1	Evaluation and Modification of ELM Seasonal Deciduous Phenology against
2	<b>Observations in a Southern Boreal Peatland Forest</b>
3 4	Lin Meng <sup>a</sup> , Jiafu Mao <sup>b,*</sup> , Daniel M. Ricciuto <sup>b</sup> , Xiaoying Shi <sup>b</sup> , Andrew D. Richardson <sup>c,d</sup> , Paul J. Hanson <sup>b</sup> , Jeffrey M. Warren <sup>b</sup> , Yuyu Zhou <sup>a</sup> , Xuecao Li <sup>e</sup> , Li Zhang <sup>f</sup> , and Christina Schädel <sup>d</sup>
5	<sup>a</sup> Department of Geological and Atmospheric Sciences, Iowa State University, Ames, IA 50011
6 7	<sup>b</sup> Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN 37830
8 9	<sup>c</sup> School of Informatics, Computing and Cyber Systems, Northern Arizona University, Flagstaff, AZ 86011
10	<sup>d</sup> Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ 86011
11 12	<sup>e</sup> College of Land Science and Technology, China Agricultural University, Beijing 100083, China
13 14	<sup>f</sup> Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing
15	
16	Corresponding author: Jiafu Mao (maoj@ornl.gov)
17	
18 19 20	This manuscript has been authored by UT-Battelle LLC under Contract No. DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license

- 21 to publish or reproduce the published form of this manuscript, or allow others to do so, for US government
- 22 purposes. DOE will provide public access to these results of federally sponsored research in accordance with the
- 23 DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

#### 24 Abstract

Phenological transitions determine the timing of changes in land surface properties and the 25 26 seasonality of exchanges of biosphere-atmosphere energy, water, and carbon. Accurate mechanistic modeling of phenological processes is therefore critical to understand and correctly 27 predict terrestrial ecosystem feedbacks with changing atmospheric and climate conditions. 28 29 However, the phenological components in the land model of the US Department of Energy's 30 (DOE) Energy Exascale Earth System Model (ELM of E3SM) were previously unable to accurately capture the observed phenological responses to environmental conditions in a well-31 studied boreal peatland forest. In this research, we introduced new seasonal-deciduous 32 phenology schemes into version 1.0 of ELM and evaluated their performance against the 33 34 PhenoCam observations at the Spruce and Peatland Responses Under Changing Environments (SPRUCE) experiment in northern Minnesota from 2015 to 2018. We found that phenology 35 simulated by the revised ELM (i.e., earlier spring onsets and stronger warming responses of 36 37 spring onsets and autumn senescence) was closer to observations than simulations from the 38 original algorithms for both the deciduous conifer (*Larix laricina*) and mixed shrub layers. 39 Moreover, the revised ELM generally produced higher carbon and water fluxes (e.g., 40 photosynthesis and evapotranspiration) during the growing season and stronger flux responses to warming than the default ELM. A parameter sensitivity analysis further indicated the significant 41 contribution of phenology parameters to uncertainty in key carbon and water cycle variables, 42 43 underscoring the importance of precise phenology parameterization. This phenological modeling effort demonstrates the potential to enhance the E3SM representation of land-climate interactions 44 at broader spatiotemporal scales, especially under anticipated elevated CO<sub>2</sub> and warming 45 conditions. 46

47 Keywords: Phenology, climate change, PhenoCam, modeling, E3SM, ELM

# 48 **1 Introduction**

Phenological events such as leaf unfolding and flowering are sensitive to variations in weather 49 50 and climate (Körner and Basler, 2010a). Multiple lines of observations have documented an earlier spring leaf-out, later autumn senescence, and longer growing season in response to 51 warming, especially for terrestrial ecosystems over the northern mid-high latitudes (Körner and 52 53 Basler, 2010b; Menzel and Fabian, 1999; Piao et al., 2019a; Richardson et al., 2018b). Phenology regulates plant feedbacks to the climate system via the timing changes of land surface 54 biophysical and biogeochemical fluxes and properties (Richardson et al., 2013; Schwartz, 1992; 55 56 Li et al., 2016). For example, the leaf emergence in spring increases transpiration and latent heat flux but decreases the Bowen ratio (Richardson et al., 2013); earlier spring leaf-out increased 57 annual gross primary production (GPP) by about 10 g  $C/m^2$  in a temperate deciduous forest 58 59 (Richardson et al., 2013); and an extension of just 1 day in growing season length may increase 60 the annual GPP of northern terrestrial ecosystems by ~0.5% (Piao et al., 2007; White et al., 2009). 61

Spring onset of seasonal deciduous ecosystems is mainly driven by temperature through underlying physiological processes (Sarvas, 1974). Based on this, the thermal time model was proposed as early as the eighteenth century to predict spring onset as the date when the accumulated growing-degree-day exceeds a threshold (Reaumur, 1735). Later, manipulative experiments revealed that plants need to be exposed to low chilling temperatures to break dormancy before responding to the rising spring temperature (Hannerz et al., 2003; Hänninen et al., 2019; Sarvas, 1972). Chilling models have since been proposed to predict the spring

69	phenology based on a balance between winter chilling requirement and accumulated spring
70	temperature forcing (Cannell and Smith, 1983; Landsberg, 1974; Sarvas, 1972). These models
71	characterize nonlinear responses of spring phenology to temperature (Cesaraccio et al., 2004),
72	and have been widely applied from local to global scales to examine the phenological effects of
73	climate change (Chuine et al., 2016; Meng et al., 2020a; Meng et al., 2020b). More recent
74	models for spring phenology have emerged by integrating the influences of winter chilling,
75	spring thermal forcing, photoperiod (i.e., daylength), and their interactions (Caffarra et al., 2011;
76	Liang, 2019), albeit resulting in greater model complexity and increased model parameters
77	(Chuine and Régnière, 2017; Hänninen et al., 2019).
78	In contrast, autumn senescence such as leaf coloration, reduced functionality of photosynthesis
79	and transpiration, and leaf fall has drawn less attention until recently (Keskitalo et al., 2005;
80	Richardson et al., 2012). Autumn senescence was found to be likely modulated by the decrease
81	in photoperiod and temperature (Fracheboud et al., 2009; Richardson et al., 2018b). Based on the
82	process of cold-degree-day (CDD) accumulation, several leaf senescence models were developed
83	(Dufrêne et al., 2005; Jeong and Medvigy, 2014); the impacts of daylength decrease on the
84	effectiveness of CDD accumulation were later introduced to improve the senescence modeling
85	(Delpierre et al., 2009); and the impact of timing of spring phenology on the autumn senescence
86	process by influencing the requirement of CDD accumulation was also proposed (Keenan and
87	Richardson, 2015). A recent study examining multiple leaf senescence assumptions reported that
88	models considering the interactions between photoperiod and temperature outperformed those
89	solely based on temperature or photoperiod thresholds (Liu et al., 2020), although further
90	mechanistic understanding of leaf senescence is needed.

91 Despite the progress in phenological modeling, the advancement of phenology scheme remains limited in many state-of-the-art land surface models, such as the land model of the US 92 Department of Energy's (DOE) Energy Exascale Earth System Model (ELM of E3SM). 93 94 Consistent with Community Land Model (CLM) version 4.5, ELM v1.0 includes three distinct phenology types that are represented by separate algorithms: an evergreen type, for which some 95 fraction of annual leaf growth persists for longer than one year; a seasonal-deciduous type with 96 distinct growing and dormant seasons once per year; and a stress-deciduous type with the 97 potential for multiple growing seasons per year (Lawrence et al., 2011). For the seasonal-98 99 deciduous model, two processes determine the length of the growing season: spring onset and autumn senescence. Spring onset is merely triggered by the growing-degree-day accumulation of 100 soil temperature, and autumn senescence is set as a constant date at any given location according 101 102 to a fixed daylength threshold (Oleson, 2013). Although this scheme captures the dominant drivers of seasonal phenology, the ELM phenological simulations were demonstrated to be 103 biased against observational phenology changes and failed to capture interannual variation of 104 105 phenology (Chen et al., 2016). For example, degraded land model performance was noticed when using the prognostic phenology instead of prescribed, satellite-derived phenology 106 (Lawrence et al., 2011); substantial biases in the intra-annual variation of the fraction of 107 absorbed photosynthetically active radiation were found to be partially caused by phenology 108 errors (Wang et al., 2013). Such uncertainties associated with phenology modeling are not 109 unique for ELM but are typical for most current-generation land surface models. One study 110 systematically evaluated 14 land surface models participating in the North American Carbon 111 Program Site Synthesis and revealed a model bias of about two weeks in the representation of the 112 113 growing season length and a poor reproduction of the observed inter-annual phenology

114 variability (Richardson et al., 2012). Another more recent multi–land surface model

intercomparison study (including CLM 4.5 and 5.0) showed an 18-day delay for the start of the

season and a 2-week advancement for the end of the season compared with satellite observations

117 (Peano et al., 2020). Limited efforts, however, have been spent improving the spring onset

models using satellite or in situ observations (Chen et al., 2016; Dahlin et al., 2015), and even

119 fewer have been focused on improving the autumn senescence process.

120 The Spruce and Peatland Responses Under Changing Environments (SPRUCE) wholeecosystem warming experiment in a southern boreal peatland forest provides a unique 121 opportunity to confront the seasonal-deciduous phenology schemes of ELMs at the site level in 122 response to altered environmental conditions. Results from this experiment revealed that 123 warming treatments of up to +9°C caused a linearly advanced spring green-up (1.97 days/°C for 124 Larix laricina and 1.99 days/°C for the mixed shrub layer) and delayed autumn green-down 125 (1.34 days/°C for Larix laricina and 4.98 days/°C for the mixed shrub layer) (Richardson et al., 126 127 2018b). These SPRUCE phenological changes have been monitored automatically every half 128 hour by cameras of the PhenoCam network, which use digital photography with high quality and 129 a standardized approach to track phenology evolutions and facilitate model development 130 (Sevednasrollah et al., 2019). In this study, we first incorporated improved spring onset and 131 autumn senescence models into the ELM as the seasonal-deciduous phenology schemes, in which the spring onset depends on both winter chilling and spring thermal forcing processes 132 133 whereas the timing of autumn senescence relies on the co-limitation of daylength and temperature. The modified ELM was then calibrated and evaluated against the SPRUCE 134 PhenoCam observations for the period beginning in autumn 2015 through 2018. We also 135 simulated and examined possible phenological feedbacks to major carbon and hydrological 136

137 variables under various warming and CO<sub>2</sub> levels using the default and updated ELM. The sensitivity of model outputs to the new phenology parameters was further analyzed. The 138 following questions were addressed in this research: (1) How different are the default and 139 140 phenology-modified ELM versions regarding the prediction of phenology patterns and phenological responses to warming? (2) To what extent are the ELM-simulated exchanges of 141 land-atmosphere carbon and water fluxes affected by the new phenological algorithms? (3) How 142 sensitive are the carbon and water cycle outputs of the updated ELM to uncertainty in major 143 phenology parameters across vegetation types? 144

### 145 **2 Materials and Methods**

### 146 **2.1 Experimental site**

The SPRUCE experiment is a large-scale ecological manipulation that evaluates the combined 147 response of a southern boreal peatland to multiple levels of whole-ecosystem warming up to 148  $+9^{\circ}$ C at both ambient CO<sub>2</sub> (aCO<sub>2</sub>) and elevated CO<sub>2</sub> (eCO<sub>2</sub>) concentrations (Hanson et al., 2017). 149 This field experiment was designed to investigate how underlying mechanistic processes respond 150 to altered climate conditions, and how the interaction between those processes scale to 151 152 ecosystem-level responses. SPRUCE consists of large open-top enclosures and is located within a *Picea-Larix—Sphagnum* spp. bog (S1 bog) at the USDA Forest Service Marcell Experimental 153 Forest in northern Minnesota (47° 30.476'N; 93° 27.162'W; 418 m above mean sea level) (Kolka 154 et al., 2011). Situated at the southern extent of the spatially expansive boreal peatland forests, 155 156 this ecosystem is considered to be near its tipping point with respect to environmental change (Hanson et al., 2020). 157

158 The bog surface has a hummock/hollow microtopography, where the tops of hummocks are typically 10–30 cm higher than the bottoms of hollows (Shi et al., 2015). There are distinct 159 hydrologic dynamics and vegetation communities between the raised hummock and sunken 160 hollow microtopography characteristic of peatland bogs. The hummock/hollow microtopography 161 allows a greater aerobic rooting profile for shrubs and trees during the wet spring but limits water 162 availability to roots during summer drying periods. At SPRUCE, the primary vegetation types 163 include *Picea mariana* (black spruce) and *Larix laricina* (larch) trees with a mixed deciduous 164 and semi-evergreen Ericaceous shrub layer. The types of mixed shrub layer at SPRUCE mainly 165 166 include Rhododendron groenlandicum (Oeder) Kron and Judd (Labrador tea) and Chamaedaphne calyculata (L.) Moench. (leatherleaf). The ground layer vegetation beneath the 167 shrubs within the S1 bog also comprises a bryophyte layer dominated by Sphagnum spp. Air and 168 soil warming at five target levels (+0°C, +2.25°C, +4.5°C, +6.75°C, and +9°C) was conducted 169 across different enclosures under aCO<sub>2</sub> or eCO<sub>2</sub> (800 to 900 ppm) conditions. The whole-170 ecosystem warming began in August 2015, and eCO<sub>2</sub> treatments began in June 2016. 171

# 172 **2.2 Phenological and environmental measurements**

PhenoCams (i.e., high-resolution digital cameras) record seasonal variation in vegetation 173 "greenness" (i.e., a proxy for vegetation phenology and associated physiological activity) in each 174 SPRUCE enclosure (Richardson et al., 2018a). These cameras were installed at a height of 6 m 175 on the south wall of each enclosure and took images automatically every half hour. A reliable 176 metric of vegetation greenness, green chromatic coordinate ( $G_{CC}$ ), was derived from each image 177 for three vegetation types: Picea mariana, Larix laricina, and the mixed shrub layer (Richardson 178 et al., 2007; Sonnentag et al., 2012). G<sub>CC</sub> increases during spring, reaches its maximum in 179 summer, and decreases during autumn. The estimated transition dates of "greenness rising" in 180

spring and "greenness falling" in autumn were derived when the 3-day smoothed  $G_{CC}$  crosses its 25% seasonal amplitude threshold for each vegetation type from each enclosure every year (Schädel et al., 2019). In this study, phenological transition dates of *Larix laricina* and the mixed shrub layer from 2015 to 2018 were used to improve the seasonal-deciduous phenological models in the ELM.

186 Related environmental observations from the SPRUCE experiment have been detailed by

Hanson et al. (2017). Briefly, the half-hour air temperature was observed at the center of each 187 plot at 0.5, 1, 2, and 4 m above the surface of the peat. Half-hour belowground soil temperatures 188 were measured at three locations (central, middle, and edge) across each plot at the following 189 depths: 0, -0.05, -0.1, -0.2, -0.3, -0.4, -0.5, -1, and -2 m, where 0 m is the peatland hollow 190 191 height (low points in an undulating surface). In this study, we used spatially averaged soil temperatures simulated at -0.05 m (which is the third soil layer in the ELM) and the observed 2 192 m air temperature within each enclosure to stimulate phenology in the default and revised ELM 193 194 as described in the next section.

# 195 **2.3 Default phenological schemes in the ELM**

The default seasonal-deciduous phenology schemes in the ELM, which include spring onset and autumn senescence models, were adapted from the phenology algorithms designed for temperate deciduous broadleaf forest (Thornton et al., 2002; White et al., 1997). The spring onset occurs when the growing-degree-day accumulation  $(GDD_{sum})$  of the third-layer soil temperature  $(T_{s,3},$ in kelvin) initiated at the winter solstice exceeds a threshold  $GDD_{sum\_crit}$ .  $GDD_{sum\_crit}$  is determined by the annual average 2 m air temperature in the preceding year  $(T_{2m,ann_{avg}}, in$ 

202 kelvin).

$$GDD_{sum} = \sum (T_{s,3} - TKFRZ) , \qquad (1)$$

204 
$$GDD_{sum\_crit} = \exp(4.8 + 0.13 \times (T_{2m,ann_{avg}} - TKFRZ)), \qquad (2)$$

where *TKFRZ* is the freezing point of water (273.15 K). 205

Because plants across regions require different degrees of warming (i.e., GDD<sub>sum crit</sub> here) to 206 initiate growth, Eq. (2) of GDD<sub>sum crit</sub> was calculated as a function of mean annual temperature 207 in the preceding year to accommodate spatial phenology simulation at a broader scale (White 208 209 et al., 1997). However, this calculation may cause incorrect responses of spring phenology to warming (see Section 4). 210

The autumn senescence is simulated to occur when the daylength is shorter than 39,300 s (about 211 10.9 h). Because daylength is calculated as a function of latitude and day of the year (Forsythe et 212 213 al., 1995), the default autumn senescence date is temporally invariant for any given location.

203

#### 2.4 Revised phenological models 214

Different spring and autumn phenology models in previous studies have been widely 215

216 intercompared and evaluated (Cannell and Smith, 1983; Delpierre et al., 2009; Landsberg, 1974;

Migliavacca et al., 2008; Murray et al., 1989; Sarvas, 1972). For application to the ELM, the 217

alternating model for spring onset (Murray et al., 1989) and the CDD model proposed by 218

219 Delpierre et al. (2009) (hereafter termed "DM") for autumn senescence were selected because of

their mechanistic assumption, performance, and simplicity. 220

For the alternating model, the spring onset is simulated through winter chilling and spring 221

thermal forcing processes. The state of chilling  $(S_c)$  or forcing  $(S_t)$  is the time integral from  $t_0$ 222

(i.e., winter solstice here) of the rate of chilling  $(R_c)$  or forcing  $(R_f)$ .  $R_f$  is the relative daily

temperature above the base temperature ( $T_{\text{base}}$ ), and  $R_c$  is 1 if the daily mean temperature T(t) is below  $T_{\text{base}}$  and is otherwise 0.

226 
$$R_{f}(t) = \begin{cases} T(t) - T_{\text{base}} & T(t) > T_{\text{base}} \\ 0 & T(t) \le T_{\text{base}} \end{cases},$$
(3)

227 
$$S_f(t) = \sum_{t_0} R_f(T(t)),$$
 (4)

228 
$$R_c(t) = \begin{cases} 1 & T(t) < T_{\text{base}} \\ 0 & T(t) \ge T_{\text{base}} \end{cases},$$
(5)

229 
$$S_c(t) = \sum_{t_0} R_c(T(t))$$
. (6)

230 Spring leaf-out occurs when

231 
$$S_f(t) \ge a + b * \exp(c * S_c(t)),$$
 (7)

where c < 0. *a*, *b*, *c*, and  $T_{\text{base}}$  are parameters to be calibrated.

# 233 For the DM, the progress of the leaf senescence is represented through a CDD sum and a

- photoperiod decrease process. The coloring state  $(S_{sen})$  is the time integral at the rate of  $R_{sen}$
- starting when daily daylength P(t) is shorter than a photoperiod threshold  $P_{start}$ , which varies

with latitudes.  $R_{sen}$  is a function of temperature and photoperiod and only accumulates when the

daily mean temperature T(t) is below the base temperature  $T_b$ . The function of  $R_{sen}$  means a

stronger CDD sum effect under short photoperiod conditions relative to  $P_{start}$ .

239 
$$If P(t) < P_{start} \quad if \begin{cases} T(t) < T_b & R_{sen}(t) = (T_b - T(t)) * (1 - \frac{P(t)}{P_{start}}) \\ T(t) > T_b & R_{sen}(t) = 0 \end{cases},$$
(8)

$$S_{sen}(t) = \sum R_{sen}(t) . \tag{9}$$

Autumn senescence is simulated to occur when  $S_{sen}(t) \ge Y_{crit}$ .  $Y_{crit}$ ,  $P_{start}$ , and  $T_b$  are parameters to be calibrated.

The spring alternating model and autumn DM were first calibrated using the PhenoCam observations from 2016 to 2017 and 2015 to 2017, respectively, across all enclosures and were then evaluated using the PhenoCam data during 2018 (Table S1, S2). The model calibration was performed based on the minimum root mean square error (RMSE) between simulated phenology and observations (20 enclosure-years of data for spring onset and 30 enclosure-years of data for autumn senescence). Independent model calibration and evaluation were separately conducted for *Larix laricina* trees and the mixed shrub layer.

# 250 **2.5 Phenology comparison between PhenoCam and the ELM**

240

251 Because the phenological transition dates from PhenoCam and the ELM are defined differently, we adjusted the PhenoCam observations to make them directly comparable with the ELM 252 phenology outputs. In the ELM, the spring onset is the timing when leaf area index (LAI) starts 253 to increase from 0, and the autumn senescence is the timing when LAI begins to decrease from 254 its maximum value. Within a fixed 30-day period after spring onset, the ELM LAI increases to 255 its maximum value and stays the same until the autumn senescence occurs; within a fixed 15-day 256 period after the start of autumn senescence, the ELM LAI then decreases to 0. In contrast, the 257 spring green-up and autumn green-down dates from PhenoCam represent the timings 258 259 corresponding to 25% seasonal amplitude of greenness. Therefore, we adjusted the PhenoCam observations by subtracting 7.5 days ( $25\% \times 30$  days) from spring green-up dates and by 260 subtracting 11.25 days ( $(1\%-25\%) \times 15$  days) from autumn green-down dates, and we performed 261

the ELM evaluation and calibration processes using the updated observational timings. This adjustment is valid based on abundant previous studies showing that the temporal development of plant LAI was synchronous with PhenoCam  $G_{CC}$ , especially for deciduous forests, and they both represent canopy development (Liu et al., 2015; Cremonese et al., 2017; Peichl et al., 2015).

#### 266 **2.6 ELM-SPRUCE**

267 In this study, we used the SPRUCE version of the ELM, which was developed mainly for the ELM improvements and simulations at SPRUCE. New functions in this ELM version include 268 269 revised hydrological parameters, representation of hummock and hollow microtopography and lateral flows, the inclusion of more mechanistic CH<sub>4</sub> processes, and Sphagnum dynamics 270 271 (Hanson et al., 2020; Ricciuto et al., 2020; Shi et al., 2015). However, until this study, the 272 representation of seasonal deciduous phenology was unchanged from CLM4.5 (Oleson, 2013). For the ELM, one land grid comprises five primary units-glacier, lake, wetland, urban, and 273 vegetation—at different fractional coverages. Within each grid cell, the vegetated portion is 274 further divided into patches of plant functional types (PFTs), with each possessing unique 275 vegetation properties but sharing the same environmental forcings and soil conditions (Oleson, 276 2013). Thus, the ELM produces both grid- and PFT-level outputs, with the latter having 277 connected carbon, water, and energy summaries across different PFTs. At SPRUCE, four PFTs 278 coexisted within one ELM grid cell (i.e., enclosure here)-boreal needleleaf evergreen tree, 279 boreal needleleaf deciduous tree, boreal broadleaf shrub, and moss (Sphagnum). We only 280 focused on two PFTs that have seasonal-deciduous characteristics in the ELM: Larix laricina 281 trees (boreal needleleaf deciduous tree) and mixed shrub layer (boreal broadleaf shrub). Also, we 282 283 examined the phenology-relevant changes at the PFT and grid-cell levels for each enclosure from the original and improved ELMs. 284

We performed the ELM simulations using the default and revised phenological algorithms 285 (hereafter referred to as "ELM Default" and "ELM New," respectively). ELM New includes 286 the alternating model for spring onset and DM for autumn senescence introduced in Section 2.4. 287 We conducted ELM simulations for three phases: accelerated decomposition spin-up, final spin-288 up, and transient. Each spin-up generated an equilibrium state to estimate the relative content of 289 carbon, nitrogen, and phosphorus in different vegetation and soil pools. For the first part of the 290 transient phase, we conducted a run from 1850 to 2010 driven by the SPRUCE-observed cycling 291 of meteorology variables for 2011 to 2018 and the historical changes of other environmental 292 293 forcings at the ambient enclosures (e.g., atmospheric CO<sub>2</sub> concentrations, nitrogen deposition). Then, for the 2011 to 2018 period, we ran 10 ELM simulations (2  $CO_2$  concentrations  $\times$  5 294 warming levels) using the SPRUCE-manipulated temperatures and CO<sub>2</sub> concentrations that were 295 observed in the enclosures; other forcings were kept the same with transient for each ELM 296 version. We systematically intercompared the PFT-level model outputs from ELM Default and 297 ELM New in terms of spring onset, autumn senescence, and major carbon and water fluxes (i.e., 298 total LAI (TLAI), GPP, canopy transpiration (QVEGT), and canopy evaporation (QVEGE)) 299 across different warming and CO<sub>2</sub> levels during 2016 to 2018 at monthly, seasonal, and annual 300 301 scales. We also examined main carbon and water variables at the grid-cell level for the hummock and hollow surfaces. 302

# 303 2.7 Sensitivity analysis

Quantifying parameter uncertainty could reveal the impact of certain parameters on key carbon and water variables. We thus used a global sensitivity analysis (GSA) to examine the parameter uncertainty in ELM\_New, including phenology- and non-phenology-related parameters, on carbon and water variables. GSA, also called "variance-based decomposition", attributes

308 variations in predictions to specific model parameters and their interactions given ranges of possible values for the parameters. Following Sargsyan et al. (2014) and Ricciuto et al. (2018), 309 we used Polynomial Chaos surrogate models via a Bayesian compressive sensing approach to 310 conducting the GSA. A surrogate model is a functional representation of a model quantity of 311 interest (QoI) that is constructed from an ensemble of simulations from ELM New and allows 312 for further exploration of responses over the complex multi-dimensional parameter space. In this 313 case, we performed a total of 10 QoIs representing 5 model output variables averaged over the 314 2011 to 2018 period for each of the two deciduous PFTs: GPP, net primary productivity (NPP), 315 316 QVEGT, total vegetation carbon (TOTVEGC), and TLAI. For this GSA, we only considered the ambient case with no experimental temperature or  $CO_2$  treatments. Bayesian compressive 317 sensing allows the construction of surrogates using a limited number of simulations, leading to 318 computational cost reduction in studies that otherwise require an infeasible number of ELM 319 simulations, such as GSA or model tuning. We employed the Uncertainty Quantification Toolkit 320 for the Bayesian compressive sensing surrogate construction and GSA (Debusschere et al., 321 2015), available at https://github.com/sandialabs/UQTk. 322

323 We used 2,000 ELM New simulations to construct the surrogate model and conducted the GSA. 324 Although we only analyzed QoIs from the deciduous PFTs, we also varied the *Sphagnum* and Picea PFT parameters to access potential interactions with these types. Therefore, 44 model 325 parameters (some of them PFT-specific), including plant trait and soil parameters detailed in Shi 326 327 et al. (2020) and 7 new phenology parameters, were varied as uniform random variables over their respective reasonable ranges (Table S3). Main effect indices were calculated for each 328 parameter, accounting for individual fractional contributions toward the overall QoI variance. 329 330 These sensitivity indices do not provide information about whether a parameter has a positive or

negative effect on a QoI. Therefore, using the same 2,000 model samples, we also derived the
linear correlation coefficient between each parameter and each QoI to further understand the
impacts of individual parameters.

334 **3 Results** 

#### **335 3.1 Phenological evaluations and improvements**

Compared with ELM Default, ELM New-which includes phenology parameters calibrated 336 against the PhenoCam observations—showed substantial improvement in predicting phenology 337 and the temperature responses of phenology for both Larix laricina and the shrub layer. For the 338 simulated spring onset, ELM New showed a higher correlation (r) with PhenoCam observations 339 (e.g., 0.34 of *Larix laricina* and 0.50 of the shrub layer) compared with that from ELM Default 340 (corresponding to 0.18 and 0.41) based on the independent model evaluation in 2018 (Table S2). 341 The RMSE of spring onset between ELM Default simulations and PhenoCam observations for 342 343 *Larix laricina* was 26 days, which decreased to 8 days when using ELM New (Table S2). However, for the shrub layer, the RMSE between simulated and observed spring onset was lower 344 for ELM Default (3.4 days) than for ELM New (7.4 days), likely because of the unusual 345 346 weather conditions in 2018. The spring frost in early April 2018 and the following abrupt increase in temperature triggered spring onset to occur later than normal but quickly across all 347 enclosures (standard deviation of spring onset across all enclosures was 3.8 days in 2018 348 compared with 6.8 days in 2016 and 2017, Fig. S1). This appears to be an abnormal phenological 349 350 response to various warming levels across enclosures, coincidently better simulated by ELM Default than ELM New. 351

For the simulated spring onset during 2016 to 2018, ELM Default (day of the year:  $137 \pm 4$ ) was 352 biased late for both *Larix laricina* (36 days) and the shrub layer (9 days) (Fig. 1a). In contrast, 353 the spring onset simulated by ELM New (day of the year:  $106 \pm 9$  for *Larix laricina* and  $130 \pm 7$ 354 for the shrub layer) was only 5 and 2 days later than the PhenoCam observations (left subfigures 355 in Fig. 1a). In terms of the temperature responses of spring onset, ELM New (-1.2 days/°C for 356 *Larix laricina* and -1.0 days/°C for the shrub, P < 0.1) well captured the observed advancement 357 of spring onset responding to warming (-1.0 days/°C and -1.5 days/°C, P < 0.1), whereas 358 ELM Default failed to capture phenology trends associated with warming (P > 0.1) (right 359 360 subfigures in Fig. 1a), consistent with previous studies (Chen et al., 2016).



PhenoCam <->
ELM\_Default <->
ELM\_New

Fig. 1. Observed and simulated responses of spring onset (a) and autumn senescence (b) to warming at SPRUCE. Linear regression lines are shown as dashed lines. The mean phenology across all warming and CO<sub>2</sub> levels and slopes of phenology against warming levels are shown in the subfigures. The error bars in the subfigures represent the standard deviations of phenology across all warming levels. Significance P < 0.1 from two-tailed Student's *t* test. Spring onset was during 2016 to 2018, and autumn senescence was during 2015 to 2018.

368

361

369 ELM\_Default simulated the same autumn senescence across the two vegetation types (day of

year:  $287 \pm 0$ ), with 6 days later for *Larix laricina* but 8 days earlier for the shrub layer than the

PhenoCam, and produced no phenological responses to warming during 2015 to 2018 (Fig. 1b).

In contrast, ELM\_New (day of year:  $284 \pm 10$  for *Larix laricina* and  $298 \pm 15$  for the shrub

layer) simulated autumn senescence only 3 days later than that of PhenoCam (day of year:  $281 \pm$ 

8 and 295  $\pm$  15) for both *Larix laricina* and the shrub layer. More notably, ELM\_New (1.72)

days/°C for *Larix laricina* and 1.74 days/°C for the shrub layer, P < 0.1) captured comparable

delay trends under warming with the PhenoCam (1.88 days/°C and 4.16 days/°C, P < 0.1), albeit

377 with underestimated magnitude for the shrub layer.

# 378 **3.2** Phenological effects on PFT-level carbon and water fluxes

379 Mean seasonal cycles of carbon and water fluxes from both ELM Default and ELM New were 380 similar in general—TLAI, GPP, and QVEGT peaked during summer, whereas canopy 381 evaporation (QVEGE) showed two peaks in May and October (Figs. S2, S3). Both versions also 382 simulated similar responses of carbon and water fluxes to temperatures and CO<sub>2</sub>—TLAI, GPP, 383 QVEGE, and QVEGT showed higher values at warmer levels, and all but QVEGT showed higher values at eCO<sub>2</sub> than aCO<sub>2</sub> conditions (Fig. S2). However, for *Larix laricina*, ELM New 384 simulated significantly higher annual TLAI, GPP, QVEGE, and QVEGT than those of 385 ELM Default (P < 0.1, Fig. S4a). Such high fluxes occurred mainly during the growing season 386 (i.e., positive difference, represented by green, Fig. 2a), indicating an increase in ecosystem 387 productivity and water fluxes through the extended period of active carbon uptake induced by the 388 early onset of spring. The magnitude of differences between ELM Default and ELM New in 389 TLAI (ΔTLAI), GPP (ΔGPP), QVEGE (ΔQVEGE), and QVEGT (ΔQVEGT) was the largest 390 during spring transition periods, especially at higher warming levels and eCO<sub>2</sub> conditions (Fig. 391 2a).  $\Delta$ TLAI was significant during most of the growing season, whereas  $\Delta$ GPP,  $\Delta$ QVEGE, and 392  $\Delta$ QVEGT were evident only during spring transition periods and a few other months at the 393 highest warming levels (P < 0.1, Fig. 2a). The largest phenology scheme–induced differences 394 occurred in May at the 9°C warming and eCO<sub>2</sub> conditions—5.27 for  $\Delta$ TLAI (P < 0.1), 200.7 395

396 gC/m<sup>2</sup>/month for  $\Delta$ GPP (P < 0.1), 9.5 mm/month for  $\Delta$ QVEGE, and 35.5 mm/month for 397  $\Delta$ QVEGT (P < 0.1) (Fig. 2a). ELM\_New simulated lower TLAI, GPP, and QVEGT in October 398 and lower QVEGE in November and December than those of ELM\_Default at relatively low 399 warming levels (i.e., negative difference, represented by brown, Figs. 2a and S4a), indicating an 400 early reduction of land flux caused by earlier autumn senescence from ELM\_New.



401 402

Fig. 2. Multi-year mean differences in simulated monthly TLAI, GPP, QVEGE, and

403 QVEGT between ELM\_New and ELM\_Default for *Larix laricina* (a) and the shrub layer
404 (b) during 2016 to 2018. Green represents higher values by ELM New, brown represents lower

values by ELM New, and gray represents no differences between ELM New and ELM Default.

- 405 Two pixels are in each grid; left and right pixels represent values under  $eCO_2$  and  $aCO_2$
- 407 conditions, respectively. Stars indicate significant differences between ELM New and
- ELM Default based on a two-tailed Student's *t* test (P < 0.1).
- 409
- For the shrub layer, the higher TLAI, GPP, and QVEGT simulated by ELM\_New occurred only
- 411 during the spring and autumn transition periods, and the higher QVEGE occurred during the
- transition periods and non-growing season under 2.25°C or higher warming levels (Fig. 2b). This

indicates the extended growing season of the shrub layer only extended the length of the active 413 land flux period but did not change the magnitude of growing-season land flux, which is 414 different from that of Larix laricina. In contrast, within the 0°C warming enclosures, the above 415 variables were lower from ELM New during the growing season. This corresponds to the 416 phenological changes in Fig. 1 (see Section 4). Such increased TLAI, GPP, QVEGE, and 417 QVEGT by ELM New were more significant at the warmer enclosures (P < 0.1). The largest 418  $\Delta$ TLAI (3.1) occurred in November, whereas the largest  $\Delta$ GPP (157.2 gC/m<sup>2</sup>/month),  $\Delta$ QVEGE 419 (3.1 mm/month), and  $\Delta QVEGT$  (30.3 mm/month) occurred in May at the 9°C warming level (P 420 < 0.1). The magnitudes of change in TLAI, GPP, QVEGE, and QVEGT were smaller for the 421 shrub layer than those for *Larix laricina*. 422

423 ELM\_New also simulated closer seasonality of TLAI and GPP to the seasonality of  $G_{CC}$ 

424 compared with ELM\_Default (normalized value, Figs. S5, S6). For example, in spring 2017, the

timing of the increase in TLAI and GPP for *Larix laricina* simulated by ELM\_New was

426 concurrent with the timing of the increase in  $G_{CC}$ , whereas that by ELM\_Default was much later 427 (Fig. S5).

# 428 **3.3** Temperature responses of carbon and water fluxes at the PFT level

The enhanced phenological responses to warming strengthened the responses of carbon and water fluxes to warming simulated by ELM\_New. Although both model versions generally produced stronger responses under the eCO<sub>2</sub> than the aCO<sub>2</sub>, ELM\_New simulated overall stronger responses of annual TLAI, GPP, QVEGE, and QVEGT to warming than those by ELM\_Default (Fig. 3). For *Larix laricina* at the eCO<sub>2</sub>, the temperature responses of TLAI and GPP by ELM New were 4.5 and 17 times stronger, respectively, than by ELM Default, though

435 the GPP response by ELM New at the aCO<sub>2</sub> was weaker (Fig. 3a). The responses of OVEGE and QVEGT to warming were strengthened approximately 2 and 1.5 times, respectively, by 436 ELM New compared with ELM Default. For the shrub layer, the temperature responses of 437 carbon and water fluxes showed similar patterns to those of Larix laricina but with weaker 438 magnitudes except for GPP (Fig. 3b). Such enhanced responses of TLAI and GPP to warming by 439 ELM New are also shown in Figs. S5 and S6; the time series of TLAI and GPP by ELM New 440 were more dispersed across warming levels during the spring and autumn transition periods, 441 which better matched  $G_{CC}$ . 442



443

Fig. 3. Temperature responses of annual TLAI, GPP, QVEGE, and QVEGT simulated by ELM\_Default and ELM\_New for *Larix laricina* (a) and the shrub layer (b). We used annual average TLAI ( $m^2/m^2$ ) and GPP (gC/ $m^2/year$ ), and annual sum QVEGE (mm/year) and QVEGT (mm/year). The temperature responses were calculated as the slope of the linear regression between annual TLAI, GPP, QVEGE, and QVEGT and warming levels. Stars indicate significance from linear regression (P < 0.1). Results for each season are shown in Figs. S7 and S8.

452 The temperature responses of the selected land surface variables displayed distinct patterns

- 453 across seasons (Figs. S7, S8, no value in winter). During the spring and autumn, ELM\_Default
- simulated small responses of TLAI and GPP to warming, but ELM\_New simulated prominent

responses for both *Larix laricina* (Fig. S7) and the shrub layer (Fig. S8). The differences in the
annual response of QVEGE and QVEGT simulated by ELM\_Default and ELM\_New (Fig. 3)
primarily came from spring, where, for example, ELM\_New (-0.93 mm/season/°C) simulated
383% higher QVEGE than that of ELM\_Default (-0.24 mm/season/°C) at eCO<sub>2</sub> for *Larix laricina* (Fig. S7). In summer, the temperature responses of the selected land surface variables
were very similar.

The magnitudes of  $\Delta$ TLAI and  $\Delta$ GPP between ELM New and ELM Default were well 461 correlated with the difference magnitudes in phenology (Fig. S9). The  $\Delta$ TLAI and  $\Delta$ GPP in May 462 were positively correlated with the differences in spring onset between ELM New and 463 ELM Default for the shrub layer (P < 0.01, Fig. S9b). However, for *Larix laricina*, only  $\Delta$ TLAI 464 at the eCO<sub>2</sub> showed a positive correlation (P = 0.02, Fig. S9a). Likewise, the difference 465 magnitude in autumn senescence was positively correlated with the magnitude of  $\Delta$ TLAI and 466  $\Delta$ GPP in October (Fig. S9c, d). This corresponds to Figs. 1 and 2; the earlier autumn senescence 467 468 at low warming levels by ELM New led to reduced magnitudes in TLAI and GPP in autumn. 469 However, the later senescence at high warming levels by ELM New caused increased 470 magnitudes in TLAI, GPP, QVEGE, and QVEGT. The responses of TLAI and GPP to changes 471 in phenology were more prominent in  $eCO_2$  than in  $aCO_2$  and for *Larix laricina* than the shrub layer (higher slopes). Interestingly, for the shrub layer, the responses of TLAI and GPP to 472 changes in spring onset were much stronger than their responses to changes in autumn 473 474 senescence, but opposite for *Larix laricina*. The enhanced temperature responses of spring onset and autumn senescence (Fig. 1) also contributed to the larger temperature responses in TLAI, 475 GPP, QVEGE, and QVEGT by ELM New (Fig. 3). 476

#### 477 **3.4** Phenological feedbacks to carbon and water fluxes at the grid-cell level

Grid average carbon and water fluxes showed distinct changes simulated by ELM Default and 478 479 ELM New (Fig. 4). Such grid-cell-level patterns reflect the combined responses of all PFTs— Larix laricina trees (boreal needleleaf deciduous tree), mixed shrub layer (boreal broadleaf 480 shrub), Picea mariana (boreal needleleaf evergreen tree), and Sphagnum (moss)-among which 481 the phenology schemes of the former two were revised and the latter two remained the same in 482 ELM New compared to ELM Default. ELM New simulated higher magnitudes of seasonal and 483 annual grid-cell NPP, ecosystem respiration (ER), GPP, TLAI, and ecosystem evapotranspiration 484 (ET) than ELM Default for both hummock and hollow (Fig. 4). The ELM New net ecosystem 485 production (NEP) was also higher than that of ELM Default during spring (e.g., 245% higher at 486 aCO<sub>2</sub>) and autumn but lower during winter and summer (e.g., 47% lower at aCO<sub>2</sub>) for hummock. 487 Runoff, water table depth (ZWT), and volumetric soil water (H2OSOI) showed various degrees 488 of decrease simulated by ELM New for both hummock and hollow, and the magnitude of 489 490 decrease was larger for hummock than hollow. The relative changes in ZWT and H2OSOI were opposite-ELM New simulated higher H2OSOI but lower ZWT in winter and summer, and 491 492 reversed patterns for spring and autumn.

	(a) Hummock														(b) Hollow										
	Mean(%) Response(%)													Me	ean	(%)			R	lesp	ons	e(%	6)		
	0	100	200	300	400	-4	00 -20	0 0	200	400	600 80	0 -4	10 -20	0 2	20 40	60	, 80 10	0	- 1000	0	10	00 2	:000		
	0	10.4	-1.5	15	-36		-13.2	-12.2	-13.4	-13.1	-12.5		1.2	4.1	0.9	5.3	-36.8		-38.7	-7.2	-41.1	-8.3	-4.4		
NEP	-2.9	4.3	-4.6	4.3	20.7		-17	-16.7	-18.1	-17.2	-16.7		1.4	0.4	1.2	-0.3	-1.3		-10.9	-8.8	-12.7	-9.4	-8.6		
	3.2	244.8	1.8	-46.5	-8.2		-6	-9	-4.4	-10.2	-10		1.1	77.7	0.5	-35.5	-7.3		-87.5	-6.1	-88.5	-7.6	-1.9		
•	17.9	2.9	17.9	2.8	4.3		16.6	-10.6	17.6	-10.8	-27.2		18.1	1.2	17.2	0.9	2.4		14.3	-13.7	14.2	-14.6	-33.4		
NPF	21.8	2.2	22.1	2	3.7		17.7	-16.9	19.1	-16.9	-70.5		21.2	0.4	20.5	0.1	1.7		12	-23.1	11.4	-23.9	-154.1		
	13.9	4	13.7	4	5.2		15	-7.1	15.6	-7.6	-14.7		14.8	2.4	13.7	2.2	3.5		17.5	-9.6	18.2	-10.7	-19		
	5.3	2.2	5.2	2.2	2.8		53.2	8.9	51.3	8.3	23.4		6.5	2.3	6.7	2.3	3.1		52.4	12.4	50.8	11.5	23.6		
R	4.6	2.5	4.5	2.5	2.9		58.7	10.5	57.5	10.5	25.5		6.6	2.7	6.9	2.7	3.4		39.2	14.1	37.9	14.1	21.6		
	6.2	1.7	6	1.7	2.7		46.2	6.2	43.6	5	20.2		6.4	1.8	6.4	1.7	2.7		72.8	10	70.2	8.2	26.3		
	8.5	2.8	8.9	2.8	3.6		28.5	-40.6	29.5	-43.8	746.3		9.2	2.4	9.5	2.4	3.3		29.5	-52.2	29.4	-56.8	310.2		
Ч	8.9	2.8	9.4	2.7	3.6		27.9	-189.2	29.1	-210.9	84.9		9.2	2.4	9.6	2.4	3.2		26.1	2646.4	25.8	1623.1	64		
0	8.1	2.9	8.3	2.9	3.7		29.3	-17.8	30.1	-19.8	-70		9.3	2.5	9.4	2.5	3.5		34.9	-21.6	35	-24.2	-80		
	6.9	5.8	7.7	5.8	6.3		-210.8	-375.9	-225.5	-360	-274		6	4.1	6.8	4.1	4.9		-283.3	1081.2	-310.2	1387.9	-750.7		
[A]	6.8	5.7	7.5	5.7	6.2		466.4	130.1	463.4	136.7	204.4		5.8	4	6.6	3.9	4.7		249.1	66.5	249.3	69.4	109.9		
н	7.1	6	7.9	6	6.5		-83.6	-68.4	-87.7	-69.5	-76.6		6.2	4.3	7	4.2	5.1		-85.5	-76.6	-90.2	-77.8	-82.2		
н	-0.1	-3.6	0.3	-3	-1.1		19.2	-5.3	11.2	-2.5	4.4		0	-0.7	0	-0.7	-0.3		-229.2	3.7	-268.8	4.4	13.7		
NOF	0.7	-2.9	0.1	-2.8	-0.8		8.5	-1.5	8.5	-0.7	2.9		0.1	-0.7	0.2	-0.9	-0.3		-135.5	-0.6	-142.9	0.7	17.8		
RU	-0.9	-4.4	0.5	-3.4	-1.4		30.9	-10.2	13.8	-5	6.1		-0.1	-0.6	-0.1	-0.6	-0.4		93	7.7	74.8	7.9	10.3		
	1.5	0.8	1.8	1	1.1		1.3	4.1	3.2	4.2	3.6		1.2	0.7	1.4	0.8	0.8		5.4	3.3	5.1	4	4		
늡	1.4	0.8	1.8	0.9	1		0.7	3.6	2	3.8	3		0.9	0.6	0.8	0.8	0.7		6.5	2.5	3.5	2.9	3.2		
_	1.5	0.9	1.9	1	1.1		1.9	4.6	4.5	4.7	4.2		1.4	0.7	1.9	0.8	0.9		4.3	4.3	6.7	5.2	4.9		
	-1.8	3.4	-1.7	3.4	1.6		39.8	1.7	37.7	1.9	12.5		0	3.8	0	4.3	2.9		-127	5.3	-146.7	6.1	14.4		
WT	-1.8	3.8	-1.8	3.9	1.9		14.2	2	12.8	1.9	5.5		-1.7	4.5	-1.8	5.2	2.9		-107	1.1	-116.9	2.3	17.8		
Ν	-1.8	3	-1.7	3	1.3		74.9	1.5	71	1.9	20.3		1.5	3.3	1.5	3.5	2.8		1661.7	9.2	208.6	9.6	11.6		
Ξ	0.5	-1.6	0.5	-1.6	-0.4		21.1	-4.4	20.2	-4.6	7.2		-0.2	-0.9	-0.2	-0.9	-0.5		-447.6	2.2	1283.3	2.8	9.1		
osc	0.5	-1.8	0.5	-1.8	-0.5		9.5	-4.9	8.7	-5.2	1.5		0	-1	0	-1.1	-0.5		-90.7	-2.3	-125.9	-1.5	4.8		
H2	0.5	-1.4	0.5	-1.4	-0.3		35.1	-3.8	34	-3.9	13.9		-0.3	-0.7	-0.3	-0.7	-0.5		168.3	6.8	99.9	7.2	12.9		
	er	b	er	lle	a		er	b	er	IE	ਯ		er	g	er	lle	a		er	b	ъ	IE	a		
	/int	prir	Ű	ш	nuc		/int	prir	Ш	ш	nu		/int	prin	ũ	Ц	nu		/int(	prin	ŭ	Ц	nu		
	\$	S	Sun		A		\$	S	Sun		Ar		\$	S	Sun		Ar		\$	S	Sun		Ar		



Fig. 4. Relative changes in mean and temperature responses in carbon and water fluxes of
hummock (a) and hollow (b) simulated by ELM\_New compared with ELM\_Default.
Relative changes = (ELM\_New - ELM\_Default) / ELM\_Default × 100%. Mean carbon and
water fluxes represent the average values across all warming levels, and temperature responses
represent the changes in carbon and water fluxes per degree warming. Green represents a higher

499 value or stronger response simulated by ELM\_New, and brown represents lower values or

500 weaker response by ELM\_New. For each variable, each of the three rows within each grid

- represents the mean under  $eCO_2$  and  $aCO_2$  condition,  $eCO_2$  condition, and  $aCO_2$  condition across 502 all warming levels during 2016 to 2018.
- 503

504 In terms of the responses of carbon and water fluxes to warming (i.e., changes in carbon and 505 water fluxes per degree warming), TLAI showed the largest changes for both hummock and 506 hollow compared with other variables, indicating evident impacts of phenology on TLAI at the 507 grid-cell level. The temperature response of NEP decreased, whereas those for ER, runoff, ET, 508 and ZWT generally increased for the hummock and hollow simulated by ELM New. The 509 temperature responses of NPP and GPP showed similar seasonal changes-the responses were 510 511 stronger in winter and summer but weaker in spring and autumn. The seasonal changes in runoff, ZWT, and H2OSOI were opposite for hummock and hollow, with larger changing magnitudes 512 for hollow. These findings indicate significant interactions among individual PFTs within a grid, 513 as well as prominent phenological impacts of the deciduous PFTs on carbon and water fluxes at 514 the grid-cell level. 515

#### 516 3.5 Sensitivity analysis of ELM New parameters

To assess the impact of phenology uncertainty on ELM simulations of carbon and water cycling, 517 we conducted a GSA of 10 model outputs to 44 parameters in ELM New (Fig. 5). Out of the 44 518 total parameters (Table S3), only 12 parameters, including 3 phenology parameters, were 519 responsible for more than 1% of the uncertainty (i.e., sensitivity > 1%, shown in Fig. 5) for at 520 521 least one of the examined carbon and water variables. For *Larix laricina*, the responses were dominated by the Larix fraction of leaf nitrogen in RuBisCO (flnr Larix), the Larix-specific leaf 522 area at the top of the canopy (*slatop Larix*), the *Larix* fine root to leaf allocation ratio 523 524 (froot leaf Larix), the phen tbfall (Tbase in equation 3), and phen c (c in equation 7) parameters. The leaf and fine root parameters were responsible for the majority of output 525

variance, whereas the phenology parameters contributed more than 20% of the variance for
QVEGT and TLAI. In particular, the base temperature for the autumn phenology parameter
(*phen\_tbfall*) was sensitive, indicating the importance of quantifying the temperature effects on
senescence accurately. Higher base temperatures were associated with smaller values of all
carbon and water variables, reflecting the control of that parameter over growing season length
(higher base temperature means earlier senescence given the same meteorological conditions).





- For the shrub layer, the responses were found to be dominated by the shrub flnr (*flnr\_Shrub*), the
- shrub stem to leaf allocation ratio (*stem\_leaf\_Shrub*), the shrub fine root to leaf allocation

544 (froot leaf Shrub), the shrub Ball-Berry stomatal conductance slope term (*mbbopt Shrub*), and the *phen tbfall* and *phen c* parameters. Perhaps unsurprisingly, parameters associated with the 545 *Picea* and *Sphagnum* were not identified as sensitive for the carbon and water variables of 546 deciduous conifer and shrub, indicating that belowground competition for water and nutrients in 547 the ELM among PFTs was not strongly affecting these QoIs. Overall, the two phenology 548 parameters (*phen tbfall* and *phen c*) contributed about 10% of the QoI variance. For both PFTs, 549 the high sensitivity of the *phen* c parameter indicates the importance of representing and 550 accurately parameterizing chilling processes in the phenology sub-model. 551

#### 552 4 Discussion

With the improved phenology processes incorporated, ELM New demonstrated several 553 554 advantages over ELM Default. ELM Default without a chilling process showed little response of spring onset to warming, whereas ELM New with a chilling process showed strong 555 temperature sensitivity, which better matched the observations. This might be different from the 556 common understanding that a chilling requirement mediates the temperature sensitivity of spring 557 phenology (Chuine et al., 2016) because the spring onset simulated by ELM Default occurs 558 when the accumulated temperature reaches a threshold, which is determined by the annual 559 average temperature in the previous year (Eq. 2). Such an algorithm was originally developed to 560 consider spatial heterogeneity of thermal requirement but may lead to a higher threshold of 561 thermal requirement at warmer enclosures in this study. Consequently, even though growing-562 degree-day accumulated faster at warmer enclosures, the net result was a similar spring onset 563 across different temperature treatments; this was also discussed in Chen et al. (2016) from a 564 565 temporal perspective of climate warming. In contrast, the earlier spring onset in warmer

566	enclosures was better reproduced by ELM_New because of the improved mechanistic processes
567	represented by the accumulation of both winter chilling and spring thermal forcing (Eqs. 3–7).
568	In terms of the autumn senescence, ELM_Default is solely based on a daylength threshold,
569	which is partially supported by much evidence that photoperiod is the dominant cue for growth
570	cessation in temperate and boreal trees (Keller et al., 2011; Way and Montgomery, 2015).
571	However, such simplification could cause obvious biases in simulations of phenology and
572	phenological response to warming (Delpierre et al., 2009), requiring more realistic and process-
573	based improvement. In this study, ELM_New brought ELM_Default one step further toward this
574	improvement by considering the temperature regulation for autumn senescence. Such
575	temperature control was shown to be important at SPRUCE across different treatment levels
576	(Fig. 1) and was also confirmed by recent studies indicating that warming could significantly
577	delay the autumn senescence (Heide, 2011; Richardson et al., 2018b). Overall, ELM_New
578	showed earlier spring onset across all warming levels and earlier autumn senescence only at
579	lower warming levels but later autumn senescence at higher warming levels compared with
580	ELM_Default. Such changes caused by phenological model selection may have divergent effects
581	on carbon and water cycles at different warming levels as shown by this study (e.g., Figs. 2 and
582	S4) and previous ones (e.g., Kim et al., 2018; Piao et al., 2019b).

583 Our modeling results demonstrated significant phenological impacts on multiple processes 584 within the ELM. The most direct and considerable phenological effects identified to occur were 585 for LAI in terms of the seasonality and magnitude during the growing season. Such altered LAI 586 further induced a series of changes in the ELM vegetation-soil system, including enhanced 587 photosynthesis and water fluxes (Figs. 2 and S4), largely consistent with observational responses 588 of carbon and water cycles to phenology across different scales (Piao et al., 2007; Zeng et al.,

589 2017). Because the rate of soil carbon decomposition increases concurrently with warming, such increased vegetation activities may not necessarily lead to a strengthened net land carbon sink 590 (Piao et al., 2019b). In this study, for example, the increase in ecosystem respiration exceeded 591 that of photosynthetic carbon gain and caused the NEP decrease in both winter and summer 592 seasons for the hummock (Fig. 4). In agreement with previous empirical findings (Kim et al., 593 2018; Lian et al., 2020; Zeng et al., 2018), our results also showed notable influences of plant 594 phenological shifts on the water cycle, such as increased ET, lower soil moisture, and modified 595 runoff and ZWT during the growing season, especially under the eCO<sub>2</sub> conditions (Fig. 4). 596 Compared with ELM Default, ELM New produced enhanced photosynthesis, increased ET, and 597 lower soil moisture for plants only at higher warming levels, but such effects reversed at lower 598 warming levels because of earlier autumn senescence. At the grid-cell level, simulated carbon 599 and water fluxes were mainly affected by the two modified PFTs (Larix laricina and the shrub 600 layer), but also by the other two unmodified ones, Picea mariana and Sphagnum. Within an 601 ELM grid-cell, different PFTs receive identical environmental forcings and share the same soil 602 603 conditions. Therefore, the modifications in phenology of Larix laricina and the shrub layer could alter the belowground soil characteristics, which further induced the carbon and water flux 604 605 changes of *Picea mariana* and *Sphagnum* (results not shown). When applying ELM New onto the studies of land-atmosphere feedbacks under future climate change, the phenological 606 influences by other environmental factors ought to be considered. For example, if droughts 607 become more pervasive or less water is available in late summer, the growing season length may 608 be reduced (because of earlier leaf senescence) instead of being lengthened, likely leading to a 609 decline in growing-season carbon fixation and ET (Lian et al., 2020). 610

611 The ELM phenology scheme will continue to be developed, including representation of phenology responses to extreme events such as spring frost and droughts and optimization of 612 more phenological parameters over various vegetation types. For example, the unusually warm 613 March followed by severe spring frost in the extreme cold April in 2016 at SPRUCE caused 614 earlier spring onset and severe tissue mortality of *Larix laricina* in the warmest enclosures 615 (Richardson et al., 2018b). Such processes are completely lacking but urgently needed in most 616 land surface models, including the ELM, and are increasingly critical as extreme events become 617 more frequent and intense. Proper simulation of the impacts from extreme events on phenology 618 619 changes could be achieved by, for example, adding additional physiological limitations (e.g., hydraulic dynamics, freezing tolerance, impacts on foliar morphology) to current phenological 620 algorithms (Jensen et al. 2020). Moreover, the evergreen phenology scheme in the ELM, which 621 is currently active during the whole year, needs to be updated in the future. A recent study 622 revealed clear seasonal changes in the color and photosynthetic activity of evergreen conifer 623 forest using the PhenoCam observation of *Picea mariana* (Sevednasrollah et al., 2020). Based on 624 these new findings, adding spring and autumn phenological transitions to a certain percentage of 625 evergreen conifer canopy could improve the ELM evergreen phenology scheme. 626

ELM\_New and associated site-level parameters could be upscaled onto broad scales through
linking the SPRUCE observations with other observational data sets and satellite images.
Regional phenological data from observational networks, such as the Pan European Phenology
Observation and USA National Phenology Network, could be examined in the future to verify
ELM\_New and improve its capability in simulating phenology across various climate and
vegetation zones (Meng et al., 2021). Different upscaling strategies and methods, such as
merging local phenology observations with high-resolution satellite products (e.g., NASA's

634 MSLSP30NA Land Surface Phenology product derived from Harmonized Landsat Sentinel-2 data), may provide new insight into the modeling of large-scale phenology responses to a 635 changing climate (Pope et al., 2013; Bolton et al., 2020). Caution should be taken for the 636 upscaling process, especially for the SPRUCE shrub types. The mixed shrub layer at SPRUCE 637 mainly includes Rhododendron groenlandicum (Oeder) Kron and Judd (Labrador tea) and 638 Chamaedaphne calyculata (L.) Moench. (leatherleaf); the latter (occurring in cool temperate and 639 subarctic regions from North America to Finland and Japan) is more representative than the 640 former (occurring only in North America) for the shrubs in northern high latitudes. However, 641 642 current PhenoCam signals for deciduous shrubs are merely a mix of these two types, thus limiting the clean upscaling for individual shrub species. 643

644 Sensitivity analysis in this study showed that parameter uncertainty of the phenology model significantly drove the prediction uncertainty of carbon and water cycles; this suggests that 645 further regional-scale observations across vegetation types and model developments remain 646 647 necessary to reduce uncertainties regarding chilling effects on spring phenology and autumn 648 temperature constraints on fall senescence. The extending time series of phenology and 649 environment observations at SPRUCE that cover a wider range of conditions over time will be 650 helpful to further reduce the model uncertainties. Future model evaluation and development 651 processes may be also guided by more comprehensive uncertainty quantification studies, 652 including examining the sensitivities of simulated carbon, water, and nutrient cycling to 653 phenological responses under extreme events, such as spring frosts and freezes, droughts, and wildfires. 654

### 655 **5 Conclusions**

Using the latest experimental observations from SPRUCE, this study evaluated and improved 656 657 seasonal-deciduous phenology components of ELM and investigated possible phenological feedbacks to major carbon and water fluxes. Selected mechanism-based phenology models of 658 spring onset and autumn senescence were introduced, in which the timing of spring onset 659 depends on both winter chilling and spring thermal forcing processes, whereas the timing of 660 autumn senescence relies on the co-limitation of daylength and temperature. In contrast to the 661 default ELM, the revised model better captured the PhenoCam observations (i.e., distinct 662 seasonal phenology of *Larix laricina* and the shrub layer and linearly extended growing season 663 length in response to warming). In addition, the improved ELM produced intensified carbon and 664 water fluxes associated with a longer growing season and stronger temperature response of 665 vegetation, especially under eCO<sub>2</sub> and warmer conditions. A model sensitivity analysis further 666 indicated that phenology parameters contributed significantly to simulated carbon and water 667 668 cycle variations over interannual timescales. This practice of model-experimental coupling highlights the importance of phenological processes in affecting complex terrestrial-climate 669 670 interactions, and it facilitates the uncertainty reduction of E3SM in predicting Earth dynamics at 671 broad spatiotemporal scales.

### 672 Acknowledgments, Samples, and Data

673 This work was supported by the Terrestrial Ecosystem Science Scientific Focus Area project

- funded by the US Department of Energy, Office of Science, Office of Biological and
- 675 Environmental Research. Oak Ridge National Laboratory is supported by the Office of Science
- of the US Department of Energy under Contract No. DE-AC05-00OR22725. This research used
- 677 resources of the Compute and Data Environment for Science (CADES) at Oak Ridge National
- 678 Laboratory. L. Meng is also supported by NASA FINESST Program (80NSSC19K1356). ADR
- acknowledges support from the NSF (DEB 1702697).

**Data and materials availability:** PhenoCam datasets and environmental measurements

681 pertaining to this study are in the online project archive at http://mnspruce.ornl.gov and for long-

term storage in the US Department of Energy's Environmental Systems Science Data

683 Infrastructure for a Virtual Ecosystem (ESS-DIVE; http://ess-dive.lbl.gov/).

#### 684 **References**

- Bolton, D.K. et al., 2020. Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2
   imagery. Remote Sensing of Environment, 240, 111685.
- Caffarra, A., Donnelly, A. and Chuine, I., 2011. Modelling the timing of Betula pubescens budburst. II. Integrating
   complex effects of photoperiod into process-based models. Climate Research, 46(2): 159–170.
- Cannell, M. and Smith, R., 1983. Thermal time, chill days and prediction of budburst in Picea sitchensis. Journal of
   applied Ecology: 951–963.
- Cesaraccio, C., Spano, D., Snyder, R.L. and Duce, P., 2004. Chilling and forcing model to predict bud-burst of crop
   and forest species. Agricultural and Forest Meteorology, 126(1–2): 1–13.
- Chen, M., Melaas, E.K., Gray, J.M., Friedl, M.A. and Richardson, A.D., 2016. A new seasonal- deciduous spring
   phenology submodel in the Community Land Model 4.5: impacts on carbon and water cycling under future
   climate scenarios. Global Change Biology, 22(11): 3675–3688.
- Chuine, I. and Régnière, J., 2017. Process-Based Models of Phenology for Plants and Animals. Annual Review of
   Ecology, Evolution, and Systematics, 48(1): 159–182.
- Chuine, I. et al., 2016. Can phenological models predict tree phenology accurately in the future? The unrevealed
   hurdle of endodormancy break. Global Change Biology, 22(10): 3444–3460.
- Cremonese, E. et al. 2017. Heat wave hinders green wave: The impact of climate extreme on the phenology of a
   mountain grassland. Agricultural and Forest Meteorology 247: 320–330.
- Dahlin, K., Fisher, R. and Lawrence, P., 2015. Environmental drivers of drought deciduous phenology in the
   Community Land Model. Biogeosciences, 12(16): 5061–5074.
- Debusschere, N., Segers, P., Dubruel, P., Verhegghe, B. and De Beule, M., 2015. A finite element strategy to
   investigate the free expansion behaviour of a biodegradable polymeric stent. Journal of Biomechanics,
   48(10): 2012–2018.
- Delpierre, N. et al., 2009. Modelling interannual and spatial variability of leaf senescence for three deciduous tree
   species in France. agricultural and forest meteorology, 149(6–7): 938–948.
- Dufrêne, E. et al., 2005. Modelling carbon and water cycles in a beech forest: Part I: Model description and
   uncertainty analysis on modelled NEE. Ecological Modelling, 185(2–4): 407–436.
- Forsythe, W.C., Rykiel Jr, E.J., Stahl, R.S., Wu, H.-i. and Schoolfield, R.M., 1995. A model comparison for
   daylength as a function of latitude and day of year. Ecological Modelling, 80(1): 87–95.
- Fracheboud, Y. et al., 2009. The control of autumn senescence in European aspen. Plant physiology, 149(4): 1982–
  1991.
- Hannerz, M., Ekberg, I. and Norell, L., 2003. Variation in chilling requirements for completing bud rest between
   provenances of Norway spruce. Silvae Genetica, 52(3–4): 161–168.

- Hänninen, H. et al., 2019. Experiments Are Necessary in Process-Based Tree Phenology Modelling. Trends in Plant
   Science, 24(3): 199–209.
- Hanson, P.J. et al., 2017. Attaining whole-ecosystem warming using air and deep-soil heating methods with an
   elevated CO2 atmosphere. Biogeosciences (Online), 14(4): 861–883.
- Hanson, P.J. et al., 2020. Rapid net carbon loss from a whole- ecosystem warmed Peatland. AGU Advances, 1(3):
   e2020AV000163.
- Heide, O.M., 2011. Temperature rather than photoperiod controls growth cessation and dormancy in Sorbus species.
   Journal of experimental botany, 62(15): 5397–5404.
- Jeong, S.J. and Medvigy, D., 2014. Macroscale prediction of autumn leaf coloration throughout the continental U
   nited S tates. Global Ecology and Biogeography, 23(11): 1245–1254.
- Keenan, T.F. and Richardson, A.D., 2015. The timing of autumn senescence is affected by the timing of spring
   phenology: implications for predictive models. Global change biology, 21(7): 2634–2641.
- Keller, S.R. et al., 2011. Climate- driven local adaptation of ecophysiology and phenology in balsam poplar,
   Populus balsamifera L.(Salicaceae). American Journal of Botany, 98(1): 99–108.
- Keskitalo, J., Bergquist, G., Gardeström, P. and Jansson, S., 2005. A cellular timetable of autumn senescence. Plant
   Physiology, 139(4): 1635–1648.
- Kim, J.H. et al., 2018. Warming- induced earlier greenup leads to reduced stream discharge in a temperate mixed
   forest catchment. Journal of Geophysical Research: Biogeosciences, 123(6): 1960–1975.
- Kolka, R., Sebestyen, S., Verry, E.S. and Brooks, K., 2011. Peatland biogeochemistry and watershed hydrology at
   the Marcell Experimental Forest. CRC Press.
- Körner, C. and Basler, D., 2010a. Phenology under global warming. Science, 327(5972): 1461–1462.
- Körner, C. and Basler, D., 2010b. Response—Warming, Photoperiods, and Tree Phenology. Science, 329(5989):
   278–278.
- Landsberg, J., 1974. Apple fruit bud development and growth; analysis and an empirical model. Annals of Botany,
   38(5): 1013–1023.
- Lawrence, D.M. et al., 2011. Parameterization improvements and functional and structural advances in version 4 of
   the Community Land Model. Journal of Advances in Modeling Earth Systems, 3(1): M03001.
- Li, H., et al. 2016. Evapotranspiration dynamics over a temperate meadow ecosystem in eastern Inner Mongolia,
   China. Environmental Earth Sciences, 75(11): 978.
- Lian, X. et al., 2020. Summer soil drying exacerbated by earlier spring greening of northern vegetation. Science
   advances, 6(1): eaax0255.
- Liang, L., 2019. A spatially explicit modeling analysis of adaptive variation in temperate tree phenology.
   Agricultural and Forest Meteorology, 266: 73–86.

- Liu, Q. et al., 2020. Modeling leaf senescence of deciduous tree species in Europe. Global Change Biology, 26(7):
   4104–4118.
- Liu, Z. et al. 2015. Relationship between leaf physiologic traits and canopy color indices during the leaf expansion
   period in an oak forest. Ecosphere 6(12): 1–9.
- Meng, L. et al. 2021. Photoperiod decelerates the advance of spring phenology of six deciduous tree species under
   climate warming. Global Change Biology.
- Meng, L. et al., 2020a. Urban warming advances spring phenology but reduces the response of phenology to
   temperature in the conterminous United States. Proceedings of the National Academy of Sciences, 117(8):
   4228–4233.
- Meng, L. et al., 2020b. Divergent responses of spring phenology to daytime and nighttime warming. Agricultural
   and Forest Meteorology, 281: 107832.
- 761 Menzel, A. and Fabian, P., 1999. Growing season extended in Europe. Nature, 397(6721): 659–659.
- Migliavacca, M. et al., 2008. European larch phenology in the Alps: can we grasp the role of ecological factors by
   combining field observations and inverse modelling? International journal of biometeorology, 52(7): 587–
   605.
- Murray, M., Cannell, M. and Smith, R., 1989. Date of budburst of fifteen tree species in Britain following climatic
   warming. Journal of Applied Ecology, 26(2): 693–700.
- Oleson, K., 2013. Technical description of version 4.5 of the Community Land Model (CLM), Natl. Cent. for
   Atmos. Res., Tech. Note NCAR/TN-503+ STR.
- Peano, D. et al., 2020. Plant phenology evaluation of CRESCENDO land surface models–Part I: Start and end of
   growing season. Biogeosciences, 18: 2405–2428.
- Peichl, M. 2015. Bringing color into the picture: using digital repeat photography to investigate phenology controls
   of the carbon dioxide exchange in a boreal mire. Ecosystems, 18.1: 115–131.
- Piao, S. et al., 2019a. Plant phenology and global climate change: Current progresses and challenges. Global change
   biology, 25(6): 1922–1940.
- Piao, S. et al., 2019b. Characteristics, drivers and feedbacks of global greening. Nature Reviews Earth &
   Environment, 1: 14–27.
- Piao, S., Friedlingstein, P., Ciais, P., Viovy, N. and Demarty, J., 2007. Growing season extension and its impact on
   terrestrial carbon cycle in the Northern Hemisphere over the past 2 decades. Global Biogeochemical
   Cycles, 21(3): GB3018.
- Pope, K.S. et al., 2013. Detecting nonlinear response of spring phenology to climate change by Bayesian analysis.
   Global Change Biology, 19(5): 1518–1525.

- Reaumur, R.d., 1735. Observations du thermomètre faites à Paris pendant l'année 1735, comparées avec celles qui
  ont été faites sous la ligne, à l'Isle de France, à Alger et quelques unes de nos iles de l'Amérique. Mémoires
  l'Académie R des Sci: 545–576.
- Ricciuto, D., Sargsyan, K. and Thornton, P., 2018. The impact of parametric uncertainties on biogeochemistry in the
   E3SM land model. Journal of Advances in Modeling Earth Systems, 10(2): 297–319.
- Ricciuto, D., X. Xu, X. Shi, Y. Wang, X. Song, C.W. Schadt, N.A. Griffiths, J. Mao, J.M. Warren, P.E. Thornton, J.
   Chanton, J.K. Keller, S. Bridgham, J. Gutknecht, S.D. Sebestyen, A. Finzi, R. Kolka, P.J. Hanson, 2020.
   An integrative model for soil biogeochemistry and methane processes: I. model structure and sensitivity
   analysis, Accepted, Journal of Geophysical Research-Biogeosciences.
- Richardson, A.D. et al., 2007. Use of digital webcam images to track spring green-up in a deciduous broadleaf
   forest. Oecologia, 152(2): 323–334.
- Richardson, A.D. et al., 2012. 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(2): 566–
   584.
- Richardson, A.D. et al., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the
   climate system. Agricultural and Forest Meteorology, 169: 156–173.
- Richardson, A.D. et al., 2018a. Tracking vegetation phenology across diverse North American biomes using
   PhenoCam imagery. Scientific Data, 5: 180028.
- Richardson, A.D. et al., 2018b. Ecosystem warming extends vegetation activity but heightens vulnerability to cold
   temperatures. Nature, 560: 368–371.
- Sargsyan, K. et al., 2014. Dimensionality reduction for complex models via Bayesian compressive sensing.
   International Journal for Uncertainty Quantification, 4(1): 63–93.
- Sarvas, R., 1972. Investigations on the annual cycle of development of forest trees. Active period. Investigations on
   the annual cycle of development of forest trees. Active period. Metsantutkimuslaitoksen Julkaisuja, 76(3).
- Sarvas, R., 1974. Investigations on the annual cycle of development of forest trees. Autumn dormancy and winter
   dormancy. Communicationes Linstinstituti Forestalis Fenniae, 84: 1–101.
- Schädel, C. et al., 2019. SPRUCE Vegetation Phenology in Experimental Plots from Phenocam Imagery, 2015 2018, Oak Ridge National Laboratory, Oak Ridge, TN (United States).
- Schwartz, M.D., 1992. Phenology and springtime surface-layer change. Monthly weather review, 120(11): 2570–
   2578.
- Seyednasrollah, B. et al., 2019. Tracking vegetation phenology across diverse biomes using Version 2.0 of the
   PhenoCam Dataset. Scientific Data, 6(1): 222.
- Seyednasrollah, B. et al.. 2020. Seasonal variation in the canopy color of temperate evergreen conifer forests. New
   Phytologist, 229: 2586–2600.

- Shi, X. et al., 2015. Representing northern peatland microtopography and hydrology within the Community Land
   Model. Biogeosciences Discussions (Online), 12(4).
- Shi, X. et al., 2020. Modeling the hydrology and physiology of Sphagnum moss in a northern temperate bog.
   Biogeosciences Discussions: 1–49.
- Sonnentag, O. et al., 2012. Digital repeat photography for phenological research in forest ecosystems. Agricultural
   and Forest Meteorology, 152: 159–177.
- Thornton, P.E. et al., 2002. Modeling and measuring the effects of disturbance history and climate on carbon and
   water budgets in evergreen needleleaf forests. Agricultural and forest meteorology, 113(1–4): 185–222.
- Wang, K. et al., 2013. Evaluation of CLM4 solar radiation partitioning scheme using remote sensing and site level
   FPAR datasets. Remote Sensing, 5(6): 2857–2882.
- Way, D.A. and Montgomery, R.A., 2015. Photoperiod constraints on tree phenology, performance and migration in
   a warming world. Plant, Cell & Environment, 38(9): 1725–1736.
- White, M.A. et al., 2009. Intercomparison, interpretation, and assessment of spring phenology in North America
   estimated from remote sensing for 1982–2006. Global Change Biology, 15(10): 2335–2359.
- White, M.A., Thornton, P.E. and Running, S.W., 1997. A continental phenology model for monitoring vegetation
   responses to interannual climatic variability. Global biogeochemical cycles, 11(2): 217–234.
- Zeng, Z. et al., 2017. Climate mitigation from vegetation biophysical feedbacks during the past three decades.
   Nature Climate Change, 7(6): 432–436.
- Zeng, Z. et al., 2018. Impact of Earth greening on the terrestrial water cycle. Journal of Climate, 31(7): 2633–2650.

836	Supplementary Material for
837	Evaluation and Modification of ELM Seasonal Deciduous Phenology against
838	<b>Observations in a Southern Boreal Peatland Forest</b>
839	
840	
841	
842	
843	
844	
845	
846	
847	
848	
849	
850	
851	
852	
853	This PDF contains:
854	Figures S1–S9
855	Tables S1–S3



Fig. S1. 2 m air temperature and spring onset dates in 2018 in comparison to in 2016 and

**2017.** Air temperature shown here is 3-day averaged temperature at 2 m on central tower at plot

859 06 (0 °C warming and aCO<sub>2</sub>), spatially averaged from two replicates. Red line represents mean

air temperature during 2016 and 2017. Blue line represents air temperature during 2018. Vertical

lines and shadow represent the mean and standard deviation of spring onset dates across all

862 enclosures, respectively.





866 ELM\_New for *Larix laricina* (a) and the shrub layer (b). The value is the inter-annual mean

during 2016 to 2018. Four pixels are within each grid. The left and right columns represent value

at  $eCO_2$  and  $aCO_2$  chambers, respectively, and the top and bottom rows represent values

869 predicted by ELM\_Default and ELM\_New, respectively.





872 ELM\_New for *Larix laricina* (a) and the shrub layer (b). The value is the inter-annual mean

during 2016 to 2018. Four pixels are within each grid. The left and right columns represent value

at  $eCO_2$  and  $aCO_2$  chambers, respectively, and the top and bottom rows represent value predicted

by ELM\_Default and ELM\_New, respectively.



877 Fig. S4. Changes in predicted seasonal TLAI, GPP, QVEGE, and QVEGT by

ELM\_Default and ELM\_New for *Larix laricina* (a) and the shrub layer (b). The value is the

inter-annual mean during 2016 to 2018. Green represents higher values by ELM\_New, brown

represents lower values by ELM\_New, and gray represents no differences between ELM\_New

and ELM\_Default. Four pixels are within each grid. The left and right columns represent value at

- eCO<sub>2</sub> and aCO<sub>2</sub> chambers, respectively. Stars indicate significant differences between
- ELM\_New and ELM\_Default based on a two-tailed Student's *t* test (P < 0.1).



884

Fig. S5. TLAI, GPP, and *G<sub>CC</sub>* of *Larix laricina* during 2016 to 2018 at aCO<sub>2</sub> across five

warming levels. The y-axis is the normalized value of each variable. The shadow of  $G_{CC}$ 

represents the range of  $G_{CC}$  at 0°C and 9°C warming chambers. Normalization is conducted for

each variable across all warming levels during 2016 to 2018, and separately for ELM\_Default
 and ELM New.



890

Fig. S6. TLAI, GPP, and *G<sub>CC</sub>* of the shrub layer during 2016 to 2018 at aCO<sub>2</sub> across five

warming levels. The y-axis is the normalized value of each variable. The shadow of  $G_{CC}$ 

represents the range of  $G_{CC}$  at 0°C and 9°C warming chambers. Normalization is conducted for each variable across all warming levels during 2016 to 2018, and separately for ELM Default

894 each variable acros895 and ELM New.



Fig. S7. Temperature responses of carbon and water fluxes of *Larix laricina* to warming and elevated CO<sub>2</sub> predicted by ELM\_Default and ELM\_New for spring (a), summer (b), and autumn (c). There is no value in winter. The temperature responses are calculated as the slope of linear regression between carbon/water variables and warming levels. Significance is shown (P < 0.1). TLAI is m<sup>2</sup>/m<sup>2</sup>; GPP is gC/m<sup>2</sup>/month; QVEGE is mm/month; and QVEGT is mm/month.



Fig. S8. Temperature responses of carbon and water fluxes of shrub layer to warming and
 elevated CO2 predicted by ELM\_Default and ELM\_New for spring (a), summer (b), and
 autumn (c). There is no value in winter. The temperature responses are calculated as the slope of

907 linear regression between carbon/water variables and warming levels. Significance is shown

908 (P < 0.1). TLAI is m<sup>2</sup>/m<sup>2</sup>; GPP is gC/m<sup>2</sup>/month; QVEGE is mm/month; and QVEGT is 909 mm/month.



Fig. S9. The relationship between differences in TLAI and GPP and differences in spring

green-up (a, b) and autumn green-down (c, d) simulated by ELM\_Default and ELM\_New.

The linear regression lines, slopes, and *P* values are shown in each figure. The  $\Delta$ TLAI,  $\Delta$ GPP,

and  $\Delta$ Autumn green-down are calculated by subtracting ELM\_Default from ELM\_New.  $\Delta$ Spring green-up is calculated by subtracting ELM New from ELM Default. a, b:  $\Delta$ TLAI and  $\Delta$ GPP are

915 green-up is calculated by subtracting ELM\_New fro 916 in May; c, d:  $\Delta$ TLAI and  $\Delta$ GPP are in October.

		Spring	g onset m	odel	Autumn	senescei	nce model	Statistic					
Vegetation type	а	b	С	$T_{base}\left(\mathrm{F} ight)$	Pstart (s)	Ycrit	<i>Tb</i> (F)	LL	AIC	AICc			
Larix laricina	9	2,112	-0.04	279.50	46,800	1,750	294.5	71.93	157.86	164.4			
Shrub layer	33	1,388	-0.02	279.05	54,600	1,600	290.15	90.46	194.92	201.5			

**Table S1. Statistics in model calibration.** 

*LL*: log likelihood, *AIC*: Akaike information criterion

- 919 Table S2. Statistics in model evaluation during 2018. There is no correlation in autumn green-
- down between ELM\_Default and PhenoCam observation, due to the unchanged in autumn
- green-down across warming levels by ELM\_Default.

Model	Vegetation type	Spring	onset model	Autumn senescence model					
		RMSE	Correlation	RMSE	Correlation				
ELM_Default	Larix laricina	26.3	0.18	9.1					
	Shrub layer	3.4	0.41	7.6					
ELM_New	Larix laricina	8.2	0.34	11.9	-0.22				
	Shrub layer	7.4	0.50	13.4	0.21				

Parameter	Description	Min	Max
flnr <sup>1</sup>	Rubisco-N fraction of leaf N	0.05	0.30
croot_stem <sup>1</sup>	Coarse root to stem allocation ratio	0.05	0.8
stem_leaf <sup>1</sup>	Stem to leaf allocation ratio	0.3	2.2
leaf_long <sup>1</sup>	Leaf longevity (yr)	0.75	2.0
slatop <sup>1</sup>	Specific leaf area at canopy top (m <sup>2</sup> gC <sup>-1</sup> )	0.004	0.04
leafcn <sup>1</sup>	Leaf C to N ratio	20	75
froot_leaf <sup>1</sup>	Fine root to leaf allocation ratio	0.15	2.0
mp <sup>1</sup>	Ball-Berry stomatal conductance slope	4.5	12
r_mort	Vegetation mortality	0.005	0.1
decomp_depth_efolding	Depth-dependence e-folding depth for decomposition (m)	0.2	0.7
<b>q</b> drai,0	Maximum subsurface drainage rate (kg m <sup>-2</sup> s <sup>-1</sup> )	0	1e-3
Q <sub>10</sub> _mr	Temperature sensitivity of maintenance respiration	1.2	3.0
br_mr	Base rate for maintenance respiration (gC gN m <sup>2</sup> s <sup>-1</sup> )	1e-6	5e-6
crit_onset_gdd	Critical growing degree days for leaf onset	20	500
lw_top_ann	Live wood turnover proportion (yr <sup>-1</sup> )	0.2	0.85
gr_perc	Growth respiration fraction	0.12	0.4
r <sub>drai,0</sub>	Coefficient for surface water runoff (kg m <sup>-4</sup> s <sup>-1</sup> )	1e-9	1e-6
phen_a	A parameter for spring onset process	1.0	30.0
phen_b	A parameter for spring onset process	1,000	2,500
phen_c	A parameter for spring onset process	-0.06	-0.01
phen_tbspring	Base temperature for spring onset process	269.15	282.15
phen_tbfall	Base temperature for autumn senescence process	272.15	308.15
phen_ycrit	Threshold to trigger autumn senescence	1,000	2,000
phen_pstart	Photoperiod threshold for autumn senescence process	12.0	16.5

# **Table S3. Parameters and ranges used in the GSA.**

924 <sup>1</sup>PFT-specifc parameters

±

		Spring	g onset m	odel	Autumn	senescei	nce model	Statistic				
Vegetation type	а	b	С	$T_{base}$ (F)	Pstart (s)	Ycrit	<i>Tb</i> (F)	LL	AIC	AICc		
Larix laricina	9	2,112	-0.04	279.50	46,800	1,750	294.5	71.93	157.86	164.4		
Shrub layer	33	1,388	-0.02	279.05	54,600	1,600	290.15	90.46	194.92	201.5		

# 1 Table S1. Statistics in model calibration.

2 *LL*: log likelihood, *AIC*: Akaike information criterion

- 3 Table S2. Statistics in model evaluation during 2018. There is no correlation in autumn green-
- 4 down between ELM\_Default and PhenoCam observation, due to the unchanged in autumn
- 5 green-down across warming levels by ELM\_Default.

Model	Vegetation type	Spring	onset model	Autumn senescence model					
		RMSE	Correlation	RMSE	Correlation				
ELM_Default	Larix laricina	26.3	0.18	9.1					
	Shrub layer	3.4	0.41	7.6					
ELM_New	Larix laricina	8.2	0.34	11.9	-0.22				
	Shrub layer	7.4	0.50	13.4	0.21				

Parameter	Description	Min	Max
flnr <sup>1</sup>	Rubisco-N fraction of leaf N	0.05	0.30
croot_stem <sup>1</sup>	Coarse root to stem allocation ratio	0.05	0.8
stem_leaf <sup>1</sup>	Stem to leaf allocation ratio	0.3	2.2
leaf_long <sup>1</sup>	Leaf longevity (yr)	0.75	2.0
slatop <sup>1</sup>	Specific leaf area at canopy top (m <sup>2</sup> gC <sup>-1</sup> )	0.004	0.04
leafcn <sup>1</sup>	Leaf C to N ratio	20	75
froot_leaf <sup>1</sup>	Fine root to leaf allocation ratio	0.15	2.0
mp <sup>1</sup>	Ball-Berry stomatal conductance slope	4.5	12
r_mort	Vegetation mortality	0.005	0.1
decomp_depth_efolding	Depth-dependence e-folding depth for decomposition (m)	0.2	0.7
<b>q</b> drai,0	Maximum subsurface drainage rate (kg m <sup>-2</sup> s <sup>-1</sup> )	0	1e-3
Q <sub>10</sub> _mr	Temperature sensitivity of maintenance respiration	1.2	3.0
br_mr	Base rate for maintenance respiration (gC gN m <sup>2</sup> s <sup>-1</sup> )	1e-6	5e-6
crit_onset_gdd	Critical growing degree days for leaf onset	20	500
lw_top_ann	Live wood turnover proportion (yr <sup>-1</sup> )	0.2	0.85
gr_perc	Growth respiration fraction	0.12	0.4
r <sub>drai,0</sub>	Coefficient for surface water runoff (kg m <sup>-4</sup> s <sup>-1</sup> )	1e-9	1e-6
phen_a	A parameter for spring onset process	1.0	30.0
phen_b	A parameter for spring onset process	1,000	2,500
phen_c	A parameter for spring onset process	-0.06	-0.01
phen_tbspring	Base temperature for spring onset process	269.15	282.15
phen_tbfall	Base temperature for autumn senescence process	272.15	308.15
phen_ycrit	Threshold to trigger autumn senescence	1,000	2,000
phen_pstart	Photoperiod threshold for autumn senescence process	12.0	16.5

# 7 Table S3. Parameters and ranges used in the GSA.

8 <sup>1</sup>PFT-specifc parameters



Fig. 1. Observed and simulated responses of spring onset (a) and autumn senescence (b) to



4 across all warming and CO<sub>2</sub> levels and slopes of phenology against warming levels are shown in

5 the subfigures. The error bars in the subfigures represent the standard deviations of phenology

6 across all warming levels. Significance P < 0.1 from two-tailed Student's *t* test. Spring onset was

7 during 2016 to 2018, and autumn senescence was during 2015 to 2018.

- 8
- 9
- 10
- 11
- 12
- 13
- 14



15 16

<sup>16</sup> Fig. 2. Multi-year mean differences in simulated monthly TLAI, GPP, QVEGE, and

- 17 **QVEGT between ELM\_New and ELM\_Default for** *Larix laricina* (a) and the shrub layer
- (b) during 2016 to 2018. Green represents higher values by ELM\_New, brown represents lower
- values by ELM\_New, and gray represents no differences between ELM\_New and ELM\_Default.
- 20 Two pixels are in each grid; left and right pixels represent values under  $eCO_2$  and  $aCO_2$
- 21 conditions, respectively. Stars indicate significant differences between ELM\_New and
- ELM\_Default based on a two-tailed Student's *t* test (P < 0.1).
- 23 24
- 25
- 26
- 27



**Fig. 3.** Temperature responses of annual TLAI, GPP, QVEGE, and QVEGT simulated by

30 ELM\_Default and ELM\_New for *Larix laricina* (a) and the shrub layer (b). We used annual

31 average TLAI  $(m^2/m^2)$  and GPP  $(gC/m^2/year)$ , and annual sum QVEGE (mm/year) and QVEGT

32 (mm/year). The temperature responses were calculated as the slope of the linear regression

between annual TLAI, GPP, QVEGE, and QVEGT and warming levels. Stars indicate

significance from linear regression (P < 0.1). Results for each season are shown in Figs. S7

35 and S8.

	(a) Hummock														(b) Hollow										
	Mean(%) Response(%)												Me	ean	(%)			R	lesp	ons	e(%	6)			
						J																			
	0	100	200	300	400	-4	00 -20	0 0	200	400	600 80	0 -4	10 -20	0 2	20 40	60	80 10	0	- 1000	0	10	00 2	000		
	0	10.4	-1.5	15	-36		-13.2	-12.2	-13.4	-13.1	-12.5		1.2	4.1	0.9	5.3	-36.8		-38.7	-7.2	-41.1	-8.3	-4.4		
NEP	-2.9	4.3	-4.6	4.3	20.7		-17	-16.7	-18.1	-17.2	-16.7		1.4	0.4	1.2	-0.3	-1.3		-10.9	-8.8	-12.7	-9.4	-8.6		
	3.2	244.8	1.8	-46.5	-8.2		-6	-9	-4.4	-10.2	-10		1.1	77.7	0.5	-35.5	-7.3		-87.5	-6.1	-88.5	-7.6	-1.9		
	17.9	2.9	17.9	2.8	4.3		16.6	-10.6	17.6	-10.8	-27.2		18.1	1.2	17.2	0.9	2.4		14.3	-13.7	14.2	-14.6	-33.4		
٩٩٧	21.8	2.2	22.1	2	3.7		17.7	-16.9	19.1	-16.9	-70.5		21.2	0.4	20.5	0.1	1.7		12	-23.1	11.4	-23.9	-154.1		
_	13.9	4	13.7	4	5.2		15	-7.1	15.6	-7.6	-14.7		14.8	2.4	13.7	2.2	3.5		17.5	-9.6	18.2	-10.7	-19		
	5.3	2.2	5.2	2.2	2.8		53.2	8.9	51.3	8.3	23.4		6.5	2.3	6.7	2.3	3.1		52.4	12.4	50.8	11.5	23.6		
Ř	4.6	2.5	4.5	2.5	2.9		58.7	10.5	57.5	10.5	25.5		6.6	2.7	6.9	2.7	3.4		39.2	14.1	37.9	14.1	21.6		
ш	6.2	1.7	6	1.7	2.7		46.2	6.2	43.6	5	20.2		6.4	1.8	6.4	1.7	2.7		72.8	10	70.2	8.2	26.3		
	8.5	2.8	8.9	2.8	3.6		28.5	-40.6	29.5	-43.8	746.3		9.2	2.4	9.5	2.4	3.3		29.5	-52.2	29.4	-56.8	310.2		
Ч	8.9	2.8	9.4	2.7	3.6		27.9	-189.2	29.1	-210.9	84.9		9.2	2.4	9.6	2.4	3.2		26.1	2646.4	25.8	1623.1	64		
G	8.1	2.9	8.3	2.9	3.7		29.3	-17.8	30.1	-19.8	-70		9.3	2.5	9.4	2.5	3.5		34.9	-21.6	35	-24.2	-80		
	6.9	5.8	7.7	5.8	6.3		-210.8	-375.9	-225.5	-360	-274		6	4.1	6.8	4.1	4.9		-283.3	1081.2	-310.2	1387.9	-750.7		
Ā	6.8	5.7	7.5	5.7	6.2		466.4	130.1	463.4	136.7	204.4		5.8	4	6.6	3.9	4.7		249.1	66.5	249.3	69.4	109.9		
F	7.1	6	7.9	6	6.5		-83.6	-68.4	-87.7	-69.5	-76.6		6.2	4.3	7	4.2	5.1		-85.5	-76.6	-90.2	-77.8	-82.2		
ш	-0.1	-3.6	0.3	-3	-1.1		19.2	-5.3	11.2	-2.5	4.4		0	-0.7	0	-0.7	-0.3		-229.2	3.7	-268.8	4.4	13.7		
10F	0.7	-2.9	0.1	-2.8	-0.8		8.5	-1.5	8.5	-0.7	2.9		0.1	-0.7	0.2	-0.9	-0.3		-135.5	-0.6	-142.9	0.7	17.8		
RUN	-0.9	-4.4	0.5	-3.4	-1.4		30.9	-10.2	13.8	-5	6.1		-0.1	-0.6	-0.1	-0.6	-0.4		93	7.7	74.8	7.9	10.3		
	1.5	0.8	1.8	1	1.1		1.3	4.1	3.2	4.2	3.6		1.2	0.7	1.4	0.8	0.8		5.4	3.3	5.1	4	4		
E	1.4	0.8	1.8	0.9	1		0.7	3.6	2	3.8	3		0.9	0.6	0.8	0.8	0.7		6.5	2.5	3.5	2.9	3.2		
ш	1.5	0.9	1.9	1	1.1		1.9	4.6	4.5	4.7	4.2		1.4	0.7	1.9	0.8	0.9		4.3	4.3	6.7	5.2	4.9		
	-1.8	3.4	-1.7	3.4	1.6		39.8	1.7	37.7	1.9	12.5		0	3.8	0	4.3	2.9		-127	5.3	-146.7	6.1	14.4		
٧T	-1.8	3.8	-1.8	3.9	1.9		14.2	2	12.8	1.9	5.5		-1.7	4.5	-1.8	5.2	2.9		-107	1.1	-116.9	2.3	17.8		
Ñ	-1.8	3	-1.7	3	1.3		74.9	1.5	71	1.9	20.3		1.5	3.3	1.5	3.5	2.8		1661.7	9.2	208.6	9.6	11.6		
_	0.5	-1.6	0.5	-1.6	-0.4		21.1	-4.4	20.2	-4.6	7.2		-0.2	-0.9	-0.2	-0.9	-0.5		-447.6	2.2	1283.3	2.8	9.1		
OSC	0.5	-1.8	0.5	-1.8	-0.5		9.5	-4.9	8.7	-5.2	1.5		0	-1	0	-1.1	-0.5		-90.7	-2.3	-125.9	-1.5	4.8		
H2C	0.5	-1.4	0.5	-1.4	-0.3		35.1	-3.8	34	-3.9	13.9		-0.3	-0.7	-0.3	-0.7	-0.5		168.3	6.8	99.9	7.2	12.9		
	<u> </u>	D	<u> </u>	=	=		<u> </u>		<u> </u>	=	_			Γ		_	_		<u> </u>	~	<u> </u>	_	_		
	nte	rinç	me	Га	nua		nte	rinç	me	Fa	nua		nte	rinç	me	Га	nua		nte	ring	me	Fal	nua		
	Ň	Sp	ш		Ani		Ň	Sp	ш		Anr		Ň	Sp	ш		Anr		Ň	Sp	Ш		Anr		
			S		-				S		-				S						S				



**Fig. 4. Relative changes in mean and temperature responses in carbon and water fluxes of** 

<sup>hummock (a) and hollow (b) simulated by ELM\_New compared with ELM\_Default.
Relative changes = (ELM New - ELM Default) / ELM Default × 100%. Mean carbon and</sup> 

<sup>40</sup> Relative changes –  $(ELW_New - ELW_Default) / ELW_Default × 100%. Mean carbon and$ 41 water fluxes represent the average values across all warming levels, and temperature responses

represent the changes in carbon and water fluxes per degree warming. Green represents a higher

value or stronger response simulated by ELM New, and brown represents lower values or

<sup>44</sup> weaker response by ELM New. For each variable, each of the three rows within each grid

represents the mean under eCO<sub>2</sub> and aCO<sub>2</sub> condition, eCO<sub>2</sub> condition, and aCO<sub>2</sub> condition across
 all warming levels during 2016 to 2018.

- 47
- 48
- 49



# 50

51 Fig. 5. Sensitivity of major carbon and water variables to model parameters for *Larix* 

52 *laricina* and the shrub layer. The height of each bar represents the percentage of the uncertainty

of the examined carbon and water variables caused by a particular parameter. We only show 12

54 out of 44 parameters (represented by colors) that have sensitivity indices large than 1%. Within

each stacked bar, if the sensitivity is greater than 5%, a + or - symbol is included to show whether positive perturbations of that parameter result in a positive or negative deviation on the

 $_{57}$  QoI. GPP is gC/m<sup>2</sup>/year; NPP is gC/m<sup>2</sup>/year; QVEGT is mm/year; TOTVEGC is gC/m<sup>2</sup>; and

- 58 TLAI is  $m^2/m^2$ .
- 59 60



Fig. S1. 2 m air temperature and spring onset dates in 2018 in comparison to in 2016 and

63 **2017.** Air temperature shown here is 3-day averaged temperature at 2 m on central tower at plot

64 06 (0 °C warming and aCO<sub>2</sub>), spatially averaged from two replicates. Red line represents mean

air temperature during 2016 and 2017. Blue line represents air temperature during 2018. Vertical

66 lines and shadow represent the mean and standard deviation of spring onset dates across all

67 enclosures, respectively.



70 Fig. S2. Predicted monthly TLAI, GPP, QVEGE, and QVEGT by ELM\_Default and

during 2016 to 2018. Four pixels are within each grid. The left and right columns represent value

- $^{73}$  at eCO<sub>2</sub> and aCO<sub>2</sub> chambers, respectively, and the top and bottom rows represent values
- 74 predicted by ELM\_Default and ELM\_New, respectively.

<sup>71</sup> ELM\_New for *Larix laricina* (a) and the shrub layer (b). The value is the inter-annual mean





77 ELM\_New for *Larix laricina* (a) and the shrub layer (b). The value is the inter-annual mean

during 2016 to 2018. Four pixels are within each grid. The left and right columns represent value

79 at eCO<sub>2</sub> and aCO<sub>2</sub> chambers, respectively, and the top and bottom rows represent value predicted

80 by ELM\_Default and ELM\_New, respectively.



82 Fig. S4. Changes in predicted seasonal TLAI, GPP, QVEGE, and QVEGT by

83 ELM\_Default and ELM\_New for *Larix laricina* (a) and the shrub layer (b). The value is the

inter-annual mean during 2016 to 2018. Green represents higher values by ELM\_New, brown

represents lower values by ELM\_New, and gray represents no differences between ELM\_New

and ELM\_Default. Four pixels are within each grid. The left and right columns represent value at

- eCO<sub>2</sub> and aCO<sub>2</sub> chambers, respectively. Stars indicate significant differences between
- ELM\_New and ELM\_Default based on a two-tailed Student's *t* test (P < 0.1).



89

90 Fig. S5. TLAI, GPP, and *G<sub>CC</sub>* of *Larix laricina* during 2016 to 2018 at aCO<sub>2</sub> across five

91 warming levels. The y-axis is the normalized value of each variable. The shadow of  $G_{CC}$ 

P2 represents the range of  $G_{CC}$  at 0°C and 9°C warming chambers. Normalization is conducted for

each variable across all warming levels during 2016 to 2018, and separately for ELM\_Default
 and ELM New.





96 Fig. S6. TLAI, GPP, and G<sub>CC</sub> of the shrub layer during 2016 to 2018 at aCO<sub>2</sub> across five

97 warming levels. The y-axis is the normalized value of each variable. The shadow of  $G_{CC}$ 

represents the range of  $G_{CC}$  at 0°C and 9°C warming chambers. Normalization is conducted for

each variable across all warming levels during 2016 to 2018, and separately for ELM\_Default

100 and ELM\_New.



Fig. S7. Temperature responses of carbon and water fluxes of *Larix laricina* to warming and elevated CO<sub>2</sub> predicted by ELM\_Default and ELM\_New for spring (a), summer (b), and autumn (c). There is no value in winter. The temperature responses are calculated as the slope of linear regression between carbon/water variables and warming levels. Significance is shown (P < 0.1). TLAI is m<sup>2</sup>/m<sup>2</sup>; GPP is gC/m<sup>2</sup>/month; QVEGE is mm/month; and QVEGT is mm/month.



109 Fig. S8. Temperature responses of carbon and water fluxes of shrub layer to warming and

elevated CO2 predicted by ELM\_Default and ELM\_New for spring (a), summer (b), and

autumn (c). There is no value in winter. The temperature responses are calculated as the slope of

112 linear regression between carbon/water variables and warming levels. Significance is shown

- 113 (P < 0.1). TLAI is m<sup>2</sup>/m<sup>2</sup>; GPP is gC/m<sup>2</sup>/month; QVEGE is mm/month; and QVEGT is
- 114 mm/month.



116 Fig. S9. The relationship between differences in TLAI and GPP and differences in spring

117 green-up (a, b) and autumn green-down (c, d) simulated by ELM\_Default and ELM\_New.

118 The linear regression lines, slopes, and *P* values are shown in each figure. The  $\Delta$ TLAI,  $\Delta$ GPP, 119 and  $\Delta$ Autumn green-down are calculated by subtracting ELM Default from ELM New.  $\Delta$ Spring

and  $\Delta$ Autumn green-down are calculated by subtracting ELM\_Default from ELM\_New.  $\Delta$ Spring green-up is calculated by subtracting ELM New from ELM Default. a, b:  $\Delta$ TLAI and  $\Delta$ GPP are

121 in May; c, d:  $\Delta$ TLAI and  $\Delta$ GPP are in October.