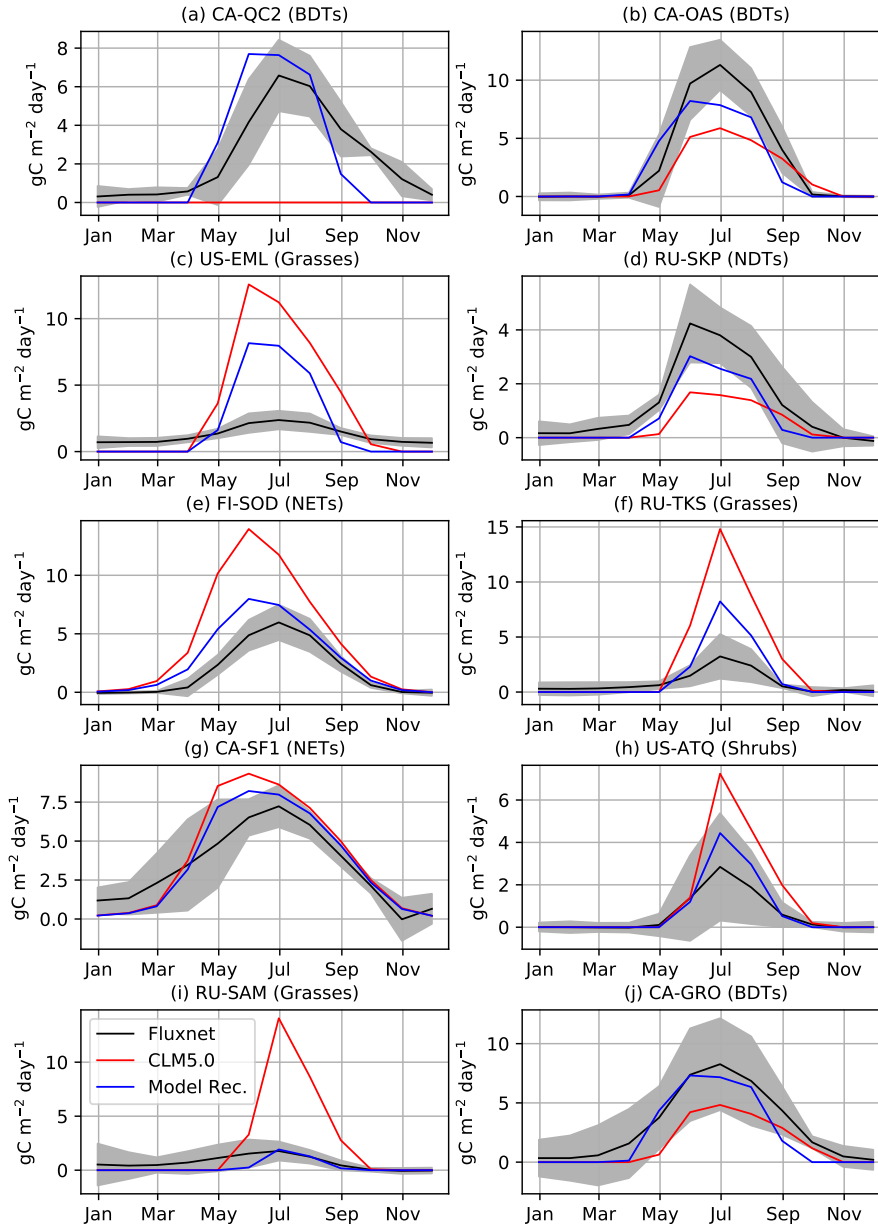
Figure 4. Comparison of GPP at Flux Tower sites to the CLM5.0 release and the results of our model development. We performed model development at CA-QC2 (mixed forest site in Quebec), CA-OAS (Aspen forest in Saskatchewan), US-EML (Eight Mile Lake Tundra location), and RU-SKP (Yakutsk Spasskaya Larch Forest). We did additional comparisons with FI-SOD (Sodankyla, Finland), RU-TKS (Tiksi grasslands), CA-SF1 (Saskatchewan boreal forest), RU-COK (Chokurdakh Tundra shrubs), RU-SAM (Samoylov grasslands), and CA-GRO (Groundhog River Boreal Forest).

16

we decrease the June spike in GPP for NET because the daylight fraction around the summer solstice is no longer greater than 1. Overall, this modification decreases the high bias in ABZ GPP by 2 gC/m^2 in the summer (Supplement Fig. S8) and generates a latitudinal gradient in the PFT specific output of trees (Supplement Fig. S5). By reverting the CLM5.0 temperature scaling scheme from (Kattge and Knorr, 2007) to Leuning (2002) as used on CLM4.5, we find that GPP is decreased for all PFTs, especially in spring when the photosynthesis ramp up had been artificially high. Our model updates improve phenology of NETs due to more realistic scaling of J$_{max}$ and V$_{cmax}$ (Fig. 3b, the light blue "Temp Scaling" simulation and Fig. S13). The GPP of grasses is also decreased (Fig. S14), but shrubs are still biased high.

After decreasing productivity of all PFTs (Fig. 3, compare the red "CLM5.0 Release" with the blue "~~Temp~~Param. Scaling"), tundra shrub and grass productivity in CLM5.0 retain a substantial high GPP bias (Fig. 3e,f), while the deciduous tree PFTs have a low GPP bias (Fig. 3c,d), which may be related to non-optimized ABZ carbon allocation parameters. Allowing for a dynamic stem-to-leaf allocation improves both the timing and magnitude of photosynthesis (Supplement Fig. S9). Additionally, we make PFT specific alterations to root-to-leaf ratios based on our findings from Section 2.4. As a result, GPP is lowered in grasses and shrubs and increased for deciduous trees (NDT and BDT), approaching FLUXCOM values (Fig. 3). The rest of our changes to the model involve schemes in LUNA that we believe are initialized incorrectly for the ABZ, but also the rest of the world. These recommended model updates include initializing the winter default values of J$_{max}$ and V$_{cmax}$ using the mean value for a given grid cell during the previous growing season (Equations 4,5 and Supplement Fig. S4) and a model error related to the calculation of 10-day leaf temperature (Supplemental Section S3).

3.3 Improved Carbon Fluxes in the Model Recommendation

Based on our changes to phenology, photosynthesis, and carbon allocation schemes, we recommend the following mechanistically based changes to CLM5.0 for a considerably improved representation of CO2 fluxes in the ABZ:

1. GPP onset is based on soil temperature, air temperature, and snow cover.

2. GPP offset is based on a latitudinal photoperiod gradient such that the high Arctic begins senescence earlier.

3. J$_{max}$ is scaled by maximum day length instead of a constant 12 hours.

4. J$_{max}$ and V$_{cmax}$ are scaled by temperature response functions parameterized by Leuning (2002).

5. J$_{max}$ and V$_{cmax}$ have initial values in spring that are based on the LUNA predictions from the previous growing season, allowing the values to vary across PFTs, time, and space.

6. The stem-to-leaf carbon allocation ratio for trees and shrubs is allowed to be dynamic throughout the season.

7. Observationally based root-to-leaf carbon allocation ratios are used. For deciduous tree PFTs, the root-to-leaf ratio is decreased, while the ratio for shrubs and grasses is increased.

We first validate that our modifications to CLM5.0 offer improvements to simulations of CO$_2$ fluxes when looking at specific flux towers (Fig. 4). For instance, productivity is increased at BDT at CA-OAS, which is further validated at CA-GRO. The

17

"dead zone" at CA-QC2 is now highly productive and much closer to observed carbon fluxes in terms of seasonality and magnitude. The new cold deciduous onset scheme causes this improvement, as the previous growing degree day scheme in CLM5.0 prevented photosynthesis from ever starting. Photosynthesis in NDT is increased due to our model development, but as shown in both gridded output and RU-SKP, the NDT photosynthesis needs to increase further. The phenology and magnitude of NET is much improved at validation sites. However, as is expected (Schaefer et al., 2012), not all of the flux tower sites have CO_2 fluxes that are reproduced within the range of observational uncertainty (Fig. 4c and Supplement Fig. S11). For example, although GPP at US-EML is reduced due to our model development, it is still biased high. However, the grasses and shrubs at other validation sites are much improved compared to flux measurements, like the grasses and shrubs at RU-SAM, RU-TKS, and US-ATQ.

We next compare our model improvements in a gridded simulation against CO_2 fluxes from FLUXCOM (Figure 2a). The high productivity bias at high latitudes is substantially reduced due to our model development by decreasing the productivity of ABZ shrubs and grasses. Our Model Recommendation for CLM5.0 produces a latitudinal gradient (Fig. 1f) in productivity, with the tundra no longer being as productive as the boreal forest (Fig. 3). The timing of photosynthesis is also improved as dormancy is reached by October in most PFTs, instead of November (Figs. 1 and 3). In examining the PFT specific output (Fig. 3), we confirm an earlier beginning to spring photosynthesis in deciduous trees, shrubs, and grasses. In contrast, our model development successfully delays the onset of productivity in NET, due to a combination of daylight scaling and the temperature scaling from Leuning (2002). As for the magnitude of photosynthesis, we find that NET photosynthesis is reduced across the ABZ, while the deciduous boreal tree PFTs experience an increase in productivity (Supplement Fig. S10).

Our modifications to CLM5.0 offer similar improvements to simulations of CO_2 fluxes when looking at specific flux towers (Fig. 4), such as the increase in productivity of BDT at CA-OAS, which is further validated at CA-GRO. Photosynthesis in NDT is increased due to our model development, but as shown in both gridded output and RU-SKP, the NDT photosynthesis needs to increase further. The phenology and magnitude of NET is much improved at validation sites. However, as is expected (Schaefer et al., 2012), not all of the flux tower sites have CO_2 fluxes that are reproduced within the range of observational uncertainty (Fig. 4c and Supplement Fig. S11). For example, although GPP at US-EML is reduced due to our model development, it is still biased high. However, the grasses and shrubs at other validation sites are much closer to flux measurements, like the grasses and shrubs at RU-SAM, RU-TKS, and US-ATQ.

Although GPP is our main focus of development due to the number of productivity biases, we find improvements in other component CO_2 fluxes. Changes in TER generally follow those of GPP, and our modifications substantially reduce the high summer respiration bias in CLM5.0 (Fig. 2), which is mostly due to the reduction of GPP and TER in grasses and shrubs at high latitudes(Supplement Fig. S6). The respiration of deciduous tree PFTs did not change substantially, but TER decreases in NET due to the decrease in GPP. Due to the improvements in GPP onset/offset, the model no longer simulates large net CO_2 emissions in spring (Supplement Fig. S7), which better matches observations . However, in of NEE. In many PFTs the respiration in fall is and thus NEE in fall are high in the gridded output. Although this does not match FLUXCOM, it is consistent with recent measurements of high fall respiration in the ABZ (Commane et al., 2017; Natali et al., 2019). The net

balance of carbon fluxes does not change substantially due to our model development. We find that the carbon sink in ABZ CLM5.0 decreases by about 12%, which is due to a smaller net sink in the summer and larger carbon release in fall.

3.4 Validation with ILAMB
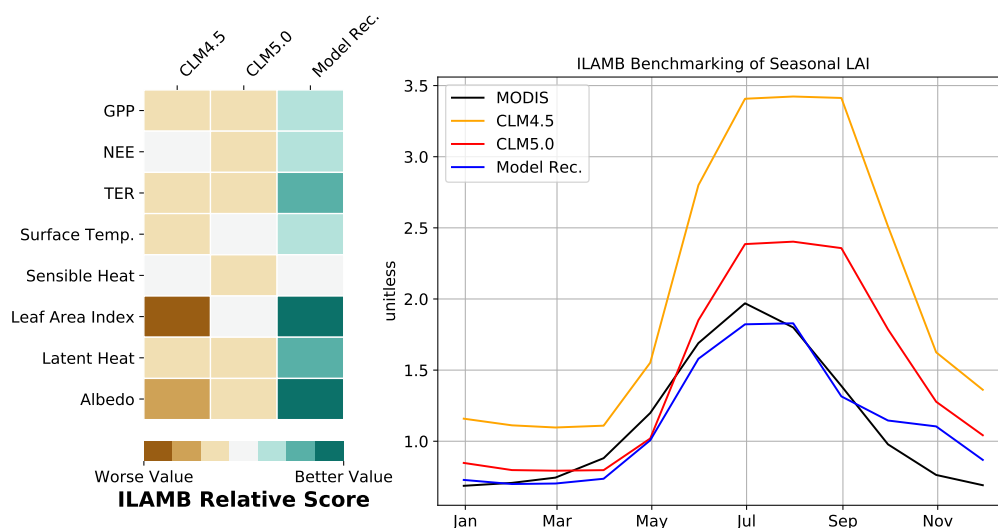
Figure 5. ILAMB benchmark score of CLM5.0 Model Development Recommendation against previous versions of CLM: 5.0 and 4.5 (left). LAI seasonal results compared to MODIS (right).

Comparing output from CLM5.0 and our Model Recommendation to observational data provided by the ILAMB framework confirms broad improvements in model fidelity (Fig. 5). This includes the CO_2 fluxes we focus development on, as well as surface fluxes (sensible heat, latent heat, albedo). ILAMB confirms that the high GPP bias and late phenology biases are reduced. LAI, in particular, has been improved greatly in the Arctic in regards to both timing and magnitude (Fig. 5). According to the ILAMB score, the implemented changes are not detrimental to any other essential land surface variables, and in fact improve their simulation according to the centralized benchmark scores. The Breaking down the overall benchmark score, our Model Recommendation has large improvements in the spatial distribution and interannual variability scores. The seasonal cycle score of GPP did not improve substantially, which make sense due to the phenology problems only becoming apparent in the PFT specific output, which is not a standard ILAMB benchmark. The NEE seasonal cycle is substantially improved according to ILAMB, which agrees with our FLUXCOM validation. The relative improvements to moisture and heat fluxes are particularly noteworthy, as these changes can feed back to the regional climate system.

4 Discussion

~~We have approached model development for CO_2 cycling in the ABZ in a mechanistic and targeted fashion, leveraging available observational data and derived products. We find that many biases are interconnected~~

ILAMB Benchmarking of Global GPP
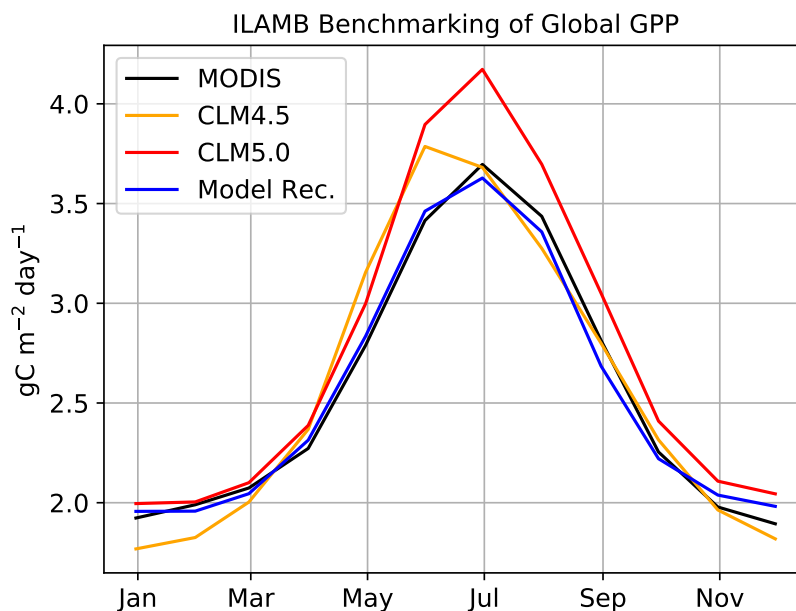
Figure 6. Seasonal Global GPP Benchmakred in ILAMB. Our model recommendation for the ABZ is applied globally and benchmarked for GPP in CLM4.5 and CLM5.0.

We are also interested in contributing to the improvement of the global CLM5.0 simulation. We confirm that a global simulation is possible and reasonable (Fig. 6) with these additions to the code base. All but two of our model improvements are limited to the ABZ, meaning that ~~mechanistic model developmentand bug fixes can improve part of the simulation, while making other aspects worse (Fig. 3). Overall, our work with~~ we do not expect significant biases to emerge at lower latitudes due to our model development, and our ILAMB validation confirms this. The constant LUNA equation is one of the global changes and the other is the change in temperature scaling from Kattge and Knorr (2007) to Leuning (2002). Additional testing at lower latitudes, which is outside of the scope of this study, is necessary to determine the effects on the global carbon budget.

4 Discussion

Through mechanistic model development, we have reduced the biases in carbon cycling in CLM5.0 ~~in the ABZ highlights the importance of regional model analysis and development~~ for the Arctic-Boreal PFTs. Many of our recommendations in Section 2.4 affect several of the biases noted in Section 3.1, indicative of the many interconnected schemes in CLM5.0. Ultimately, we believe our modifications to be reasonable, based in observations, and a step towards a more accurate simulation of carbon cycling in high latitude terrestrial ecosystems, but we also discuss the limitations of our model development choices and identify avenues of research that could continue to improve CLM5.0. ~~We find the extrapolation of model schemes developed for temperate latitudes to the high latitudes to be the root of many biases, as has also been noted elsewhere (Rogers et al., 2017).~~

Having more accurate phenology in CLM5.0 is critical for understanding the recent and future changes in biogeochemistry in the ABZ, as global change drivers during the shoulder seasons may be drivers of carbon cycle changes. ~~The growing degree threshold in the CLM5.0 release is not met until late spring for deciduous plants (Figure 3), consistently delaying the spring onset. Using air and soil temperatures combined with snow depth as a metric for spring onset more accurately captures spring productivity timing in deciduous trees , shrubs and grasses across the ABZ (Fig. 4). We find that the freeze-thaw transition simulated by CLM5.0 matches the timing of photosynthesis across vegetation types and latitudes from other studies (Botta et al., 2000; White et al., 2009) as well. Our new metric thresholds account for the key environmental conditions that signal the end of winter, as deciduous trees~~ Deciduous trees have protective mechanism to avoid the onset of growth when there is a strong probability of cold snaps (McMILLAN et al., 2008). ~~By combining a snow threshold with our air and soil temperatures metrics, we prevent the more southern boreal forest from beginning photosynthesis too early (model sensitivity tests not pictured). This new onset scheme in our model recommendation also causes some areas with dead PFTs to begin photosynthesizing. The previous~~, and our new thresholds are key environmental conditions designed to mimic the end of winter signals. The use of threshold values in onset schemes is a well established phenology method (Jolly et al., 2005; Arora and Boer, 2005) that is used in other models, like LPJ (Forkel et al., 2014) and the Canadian Terrestrial Ecosystem Model. These schemes can be fairly simple or more complicated, much like growing degree day ~~scheme in CLM5.0 required soil temperatures to sum to a threshold before the summer solstice, meaning cold soil temperatures prevented photosynthesis from ever starting, causing some of the "dead" PFTs~~ models (GDD). We choose to focus on environmental thresholds due to the uniqueness of the ABZ environment, which is dominating by freeze-thaw dynamics, making a threshold approach effective. We did investigate other GDD schemes, but we found that many GDD models perform similarly (Hufkens et al., 2018) and in the ABZ have a documented late onset (Botta et al., 2000; Fu et al., 2014), just like the current scheme in CLM5.0. ~~The deciduous (needleleaf and broadleaf) trees in CLM5.0 become more productive in summer as their growing season is allowed to begin earlier, which is desirable due to their low GPP bias. The grasses and shrubs increase further in productivity after more mechanistic phenology drivers are applied, meaning the previous phenology parameterization is not the cause of their high bias~~We did not identify a scheme well validated in the ABZ, as most observational validation studies are focused on the lower latitudes, and those studies also identify that gridded phenology products tend to be produced using optical imagery, which often does

21

530   not correspond well with CO2 fluxes in shoulder seasons (Fisher et al., 2007; Parazoo et al., 2018a). Ultimately, using another GDD scheme from the known literature would be trading one extrapolated temperate scheme for another. This threshold approach may have limitations for a warming world, which is why an ABZ focused phenology study, using novel datasets like PhenoCam (Richardson et al., 2018; Hufkens et al., 2018), have potential to uncover more complicated vegetations processes and filling in gaps from satellite based studies (Fisher et al., 2007).

535   ~~We implement~~ Implementing an offset scheme with a latitudinal gradient in CLM5.0 ~~, where the offset day length decreases linearly from15 hours at 65° N to 11 hours at 45° N. This~~ is a first step towards more realistic timing of fall senescence in CLM5.0, and additional work is needed to understand how temperature and other climate drivers impact the timing of dormancy. Studies are divided on the issue of whether temperature or photoperiod are driving offset in the ABZ. Field experiments have found photoperiod to be a likely driver at high latitudes (Eitel et al., 2019), but temperature and even precipitation may be

540 more important at lower latitudes. Experiments of high arctic tundra have shown that senescence can be delayed by up to 15 days due to warming (Marchand et al., 2004). Given this uncertainty, we suggest that photoperiod is a sufficient and justified approximation for fall dormancy, until better mechanistic relationships can be derived from observational data. The late bias in fall phenology has also been noted in other TBMs (Richardson et al., 2012), also likely due to extrapolation of temperate schemes to the high latitudes. Thus, our simple daylight threshold here could be applicable to other models.

545   One of the most impactful changes to CLM5.0 GPP is the use of maximum day length to scale J_{max}, rather than 12 hours. Equation 1 would previously scale the prediction of J_{max} high, particularly in June with the summer solstice and day length at its maximum. The previous use of 12 hours for maximum day length is likely a holdover from using LUNA in the tropics, as (Bauerle et al., 2012) was the cited basis for Equation 1, which used maximum day length appropriately. After we fix this substantial scaling bias, we find other algorithms and parameterizations are more sensitive to model changes. This opens up

550 many avenues of model development, some of which we are able to accomplish in this study, like more realistic allocation parameters and scaling changes to J_{max} and V_{cmax}. For future steps, we argue that a re-parameterization CLM is necessary, as previous model tunings attempted to bring down the high GPP bias through parameter choices rather than this bug fix.

  As with the length of day scaling described above, the photosynthetic temperature acclimation in scaling of J_{max} and V_{cmax} were not created for ABZ latitudes and tend to exacerbate model biases. We, therefore, recommend not using these functions

555 from (Kattge and Knorr, 2007), and stress the need for further research on photosynthetic temperature acclimation in the ABZ, especially for projecting responses to future climate. The current implementation generates unrealistic seasonal temperature response functions for GPP resulting in model biases (Smith et al., 2017). ~~The temperature acclimation equations are strictly only valid between 11° and 35° C (Kattge and Knorr, 2007), and below that value, J_{max} and V_{cmax} have a scaling discontinuity that biases GPP (Supplement Fig. S3).~~ Previous work by Rogers et al. (2017) advocated for removing ~~this~~ the 11° C limit from

560 (Kattge and Knorr, 2007), but our tests of this did not decrease productivity of the grasses and shrubs at high latitudes (not shown) or offer any other improvements, at least not without a re-parameterization of the acclimation scheme. We do advise caution in using Recently, Kumarathunge et al. (2019) have created an acclimation parameterization that included ground measurement sites from Utqiagvik (Barrow), Alaska and Finland. Though most of Canada and Siberia are not represented in the parameterization, including this recent observationally-based acclimation function in CLM5.0 is a logical next step for a

565 better ABZ simulation that includes temperature acclimation, ~~critical for simulating GPP under warming climate conditions~~ a critical process for projecting carbon budges into the future.

The initialization of J_{max} and V_{cmax} in spring is a highly sensitive choice (Supplement Fig. S4). We find improvements in GPP when using the mean predicted value of J_{max} and V_{cmax} from the previous year as an initial spring value. The default values for spring J_{max} and V_{cmax} in CLM5.0 are high for ABZ PFTs (Lawrence et al., 2019). Now with a default value that

570 considers the PFT and climatological conditions of J_{max} and V_{cmax}, the simulated seasonal cycle of J_{max} and V_{cmax} mimic the timing of GPP more closely. We observe only very small temporal fluctuation in these average values of J_{max} and V_{cmax}, indicating that the LUNA predictions are relatively stable for each geographic-climatological region. The productivity for most ABZ PFTs is decreased throughout the whole summer due to this change in CLM5.0. Though we note that LUNA simulates lower values of J_{max} and V_{cmax} in most Arctic-boreal PFTs, this implementation of average J_{max} and V_{cmax} allows for

575 spatial and temporal variability, including increases of J_{max} and V_{cmax} in some locations. Therefore, there is potential for this scheme to improve future predictions, as the sensitive initialized spring values can now adjust in a warming climate.

Finally, we examine the carbon allocation parameters in CLM5.0, which are a considerable and long-standing source of uncertainty in TBMs. Franklin et al. (2012) appropriately called carbon allocation the "Achilles' heel of most forest models". We argue that CLM5.0 does not use ideal carbon allocation values for the ABZ, but there are multiple diverging development

580 paths for carbon allocation (Fisher et al., 2019) that could lead to a better simulation. During model development and testing of carbon allocation ratios for roots, leaves, and stems, we attempt to balance static and dynamic allocation schemes, but acknowledge there is always room for further development and improvement. ~~Our results suggest that altering carbon allocation ratios, particularly stem-to-leaf and root-to-leaf, to values more consistent with observations improves CO_2 fluxes across the ABZ .~~ The default parameters in CLM5.0 do not utilize dynamic stem-to-leaf carbon allocation due to exponential increases in

585 biomass in the tropics (Negrón-Juárez et al., 2015), but not necessarily for the ABZ. We find that simulations of productivity are improved across all ABZ PFTs when this ratio is dynamic. In the model, the previous year's NPP is used to set this ratio, based on the assumption that NPP can serve as a proxy for environmental conditions that send resources to either roots or leaves, thereby increasing woody allocation in favorable growth environments (Vanninen and Mäkelä, 2005). In the previous version of CLM, CLM4.5, dynamic stem-to-leaf allocation generally lowered carbon allocation, sending more carbon to leaves

590 than stems, agreeing with observational comparisons (Montané et al., 2017), which is the same pattern we see in ABZ PFTs in CLM5.0.

The final improvements to the deciduous tree PFTs, grasses, and shrubs come from changing the root-to-leaf ratio, which is currently the same constant value for all PFTs ~~. Increasing allocation to roots for the grasses and shrubs was advocated in the CLM5.0 documentation for lower latitudes and field observations agree with this choice for the high latitudes (Iversen et al., 2015).~~

595 ~~Observations also support decreasing the allocation to roots and increasing allocation to stems for deciduous trees in the ABZ (Gower et al., 2001). Ultimately, these~~ and does not make sense ecologically (Iversen et al., 2015). We test a range of allocation values for each PFT, using observational data to guide our decisions. These static carbon allocation ratios are our best ~~estimation for~~ estimate for mimicking emergent plant processes, ~~but~~ and we find that the allocation parameters relative to other PFTs may be more important for model simulations than matching observations exactly. The success of focusing on

600 allocation parameters relative to other PFTs may be of guidance to other model schemes. The NDT allocation that we find does align with observations, is used in other models (Arora and Boer, 2005). The grass allocation parameter is low compared to some observations, but we find increasing the root-to-leaf ratio for grass PFTs in CLM5.0 only succeeds in killing productivity completely. We did consider changing the allocation of NETs, but had little success in finding parameters that made sense (Fig. S14). With the difficulty in relating allocation parameters to model counterparts, a more dynamic scheme may be needed for

605 future model versions. For instance, the carbon allocation to roots may saturate when LAI=1 in Arctic shrubs (Sloan et al., 2013). Allowing TBMs to adjust these carbon ratios based on LAI, may be the next step in CLM5.0 model development, approximating the realistic behavior of plant species.

The "dead zones" in CLM5.0 are caused by a combination of issues. ~~For instance, modified onset and offset formulations allowed for some "dead zones" to become productive in our simulations, due to a previously long growing season using~~

610 ~~all available resources and/or the onset threshold not being reached before the summer solstice. Other non-productive areas were improved by changing the deciduous tree PFT allocation parametersto send more resources to leaves, agreeing with observations and other models (Arora and Boer, 2005),~~ such as late onset and allocation parameters. However, these increases in productivity do not clearly and substantially ~~improved~~ improve the dead PFTs in CLM5.0 as a whole, particularly the non-productive area of Siberia (Fig. 1c). These areas of non-productivity are particularly problematic for future CLM5.0 simulations

615 due to the suspected importance of Siberian CO_2 fluxes for current and future seasonal carbon balance (Zimov et al., 1996, 1999; Lin et al., 2020). Although we did not cause any additional areas to die, we also did not succeed in increasing productivity in the larch forests of Siberia. It is also worth noting that areas with "dead zones" clearly visible in the gridded product have all PFTs dead, not just NDT. We hypothesize that there may be thresholds for climatic drivers that inhibit photosynthesis. For instance, there is a minimum relative humidity threshold for nitrogen allocation in LUNA. This threshold appears to be

620 somewhat arbitrary and the ABZ often experiences a dry continental climate. The larch forests of Siberia could benefit from a lower relative humidity threshold to raise their low productivity and potentially improve "dead zones". A test of this hypothesis does raise productivity for deciduous tree PFTs, but the "dead zones" do not become productive. We require more information on relative humidity and nitrogen allocation at low temperatures to determine if this parameter should be changed in the model. Here we focus on mechanistic model development, but a parameter specific to ABZ vegetation may lead to further

625 improvements.

The relative improvements to each PFT are also significant. For instance, as noted previously, our model recommendation generates a latitudinal gradient in CLM5.0 (Fig. 1f), which is due to tundra vegetation being less productive. The Arctic grasses and shrubs are now less productive than the boreal trees, which is consistent with observations. The deciduous tree PFTs in our model recommendation are more productive than the CLM5.0 Release, with NDT productivity increasing by 20% and BDT

630 increasing by 50% (not shown). Additional work is needed for the deciduous schemes, as deciduous trees are observed to be more productive than evergreen trees (McMILLAN et al., 2008). We did decrease the high bias of productivity in NETs by 20%, but NET are still causing a high bias in the simulation of GPP. The evergreen scheme has long been noted to be relatively simple in CLM5.0 (Lawrence et al., 2019) and a key next step for model development.

As with most studies of the high latitudes, we are limited by the availability of ground observations (Arora and Boer, 2005; Richardson et

635 For instance, we only have one flux tower corresponding to NDT (RU-SKP), which is also unfortunately a PFT in CLM5.0 that needs substantial work, particularly in the "dead-zone" in Siberia (Supplement Fig. S5). For other PFTs we restrict our tower choices to have the best data available. We only include EC flux observations where the vegetation classes corresponded with the CLM5.0 PFTs that occur within the ABZ. Metadata indicating a mixed forest or a combination of short stature vegetation and trees causes comparisons to not be clean (Supplement Fig. S11), meaning we withhold these kind of sites for extra vali-

640 dation, but they . These flux towers are not ideal sites for comparison with CLM5.0model development, but they do generally confirm a reduction in GPP due to our model development efforts. The EC record must also span over at least three consecutive years before 2014 due to the forcing dataset used by CLM ending in 2014, which does not allow us to leverage the most recent ABZ ground observations. These criteria restrict our point-based analysis to only a few sites (Supplement Table S1), which does further limit our ability to make statistically robust conclusions. We attempt to mitigate this lack of data by separating the

645 flux towers into model development sites and validation sites, while also including gridded simulationsin the validation of our validating carbon flux results with point simulations and gridded simulations. Then, we use ILAMB to confirm improvements in other independent variables that were no the focus of model development. More comprehensive model improvements for the ABZ may be possible through an increase in the availability and spatial representativeness of tower EC data. We look forward to the increased emphasis on data archiving, standardization, and synthesis, as well as more detailed examination of functional

650 relationships and PFT-specific parameters. As our understanding evolves, observational networks improve, and long-term data archives grow, we stress the need for continued development and model fidelity for high latitude terrestrial ecosystems given their importance in the global climate system.

5 Conclusions

Through mechanistic model development , we have reduced the biases in carbon cycling in We have approached model

655 development for CO$_2$ cycling in the ABZ in a mechanistic and targeted fashion, leveraging available observational data and derived products. We find that many biases are interconnected, meaning that mechanistic model development and bug fixes can improve part of the simulation, while making other aspects worse (Fig.3). Overall, our work with CLM5.0 for the Arctic-Boreal PFTs. Many of our recommendations in Section 2.4 affect several of the biases noted in Section 3.1, indicative of the many interconnected schemes in CLM5.0.Ultimately, we believe our modifications to be reasonable, based in observations, and a

660 step towards a more accurate simulation of carbon cycling in high latitude terrestrial ecosystems.0 in the ABZ highlights the importance of regional model analysis and development. We find the extrapolation of model schemes developed for temperate latitudes to the high latitudes to be the root of many biases, as has also been noted elsewhere (Rogers et al., 2017).

We find that a physically-based phenology formulation using soil temperatures, air temperature, and snow depth is more accurate than the existing parameterization developed for temperate latitudes. Allowing offset timing to vary with latitude

665 instead of a single global value improved circumpolar leaf offset. We improve the scaling of day light in the maximum rate of electron transport (J_{max}) using the maximum day length. Additionally, we remove a global initialization of the maximum rate of

electron transport (J_{max}) and the maximum rate of carboxylation (V_{cmax}) that biases prediction of these critical photosynthesis components each spring. We also recommend the Leuning (2002) temperature scaling of J_{max} and V_{cmax} for the ABZ, as Kattge and Knorr (2007) is not optimized for the ABZ, biasing both the maximum productivity and phenology. Finally, we

670 adjust carbon allocation ratios for ABZ PFTs to levels that better match observations and result in more realistic simulations of GPP.

~~We are also interested in contributing to the improvement of the global CLM5.0 simulation. We confirm~~ Finally, we assess the performance of our model recommendation in a global simulation, confirming that a global simulation is possible and ~~reasonable (Supplement Fig. S12) with these additions to the code base. Additional testing at lower latitudes, which is outside of~~

675 ~~the scope of this study, is necessary to determine the effects on the global carbon budget, particularly the change in temperature scaling from Kattge and Knorr (2007) to Leuning (2002) that would be applied globally. However, for a regional simulation of the ABZ we argue that our model improvements can be used for a more accurate simulation of the high latitudes~~ yields reasonable carbon fluxes. Furthermore, we actually identify improvements in the global carbon cycle and budget according to ILAMB metrics. The reduction of biases in the ABZ carbon cycle has implications for future projections with models that

680 overestimate GPP. The modifications we implement here illustrate that previous extrapolations of temperate or even tropical observations ~~can be problematic~~ cause significant biases. We advocate for more regional ABZ focused development to ensure accuracy in the ABZ when implemented in global simulations, as the high latitudes are a critical component of the rapidly changing climate system.

Code availability. Our model recommendations are currently on Github for incorporation into the CESM master branch and CLM5.1 (

685 https://github.com/ESCOMP/CTSM/pull/947).

Data availability. For any questions or interest in our model data, please reach out to the corresponding authors Leah Birch and Brendan Rogers.

Author contributions. LB and BMR designed the study and model development. LB performed benchmarking, model development, and validation. All authors contributed to the manuscript.

690 *Competing interests.* The authors declare no competing interests.

References

Alexander, H. D. and Mack, M. C.: A canopy shift in interior Alaskan boreal forests: consequences for above-and belowground carbon and nitrogen pools during post-fire succession, Ecosystems, 19, 98–114, 2016.

705 Ali, A. A., Xu, C., Rogers, A., Fisher, R. A., Wullschleger, S. D., Massoud, E., Vrugt, J. A., Muss, J. D., McDowell, N. G., Fisher, J. B., et al.: A global scale mechanistic model of photosynthetic capacity (LUNA V1. 0), Geoscientific Model Development, 9, 587–606, 2016.

Amiro, B.: FLUXNET2015 CA-SF1 Saskatchewan-Western Boreal, forest burned in 1977, Tech. rep., FluxNet; University of Manitoba, 2016.

Archer, S. and Tieszen, L.: Effects of simulated grazing on foliage and root production and biomass allocation in an arctic tundra sedge 710 (Eriophorum vaginatum), Oecologia, 58, 92–102, 1983.

Arora, V. K. and Boer, G. J.: A parameterization of leaf phenology for the terrestrial ecosystem component of climate models, Global Change Biology, 11, 39–59, 2005.

Aurela, M.: FLUXNET2015 RU-Tks Tiksi, Tech. rep., FluxNet; Finnish Meteorological Institute-Helsinki, 2016.

Aurela, M., Tuovinen, J.-P., Hatakka, J., Lohila, A., Mäkelä, T., Rainne, J., and Lauria, T.: FLUXNET2015 FI-Sod Sodankyla, Tech. rep., 715 FluxNet; Finnish Meteorological Institute, 2016.

Bala, G., Caldeira, K., Wickett, M., Phillips, T., Lobell, D., Delire, C., and Mirin, A.: Combined climate and carbon-cycle effects of large-scale deforestation, Proceedings of the National Academy of Sciences, 104, 6550–6555, 2007.

Bauerle, W. L., Oren, R., Way, D. A., Qian, S. S., Stoy, P. C., Thornton, P. E., Bowden, J. D., Hoffman, F. M., and Reynolds, R. F.: Photoperiodic regulation of the seasonal pattern of photosynthetic capacity and the implications for carbon cycling, Proceedings of the 720 National Academy of Sciences, 109, 8612–8617, 2012.

Beck, P. S., Juday, G. P., Alix, C., Barber, V. A., Winslow, S. E., Sousa, E. E., Heiser, P., Herriges, J. D., and Goetz, S. J.: Changes in forest productivity across Alaska consistent with biome shift, Ecology letters, 14, 373–379, 2011.

Belshe, E., Schuur, E., Bolker, B., and Bracho, R.: Incorporating spatial heterogeneity created by permafrost thaw into a landscape carbon estimate, Journal of Geophysical Research: Biogeosciences, 117, 2012.

725 Biancamaria, S., Cazenave, A., Mognard, N. M., Llovel, W., and Frappart, F.: Satellite-based high latitude snow volume trend, variability and contribution to sea level over 1989/2006, Global and Planetary Change, 75, 99–107, 2011.

Black, T. A.: FLUXNET2015 CA-Obs Saskatchewan-Western Boreal, Mature Black Spruce, Tech. rep., FluxNet; The University of British Columbia, 2016.

Bonan, G. B., Pollard, D., and Thompson, S. L.: Effects of boreal forest vegetation on global climate, Nature, 359, 716–718, 1992.

730 Borner, A. P., Kielland, K., and Walker, M. D.: Effects of simulated climate change on plant phenology and nitrogen mineralization in Alaskan Arctic tundra, Arctic, Antarctic, and Alpine Research, 40, 27–38, 2008.

Botta, A., Viovy, N., Ciais, P., Friedlingstein, P., and Monfray, P.: A global prognostic scheme of leaf onset using satellite data, Global Change Biology, 6, 709–725, 2000.

Buchwal, A., Rachlewicz, G., Fonti, P., Cherubini, P., and Gärtner, H.: Temperature modulates intra-plant growth of Salix polaris from a high 735 Arctic site (Svalbard), Polar Biology, 36, 1305–1318, 2013.

Callaghan, T. V., Johansson, M., Brown, R. D., Groisman, P. Y., Labba, N., Radionov, V., Barry, R. G., Bulygina, O. N., Essery, R. L., Frolov, D., et al.: The changing face of Arctic snow cover: A synthesis of observed and projected changes, Ambio, 40, 17–31, 2011.

Carroll, M. and Loboda, T.: Multi-decadal surface water dynamics in north american tundra, Remote Sensing, 9, 497, 2017.

Carroll, M. L., Townshend, J., DiMiceli, C., Loboda, T., and Sohlberg, R.: Shrinking lakes of the Arctic: Spatial relationships and trajectory of change, Geophysical Research Letters, 38, 2011.

CESM2.0: Community Earth System Model, http://www.cesm.ucar.edu/models/cesm2/, accessed 2020-03-27.

Chapin, F. S., Sturm, M., Serreze, M. C., McFadden, J. P., Key, J., Lloyd, A. H., McGuire, A., Rupp, T. S., Lynch, A. H., Schimel, J. P., et al.: Role of land-surface changes in Arctic summer warming, science, 310, 657–660, 2005.

Chapin, F. S., Woodwell, G. M., Randerson, J. T., Rastetter, E. B., Lovett, G. M., Baldocchi, D. D., Clark, D. A., Harmon, M. E., Schimel, D. S., Valentini, R., et al.: Reconciling carbon-cycle concepts, terminology, and methods, Ecosystems, 9, 1041–1050, 2006.

Chapin III, F. S.: Nutrient allocation and responses to defoliation in tundra plants, Arctic and Alpine Research, 12, 553–563, 1980.

Chapin III, F. S. and Shaver, G. R.: Physiological and growth responses of arctic plants to a field experiment simulating climatic change, Ecology, 77, 822–840, 1996.

Chen, Z., Yu, G., Zhu, X., Wang, Q., Niu, S., and Hu, Z.: Covariation between gross primary production and ecosystem respiration across space and the underlying mechanisms: a global synthesis, Agricultural and Forest Meteorology, 203, 180–190, 2015.

Ciais, P., Tans, P., Trolier, M., White, J., and Francey, R.: A large northern hemisphere terrestrial CO_2 sink indicated by the 13C/12C ratio of atmospheric CO_2, Science, 269, 1098–1102, 1995.

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., Mu, M., and Randerson, J. T.: The International Land Model Benchmarking (ILAMB) system: design, theory, and implementation, Journal of Advances in Modeling Earth Systems, 10, 2731–2754, 2018.

Commane, R., Lindaas, J., Benmergui, J., Luus, K. A., Chang, R. Y.-W., Daube, B. C., Euskirchen, E. S., Henderson, J. M., Karion, A., Miller, J. B., Parazoo, N. C., Randerson, J. T., Sweeney, C., Tans, P., Thoning, K., Veraverbeke, S., Miller, C. E., and Wofsy, S. C.: Carbon dioxide sources from Alaska driven by increasing early winter respiration from Arctic tundra, Proceedings of the National Academy of Sciences, 114, 5361–5366, 2017.

Computational and Laboratory, I. S.: Cheyenne: HPE/SGI ICE XA System (University Community Computing), doi:10.5065/D6RX99HX, 2017.

Corradi, C., Kolle, O., Walter, K., Zimov, S., and Schulze, E.-D.: Carbon dioxide and methane exchange of a north-east Siberian tussock tundra, Global Change Biology, 11, 1910–1925, 2005.

Duncan, B. N., Ott, L. E., Abshire, J. B., Brucker, L., Carroll, M. L., Carton, J., Comiso, J. C., Dinnat, E. P., Forbes, B. C., Gonsamo, A., et al.: Space-Based Observations for Understanding Changes in the Arctic-Boreal Zone, Reviews of Geophysics, 58, e2019RG000 652, 2020.

Eitel, J. U., Maguire, A. J., Boelman, N., Vierling, L. A., Griffin, K. L., Jensen, J., Magney, T. S., Mahoney, P. J., Meddens, A. J., Silva, C., et al.: Proximal remote sensing of tree physiology at northern treeline: Do late-season changes in the photochemical reflectance index (PRI) respond to climate or photoperiod?, Remote sensing of environment, 221, 340–350, 2019.

Elmendorf, S. C., Henry, G. H., Hollister, R. D., Björk, R. G., Bjorkman, A. D., Callaghan, T. V., Collier, L. S., Cooper, E. J., Cornelissen, J. H., Day, T. A., et al.: Global assessment of experimental climate warming on tundra vegetation: heterogeneity over space and time, Ecology letters, 15, 164–175, 2012.

Euskirchen, E. S., Edgar, C., Turetsky, M., Waldrop, M. P., and Harden, J. W.: Differential response of carbon fluxes to climate in three peatland ecosystems that vary in the presence and stability of permafrost, Journal of Geophysical Research: Biogeosciences, 119, 1576–1595, 2014.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development (Online), 9, 2016.

Fisher, J., Sikka, M., Oechel, W., Huntzinger, D. N., Melton, J., Koven, C., Ahlström, A., Arain, A., Baker, I., Chen, J., et al.: Carbon cycle uncertainty in the Alaskan Arctic, Biogeosciences, 11, 4271–4288, 2014.

780   Fisher, J. I., Richardson, A. D., and Mustard, J. F.: Phenology model from surface meteorology does not capture satellite-based greenup estimations, Global Change Biology, 13, 707–721, 2007.

Fisher, R. A., Wieder, W. R., Sanderson, B. M., Koven, C. D., Oleson, K. W., Xu, C., Fisher, J., Shi, M., Walker, A. P., and Lawrence, D. M.: Parametric controls on vegetation responses to biogeochemical forcing in the CLM5, Journal of Advances in Modeling Earth Systems, 2019.

785   Forkel, M., Carvalhais, N., Schaphoff, S., Migliavacca, M., Thurner, M., Thonicke, K., et al.: Identifying environmental controls on vegetation greenness phenology through model–data integration, Biogeosciences, 11, 7025–7050, 2014.

Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R., Heimann, M., Thonicke, K., Zaehle, S., and Reichstein, M.: Enhanced seasonal CO_2 exchange caused by amplified plant productivity in northern ecosystems, Science, 351, 696–699, 2016.

Franklin, O., Johansson, J., Dewar, R. C., Dieckmann, U., McMurtrie, R. E., Brännström, Å., and Dybzinski, R.: Modeling carbon allocation

790   in trees: a search for principles, Tree Physiology, 32, 648–666, 2012.

Friedlingstein, P., Joel, G., Field, C. B., and Fung, I. Y.: Toward an allocation scheme for global terrestrial carbon models, Global Change Biology, 5, 755–770, 1999.

Fu, Y., Zhang, H., Dong, W., and Yuan, W.: Comparison of phenology models for predicting the onset of growing season over the Northern Hemisphere, PloS one, 9, e109 544, 2014.

795   Goulden, M., Wofsy, S., Harden, J., Trumbore, S. E., Crill, P., Gower, S., Fries, T., Daube, B., Fan, S.-M., Sutton, D., et al.: Sensitivity of boreal forest carbon balance to soil thaw, Science, 279, 214–217, 1998.

Gower, S., Vogel, J., Norman, J., Kucharik, C., Steele, S., and Stow, T.: Carbon distribution and aboveground net primary production in aspen, jack pine, and black spruce stands in Saskatchewan and Manitoba, Canada, Journal of Geophysical Research: Atmospheres, 102, 29 029–29 041, 1997.

800   Gower, S., Krankina, O., Olson, R., Apps, M., Linder, S., and Wang, C.: Net primary production and carbon allocation patterns of boreal forest ecosystems, Ecological applications, 11, 1395–1411, 2001.

Graven, H., Keeling, R., Piper, S., Patra, P., Stephens, B., Wofsy, S., Welp, L., Sweeney, C., Tans, P., Kelley, J., et al.: Enhanced seasonal exchange of CO_2 by northern ecosystems since 1960, Science, 341, 1085–1089, 2013.

Hanes, C. C., Wang, X., Jain, P., Parisien, M.-A., Little, J. M., and Flannigan, M. D.: Fire-regime changes in Canada over the last half century,

805   Canadian Journal of Forest Research, 49, 256–269, 2019.

Holl, D., Wille, C., Sachs, T., Schreiber, P., Runkle, B. R., Beckebanze, L., Langer, M., Boike, J., Pfeiffer, E.-M., Fedorova, I., et al.: A long-term (2002 to 2017) record of closed-path and open-path eddy covariance CO_2 net ecosystem exchange fluxes from the Siberian Arctic, Earth System Science Data, 11, 221–240, 2019.

Høye, T. T., Post, E., Meltofte, H., Schmidt, N. M., and Forchhammer, M. C.: Rapid advancement of spring in the High Arctic, Current

810   Biology, 17, R449–R451, 2007.

Hufkens, K., Basler, D., Milliman, T., Melaas, E. K., and Richardson, A. D.: An integrated phenology modelling framework in R, Methods in Ecology and Evolution, 9, 1276–1285, 2018.

Huntzinger, D. N., Schwalm, C., Michalak, A., Schaefer, K., King, A., Wei, Y., Jacobson, A., Liu, S., Cook, R., Post, W., et al.: The north american carbon program multi-scale synthesis and terrestrial model intercomparison project–part 1: Overview and experimental design, Geoscientific Model Development, 6, 2121–2133, 2013.

Ito, A., Inatomi, M., Huntzinger, D. N., Schwalm, C., Michalak, A. M., Cook, R., King, A. W., Mao, J., Wei, Y., Post, W. M., et al.: Decadal trends in the seasonal-cycle amplitude of terrestrial CO_2 exchange resulting from the ensemble of terrestrial biosphere models, Tellus B: Chemical and Physical Meteorology, 68, 28 968, 2016.

Iversen, C. M., Sloan, V. L., Sullivan, P. F., Euskirchen, E. S., McGuire, A. D., Norby, R. J., Walker, A. P., Warren, J. M., and Wullschleger, S. D.: The unseen iceberg: plant roots in arctic tundra, New Phytologist, 205, 34–58, 2015.

Jeong, S.-J., Bloom, A. A., Schimel, D., Sweeney, C., Parazoo, N. C., Medvigy, D., Schaepman-Strub, G., Zheng, C., Schwalm, C. R., Huntzinger, D. N., et al.: Accelerating rates of Arctic carbon cycling revealed by long-term atmospheric CO_2 measurements, Science advances, 4, eaao1167, 2018.

Jolly, W. M., Nemani, R., and Running, S. W.: A generalized, bioclimatic index to predict foliar phenology in response to climate, Global Change Biology, 11, 619–632, 2005.

Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., Arneth, A., Camps-Valls, G., Ciais, P., Friedlingstein, P., et al.: Compensatory water effects link yearly global land CO 2 sink changes to temperature, Nature, 541, 516, 2017.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., et al.: Scaling carbon fluxes from eddy covariance sites to globe: Synthesis and evaluation of the FLUXCOM approach, Biogeosciences Discussions, 2019.

Kajimoto, T., Matsuura, Y., Sofronov, M., Volokitina, A., Mori, S., Osawa, A., and Abaimov, A.: Above-and belowground biomass and net primary productivity of a Larix gmelinii stand near Tura, central Siberia, Tree physiology, 19, 815–822, 1999.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., et al.: The NCEP/NCAR 40-year reanalysis project, Bulletin of the American meteorological Society, 77, 437–471, 1996.

Kasischke, E. S., Verbyla, D. L., Rupp, T. S., McGuire, A. D., Murphy, K. A., Jandt, R., Barnes, J. L., Hoy, E. E., Duffy, P. A., Calef, M., et al.: Alaska's changing fire regime—implications for the vulnerability of its boreal forests, Canadian Journal of Forest Research, 40, 1313–1324, 2010.

Kattge, J. and Knorr, W.: Temperature acclimation in a biochemical model of photosynthesis: a reanalysis of data from 36 species, Plant, cell & environment, 30, 1176–1190, 2007.

Keeling, C. D., Chin, J., and Whorf, T.: Increased activity of northern vegetation inferred from atmospheric CO_2 measurements, Nature, 382, 146, 1996.

Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., Lola da Costa, A. C., and Gentine, P.: Implementing plant hydraulics in the community land model, version 5, Journal of Advances in Modeling Earth Systems, 11, 485–513, 2019.

Kim, Y., Kimball, J. S., Zhang, K., and McDonald, K. C.: Satellite detection of increasing Northern Hemisphere non-frozen seasons from 1979 to 2008: Implications for regional vegetation growth, Remote Sensing of Environment, 121, 472–487, 2012.

Kobak, K., Turcmnovich, I. Y., Kondrasiheva, N. Y., Schulze, E.-D., Schulze, W., Koch, H., and Vygodskaya, N.: Vulnerability and adaptation of the larch forest in eastern Siberia to climate change, Water, Air, and Soil Pollution, 92, 119–127, 1996.

Köhler, P., Guanter, L., and Joiner, J.: A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIA-MACHY data, Atmospheric Measurement Techniques, 8, 2589–2608, 2015.

850   Koven, C. D., Lawrence, D. M., and Riley, W. J.: Permafrost carbon- climate feedback is sensitive to deep soil carbon decomposability but not deep soil nitrogen dynamics, Proceedings of the National Academy of Sciences, 112, 3752–3757, 2015.

Kumarathunge, D. P., Medlyn, B. E., Drake, J. E., Tjoelker, M. G., Aspinwall, M. J., Battaglia, M., Cano, F. J., Carter, K. R., Cavaleri, M. A., Cernusak, L. A., et al.: Acclimation and adaptation components of the temperature dependence of plant photosynthesis at the global scale, New Phytologist, 222, 768–784, 2019.

855   Kutzbach, L., Sachs, T., Boike, J., Wille, C., Schreiber, P., Langer, M., and Pfeiffer, E.-M.: FLUXNET2015 RU-Sam Samoylov, Tech. rep., FluxNet; GFZ German Research Centre for Geosciences, https://doi.org/10.18140/FLX/1440185, 2002-2014.

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., et al.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, 2010.

Lawrence, D. M., Hurtt, G. C., Arneth, A., Brovkin, V., Calvin, K. V., Jones, A. D., Jones, C. D., Lawrence, P. J., de Noblet-Ducoudré, N.,
860   Pongratz, J., et al.: The Land Use Model Intercomparison Project (LUMIP) contribution to CMIP6: rationale and experimental design, Geoscientific Model Development, 9, 2973–2998, 2016.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., et al.: The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty, Journal of Advances in Modeling Earth Systems, 2019.

865   Leuning, R.: Temperature dependence of two parameters in a photosynthesis model, Plant, Cell & Environment, 25, 1205–1210, 2002.

Li, H., Wigmosta, M. S., Wu, H., Huang, M., Ke, Y., Coleman, A. M., and Leung, L. R.: A physically based runoff routing model for land surface and earth system models, Journal of Hydrometeorology, 14, 808–828, 2013.

Lin, X., Rogers, B. M., Sweeney, C., Chevallier, F., Arshinov, M., Dlugokencky, E., Machida, T., Sasakawa, M., Tans, P., and Keppel-Aleks, G.: Siberian and temperate ecosystems shape Northern Hemisphere atmospheric CO_2 seasonal amplification, Proceedings of the National
870   Academy of Sciences, 117, 21 079–21 087, 2020.

Liptak, J., Keppel-Aleks, G., and Lindsay, K.: Drivers of multi-century trends in the atmospheric CO_2 mean annual cycle in a prognostic ESM, Biogeosciences, 14, 1383–1401, 2017.

Lloyd, A. H. and Fastie, C. L.: Recent changes in treeline forest distribution and structure in interior Alaska, Ecoscience, 10, 176–185, 2003.

Marchand, F. L., Nijs, I., Heuer, M., Mertens, S., Kockelbergh, F., Pontailler, J.-Y., Impens, I., and Beyens, L.: Climate warming postpones
875   senescence in High Arctic tundra, Arctic, Antarctic, and Alpine Research, 36, 390–394, 2004.

Margolis, H.: AmeriFlux CA-Qc2 Quebec-1975 Harvested Black Spruce (HBS75), Tech. rep., AmeriFlux; Laval University, 2018.

Maximov, T.: FLUXNET2015 RU-SkP Yakutsk Spasskaya Pad larch, Tech. rep., FluxNet; IBPC, Russia, 2016.

McCaughey, H.: FLUXNET2015 CA-Gro Ontario-Groundhog River, Boreal Mixedwood Forest, Tech. rep., FluxNet; Queen's University, 2016.

880   McGuire, A., Hayes, D. J., Kicklighter, D. W., Manizza, M., Zhuang, Q., Chen, M., Follows, M. J., Gurney, K., Mcclelland, J. W., Melillo, J., et al.: An analysis of the carbon balance of the Arctic Basin from 1997 to 2006, Tellus B: Chemical and Physical Meteorology, 62, 455–474, 2010.

McGuire, A. D., Anderson, L. G., Christensen, T. R., Dallimore, S., Guo, L., Hayes, D. J., Heimann, M., Lorenson, T. D., Macdonald, R. W., and Roulet, N.: Sensitivity of the carbon cycle in the Arctic to climate change, Ecological Monographs, 79, 523–555, 2009.

885   McGuire, A. D., Lawrence, D. M., Koven, C., Clein, J. S., Burke, E., Chen, G., Jafarov, E., MacDougall, A. H., Marchenko, S., Nicolsky, D., et al.: Dependence of the evolution of carbon dynamics in the northern permafrost region on the trajectory of climate change, Proceedings of the National Academy of Sciences, 115, 3882–3887, 2018.

McMILLAN, A. M., Winston, G. C., and Goulden, M. L.: Age-dependent response of boreal forest to temperature and rainfall variability, Global Change Biology, 14, 1904–1916, 2008.

890 Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V., Crous, K. Y., De Angelis, P., Freeman, M., and Wingate, L.: Reconciling the optimal and empirical approaches to modelling stomatal conductance, Global Change Biology, 17, 2134–2144, 2011.

Montané, F., Fox, A. M., Arellano, A. F., MacBean, N., Alexander, M. R., Dye, A., Bishop, D. A., Trouet, V., Babst, F., Hessl, A. E., et al.: Evaluating the effect of alternative carbon allocation schemes in a land surface model (CLM4. 5) on carbon fluxes, pools, and turnover in 895 temperate forests, Geoscientific Model Development (Online), 10, 2017.

Myers-Smith, I. H., Forbes, B. C., Wilmking, M., Hallinger, M., Lantz, T., Blok, D., Tape, K. D., Macias-Fauria, M., Sass-Klaassen, U., Lévesque, E., et al.: Shrub expansion in tundra ecosystems: dynamics, impacts and research priorities, Environmental Research Letters, 6, 045 509, 2011.

Myers-Smith, I. H., Elmendorf, S. C., Beck, P. S., Wilmking, M., Hallinger, M., Blok, D., Tape, K. D., Rayback, S. A., Macias-Fauria, M., 900 Forbes, B. C., et al.: Climate sensitivity of shrub growth across the tundra biome, Nature Climate Change, 5, 887–891, 2015.

Natali, S. M., Schuur, E. A., Webb, E. E., Pries, C. E. H., and Crummer, K. G.: Permafrost degradation stimulates carbon loss from experimentally warmed tundra, Ecology, 95, 602–608, 2014.

Natali, S. M., Watts, J. D., Rogers, B. M., Potter, S., Ludwig, S. M., Selbmann, A. K., Sullivan, P. F., Abbott, B. W., Arndt, K. A., Birch, L., et al.: Large loss of CO_2 in winter observed across the northern permafrost region, Nature Climate Change, 9, 852–857, 2019.

905 Negrón-Juárez, R. I., Koven, C. D., Riley, W. J., Knox, R. G., and Chambers, J. Q.: Observed allocations of productivity and biomass, and turnover times in tropical forests are not accurately represented in CMIP5 Earth system models, Environmental Research Letters, 10, 064 017, 2015.

Oberbauer, S., Elmendorf, S., Troxler, T., Hollister, R., Rocha, A., Bret-Harte, M., Dawes, M., Fosaa, A., Henry, G., Høye, T. T., et al.: Phenological response of tundra plants to background climate variation tested using the International Tundra Experiment, Philosophical 910 Transactions of the Royal Society B: Biological Sciences, 368, 20120 481, 2013.

Oechel, W. C., Laskowski, C. A., Burba, G., Gioli, B., and Kalhori, A. A.: Annual patterns and budget of CO2 flux in an Arctic tussock tundra ecosystem, Journal of Geophysical Research: Biogeosciences, 119, 323–339, 2014.

Parazoo, N. C., Arneth, A., Pugh, T. A., Smith, B., Steiner, N., Luus, K., Commane, R., Benmergui, J., Stofferahn, E., Liu, J., et al.: Spring photosynthetic onset and net CO_2 uptake in Alaska triggered by landscape thawing, Global change biology, 24, 3416–3435, 2018a.

915 Parazoo, N. C., Koven, C. D., Lawrence, D. M., Romanovsky, V., and Miller, C. E.: Detecting the permafrost carbon feedback: talik formation and increased cold-season respiration as precursors to sink-to-source transitions, The Cryosphere (Online), 12, 2018b.

Peng, S., Ciais, P., Chevallier, F., Peylin, P., Cadule, P., Sitch, S., Piao, S., Ahlström, A., Huntingford, C., Levy, P., et al.: Benchmarking the seasonal cycle of CO2 fluxes simulated by terrestrial ecosystem models, Global Biogeochemical Cycles, 29, 46–64, 2015.

Phoenix, G. K. and Bjerke, J. W.: Arctic browning: extreme events and trends reversing arctic greening, Global Change Biology, 22, 2960–920 2962, 2016.

Piao, S., Ciais, P., Friedlingstein, P., Peylin, P., Reichstein, M., Luyssaert, S., Margolis, H., Fang, J., Barr, A., Chen, A., et al.: Net carbon dioxide losses of northern ecosystems in response to autumn warming, Nature, 451, 49–52, 2008.

Piao, S., Sitch, S., Ciais, P., Friedlingstein, P., Peylin, P., Wang, X., Ahlström, A., Anav, A., Canadell, J. G., Cong, N., et al.: Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO2 trends, Global change biology, 19, 2117–2132, 2013.

925 Randerson, J., Field, C., Fung, I., and Tans, P.: Increases in early season ecosystem uptake explain recent changes in the seasonal cycle of atmospheric CO2 at high northern latitudes, Geophysical Research Letters, 26, 2765–2768, 1999.

Randerson, J. T., Thompson, M. V., Conway, T. J., Fung, I. Y., and Field, C. B.: The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, Global Biogeochemical Cycles, 11, 535–560, 1997.

Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., Chen, J. M., Ciais, P., Davis, K. J., Desai, A. R., et al.:
930 Terrestrial biosphere models need better representation of vegetation phenology: results from the N orth A merican C arbon P rogram S ite S ynthesis, Global Change Biology, 18, 566–584, 2012.

Richardson, A. D., Hufkens, K., Milliman, T., Aubrecht, D. M., Chen, M., Gray, J. M., Johnston, M. R., Keenan, T. F., Klosterman, S. T., Kosmala, M., et al.: Tracking vegetation phenology across diverse North American biomes using PhenoCam imagery, Scientific data, 5, 1–24, 2018.

935 Rogers, A., Serbin, S. P., Ely, K. S., Sloan, V. L., and Wullschleger, S. D.: Terrestrial biosphere models underestimate photosynthetic capacity and CO2 assimilation in the Arctic, New Phytologist, 216, 1090–1103, 2017.

Rogers, B., Randerson, J., and Bonan, G.: High-latitude cooling associated with landscape changes from North American boreal forest fires, Biogeosciences, 10, 699–718, 2013.

Rogers, B. M., Soja, A. J., Goulden, M. L., and Randerson, J. T.: Influence of tree species on continental differences in boreal fires and
940 climate feedbacks, Nature Geoscience, 8, 228–234, 2015.

Rogers, B. M., Solvik, K., Hogg, E. H., Ju, J., Masek, J. G., Michaelian, M., Berner, L. T., and Goetz, S. J.: Detecting early warning signals of tree mortality in boreal North America using multiscale satellite data, Global change biology, 24, 2284–2304, 2018.

Runkle, B. R., Sachs, T., Wille, C., Pfeiffer, E.-M., and Kutzbach, L.: Bulk partitioning the growing season net ecosystem exchange of CO 2 in Siberian tundra reveals the seasonality of its carbon sequestration strength., Biogeosciences, 10, 2013.

945 Salmon, V. G., Soucy, P., Mauritz, M., Celis, G., Natali, S. M., Mack, M. C., and Schuur, E. A.: Nitrogen availability increases in a tundra ecosystem during five years of experimental permafrost thaw, Global Change Biology, 22, 1927–1941, 2016.

Schaefer, K., Schwalm, C. R., Williams, C., Arain, M. A., Barr, A., Chen, J. M., Davis, K. J., Dimitrov, D., Hilton, T. W., Hollinger, D. Y., et al.: A model-data comparison of gross primary productivity: Results from the North American Carbon Program site synthesis, Journal of Geophysical Research: Biogeosciences, 117, 2012.

950 Schwalm, C. R., Williams, C. A., Schaefer, K., Anderson, R., Arain, M. A., Baker, I., Barr, A., Black, T. A., Chen, G., Chen, J. M., et al.: A model-data intercomparison of CO2 exchange across North America: Results from the North American Carbon Program site synthesis, Journal of Geophysical Research: Biogeosciences, 115, 2010.

Searle, E. B. and Chen, H. Y.: Persistent and pervasive compositional shifts of western boreal forest plots in Canada, Global Change Biology, 23, 857–866, 2017.

955 Semenchuk, P. R., Gillespie, M. A., Rumpf, S. B., Baggesen, N., Elberling, B., and Cooper, E. J.: High Arctic plant phenology is determined by snowmelt patterns but duration of phenological periods is fixed: An example of periodicity, Environmental Research Letters, 11, 125 006, 2016.

Serreze, M. C. and Barry, R. G.: Processes and impacts of Arctic amplification: A research synthesis, Global and planetary change, 77, 85–96, 2011.

960 Serreze, M. C. and Francis, J. A.: The Arctic amplification debate, Climatic change, 76, 241–264, 2006.

Shaver, G. R., Billings, W. D., Chapin III, F. S., Giblin, A. E., Nadelhoffer, K. J., Oechel, W. C., and Rastetter, E.: Global change and the carbon balance of arctic ecosystems: Carbon/nutrient interactions should act as major constraints on changes in global terrestrial carbon cycling, Bioscience, 42, 433–441, 1992.

Sloan, V. L., Fletcher, B. J., Press, M. C., Williams, M., and Phoenix, G. K.: Leaf and fine root carbon stocks and turnover are coupled across
965   Arctic ecosystems, Global Change Biology, 19, 3668–3676, 2013.

Smith, N. G., Lombardozzi, D., Tawfik, A., Bonan, G., and Dukes, J. S.: Biophysical consequences of photosynthetic temperature acclimation for climate, Journal of Advances in Modeling Earth Systems, 9, 536–547, 2017.

Starr, G. and Oberbauer, S. F.: Photosynthesis of arctic evergreens under snow: implications for tundra ecosystem carbon balance, Ecology, 84, 1415–1420, 2003.

970 Stöckli, R., Lawrence, D., Niu, G.-Y., Oleson, K., Thornton, P. E., Yang, Z.-L., Bonan, G., Denning, A., and Running, S. W.: Use of FLUXNET in the Community Land Model development, Journal of Geophysical Research: Biogeosciences, 113, 2008a.

Stöckli, R., Rutishauser, T., Dragoni, D., O'keefe, J., Thornton, P., Jolly, M., Lu, L., and Denning, A.: Remote sensing data assimilation for a prognostic phenology model, Journal of Geophysical Research: Biogeosciences, 113, 2008b.

Sulla-Menashe, D. and Friedl, M. A.: MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006, DOI:
975   10.5067/MODIS/MCD12Q1.006, accessed 2020-03-27, 2019.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological Society, 93, 485–498, 2012.

Tramontana, G., Jung, M., Camps-Valls, G., Ichii, K., Ráduly, B., Reichstein, M., Schwalm, C. R., Arain, M. A., Cescatti, A., Kiely, G., et al.: Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms, Biogeosciences Discussions,
980   2016.

Turetsky, M. R., Kane, E. S., Harden, J. W., Ottmar, R. D., Manies, K. L., Hoy, E., and Kasischke, E. S.: Recent acceleration of biomass burning and carbon losses in Alaskan forests and peatlands, Nature Geoscience, 4, 27–31, 2011.

Ueyama, M., Iwata, H., Harazono, Y., Euskirchen, E. S., Oechel, W. C., and Zona, D.: Growing season and spatial variations of carbon fluxes of Arctic and boreal ecosystems in Alaska (USA), Ecological Applications, 23, 1798–1816, 2013.

985 Van den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., et al.: LS3MIP (v1. 0) contribution to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project-aims, setup and expected outcome, Geoscientific Model Development, 9, 2809–2832, 2016.

Vanninen, P. and Mäkelä, A.: Carbon budget for Scots pine trees: effects of size, competition and site fertility on growth allocation and production, Tree Physiology, 25, 17–30, 2005.

990 Verbyla, D.: Browning boreal forests of western North America, Environmental Research Letters, 6, 041 003, 2011.

Virkkala, A.-M., Virtanen, T., Lehtonen, A., Rinne, J., and Luoto, M.: The current state of CO2 flux chamber studies in the Arctic tundra: A review, Progress in Physical Geography: Earth and Environment, 42, 162–184, 2018.

Virkkala, A.-M., Abdi, A. M., Luoto, M., and Metcalfe, D. B.: Identifying multidisciplinary research gaps across Arctic terrestrial gradients, Environmental Research Letters, 14, 124 061, 2019.

995 Walker, X. and Johnstone, J. F.: Widespread negative correlations between black spruce growth and temperature across topographic moisture gradients in the boreal forest, Environmental Research Letters, 9, 064 016, 2014.

Walker, X. J., Mack, M. C., and Johnstone, J. F.: Stable carbon isotope analysis reveals widespread drought stress in boreal black spruce forests, Global Change Biology, 21, 3102–3113, 2015.

35

Welp, L. R., Patra, P. K., Rödenbeck, C., Nemani, R., Bi, J., Piper, S. C., and Keeling, R. F.: Increasing summer net CO_2 uptake in high northern ecosystems inferred from atmospheric inversions and comparisons to remote-sensing NDVI, Atmospheric Chemistry and Physics, 16, 9047–9066, 2016.

White, M. A., Thornton, P. E., and Running, S. W.: A continental phenology model for monitoring vegetation responses to interannual climatic variability, Global biogeochemical cycles, 11, 217–234, 1997.

White, M. A., de Beurs, K. M., Didan, K., Inouye, D. W., Richardson, A. D., Jensen, O. P., O'KEEFE, J., Zhang, G., Nemani, R. R., van Leeuwen, W. J., et al.: Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006, Global Change Biology, 15, 2335–2359, 2009.

Wieder, W. R., Lawrence, D. M., Fisher, R. A., Bonan, G. B., Cheng, S. J., Goodale, C. L., Grandy, A. S., Koven, C. D., Lombardozzi, D. L., Oleson, K. W., et al.: Beyond static benchmarking: Using experimental manipulations to evaluate land model assumptions, Global Biogeochemical Cycles, 2019.

Zhang, X., Friedl, M. A., Schaaf, C. B., and Strahler, A. H.: Climate controls on vegetation phenological patterns in northern mid-and high latitudes inferred from MODIS data, Global change biology, 10, 1133–1145, 2004.

Zhao, F. and Zeng, N.: Continued increase in atmospheric CO 2 seasonal amplitude in the 21st century projected by the CMIP5 Earth system models, Earth System Dynamics, 5, 423–439, 2014.

Zhao, F., Zeng, N., Asrar, G., Friedlingstein, P., Ito, A., Jain, A., Kalnay, E., Kato, E., Koven, C., Poulter, B., et al.: Role of CO_2, climate and land use in regulating the seasonal amplitude increase of carbon fluxes in terrestrial ecosystems: a multimodel analysis, Biogeosciences, 13, 5121–5137, 2016.

Zimov, S., Davidov, S., Voropaev, Y. V., Prosiannikov, S., Semiletov, I., Chapin, M., and Chapin, F.: Siberian CO 2 efflux in winter as a CO 2 source and cause of seasonality in atmospheric CO 2, Climatic Change, 33, 111–120, 1996.

Zimov, S., Davidov, S., Zimova, G., Davidova, A., Chapin, F., Chapin, M., and Reynolds, J.: Contribution of disturbance to increasing seasonal amplitude of atmospheric CO_2, Science, 284, 1973–1976, 1999.