

Exploring Alternative SMAP Level-4 Carbon Model Formulations for the North American Arctic–Subarctic Growing Season

Rémi Madelon^{1,2,3}, K. Arthur Endsley⁴, John S. Kimball⁴, Gabriëlle J. M. De Lannoy⁵, Oliver Sonnentag^{3,6}, Haley Alcock⁶, Alex Mavrovic^{1,3,7}, Scott N. Williamson⁸, Vincent Maire^{1,3}, Arnaud Mialon², and Alexandre Roy^{1,3}

¹Université du Québec à Trois-Rivières (UQTR), Recherche en modélisation et télédétection des environnements nordiques (ReMoTE-Nord) group, Trois-Rivières, QC, Canada

²Univ Toulouse, CNES/IRD/CNRS/INRAe, CESBIO, Toulouse, France

³Centre d'Etudes Nordiques (CEN), Québec, QC, Canada

⁴University of Montana, Numerical Terradynamic Simulation Group (NTSG), Missoula, MT, United States

⁵Katholieke Universiteit Leuven (KU Leuven), Department of Earth and Environmental Sciences, Division Soil and Water Management, Leuven, Belgium

⁶Université de Montréal (UdeM), Département de Géographie, Montréal, QC, Canada

⁷Cégep de Sherbrooke, Physics Department, Sherbrooke, QC, Canada

⁸Polar Knowledge Canada, Cambridge Bay, NU, Canada

Correspondence: Rémi Madelon (remi.madelon@uqtr.ca or remi.madelon@utoulouse.fr)

1 **Abstract.** The Soil Moisture Active Passive Level-4 Terrestrial Carbon Flux model (hereafter referred to as the L4C model)
2 provides daily estimates of net ecosystem CO₂ exchange (NEE), gross primary production (GPP), and ecosystem respiration
3 (ER) at a global scale. The model is based on direct mechanistic forcing–response relationships between CO₂ fluxes and energy
4 proxies (absorbed photosynthetically active radiation and temperature) and moisture proxies (soil moisture and vapor pressure
5 deficit). Although the L4C model aims to provide a representative estimation of the CO₂ budget of Arctic and Subarctic (AS)
6 environments, a deeper understanding of carbon cycle processes and targeted refinements are needed to improve its accuracy.
7 In this study, alternative model formulations are proposed for the North American AS regions during the growing season.
8 These formulations are calibrated and evaluated using NEE-derived GPP and ER from 20 eddy covariance towers across
9 western Canada and Alaska, covering the period from 2015 to 2022. Refinements in the representation of energy proxies
10 resulted in greater improvements in model performance than adjustments to moisture proxies. Specifically, implementing a
11 light-response curve in GPP estimation reduced unbiased root mean squared error and bias, while incorporating growing degree
12 days improved correlation. Adjustments to rootzone and surface soil moisture in GPP and ER estimation, respectively, did not
13 yield conclusive performance improvements. Vapor pressure deficit showed limited importance as a driver of GPP in upland
14 tundra and wetlands, whereas it had a stronger impact in taiga forests. Finally, the litterfall scheme used to represent SOC
15 dynamics in the L4C ER model formulation in version 8 demonstrated improved performance relative to version 7. Although
16 some adjustments in ER and GPP formulations yielded strong performance gains, improvements in NEE were more modest
17 than for the individual components. Overall, the results highlight opportunities to enhance the accuracy of the L4C model for
18 the North American AS growing season and underscore the need for further research on CO₂ flux modeling.

19 1 Introduction

20 Arctic and Subarctic (AS) environments store nearly half of the global soil organic carbon (SOC) pool (Tarnocai et al., 2009;
21 Hugelius et al., 2014; Mishra et al., 2021) and are experiencing accelerated warming (Rantanen et al., 2022). Rising temper-
22 atures increase photosynthetic activity and extend the growing season, leading to higher CO₂ uptake by vegetation (Myneni
23 et al., 1997; Jia et al., 2003; Euskirchen, E. S. et al., 2009; Natali et al., 2012; Forkel et al., 2016; Fisher et al., 2018). In
24 addition, rising temperatures enhance autotrophic respiration (AR) as well as heterotrophic respiration (HR) in two pathways:
25 directly, by stimulating microbial activity, and indirectly, by thawing permafrost and exposing frozen SOC to decomposition.
26 The combined increase in AR and HR intensifies CO₂ release to the atmosphere (Natali et al., 2019; Turetsky et al., 2020;
27 Virkkala et al., 2024). Consequently, estimating the net CO₂ budget of AS regions is essential for understanding their role in
28 global climate system feedbacks (Oechel et al., 1993; Hayes et al., 2011; Turetsky et al., 2011; Bell et al., 2013; Schuur et al.,
29 2013; Schaefer et al., 2014; Zona et al., 2016). Nevertheless, our understanding of CO₂ fluxes in the AS environments remains
30 limited. This is due to the inherent complexity and high cost of measuring CO₂ fluxes, the scarcity of such measurements, and
31 the seasonal variability in the dominant processes controlling CO₂ fluxes (Baldocchi et al., 2007; Fisher et al., 2018; Pallandt
32 et al., 2022; Mavrovic et al., 2023b).

33 Net ecosystem CO₂ exchange (NEE) represents the overall balance between CO₂ uptake by photosynthesis, called gross
34 primary production (GPP), and CO₂ release through ecosystem respiration (ER), as follows (Chapin et al., 2006)

$$35 \text{ NEE} = \text{HR} + \text{AR} - \text{GPP} = \text{ER} - \text{GPP} \quad (1)$$

36 GPP is a light-driven process whose efficiency is modulated by air temperature, soil moisture availability within the plant root
37 zone and vapor pressure deficit, which can induce stomatal closure and thereby reduce CO₂ uptake (Davis et al., 2014; Bao
38 et al., 2022). Comparatively, HR is governed by SOC availability, soil temperature, and surface soil moisture, whereas AR
39 primarily depends on air temperature, plant metabolic activity and GPP rate (Reichstein et al., 2005; Davis et al., 2014; Zona
40 et al., 2023). When soil temperature drops near 0 °C, the soil starts freezing and GPP and AR progressively ceases, following
41 a soil freezing characteristic curve (Salmabadi et al., 2025). Under fully frozen conditions, NEE is equal to HR, which is
42 controlled by soil temperature and SOC availability (Natali et al., 2019; Mavrovic et al., 2023b).

43 Although global terrestrial carbon flux (TCF) models, atmospheric inversions (which infer surface CO₂ fluxes from at-
44 mospheric CO₂ concentrations), and data-driven flux-upscaling approaches are available to estimate the CO₂ budget of AS
45 regions, they often disagree on whether these regions are CO₂ sources or sinks (McGuire et al., 2012; Fisher et al., 2018;
46 López-Blanco et al., 2019; Virkkala et al., 2021; Ramage et al., 2024; Virkkala et al., 2024; Foster et al., 2024). In recent
47 decades, satellite-based microwave remote sensing (300 MHz – 100 GHz) has provided a valuable approach for monitoring
48 land–atmosphere interactions and carbon cycle dynamics through the retrieval of key surface variables (Fisher et al., 2018;
49 Lees et al., 2018; Mavrovic et al., 2023a; Pulliainen et al., 2024) such as soil moisture (Kerr et al., 2012; Colliander et al.,
50 2017), snow properties (Lievens et al., 2019), aboveground biomass (Mialon et al., 2020), and freeze-thaw state (Rautiainen
51 et al., 2016; Derksen et al., 2017; Prince et al., 2019). In 2015, the Soil Moisture Active Passive (SMAP) satellite was launched
52 to monitor surface soil moisture and freeze-thaw dynamics using L-band brightness temperature observations (Entekhabi et al.,

2010). One of the science objectives of the SMAP mission is to improve our understanding of interconnected water, energy, and carbon cycles, as well as to quantify the boreal landscape CO₂ budget (Entekhabi et al., 2014). In this context, the SMAP Level-4 Global Daily 9-km EASE-Grid Carbon Net Ecosystem Exchange (SPL4CMDL) product currently provides global, daily estimates of NEE and GPP, as well as ER (indirectly derived from HR, GPP and NEE). These estimates are derived from a TCF model (hereafter referred to as the L4C model), which is notably informed by the SMAP Level-4 Global 9-km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data (SPL4SMGP) product (Jones et al., 2017; Endsley et al., 2022; Kimball et al., 2025; Reichle et al., 2025). Although the L4C model achieves an unbiased root-mean-square error of NEE within the targeted accuracy of 1.6 gCm⁻²d⁻¹ in AS environments, recent studies have reported that it fails to capture the amplitude of GPP and ER during the green-up phase (Endsley et al., 2022; Madelon et al., 2025). The authors also reported discrepancies in annual CO₂ budgets when compared with eddy covariance (EC) measurements, leading to uncertainties in classifying AS environments as net CO₂ sources or sinks (Madelon et al., 2025). From April to July 2025, the L4C model transitioned from version 7 to version 8 (Kimball et al., 2025), featuring a major update partly due to (i) the upgrade of the SMAP SPL4SMGP product, which transitioned from its own version 7 to version 8 (Reichle et al., 2025), and (ii) changes to the litterfall estimation scheme used for modeling SOC dynamics and ER (Section 3).

The goal of this study is to better characterize how key environmental drivers influence GPP and ER, and to refine their modeling for the North American AS growing season. To achieve this, we:

- explore alternative formulations of the L4C model (hereafter referred to as the AS-adapted models) that adjust GPP and ER responses to absorbed photosynthetically active radiation, air and soil temperature, rootzone and surface soil moisture, and vapor pressure deficit;
- calibrate and evaluate these formulations using GPP and ER data from 20 EC towers across western Canada and Alaska from 2015 to 2022;
- identify and interpret the specific model adjustments that yield the greatest performance improvements in terms of Pearson correlation, unbiased root-mean-square error, and bias;
- provide recommendations for more accurate satellite-derived estimates of GPP and ER, with indirect benefits for NEE and CO₂ budget estimation.

2 Eddy covariance measurements

NEE, GPP, and ER data were collected from 20 EC towers located in AS environments (Figure 1), all within the NASA Arctic-Boreal Vulnerability Experiment (ABoVE) study domain (<https://above.nasa.gov/sites.html>). The dataset spans from April 2015 through December 2022 (Table 1) and includes

- half-hourly fluxes from 13 EC towers in Alaska, downloaded from the AmeriFlux Network website (<https://ameriflux.lbl.gov>).

83 – half-hourly fluxes from 7 EC towers in western Canada, provided directly by the principal investigators to ensure the use
84 of the most up-to-date records; some of these sites are not yet available on the AmeriFlux Network website.

85 EC towers measure NEE by quantifying the turbulent vertical exchange of CO₂ in the surface layer of the atmospheric
86 boundary layer (typically within the lowest tens of meters), where turbulence dominates the airflow (Aubinet et al., 2012;
87 Burba, 2013, 2022). The spatial footprint can extend up to 1 km or more, but remains complex to estimate and varies with
88 wind direction, wind speed, and tower height (Leclerc and Thurtell, 1990; Schuepp et al., 1990; Aubinet et al., 2012; Webb
89 et al., 2016). GPP and ER are derived from NEE using flux-partitioning methods selected by the investigators at each EC tower.
90 In this study, we placed confidence in these flux-partitioning methods and considered the resulting partitioned GPP and ER
91 values to be credible representations of the underlying processes and suitable for use as reference data for model calibration
92 and evaluation. Additional information on NEE, and the derived GPP and ER is provided in Appendix A. EC NEE, GPP, and
93 ER were averaged from 30-minute intervals to daily time steps, using at least 24 out of the theoretical 48 data points available
94 per day. Hereafter, NEE_{EC}, GPP_{EC}, and ER_{EC} refer to the daily means of EC NEE, GPP, and ER.

95 All EC towers considered in this study are located in the tundra and taiga biomes, in areas underlain by sporadic, discontin-
96 uous, or continuous permafrost (Figure 1). Permafrost (ground that has remained frozen for more than 2 years) lies beneath an
97 active layer that thaws during the growing season, enabling plant growth, as roots can only establish in thawed soil (Blume-
98 Werry et al., 2019). Permafrost also restricts surface drainage, promoting water saturation and slowing decomposition rates
99 (Wania et al., 2009; Robinson and Moore, 2000; Rouse et al., 1997; Maltby and Immerzi, 1993). These conditions can favor
100 the formation of wetlands, including peatlands as well as other seasonally or permanently waterlogged ecosystems, across both
101 tundra and taiga biomes (Treat et al., 2022). Based on site descriptions, EC towers were grouped into three distinct ecosystem
102 types (Table 1):

- 103 – **Taiga forests:** 7 EC towers are located in taiga forests, characterized by a vertically stratified vegetation structure, with
104 an open canopy of coniferous trees and an understory of shrubs, mosses, and lichens (Crawford, 2013; Juday, 2025).
- 105 – **Upland tundra:** 5 EC towers are located in upland tundra, which may exhibit lower vegetation density and diversity,
106 as well as reduced soil biological activity, compared with taiga forests (Crawford, 2013; Hagedorn et al., 2025). The
107 landscape is treeless and dominated by dwarf shrubs, grasses, sedges, mosses, and lichens, as plant growth is constrained
108 by cold temperatures, short growing seasons, and the shallow depth of the permafrost active layer (Crawford, 2013; Hu
109 and Bliss, 2025; Juday, 2025; Péwé, 2025).
- 110 – **Wetlands:** 8 EC towers are located in wetlands, where the term “wetland” refers to a wide range of types, including
111 peatlands, bogs, fens, marshes, wet meadows, and shrub swamps, present in both tundra and taiga biomes. Compared
112 with taiga forests and upland tundra, wetlands may exhibit higher species richness (McPartland et al., 2019) and localized
113 microtopography, such as hummocks and hollows, whose characteristics depend on water table depth (Rouse et al., 1997;
114 Zhang et al., 2024).

115 3 L4C model

116 The L4C model provides global, daily estimates of NEE, GPP, and ER at 9-km resolution from March 31, 2015, to the present
117 (Jones et al., 2017; Kimball et al., 2025). It takes as inputs

- 118 – A static global plant functional type (PFT) classification at 500-m resolution, retrieved from the Moderate Resolution
119 Imaging Spectroradiometer (MODIS) MCD12Q1 Type 5 product (Friedl and Sulla-Menashe, 2019).
- 120 – Eight-day fraction of photosynthetically active radiation (canopy-intercepted FPAR) and leaf area index (LAI) data at
121 500-m resolution, retrieved from the Visible Infrared Imaging Radiometer Suite (VIIRS) VNP15A2H product (Myneni
122 and Knyazikhin, 2018).
- 123 – Daily means of three-hourly data at 9-km resolution, retrieved from the SMAP SPL4SMGP product version 8 (Reichle
124 et al., 2025), including 10-cm deep soil temperature (ST_{10}), surface skin temperature, incident shortwave solar radiation
125 (SW_{in}), surface soil moisture (SSM), and rootzone soil moisture (RZSM). SSM and RZSM estimates are obtained by
126 assimilating SMAP L-band brightness temperature observations into the Goddard Earth Observing System Version 5
127 Catchment Land Surface Model (GEOS-5 CLSM) (Reichle et al., 2019).
- 128 – Daily vapor pressure deficit (VPD) and minimum air temperature (MNT) at 0.25° (approximately 25-km resolution),
129 retrieved from the GEOS-5 Forward Processing (FP) product (Lucchesi, 2018).

130 SW_{in} is combined with FPAR to compute canopy-absorbed photosynthetically active radiation (APAR), assuming that PAR
131 constitutes 45 % of SW_{in} , as follows:

$$132 \text{ APAR} = 0.45 \cdot SW_{in} \cdot \text{FPAR} = \text{PAR} \cdot \text{FPAR} \quad (2)$$

133 RZSM is rescaled using a normalized logarithmic transformation (Jones et al., 2017), and SSM is converted from volumetric
134 units to relative wetness units. With the exception of APAR, the variables MNT, VPD, RZSM, ST , and SSM affect GPP and
135 ER estimation only after being converted into stress scalars, denoted as S_{MNT} , S_{VPD} , S_{RZSM} , S_{ST} , and S_{SSM} . The derivation of
136 these stress scalars is described later in Equations 7a-e.

137 The L4C model runs at a daily time step and is defined as follows:

$$138 \text{ NEE}(t) = \text{ER}(t) - \text{GPP}(t) \quad (3a)$$

$$139 \text{ GPP}(t) = \epsilon_{\max} \cdot \text{APAR}(t) \cdot S_{MNT}(t) \cdot S_{VPD}(t) \cdot S_{RZSM}(t) \quad (3b)$$

$$140 \text{ ER}(t) = \text{AR}(t) + \text{HR}(t) = \alpha \cdot \text{GPP}(t) + [k_1 \cdot \text{SOC}_1(t) + (1 - \eta) \cdot k_2 \cdot \text{SOC}_2(t) + k_3 \cdot \text{SOC}_3(t)] \cdot S_{ST}(t) \cdot S_{SSM}(t) \quad (3c)$$

141 GPP is modeled using a light-use efficiency approach (Jones et al., 2017; Xiao et al., 2013), where $\epsilon_{\max}$ represents the bulk en-
142 vironmental reduction in PAR conversion efficiency. AR is modeled as a fixed proportion of GPP, determined by the coefficient
143 α . HR is estimated using a cascading three-pool SOC decomposition model (Ise and Moorcroft, 2006; Kimball et al., 2008;

144 Jones et al., 2017), assuming that carbon fixed from atmospheric CO₂ through GPP enters the SOC pools as litterfall (L_{fall}).

145 The daily SOC change for each of the three SOC pools is specified as:

$$146 \text{ SOC}_1(t) = \text{SOC}_1(t-1) + [\lambda \cdot L_{\text{fall}}(t) - k_1 \cdot \text{SOC}_1(t-1) \cdot S_{\text{ST}}(t) \cdot S_{\text{SSM}}(t)] \cdot dt \quad (4a)$$

$$147 \text{ SOC}_2(t) = \text{SOC}_2(t-1) + [(1-\lambda) \cdot L_{\text{fall}}(t) - k_s \cdot \text{SOC}_2(t-1) \cdot S_{\text{ST}}(t) \cdot S_{\text{SSM}}(t)] \cdot dt \quad (4b)$$

$$148 \text{ SOC}_3(t) = \text{SOC}_3(t-1) + [\eta \cdot k_s \cdot \text{SOC}_2(t) \cdot S_{\text{ST}}(t) \cdot S_{\text{SSM}}(t) - k_r \cdot \text{SOC}_3(t-1) \cdot S_{\text{ST}}(t) \cdot S_{\text{SSM}}(t)] \cdot dt \quad (4c)$$

149 SOC₁, SOC₂, and SOC₃ represent the labile, structural, and recalcitrant SOC pools, respectively, with corresponding decay
 150 rates k₁, k₂ = 0.4 · k₁, and k₃ = 0.01 · k₁. The model parameters λ and η account for the fraction of L_{fall} allocated to the SOC₁
 151 and SOC₂ pools, and the fraction of material transferred from the SOC₂ pool to the SOC₃ pool, respectively. The parameters
 152 ε_{max}, k₁, λ, and η are treated as free parameters estimated during the optimization process. The model integration step dt is set
 153 to one day.

154 In the L4C model version 7, L_{fall} was derived as a constant daily fraction of the mean annual estimated net primary produc-
 155 tivity (NPP), as follows:

$$156 L_{\text{fall}}(t) = \overline{\text{NPP}_{\text{annual}}} \cdot 1/365 = (1-\alpha) \cdot \overline{\text{GPP}_{\text{annual}}} \cdot 1/365 \quad (5)$$

157 $\overline{\text{NPP}_{\text{annual}}}$ and $\overline{\text{GPP}_{\text{annual}}}$ denote the mean annual NPP and GPP, respectively, and the daily fraction is set to 1/366 for leap
 158 years. Instantaneous NPP is initially derived as NPP(t) = GPP(t) – AR(t). In the L4C model version 8, the allocation timing
 159 was changed from constant to dynamic, and is now determined using a leaf-loss function (L_{loss}), derived from climatological
 160 LAI, as follows:

$$161 L_{\text{fall}}(t) = \overline{\text{NPP}_{\text{annual}}} \cdot [f_E \cdot dt + (1-f_E) \cdot \frac{L_{\text{loss}}(t)}{\sum L_{\text{loss}}(t)}] \quad \text{with} \quad f_E = \frac{\min(\text{LAI})}{\max(\text{LAI})} \quad (6)$$

162 Here, f_E represents the proportion of the canopy that is evergreen. L_{loss} is computed using a triangular moving average centered
 163 on the current time step, where weights increase linearly toward the center. It represents the difference between lagged and
 164 leading climatological LAI values (Endsley et al., 2022).

165 In equations 3a-b and 4a-c, the stress scalars S_{MNT} , S_{VPD} , S_{RZSM} , S_{ST} , and S_{SSM} represent the ecosystem responses to their
 166 respective environmental variables and are defined as follows:

$$167 \quad S_{\text{MNT}}(t) = \min(1, \max(0, \frac{\text{MNT}(t) - \text{MNT}_{\min}}{\text{MNT}_{\max} - \text{MNT}_{\min}})) \quad (7a)$$

$$168 \quad S_{\text{VPD}}(t) = \min(1, \max(0, 1 - \frac{\text{VPD}(t) - \text{VPD}_{\min}}{\text{VPD}_{\max} - \text{VPD}_{\min}})) \quad (7b)$$

$$169 \quad S_{\text{RZSM}}(t) = \min(1, \max(0, \frac{\text{RZSM}(t) - \text{RZSM}_{\min}}{\text{RZSM}_{\max} - \text{RZSM}_{\min}})) \quad (7c)$$

$$170 \quad S_{\text{SSM}}(t) = \min(1, \max(0, \frac{\text{SSM}(t) - \text{SSM}_{\min}}{\text{SSM}_{\max} - \text{SSM}_{\min}})) \quad (7d)$$

$$171 \quad S_{\text{ST}}(t) = \min(1, \exp(\beta_0 \cdot (\frac{1}{\beta_1} - \frac{1}{\text{ST}(t) - \beta_2}))) \quad (7e)$$

172 Each stress scalar ranges from 0 to 1, where a value of 0 indicates that the environmental variable fully constrains model esti-
 173 mates, while a value of 1 indicates no constraint. The thresholds $\text{MNT}_{\min}$, $\text{MNT}_{\max}$, $\text{VPD}_{\min}$, $\text{VPD}_{\max}$, $\text{RZSM}_{\min}$, $\text{RZSM}_{\max}$,
 174 $\text{SSM}_{\min}$, and $\text{SSM}_{\max}$ are free parameters estimated during the optimization process. These are used as model thresholds and
 175 do not correspond to the actual minimum or maximum values within the time series. Similarly, β_0 is a free parameter, while
 176 β_1 and β_2 are fixed at 66.02 K and 227.13 K, respectively (Kimball et al., 2025). The behavior of the ecosystem response
 177 functions is shown in Appendix B, Figure B1.A1-C1. The GPP formulation originally includes a stress scalar based on the
 178 freeze-thaw state, computed using surface skin temperature from the SMAP SPL4SMGP product (Jones et al., 2017; Kimball
 179 et al., 2025). However, it is not shown here, as this study focuses on the growing season.

180 L4C model estimates are initially derived at a 1-km sub-grid resolution for up to 8 MODIS MCD12Q1 PFT classes, and
 181 then averaged to 9-km resolution. In this study, only the L4C model estimates corresponding to the PFT class in which the
 182 EC towers are located were considered (Table 1). A total of two towers are located in the grassland (GRA) class, 8 in the
 183 shrubland (SHR) class, and 10 in the evergreen needleleaf forest (ENF) class. Hereafter, NEE_{L4C} , GPP_{L4C} , and ER_{L4C} refer to
 184 daily estimates derived from the L4C model, which were retrieved from the SMAP SPL4CMDL product version 8.

185 4 Method

186 4.1 Arctic-Subarctic adapted model formulations

187 In this study, we explored alternative GPP and ER model formulations that retain the core GPP and ER equations of the original
 188 L4C model version 8. We aimed to preserve the structure and variable set of the original L4C model while enabling the incor-
 189 poration of constraints or additional flexibility guided by literature-based findings on ecosystem responses and flux-partitioning
 190 methods. Five different formulations are presented for both GPP and ER, with each formulation building incrementally on the

191 previous one by incorporating earlier modifications along with additional adjustments. Testing modifications incrementally,
 192 rather than independently, allowed us to determine whether their interactions improved or degraded model performance.
 193 The GPP formulations, labeled GPP₁ through GPP₅, mainly adjust ecosystem responses to environmental variables and are
 194 defined as follows (Table 3):

- 195 – GPP₁: Under sub-freezing air temperatures, photosynthetic activity is expected to be severely reduced, approaching
 196 cessation (Schaefer et al., 2012; Ensminger et al., 2004; Bowling et al., 2018; Parazoo et al., 2018). This behavior is not
 197 well represented in the original L4C model, where S_{MNT} still remains near 0.5 at 270 K (Figures 2, 3, 4.A2), indicating
 198 that GPP capacity is reduced by only half at this temperature. In the proposed formulation, GPP is ensured to cease when
 199 MNT is equal to or below 273.15 K by fixing the minimum threshold (MNT_{min}) of S_{MNT} to 273.15 K (Equation 7a).
- 200 – GPP₂ (defined as GPP₁ with additional adjustments): Some flux-partitioning methods use a nonlinear light-response
 201 curve to partition NEE_{EC} into GPP_{EC} and ER_{EC} (Lasslop et al., 2010b; Runkle et al., 2013), capturing the saturation of
 202 leaf-level photosynthesis at high solar irradiance. In the original L4C model, APAR directly scales the dynamic range
 203 of GPP and is not transformed through a transfer function into a stress scalar, as is the case for the other environmental
 204 variables (Equation 3a). In GPP₂, this linear dependence is replaced by a nonlinear stress scalar, S_{APAR} , inspired by the
 205 light-response curve, which is defined as follows:

$$206 \quad GPP(t) = GPP_{max} \cdot S_{APAR}(t) \cdot S_{MNT}(t) \cdot S_{VPD}(t) \cdot S_{RZSM}(t) \quad (8a)$$

$$207 \quad S_{APAR}(t) = \frac{APAR(t)}{APAR(t) + APAR_{crit}} \quad (8b)$$

208 At low APAR, GPP increases rapidly, but as APAR increases, the rate of increase slows down, and GPP asymptotically
 209 approaches a maximum value, GPP_{max} . This behavior is shown in Appendix B, Figure B1.D1. GPP_{max} and $APAR_{crit}$ are
 210 free parameters estimated during the optimization process.

- 211 – GPP₃ (defined as GPP₂ with additional adjustments): In the original L4C model, GPP responds linearly to MNT, VPD,
 212 and RZSM (Equations 7a,b,c). In GPP₃, GPP responses to these variables are modeled with increased flexibility, with no
 213 assumed shape other than monotonicity, to allow varying rates of change across different ranges. To achieve this, they
 214 are redefined as logistic ramp functions:

$$215 \quad S_{MNT}(t) = \frac{g(MNT(t)) - g(MNT_{min})}{g(MNT_{max}) - g(MNT_{min})} \quad \text{with} \quad g(MNT(t)) = \frac{1}{1 + \exp(-\gamma_{MNT} \cdot (MNT(t) - MNT_{crit}))} \quad (9a)$$

$$216 \quad S_{VPD}(t) = \frac{g(VPD(t)) - g(VPD_{max})}{g(VPD_{min}) - g(VPD_{max})} \quad \text{with} \quad g(VPD(t)) = \frac{1}{1 + \exp(\gamma_{VPD} \cdot (VPD(t) - VPD_{crit}))} \quad (9b)$$

$$217 \quad S_{RZSM}(t) = \frac{g(RZSM(t)) - g(RZSM_{min})}{g(RZSM_{max}) - g(RZSM_{min})} \quad \text{with} \quad g(RZSM(t)) = \frac{1}{1 + \exp(-\gamma_{RZSM} \cdot (RZSM(t) - RZSM_{crit}))} \quad (9c)$$

In this formulation, the thresholds MNT_{min} , MNT_{max} , VPD_{min} , VPD_{max} , $RZSM_{min}$, and $RZSM_{max}$ are used to scale the stress scalars between 0 and 1, and are fixed to 273.15 K, 293.15 K, 0 kPa, 2.5 kPa, 0 m^3m^{-3} , and 1 m^3m^{-3} , respectively. The parameters MNT_{crit} , γ_{MNT} , VPD_{crit} , γ_{VPD} , $RZSM_{crit}$, and γ_{RZSM} are treated as free parameters and are estimated during the optimization process. The behavior of the logistic ramp functions is illustrated in Appendix B, Figure B1.A2,B2.

- GPP_4 (defined as GPP_3 with additional adjustments): Growing degree days (GDD) are widely used in agricultural and ecological studies as a proxy to plant development (Fotouo Makouate and Zude-Sasse, 2025), and have recently been used to develop a phenology scheme that improved GPP modeling in a temperate bog (He et al., 2025). In the present formulation, GDD is incorporated into GPP modeling to capture the vegetation green-up and senescence phases through an additional stress scalar (S_{GDD}). GDD is first derived from mean air temperature (AT) using a base temperature of 273.15 K. It is then normalized for each site and each year using the annual minimum and maximum values, resulting in a normalized range from 0 to 1. This normalization ensures that GDD acts as a seasonal shape or trend driver, rather than a magnitude driver, which is instead represented by the instantaneous variables (APAR, MNT, VPD, and RZSM). Hereafter, GDD refers to normalized GDD and is used to derive S_{GDD} , as follows:

$$GPP_4(t) = GPP_{max} \cdot S_{APAR}(t) \cdot S_{MNT}(t) \cdot S_{VPD}(t) \cdot S_{RZSM}(t) \cdot S_{GDD}(t) \quad (10a)$$

$$S_{GDD}(t) = \frac{g(GDD(t))}{g\left(\frac{a}{a+b}\right)} \quad \text{with} \quad g(GDD(t)) = GDD(t)^a \cdot (1 - GDD(t))^b \quad (10b)$$

S_{GDD} is defined as a beta-like, bell-shaped function, normalized between 0 and 1 (Appendix B, Figure B1.C2). The parameters a and b are treated as free parameters and are estimated during the optimization process.

- GPP_5 (defined as GPP_4 with additional adjustments): Some studies have reported that water-saturated soil conditions limit oxygen and nutrient availability to plant roots, restrict cellular respiration, and consequently hinder photosynthetic activity (Kreuzwieser et al., 2004; Nawaz et al., 2025). Additionally, Peng et al. (2024) showed that the relationship between GPP and soil moisture may follow a bell-shaped curve at EC tower sites with PFTs similar to those in the present study (e.g., ENF, SHR, GRA; Table 1). A similar response has also been reported for peatlands (Valkenburg et al., 2023). In GPP_5 , it is similarly assumed that GPP peaks at an optimal RZSM level, beyond which excessive moisture reduces efficiency. This assumption is tested by redefining S_{RZSM} as a beta-like, bell-shaped function, analogous to S_{GDD} (Equation 10b and Figure B1.C2 in Appendix B):

$$S_{RZSM}(t) = \frac{g(RZSM(t))}{g\left(\frac{a}{a+b}\right)} \quad \text{with} \quad g(RZSM(t)) = RZSM(t)^a \cdot (1 - RZSM(t))^b \quad (11)$$

The parameters a and b are treated as free parameters and are estimated during the optimization process.

Regarding the ER modeling, we tested different formulations for the HR component while leaving the AR component unchanged. The ER formulations, labeled ER₁ through ER₅, are described below (Table 3):

- ER₁: Rather than using the the L_{fall} estimation scheme from the baseline L4C model version 8, ER₁ instead adopts the one from version 7 (Equation 5). The L4C model transitioned from version 7 to version 8 over the course of the present study was conducted, during which the L_{fall} estimation scheme was modified. The version 7 scheme was retained to enable comparison with the one introduced in version 8. Additionally, HR response to SSM (S_{SSM}) is redefined as a logistic ramp, analogous to S_{RZSM} in GPP₃ (Equation 9c and Figure B1.A2). The original linear response (Equation 7d) was directly replaced because the logistic ramp can reproduce a linear behavior if the relationship between SSM and HR is actually linear. The ER response function to ST is unchanged but was recalibrated (Equation 7e).

- ER₂ (defined as ER₁ with additional adjustments): The L_{fall} estimation scheme is reverted to that of the baseline L4C model version 8. Hence, ER₁ and ER₂ differ only in the L_{fall} estimation scheme (version 7 vs. 8), isolating its impact.

- ER₃ (defined as ER₂ with additional adjustments): None of the established flux-partitioning methods requires SOC data to derive GPP and ER from NEE (Reichstein et al., 2005; Lasslop et al., 2010b; Runkle et al., 2013; Helbig et al., 2017a). Consequently, in ER₃, we aimed to capture the added value of incorporating SOC dynamics in ER modeling. To this end, SOC dynamics are replaced by a single constant representing a baseline heterotrophic respiration rate, (R_{base}), as follows:

$$ER(t) = \alpha \cdot GPP(t) + R_{base} \cdot S_{ST}(t) \cdot S_{SSM}(t) \quad (12)$$

The parameter R_{base} is treated as a free parameter and is estimated during the optimization process.

- ER₄ (defined as ER₃ with additional adjustments): In the present formulation, we aimed to mimic the flux-partitioning methods that derive R_{base} every few days (Reichstein et al., 2005; Lasslop et al., 2010b; Runkle et al., 2013; Helbig et al., 2017a). R_{base} is then redefined as:

$$R_{base}(t) = (1 - \omega) \cdot R_{base}(t - 1) + \omega \cdot R_0 \cdot \overline{S_{ST}(t) \cdot S_{SSM}(t)} \quad (13)$$

This follows a first-order auto-regressive (AR1) approach, in which the current value depends on the previous one and the 7-day backward mean of the product of temperature and moisture stress scalars ($\overline{S_{ST}(t) \cdot S_{SSM}(t)}$), weighted by the parameter ω (ranging from 0 to 1). This approach introduces acclimation behavior by smoothing short-term variability in environmental conditions. The parameters R₀ and ω are treated as free parameters and are estimated during the optimization process.

- ER₅ (defined as ER₄ with additional adjustments): Endsley et al. (2022) showed that including an O₂ diffusion limitation in the HR response to SSM, which penalizes HR rates under high SSM conditions, improved seasonal ER performance. In ER₅, we adopted the beta-like, bell-shaped function, analogous to S_{RZSM} in GPP₅ (Equation 11), to model HR response to SSM. This avoids to collect or estimate O₂ concentration data while still representing diminishing returns on HR under high SSM conditions.

278 In this study, daily VPD and MNT were retrieved from the Modern-Era Retrospective Analysis for Research and Applica-
 279 tions, Version 2 (MERRA-2), M2T1NXSLV Version 5.12.4 product (Gelaro et al., 2017), instead of the GEOS-5 FP product,
 280 because it is sparsely documented. MERRA-2 re-analysis dataset is better constrained by observations, exhibits a climatology
 281 comparable to GEOS-5 FP and is used in the L4C model calibration due to its longer period of record. AT, required to derive
 282 GDD, was also retrieved from the MERRA-2 M2T1NXSLV product. Finally, RZSM and SSM from the SMAP SPL4SMGP
 283 product were retained in volumetric units (m^3m^{-3}), unlike in the original L4C model.

284 4.2 Growing season timing and data filtering

285 GPP_{EC} was used as an indicator to identify the growing season (Gonsamo et al., 2013). For each EC tower, GPP_{EC} values
 286 below the noise threshold of $0.05 \text{ gCm}^{-2}\text{d}^{-1}$ were first attributed to the winter season and removed. Among the remaining
 287 values, those below the 10th percentile were considered part of the shoulder seasons (i.e., transitional periods between fully
 288 frozen and fully thawed states) and were excluded. The growing season is more commonly defined using fixed fractions
 289 of annual maximum GPP_{EC} rather than percentile-based thresholds (Panwar et al., 2023; Luo et al., 2025). However, such
 290 approaches rely on the assumption that the annual maximum GPP_{EC} is robustly captured each year, which may not hold due
 291 to data gaps for instance. In addition, a previous study has shown that the growing season identification is sensitive to the
 292 chosen fraction of annual maximum GPP_{EC} , indicating a lack of consensus on an optimal threshold (Panwar et al., 2023). In
 293 contrast, the percentile-based approach used here provides a distribution-based criteria that is less sensitive to extreme values
 294 and inter-annual variability, ensuring consistency across years. For instance, two years with the same growing season timing
 295 (i.e., identical start and end dates) but different photosynthetic peaks are treated identically, which would not be the case when
 296 using the annual maximum GPP_{EC} . Although the 10th percentile threshold may not be optimal for each GPP_{EC} time series, it
 297 is consistent with the structure of the mean seasonal cycle of GPP_{EC} , where it separates winter and shoulder seasons from the
 298 growing season (Figure B2). In addition, similar separations are obtained when using thresholds within the 5th–15th percentile
 299 range, indicating low sensitivity to the chosen percentile. Finally, GPP_{EC} and ER_{EC} values above the 99th percentile were
 300 treated as outliers and removed.

301 Complementary filtering flags were applied to ensure biophysical plausibility of root-level soil activity and photosynthesis
 302 during the growing season from a modeling perspective. The specific criteria for these flags are as follows:

- 303 – $\text{ST}_{10\text{cm}} \geq 275.15 \text{ K}$ (i.e., $2 \text{ }^\circ\text{C}$ above freezing)
- 304 – $\text{ST}_{20\text{cm}} \geq 275.15 \text{ K}$
- 305 – $\text{ST}_{39\text{cm}} \geq 275.15 \text{ K}$ (applied to ENF sites only, see Table 1)
- 306 – $\text{MNT} \geq 275.15 \text{ K}$

307 $\text{ST}_{20\text{cm}}$ and $\text{ST}_{39\text{cm}}$ refer to soil temperature at 20 and 39 cm depths, respectively, and were retrieved from the SMAP
 308 SPL4SMGP product. For the remainder of this study, ST refers to $\text{ST}_{10\text{cm}}$ as $\text{ST}_{20\text{cm}}$ and $\text{ST}_{39\text{cm}}$ are not used further.

309 After filtering, a total of 1,650 data points (23 %) remained for the upland tundra ecosystem, 4,632 data points (33 %) for
310 the taiga forest ecosystem, and 3,653 data points (30 %) for the wetland ecosystem (Table 1).

311 4.3 Model formulation calibration

312 The AS-adapted GPP and ER formulations were calibrated separately for each ecosystem type (upland tundra, taiga forests,
313 and wetlands), using GPP_{EC} and ER_{EC} as reference targets. The optimization framework for calibration used least-squares
314 minimization via the MATLAB lsqcurvefit function (MathWorks, Inc., 2023), which minimizes the sum of squared differences
315 between the model outputs and target values. This is an unconstrained optimization, with no additional penalty terms applied.
316 To mitigate overfitting, the optimization process was repeated 100 times using different random training (70 %) and testing
317 (30 %) subsets of the data for each ecosystem type: 1,155 (495) data points for upland tundra, 3,243 (1,389) for taiga forests and
318 2,557 (1,096) for wetlands. ~~Final model~~ In each optimization run, model parameters were estimated using the corresponding
319 training subset, and final parameter values were ~~taken defined~~ as the median across all runs of the parameter estimates across
320 all 100 runs. This approach aimed to capture diverse data combinations and promote a more stable and representative parame-
321 terization by smoothing out the influence of outliers or any individual biased subset. A 70 % subset size was arbitrarily chosen
322 to balance between providing sufficient data for robust model calibration and retaining enough data variability across the 100
323 iterations (Martinez Molera, 2025). Contrary to the global calibration of the original L4C model, no reference SOC data were
324 used to constrain the recalibration over the AS environments. As a result, in ER_1 and ER_2 , only the SOC_1 pool was modeled
325 to avoid potential parameter compensation and to prevent unrealistic SOC distribution across the original three pools. This re-
326 duction in the number of SOC pools constitutes an additional adjustment for ER_1 and ER_2 that is confounded with the change
327 in L_{fall} estimation scheme (Section 4.1). An initial guess was necessary for SOC_1 on March 31, 2015, to explicitly solve the
328 SOC dynamics, since the system is formulated recursively and requires a starting value to iterate forward in time (Equation 4).
329 March 31, 2015, was chosen as the start of the simulation because it precedes the first date of the period of study. The initial
330 guess was set to 0 gCm^{-2} for the first spin-up iteration, providing a neutral starting point to avoid biasing the early simulation.
331 It was subsequently updated using the SOC value on March 31, 2022, corresponding to the last March 31 within the study
332 period. A total of 20 spin-up iterations were performed.

333 4.4 Model formulation evaluation

334 For each ecosystem type, the AS-adapted GPP and ER formulations were evaluated spatiotemporally, capturing the combined
335 effects of spatial and temporal variability, with GPP_{EC} and ER_{EC} used as reference targets. For each of the 100 optimization
336 runs (Section 4.3), the Pearson correlation coefficient (r), the unbiased root mean square error (ubRMSE), and the bias (B)
337 were computed separately for the training and testing subsets, enabling the assessment of model calibration and generalization,
338 respectively. Median values across the 100 runs were reported. The trade-off between goodness of fit and model complexity
339 was quantified using the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC), which were applied on
340 the training subsets. The best-performing ER and GPP formulations were identified based on the lowest ΔAIC and ΔBIC
341 values, defined as $AIC - AIC_{L4C}$ and $BIC - BIC_{L4C}$, respectively. Median values across the 100 runs were reported. As a

complementary analysis, site-level (temporal-only) performance was evaluated for each EC tower by computing r , ubRMSE, and B using model outputs obtained with the median parameter values across the 100 runs. For this analysis, training and testing data were pooled, and median metrics across EC towers were reported.

As the ER formulations require GPP as an input, all 25 possible ER configurations were systematically evaluated (ER₁ through ER₅ combined with GPP₁ through GPP₅). This approach also yields 25 corresponding NEE configurations, as NEE is defined as the difference between ER and GPP (Equation 3). Evaluating all combinations enables assessment of whether the best-performing GPP formulation is associated with the best-performing ER formulation, and whether the combination of the best-performing ER and GPP formulations also results in the most accurate NEE configuration. It additionally enables identification of cases where GPP and ER formulations that perform less well individually (relative to ER_{EC} and GPP_{EC}) nevertheless yield a more accurate NEE configuration (relative to NEE_{EC}). Such behavior may arise because errors in modeled ER and GPP can either compensate or accumulate when computing NEE.

The 25 NEE configurations were evaluated following the same approach as for the GPP and ER formulations, using NEE_{EC} as the reference target. Although NEE configurations were not directly calibrated, the same training and testing subsets were retained for consistency with the GPP and ER formulation evaluations, ensuring that each NEE_{EC} value was paired with its corresponding GPP_{EC} and ER_{EC} values.

Because the evaluation of 25 ER formulations produces many metrics, only ER₁ through ER₅ using the GPP input that yields the best performance are presented in the main text. Results for all configurations are provided in Appendix D. This avoids bias arising from the choice of GPP input when comparing ER formulations. For NEE, summarized performance for all configurations is presented in the main text, while detailed results are provided in Appendix E. Hereafter, NEE_{ij} denotes modeled NEE computed as ER_i - GPP_j.

5 Results

5.1 Gross primary production

5.1.1 Upland tundra

Based on the spatiotemporal evaluation of the testing splits for upland tundra (Table 4), GPP₁ performs better than GPP_{L4C} in terms of B and ubRMSE (-0.33 vs. 0.41 gCm⁻²d⁻¹ and 1.19 vs. 1.46 gCm⁻²d⁻¹, respectively). However, GPP_{L4C} achieves a higher r value (0.56 vs. 0.64). Introducing a nonlinear light-response in GPP₂ (Equation 8) leads to better performance compared with GPP₁, reducing B from -0.33 to -0.02 gCm⁻²d⁻¹. In addition, ubRMSE decreases, and r increases, approaching the r observed for GPP_{L4C}. Replacing linear ramps with logistic ramps to simulate ecosystem responses in GPP₃ (Equation 9) increases model complexity but provides limited improvement over GPP₂. GPP₄ incorporates GDD through an additional stress scalar (Equation 10). This results in improved r and ubRMSE compared with GPP₃ (0.75 vs. 0.64, and 0.84 vs. 0.96 gCm⁻²d⁻¹, respectively). In GPP₅, a bell-shaped function is used to simulate the influence of RZSM (Equation 11). This further improves r and ubRMSE, though B slightly increases (-0.03 vs. -0.01 gCm⁻²d⁻¹).

374 Considering the site-level evaluation (Table 4), GPP₄ and GPP₅ exhibit the highest r (0.77 vs. 0.76) and the lowest ubRMSE
375 (0.65 vs. 0.71 gCm⁻²d⁻¹). In contrast, GPP₅, GPP₃, and GPP₂ show the lowest B with 0.04, 0.05, and 0.08 gCm⁻²d⁻¹,
376 respectively.

377 Across all formulations, S_{VPD} remains equal to 1 throughout the entire range of VPD variability (Figure 2.B3–F3). The use
378 of a non linear light-response in GPP₂ introduces an early saturation (Figure 2.B6–F6), where modeled GPP peaks are lower
379 than those of GPP_{EC}. This premature flattening is progressively reduced in GPP₄ and GPP₅, due to the incorporation of GDD
380 and the use of a bell-shaped function for simulating RZSM influence.

381 The lowest ΔAIC and ΔBIC are obtained for GPP₅, followed by GPP₄, GPP₃, GPP₂ and GPP₁ in last place (Table C1). For
382 GPP₂ and GPP₃, ΔAIC and ΔBIC are equal within each subset (training or testing) because both formulations have the same
383 number of free parameters as GPP_{L4C}.

384 5.1.2 Taiga forests

385 The spatiotemporal evaluation of the testing splits for taiga forests (Table 4) indicates that GPP₁ performs better than GPP_{L4C} in
386 terms of r (0.62 vs. 0.58) and ubRMSE (1.75 vs. 1.93 gCm⁻²d⁻¹), although GPP₁ shows higher B (-0.44 vs. -0.37 gCm⁻²d⁻¹).
387 As in upland tundra (Section 5.1.1), introducing a nonlinear light-response in GPP₂ (Equation 9) leads to overall better perfor-
388 mance compared with GPP₁, notably reducing B from -0.44 gCm⁻²d⁻¹ to -0.05 gCm⁻²d⁻¹. GPP₃ does not offer improve-
389 ment over GPP₂, aside from a slight reduction in B (-0.02 gCm⁻²d⁻¹ vs. -0.05 gCm⁻²d⁻¹). Due to the inclusion of GDD
390 through an additional stress scalar (Equation 10), GPP₄ outperforms GPP₃ (0.73 vs. 0.67 for r , and 1.30 vs. 1.41 gCm⁻²d⁻¹
391 for ubRMSE). Switching to a bell-shaped function for simulating RZSM influence in GPP₅ (Equation 11) does not result in
392 improved performance compared with GPP₄.

393 Considering the site-level evaluation (Table 4), the highest r are obtained for GPP₄ and GPP₅ (0.79 vs. 0.76). These two
394 formulations also achieve the lowest ubRMSE, with 1.12 gCm⁻²d⁻¹ for GPP₅ and 1.15 gCm⁻²d⁻¹ for GPP₄. GPP₄ and
395 GPP₄ exhibits the lowest B (-0.13 vs. -0.14 gCm⁻²d⁻¹).

396 RZSM appears to be a negligible input in GPP_{L4C}, as S_{RZSM} remains equal to 1 throughout the entire range of RZSM
397 variability (Figure 3.A4). However, RZSM gains more effect in GPP₁ through GPP₅, although its effect remains weaker than
398 that of MNT, VPD, and GDD (Figure 3.B4–F4). As in upland tundra (Section 5.1.1), the use of a nonlinear light-response
399 in GPP₂ introduces an early saturation, underestimating GPP_{EC} peaks (Figure 3.B6–F6). However, this premature flattening
400 persists in GPP₄ and GPP₅, despite the incorporation of GDD and the use of a bell-shaped function for simulating RZSM
401 influence.

402 The lowest ΔAIC and ΔBIC are obtained for GPP₅, closely followed by GPP₄, then GPP₃, GPP₂ and GPP₁ in last place
403 (Table C1). For GPP₂ and GPP₃, ΔAIC and ΔBIC are equal within each subset (training or testing) because both formulations
404 have the same number of free parameters as GPP_{L4C}.

405 5.1.3 Wetlands

406 Based on the spatiotemporal evaluation of the testing splits for wetlands (Table 4), GPP_1 outperforms GPP_{L4C} , notably exhibit-
407 ing reduced ubRMSE and B (1.33 vs. 2.18 $gCm^{-2}d^{-1}$, and -0.39 vs. 1.32 $gCm^{-2}d^{-1}$, respectively). As seen in upland tundra
408 and taiga forests (Sections 5.1.1 and 5.1.2), introducing a nonlinear light-response in GPP_2 (Equation 8) results in improved
409 r (0.63 vs. 0.53), reduced ubRMSE (1.04 vs. 1.33 $gCm^{-2}d^{-1}$), and reduced B (-0.05 vs. -0.39 $gCm^{-2}d^{-1}$), compared with
410 GPP_1 . GPP_3 shows only a minor improvement over GPP_3 . Due to the inclusion of GDD through an additional stress scalar
411 (Equation 10), GPP_4 outperforms GPP_3 (0.72 vs. 0.65 for r , and 0.92 vs. 1.01 $gCm^{-2}d^{-1}$ for ubRMSE). In contrast to up-
412 land tundra and taiga forests (Sections 5.1.1 and 5.1.2), using a bell-shaped function for simulating RZSM influence in GPP_5
413 (Equation 11), result in degraded performance compared with GPP_4 .

414 Considering the site-level evaluation (Table 4), GPP_4 and GPP_5 exhibit the highest r (0.79 vs. 0.76). GPP_4 and GPP_3
415 exhibits the lowest ubRMSE (0.76 vs. 0.79 $gCm^{-2}d^{-1}$, respectively), closely followed by GPP_5 (0.82 $gCm^{-2}d^{-1}$) and GPP_3
416 (0.85 $gCm^{-2}d^{-1}$). B is similar for GPP_2 through GPP_5 , with -0.22, -0.22, -0.23, and -0.19 $gCm^{-2}d^{-1}$, respectively.

417 Similar to upland tundra (Section 5.1.1), S_{VPD} remains equal to 1 throughout the entire range of VPD variability across all
418 formulations (Figure 4.B3–F3). Likewise, as observed in taiga forests (Section 5.1.2), RZSM appears to be a negligible input
419 in GPP_{L4C} , with S_{RZSM} staying equal to 1 throughout the entire range of RZSM variability (Figure 4.A4). However, RZSM
420 gains a considerable effect in GPP_1 through GPP_5 , especially under dry conditions (Figure 4.B4–F4). The introduction of the
421 nonlinear light-response in GPP_2 triggers an early saturation (Figure 4.B6–F6), that persists throughout GPP_5 .

422 In contrast to upland tundra and taiga forests (Sections 5.1.1 and 5.1.2), the lowest ΔAIC and ΔBIC are obtained for GPP_4 ,
423 followed by GPP_5 , GPP_3 , GPP_2 , and GPP_1 in last place (Table C1). For GPP_2 and GPP_3 , ΔAIC and ΔBIC are equal within
424 each subset (training or testing) because both formulations have the same number of free parameters as GPP_{L4C} .

425 5.2 Ecosystem respiration

426 For upland tundra and taiga forests, ER formulations using GPP_5 as input are compared, as GPP_5 yielded the lowest ΔAIC and
427 ΔBIC (Sections 5.1.1, 5.1.2). For wetlands, the set using GPP_4 as input is compared instead, because GPP_4 achieved the lowest
428 ΔAIC and ΔBIC for this ecosystem type (Section 5.1.3). Results for all 25 ER formulations are provided in Tables D1–D6.

429 5.2.1 Upland tundra

430 Based on the spatiotemporal evaluation of the testing splits for upland tundra (Table 5), ER_1 , which uses the approach where
431 mean annual NPP is allocated uniformly across the year to L_{fall} (Equation 5), performs better than ER_{L4C} . It exhibits higher
432 r (0.52 vs. 0.44), reduced ubRMSE (0.72 vs. 0.99 $gCm^{-2}d^{-1}$), and reduced B (-0.12 vs. 0.37 $gCm^{-2}d^{-1}$). Switching to the
433 dynamic allocation in ER_2 (Equation 6) leads to enhanced r (0.61) and ubRMSE (0.64 $gCm^{-2}d^{-1}$) compared with ER_1 . Using
434 a constant R_{base} term instead of a SOC model in ER_3 (Equation 12) results in r and ubRMSE close to those of ER_2 , and lower
435 B (-0.01 vs. -0.12 $gCm^{-2}d^{-1}$). Using a dynamic R_{base} in ER_4 (Equation 13) shows better performance than ER_3 with greater

436 r (0.72 vs. 0.60) and reduced ubRMSE (0.53 vs. 0.62 $\text{gCm}^{-2}\text{d}^{-1}$). In ER₅, the use of a bell-shaped function to simulate SSM
437 influence does not provide a clear benefit.

438 Considering the site-level evaluation (Table 5), the highest r are obtained for ER₂ and ER₅ (0.66 vs. 0.60). The low-
439 est ubRMSE is observed for ER₄ and ER₅ (0.41 vs. 0.42 $\text{gCm}^{-2}\text{d}^{-1}$), closely followed by the other formulations (up to
440 0.51 $\text{gCm}^{-2}\text{d}^{-1}$ for ER₁). In terms of B, ER₅, ER₄, and ER₂ show the smallest values with 0.00, -0.01, and 0.03 $\text{gCm}^{-2}\text{d}^{-1}$.

441 ST appears to have a stronger effect than SSM in ER₁ through ER₄ (Figure 5.B1–E1 vs. B2–E2). S_{ST} mainly oscillates
442 around 0.5, while S_{SSM} rapidly reaches 1 under dry conditions, and wet conditions do not constrain model outputs. In ER₅, the
443 use of a bell-shaped function to simulate SSM increases its effect (Figure 5.F2), but without any improvement in performance,
444 as previously observed.

445 The lowest ΔAIC and ΔBIC are obtained for ER₄, closely followed by ER₅, then ER₃ and ER₂, and ER₁ in last place (Ta-
446 ble D6). Overall, a similar pattern is observed regardless of the GPP input, and the performance of each ER formulation remains
447 comparable or improves as the GPP input increases in complexity (lowest with GPP₁, highest with GPP₅; Tables D1–D6).

448 **5.2.2 Taiga forests**

449 Based on the spatiotemporal evaluation of the testing splits for taiga forests (Table 5), ER₁ outperforms ER_{L4C}, showing en-
450 hanced r (0.52 vs. 0.34), reduced ubRMSE (1.28 vs. 1.71 $\text{gCm}^{-2}\text{d}^{-1}$), but slightly higher B (-0.17 vs. -0.12 $\text{gCm}^{-2}\text{d}^{-1}$).
451 Switching from a constant to dynamic allocation of mean annual NPP to L_{fall} in ER₂ (Equation 6) results in a minor improve-
452 ment in performance over ER₁. Using a constant R_{base} term instead of a SOC model in ER₃ (Equation 12) exhibits similar r
453 (0.54 vs. 0.54) and ubRMSE (1.23 vs. 1.25 $\text{gCm}^{-2}\text{d}^{-1}$), but lower B (0.02 vs. -0.12 $\text{gCm}^{-2}\text{d}^{-1}$), compared with ER₂. Using
454 a dynamic R_{base} in ER₄ (Equation 13) shows better performance than ER₃ with greater r (0.59 vs. 0.54) and reduced ubRMSE
455 (1.17 vs. 1.23 $\text{gCm}^{-2}\text{d}^{-1}$). In ER₅, the use of a bell-shaped function to simulate SSM influence does not provide a clear
456 benefit.

457 Considering the site-level evaluation (Table 5), r is similar across formulations, ranging from 0.61 to 0.65. The same pattern
458 is observed for ubRMSE, which ranges from 0.95 to 1.01 $\text{gCm}^{-2}\text{d}^{-1}$. The lowest B are obtained for ER₃, ER₄, and ER₂, with
459 0.00, -0.05, and -0.09 $\text{gCm}^{-2}\text{d}^{-1}$, respectively.

460 SSM appears to be a negligible input in ER_{L4C}, as S_{SSM} remains equal to 1 throughout the entire range of SSM variability
461 (Figure 6.A2). However, SSM gains more effect in ER₁ through ER₅, although its effect remains weaker than that of ST
462 (Figure 6.B2–F2).

463 The lowest ΔAIC and ΔBIC are obtained for ER₅, nearly tied with ER₄, then ER₃, ER₂, and ER₁ in last place. As observed
464 in upland tundra (Section 5.2.1), this ranking remains consistent regardless of the GPP input, and the performance of each
465 ER formulation is comparable or improves as the GPP input increases in complexity (lowest with GPP₁, highest with GPP₅;
466 Tables D1–D6).

467 5.2.3 Wetlands

468 Based on the spatiotemporal evaluation of the testing splits for taiga forests (Table 5), ER_1 outperforms ER_{L4C} , exhibiting
469 enhanced r (0.49 vs. 0.24), reduced ubRMSE (0.81 vs. $1.80 \text{ gCm}^{-2}\text{d}^{-1}$), and reduced B (-0.05 vs. $1.77 \text{ gCm}^{-2}\text{d}^{-1}$). Switching
470 from a constant to dynamic allocation of mean annual NPP to L_{fall} in ER_2 (Equation 6) slightly increases r (0.53 vs. 0.49), and
471 reduces ubRMSE and B (0.78 vs. $0.81 \text{ gCm}^{-2}\text{d}^{-1}$, -0.03 vs. $-0.05 \text{ gCm}^{-2}\text{d}^{-1}$, respectively). Using a constant or dynamic
472 R_{base} term instead of a SOC model in ER_3 and in ER_4 (Equations 12 and 13) does not result in enhanced performance. In ER_5 ,
473 the use of a bell-shaped function to simulate SSM influence does not provides any benefits either.

474 Considering the site-level evaluation (Table 5), r is similar across formulations, ranging from 0.63 to 0.67. The same pattern
475 is observed for ubRMSE, with the highest value assigned to ER_1 ($0.56 \text{ gCm}^{-2}\text{d}^{-1}$) and the lowest to ER_4 ($0.48 \text{ gCm}^{-2}\text{d}^{-1}$).
476 ER_1 and ER_2 exhibiting the lowest B , with -0.12 and $-0.09 \text{ gCm}^{-2}\text{d}^{-1}$, respectively (Figure 8.C2).

477 As observed in taiga forests (Section 5.2.2), SSM appears to be a negligible input in ER_{L4C} , with S_{SSM} staying equal to 1
478 throughout the entire range of SSM variability (Figure 7.A2). However, SSM gains a considerable effect in ER_1 through ER_5 ,
479 especially under dry conditions (Figure 7.B2-F2). ER_{L4C} may overestimate ER_{EC} by up to a factor of two or three (Figure 7.A3).
480 This magnitude discrepancy is largely removed in ER_1 through ER_5 , but a pattern persists, in which ER_{EC} are systematically
481 overestimated at low values (approximately 0 to $1 \text{ gCm}^{-2}\text{d}^{-1}$) across all formulations (Figure 7.B3-F3).

482 The lowest ΔAIC and ΔBIC are obtained for ER_2 , then ER_3 , ER_4 , ER_5 , and ER_1 in last place. Nevertheless, all formu-
483 lations are nearly tied. A consistent tied ranking is observed regardless of which GPP is used as input (Tables D1–D6). The
484 performance of each ER formulation (ER_1 through ER_5) remains comparable across formulations, although using GPP_1 as
485 input lead to the lowest performance.

486 5.3 Net ecosystem CO_2 exchange

487 This subsection presents the performance of the AS-adapted NEE formulations (ER_1 through ER_5 combined with GPP_1
488 through GPP_5) and NEE_{L4C} relative to NEE_{EC} .

489 5.3.1 Upland tundra

490 Based on the spatiotemporal evaluation for upland tundra, $NEE_{1-5,4-5}$ exhibit higher r than the other NEE configurations,
491 with values above 0.6 (Figure 8.A1). In contrast, NEE_{L4C} shows a r of 0.57. The lowest ubRMSE are obtained for $NEE_{2-5,5}$
492 and $NEE_{2-3,4}$ (Figure 8.B1), ranging from 0.65 to $0.70 \text{ gCm}^{-2}\text{d}^{-1}$ (Figure 8.B1). By comparison, NEE_{L4C} yields a value of
493 $0.84 \text{ gCm}^{-2}\text{d}^{-1}$. $NEE_{3-5,2-5}$ and $NEE_{1,1}$ show the lowest absolute B , with values below $0.05 \text{ gCm}^{-2}\text{d}^{-1}$, comparable to that
494 of NEE_{L4C} ($0.03 \text{ gCm}^{-2}\text{d}^{-1}$; Figure 8.C1).

495 Overall, $NEE_{2-5,4-5}$ appears to perform best in terms of higher r , and lowest ubRMSE and B . This set of configurations
496 includes the ER and GPP formulations associated with the lowest ΔAIC and ΔBIC (ER_4 , ER_4 , GPP_4 , and GPP_5 ; Sections 5.1.1
497 and 5.2.1). Although their performance is broadly comparable, some configurations may perform better depending on the
498 metric considered (r , ubRMSE or B). A similar pattern is observed considering the site-level evaluation (Tables E1–E5).

499 5.3.2 Taiga forests

500 For taiga forests, the spatiotemporal evaluation reveals that $NEE_{4-5,4-5}$ achieve the highest r with values between 0.60 and 0.65,
501 whereas NEE_{L4C} shows a r of 0.55 (Figure 8.A2). $NEE_{4-5,4}$ and $NEE_{5,5}$ exhibit the lowest ubRMSE, ranging from 1.05 to
502 $1.10 \text{ gCm}^{-2}\text{d}^{-1}$ (Figure 8.B2). $NEE_{4,5}$, $NEE_{3-5,3}$, and $NEE_{3,4}$ closely follow, with values between 1.10 and $1.15 \text{ gCm}^{-2}\text{d}^{-1}$.
503 In contrast, NEE_{L4C} shows a ubRMSE of $1.17 \text{ gCm}^{-2}\text{d}^{-1}$). $NEE_{3-4,2-5}$ and $NEE_{5,3-5}$ exhibit absolute B below $0.05 \text{ gCm}^{-2}\text{d}^{-1}$,
504 lower than both the other NEE configurations and NEE_{L4C} ($0.26 \text{ gCm}^{-2}\text{d}^{-1}$; Figure 8.C2).

505 Overall, $NEE_{4-5,4-5}$ appears to perform best in terms of higher r , and lowest ubRMSE and B. This set of configurations
506 includes the ER and GPP formulations associated with the lowest ΔAIC and ΔBIC (ER_4 , ER_4 , GPP_4 , and GPP_5 ; Sections 5.1.2
507 and 5.2.2). A similar pattern is observed considering the site-level evaluation (Tables E1-E5).

508 5.3.3 Wetlands

509 Based on the spatiotemporal evaluation for wetlands, $NEE_{3-5,4}$ exhibit the highest r , ranging from 0.45 and 0.5, comparable to
510 that of NEE_{L4C} (0.47; (Figure 8.A3)). $NEE_{2-5,4}$, $NEE_{5,5}$, $NEE_{2,5}$, and $NEE_{2,3}$ exhibit the lowest ubRMSE with values between
511 0.90 and 0.95

512 Overall, $NEE_{1-5,4}$ appears to perform best in terms of higher r , and lowest ubRMSE and B. This set of configurations includes
513 the ER and GPP formulations associated with the lowest ΔAIC and ΔBIC (ER_2 , ER_3 , and GPP_4 ; Sections 5.1.3 and 5.2.3). A
514 similar pattern is observed considering the site-level evaluation (Tables E1-E5).

515 6 Discussion

516 This section first discusses how the modifications implemented in the AS-Adapted GPP and ER model formulations affected
517 their performance relative to GPP_{EC} and ER_{EC} . We specifically identify candidate GPP model adjustments for operational
518 implementation, evaluate the contribution of incorporating SOC dynamics into ER modeling, and examine the influence of
519 input variables. We then outline the implications for NEE estimation and highlight the limitations of our study.

520 6.1 Candidate GPP model adjustments

521 Implementing a nonlinear light-response function to represent the influence of APAR on GPP (GPP_2 , Equation 8) appears
522 to be the most effective adjustment tested for reducing both ubRMSE and B across the three ecosystem types (Section 5.1,
523 and Table 4, Figure 8). However, our results indicate that this adjustment requires careful parametrization, as it can lead to
524 underestimation of GPP peaks (Figures 2–4).

525 Adding GDD into the GPP modeling (Equation 10) complements the nonlinear light-response adjustment by further reducing
526 ubRMSE and predominantly improving r across the three ecosystem types (Section 5.1, Table 4, Figure 8). These results
527 suggest that the current L4C model may lack a phenological proxy that accounts for the progressive functional adjustment
528 of vegetation to environmental conditions over time (Maire et al., 2012), thereby complementing the instantaneous proxies

currently used as inputs. Vegetation indices are assumed to capture vegetation phenology by tracking seasonal changes in canopy structure and greenness. Because GDD is derived from mean air temperature and minimum air temperature is used as a proxy for the instantaneous temperature response of GPP, replacing GDD with a vegetation index may help reduce redundancy (Huang et al., 2019; Pulliainen et al., 2024). In this context, LAI or normalized difference vegetation index (NDVI) could potentially serve this role. However, LAI is likely to introduce additional redundancy, as VIIRS LAI retrievals are already used to derive FPAR, which represents canopy phenology in the GPP formulation (Equation 2). In contrast, NDVI may provide a more independent phenological proxy without duplicating existing model inputs. Nonetheless, MODIS and VIIRS vegetation indices exhibit large uncertainties at high northern latitudes, particularly during shoulder seasons, due to extensive cloud cover and snow contamination (Xu et al., 2018; Pu et al., 2023).

Overall, the nonlinear light-response adjustment appears to be the strongest candidate for correcting GPP magnitude discrepancies, while incorporating GDD emerges as the most effective adjustment to improve GPP seasonal dynamics. These two findings are consistent with McCallum et al. (2013), where the authors reported that the inclusion of temperature acclimation and nonlinear light-response in GPP modeling in Russian boreal forests improved model performance. Future studies could explore more complex light-response functions, such as a rectangular hyperbolic function, where the parameters may vary temporally with temperature, as suggested by Wang et al. (2014a). However, testing this approach would imply departing from the current multiplicative structure of the L4C model, in which direct mechanistic forcing–response behaviors are represented. Prior work also suggests that vegetation green-up onset is influenced by winter chilling accumulation and precipitation (Fu et al., 2014; Elmendorf and Hollister, 2023). Greater accumulation of chilling days may lead to earlier green-up, as vegetation exposed to colder winter temperatures requires less thermal accumulation (less GDDs) to initiate spring growth. In contrast, higher winter precipitation may contribute to delayed green-up through thicker snowpacks, cooler soil temperatures, and increased cloud cover that reduces incoming radiation. Future improvements could therefore integrate winter chilling days and precipitation into the normalized GDD-based phenological proxy to better represent early-season GPP dynamics.

Finally, it is noteworthy that GDD was normalized using annual minimum and maximum values, implying that early-season GDD values implicitly depend on values from later in the same year. This approach therefore requires the full annual air temperature cycle to be known, introducing a one-year lag in the computation of model outputs that is not appropriate for real-time forecasting. However, the primary objective of this study is not prediction in an operational setting, but rather the retrospective reconstruction of CO₂ fluxes and the estimation of CO₂ budgets. For potential forecasting applications, GDD can instead be normalized using long-term minimum and maximum values across all years, resulting in comparable or slightly lower performance (see GPP₄ and GPP₅ in Table 4 vs. Table C2). This alternative approach avoids any within-year information leakage and suggests that GDD normalization has a limited influence on the overall performance improvements associated with adding GDD as an input.

6.2 Comparison of SOC-based and empirical approaches for ER modeling

Updating the allocation of mean annual NPP to L_{fall} from a constant to a LAI-based formulation to represent SOC dynamics (ER₁ vs. ER₂; Equations 5 and 6) improves ER model performance. Both the spatiotemporal evaluation and the median

metrics across EC towers indicate higher r and lower ubRMSE and B, with stronger improvements for upland tundra and weaker improvements for taiga forests and wetlands (Table 5). The benefits are limited relative to the added model complexity, especially in taiga forests and wetlands, compared with the simpler approach that replaces SOC dynamics with a single constant R_{base} (ER₃, Equation 12). Based on the spatiotemporal evaluation, introducing temporal variability in R_{base} (ER₄, Equation 13) leads to improved performance in upland tundra and taiga forests (Table 5). However, the median metrics across EC towers do not indicate a clear improvement across the three ecosystems (Table 5). Compared to upland tundra and taiga forests, all ER formulations are highly similar for wetlands, regardless of the metrics considered (r , ubRMSE, B, ΔAIC , and ΔBIC) and the type of evaluation (spatiotemporal vs. site-level; Tables 5 and D6).

Overall, using SOC dynamics with the L_{fall} estimation scheme from the L4C model version 8 to model ER appears to be the most suitable approach, as it performs better than version 7 and is physically grounded and mechanistically interpretable compared with the two empirical approaches. Unfortunately, the improved performance of ER₁ and ER₂ compared to ER_{L4C} is difficult to interpret, as it reflects the combined effects of changes in the L_{fall} estimation scheme, SOC pool structure, and GPP input (Sections 4.1, 4.3). Therefore, this comparison does not constitute a clean test of the added-value of the L_{fall} allocation scheme version 8 compared to version 7 for the study regions.

Continuing to explore alternative ways to estimate L_{fall} may be a promising direction for future research. However, the assumption that mean annual NPP can serve as a proxy for the magnitude of L_{fall} may not be realistic (Sierra et al., 2022). In addition, the timing of NPP allocation to L_{fall} may not accurately represent litter production dynamics, particularly given the large uncertainties in LAI and FPAR retrievals at high northern latitudes (Xu et al., 2018; Pu et al., 2023). Furthermore, because NPP is derived from modeled GPP, any inaccuracies in GPP propagate directly into modeled NPP, L_{fall} , SOC, and ultimately ER. Finally, recent work in Alaska has shown that implementing vertical SOC transport to simulate depth-dependent L_{fall} , SOC distribution, and corresponding HR rates may further improve ER estimates (Yi et al., 2020).

6.3 Tested but unretained GPP and ER model adjustments

Implementing logistic ramps to represent GPP responses to MNT, VPD, and RZSM stress (GPP₃, Equation 9) provides limited benefits based on both the spatiotemporal and site-level evaluation (GPP₂ vs. GPP₃ in Table 4). This suggests that MNT, VPD, and RZSM may exhibit nonlinear interactions with GPP, but this adjustment appears to be of secondary importance compared to the implementation of a light-response curve and the incorporation of GDD (Section 6.1).

The use of a bell-shaped function to represent RZSM influence on GPP (GPP₅, Equation 11) provides only a limited performance improvement in upland tundra and no improvement in taiga forests (Table 4). One possible explanation is that, in upland tundra, RZSM exhibits both dry and wet conditions across years and EC tower sites (grey histogram in Figure 2.F4), whereas in taiga forests, conditions remain mostly dry with less seasonal variation (Figure 3.F4). This pattern is supported by the bimodal distribution of RZSM in upland tundra, in contrast to the unimodal distribution in taiga forests. Nevertheless, the RZSM distribution in wetlands is bimodal (Figure 4.F4), with both dry and wet conditions, but the bell-shaped function worsens the performance of modeled GPP (Table 4). The diminishing returns under wet conditions appear to penalize model calibration, indicating a different ecosystem response to RZSM in wetlands compared with upland tundra and taiga forests. This may

reflect the adaptation of wetland vegetation to anaerobic conditions, where excessive RZSM does not hinder photosynthesis activity. Although several studies show that wetlands and peatlands are more sensitive to drought than to flooding (Churchill et al., 2015; Olefeldt et al., 2017; Heinzelmann et al., 2025), there is no clear evidence in the literature on whether GPP in wetlands continues to scale or levels off under high RZSM conditions. It is also noteworthy that the bell-shaped function does not improve model performance for any ecosystem when normalized RZSM is used (not shown), as is the case in the original L4C model (Section 3). Finally, the L4C model methodology focuses on direct mechanistic forcing–response behavior, where instantaneous RZSM data are used as input. However, a temporal lag in GPP response to RZSM saturation may be expected, as it can take several days to weeks for soil oxygen levels to become depleted to the point of restricting aerobic processes under saturation. A larger number of EC towers should also be included to increase RZSM variability during calibration before drawing conclusions about the value of this adjustment for North American AS regions.

As in the GPP modeling, the use of a bell-shaped function to represent SSM influence on ER provides no clear improvement, regardless of ecosystem type or whether dry and wet SSM conditions are included during calibration (Table 5 and Figures 5-7). These results indicate that an unidirectional function ramp is more appropriate, with dry conditions limiting ER rates, and no diminishing returns under wet conditions. The same conclusion is drawn when SSM expressed in relative wetness units is used (not shown), as in the original L4C model (Section 3). This finding partly contrasts with Endsley et al. (2022), who reported improved seasonal ER performance after adding an O₂ diffusion limitation (also based on SSM) to the original monotonic linear response, thereby penalizing ER rates under high SSM. The differing behavior between studies may be attributed first to differences in SSM response functions and, second, to the fact that in Endsley et al. (2022), the SPL4SMGP product did not yet account for peatland hydrology (Reichle et al., 2023).

6.4 Key drivers in shaping model performance

As VPD rises, indicating atmospheric dryness, plants typically show stomatal closure to minimize water loss, which in turn reduces their photosynthetic activity (López et al., 2021). However, in upland tundra and wetlands, GPP appears insensitive to VPD as the corresponding stress scalar S_{VPD} remains equal to 1 across all five AS-adapted formulations (Figures 2.B3-F3 and 4.B3-F3). This suggests that either the VPD response is inadequately represented in the formulations, or that vegetation in these areas is inherently less responsive to stomatal closure than in taiga forests, where S_{VPD} strongly constrains the modeling (Figure 3.B3-F3). Indeed, VPD distributions are similar across the three ecosystem types, which supports the idea that the observed differences in S_{VPD} are not due to differing environmental conditions, but rather to ecosystem-specific sensitivity. In other words, for the same VPD values, the model applies a stronger constraint to vegetation photosynthetic activity in taiga forests, while in upland tundra and wetlands it remains unconstrained. These findings are consistent with those of Chen et al. (2023), where the authors observed that increasing VPD did not hinder vegetation growth in northern peatlands. Additionally, Zona et al. (2023) reported that VPD was not correlated to GPP at monthly scales in Arctic tundra, while Mirabel et al. (2023) found that tree growth in the Canadian boreal forest responded negatively to rising VPD.

Across all three ecosystem types, the most notable model improvement arises from revising the influence of APAR and AT (through GDD) on GPP (Table 4, Figure 8, Section 6.1). Interestingly, APAR and AT are also the two drivers primarily

used to partition NEE_{EC} into GPP_{EC} and ER_{EC} (Section 2). This indicates that model performance is inherently entangled with these two drivers, rather than to VPD, RZSM, SSM, SOC, or ST. It is important to note that drivers used in the L4C model formulations are provided at 9-km and 25-km resolution (except FPAR), which is coarse relative to EC tower footprints (Sections 2 and 3). Some of the discrepancies between EC measurements and model estimates may therefore be attributed to representativeness errors, as the coarse model resolution is expected to smooth spatial variability that is captured by the EC measurements. Coupled with the candidate GPP model adjustments (Section 6.1), using higher spatial resolution PAR and AT inputs could represent a promising avenue for improving GPP estimates and, consequently, ER estimates in future studies.

6.5 Implications for NEE estimation

Overall, the individually best-performing ER and GPP formulations, when combined, yield the most, or among the most, accurate NEE configurations (Section 5.3, Figure 8, and Tables E1–E5). Among the 25 tested NEE configurations, those using GPP_4 and GPP_5 have the strongest influence on NEE performance, particularly for r , and to a lesser extent for ubRMSE. This is consistent with the inclusion of GDD as a seasonal driver into the GPP modeling (Equation 10). In contrast, no clear and consistent pattern emerges regarding the impact of ER formulations on NEE performance, even though they exhibit distinct performance when evaluated against ER_{EC} (Section 5.2). The main exception is ER_1 , which systematically leads to higher ubRMSE and absolute B in NEE. Finally, although the best-performing GPP and ER formulations showed strong performance gains relative to the original L4C model, particularly for GPP, the resulting improvement in NEE performance is more modest (Tables 4, 5 vs. Tables E1-E5 and Figure 8). These results suggest that improvements in the representation of ER and GPP do not necessarily lead to comparable improvements in NEE (Figure B3). Rather than compensating for each other, errors in modeled ER and GPP may accumulate when computing NEE, at least for the formulations tested.

6.6 Influence of temporal autocorrelation on model performance

6.6 Influence of temporal and spatial autocorrelation on model performance

The current model validation approach ~~uses random-70~~used random 70–30 % ~~–30~~ training–testing subsets for model calibration and evaluation (Sections 4.3 and 4.4). ~~As a result~~The 70–30 % splits were applied to individual data points. Consequently, model performance may be partly influenced by temporal autocorrelation, since GPP_{EC} and ER_{EC} data points from the same EC tower site may be separated by only a few days while belonging to the training and testing subsets, respectively. ~~Consequently, a complementary sensitivity analysis of model performance was conducted using a validation approach that provides a stronger temporal separation~~Model performance may also be influenced by spatial autocorrelation, as data from the same site are included in both the training and testing subsets. ~~Consequently, two complementary sensitivity analyses were conducted using alternative validation approaches that provide stronger temporal and spatial separation, respectively,~~ between the data subsets used for calibration and evaluation. ~~Instead of using random-70~~

662 The first alternative validation approach consisted of withholding complete years from model calibration instead of using
663 random 70–30 % ~~–30~~ training–testing subsets; ~~complete years were withheld from model calibration~~. For each ecosystem, the
664 available data were assigned to 6-year training and 2-year testing periods. Because the study period spans 8 years, this yielded
665 28 possible unique train–test runs. The sizes of the training and testing subsets vary-varied among runs because data availability
666 differs among years and sites. This ~~alternative~~ validation approach reduces the potential influence of temporal autocorrelation
667 but introduces variability in ~~sample-subset~~ size among runs, which may itself affect performance metrics. Nevertheless, results
668 obtained with this complementary approach were comparable to those obtained with the random ~~7070–30 % –30 splits splits~~
669 applied to the individual data points, indicating that the main findings of this study are robust to ~~the validation strategy employed~~
670 temporal autocorrelation (Tables C3 and D7 vs. Tables 4 and 5).

671 In the second alternative validation approach, the 70–30 % training–testing subsets were applied to the set of sites within
672 each ecosystem rather than to the individual data points, such that all data from a given site were assigned exclusively to
673 either the training or the testing subset. Because the number of sites differs among ecosystems, the number of possible unique
674 train–test runs was 5, 21, and 28 for upland tundra, taiga forests, and wetlands, respectively. This validation approach reduces
675 the potential influence of spatial autocorrelation while providing a stricter assessment of model ability to generalize to unseen
676 sites. However, as in the first alternative approach, the sizes of the training and testing subsets varied among runs. Overall,
677 results obtained with this complementary approach were comparable to those obtained with the other two validation approaches
678 (Tables ~~?? and ??~~. C4 and D8 vs. Tables 4 and 5 and vs. Tables C3 and D7). The most notable difference was that absolute
679 B was higher for the testing splits in some cases. However, this does not alter the main findings of the study. This difference
680 may reflect variations in the spatial representativeness of the training subsets, in subset size, in site composition among runs,
681 or a combination of these factors. Nevertheless, these effects cannot be disentangled with the present dataset, particularly for
682 upland tundra, where only five sites were available (Table 1).

683 6.7 Limitations

684 The reference GPP_{EC} and ER_{EC} used for calibration and evaluation are derived from NEE_{EC} partitioning, meaning they are
685 not direct measurements but modeled outputs based on NEE_{EC} and structural assumptions (Appendix A). This creates a po-
686 tential circularity in model evaluation, as the AS-adapted formulations may share the same structural assumptions as the
687 flux-partitioning algorithms. Consequently, improvements in modeled ER and GPP may partly reflect the model formulations
688 reproducing the behavior of the flux-partitioning algorithms, rather than independently improving the representation of car-
689 bon dynamics. In some cases, GPP_{EC} is constrained to follow a light-response curve, which is why a similar adjustment was
690 tested in GPP_2 (Equation 8). At first glance, this structural similarity likely explains why the adjustment enhanced performance
691 (Section 6.1). However, in other cases, GPP_{EC} is not directly modeled, but derived as the residual between NEE_{EC} and ER_{EC} .
692 Therefore, it is difficult to determine whether improvements from GPP_2 reflect better reproduction of specific flux-partitioning
693 algorithms, a more accurate representation of the true flux dynamics, or a combination of both. In contrast, GDD is not used at
694 all in flux-partitioning algorithms (Appendix A). Therefore, the improvements resulting from the inclusion of GDD in GPP_4
695 may capture true ecosystem state changes that are also well represented in GPP_{EC} .

696 Regarding ER, the situation is more complex. ER_{EC} is typically derived from a fitted power-based or exponential-based
697 function (Appendix A). These functions depend solely on temperature and estimate the combined contribution of AR and HR
698 as a single inseparable flux. In contrast, the L4C model and the tested AS-adapted formulations explicitly represent ER as the
699 sum of AR and HR, with each component estimated separately using multiple drivers, including APAR, GDD, MNT, VPD,
700 RZSM, ST, and SSM. This approach relies on assumed linkages between GPP and AR, and between GPP, L_{fall} , SOC, and
701 HR (Kimball et al., 2008), resulting in a more mechanistic, interaction-rich framework than the flux-partitioning algorithms.
702 Consequently, calibrating ER formulations is challenging, because the reference ER_{EC} is obtained using a simpler empirical
703 approach, which may limit model performance. If the ultimate goal is to estimate the CO_2 budget accurately, rather than to
704 predict the underlying GPP and ER components, it may be advantageous to calibrate the L4C model using NEE_{EC} as the
705 reference, rather than relying on GPP_{EC} and ER_{EC} as intermediate targets. However, this approach prevents validating whether
706 the modeled GPP and ER truly reflect the underlying processes and strongly limits the number of free parameters that can be
707 estimated, since only a single reference is available instead of two. For future research, it could also be valuable to partition
708 NEE_{EC} into GPP_{EC} and ER_{EC} using a more mechanistic approach similar to the L4C model, explicitly distinguishing between
709 AR and HR.

710 In summary, when calibrating TCF models using GPP_{EC} and ER_{EC} as references, one attempts to explain variability in fluxes
711 that originate from a reference framework with a relatively simple structure, a limited number of drivers, and parameters that
712 may vary in space and time (Appendix A). In contrast, TCF models, such as the L4C model, rely on a more complex process
713 representation, a larger set of environmental drivers, and parameters that are assumed to be constant in space and time within
714 a given ecosystem. These fundamental differences inherently complicate model calibration, hinder the interpretation of model
715 performance, and limit our ability to determine whether the underlying processes of ER and GPP are realistically represented
716 when extrapolated to larger spatial and temporal scales.

717 Several studies have also shown that GPP responds to the ratio of leaf-internal to ambient CO_2 concentration (Wang et al.,
718 2014b, 2017). Although this ratio is regulated by environmental conditions such as temperature and VPD, neither the original
719 L4C model nor the tested AS-adapted formulations and flux-partitioning algorithms explicitly accounts for the response of
720 GPP to changes in ambient CO_2 concentration. Because ambient CO_2 varies over time and may continue to increase in the
721 future, this omission may limit the ability of the L4C model to accurately predict GPP over long temporal scales.

722 Finally, it is important to note that this study focuses exclusively on the growing season, whereas the ultimate objective of
723 improving the L4C model for the North American AS is to better estimate the full annual CO_2 budget over recent years by
724 integrating modeled NEE since 2015. Although CO_2 flux magnitudes are highest during the growing season, GPP and AR
725 become minimal or absent during the shoulder seasons and winter, while HR persists, even under snow-covered and frozen
726 soil conditions. Several studies have shown that the winter and shoulder seasons play a critical role in shaping the annual CO_2
727 budget (Kim et al., 2013; Natali et al., 2019), as they primarily constitute a CO_2 source due to the dominance of HR. Therefore,
728 future work will focus on improving the L4C model for these periods to provide more reliable year-round estimates of NEE,
729 GPP, and ER, as well as more representative annual CO_2 budgets.

730 7 Conclusions

731 The goal of this study was to refine the integration of energy and moisture proxies into the SMAP L4C GPP and ER modeling
732 for the North American AS growing season. To this end, alternative GPP and ER model formulations were calibrated and
733 evaluated against GPP_{EC} and ER_{EC} across upland tundra, taiga forests and wetlands, covering the period from 2015 to 2022.

734 Ultimately, we recommend two key adjustments related to energy proxies to enhance the L4C model ability to monitor the
735 GPP process:

- 736 – Implementing a nonlinear light-response, particularly to reduce ubRMSE and B;
- 737 – Incorporating GDD to reflect vegetation green-up and senescence phases, thereby improving seasonal dynamics.

738 In contrast, model adjustments related to moisture proxies (VPD, RZSM, SSM) for both ER and GPP modeling do not
739 currently emerge as essential for future operational implementation. Moreover, evaluating the benefits of integrating SOC
740 dynamics into ER modeling remains challenging, even though the L4C version 8 approach represents an improvement over
741 that of version 7, and therefore further research into ER modeling is recommended.

742 Overall, combining the best-performing ER and GPP formulations yields the most, or among the most, accurate NEE con-
743 figurations. However, improvements in NEE are generally more modest than for the individual components.

744 We further encourage the scientific community to harmonize strategies between flux-partitioning algorithms and mechanistic
745 modeling frameworks (such as the L4C model), particularly for estimating ER and its underlying HR and AR components.
746 The alignment between partitioning and modeling frameworks is essential to enhance the reliability of spatial and temporal
747 extrapolation of GPP_{EC} and ER_{EC} using satellite-based TCF models.

748 Finally, future work will aim to improve the representation of shoulder and winter seasons in the L4C model, as these periods
749 are critical for accurately capturing year-round CO_2 fluxes and annual CO_2 budgets in North American AS regions.

750 *Author contributions.* RM, ArM, and AR designed and conducted the study, and wrote the first draft of the manuscript. KAE and JSK
751 provided scientific support throughout the study and contributed to the second version of the manuscript. OS, HA, SNW, AM provided some
752 of the eddy covariance data. All authors contributed to the final version of the manuscript.

753 *Competing interests.* The authors declare no conflict of interest.

754 *Acknowledgements.* R. Madelon, A. Mialon and A. Roy acknowledge fundings from the Université of Toulouse 3 Paul Sabatier in France,
755 and the Université du Québec à Trois-Rivières in Canada through the Universalis Causa, Samuel-de-Champlain, Programme d'Aide à
756 L'Internationalisation de la Recherche (PAIR), Fonds Canadien de l'Innovation (FCI) and Conseil de Recherches en Sciences Naturelles
757 et en Génie du Canada (CRSNG), and Fonds de recherche du Québec (FRQ) grants. J.S. Kimball and K.A. Endsley acknowledge fundings

758 from NASA (grant no. NX14AI50G). G. De Lannoy acknowledges funding from KU Leuven C1 (C14/21/057). O. Sonnentag acknowl-
759 edges financial support through the Canada Research Chair and NSERC Discovery Grants programs, ArcticNet, a Network of Centres of
760 Excellence Canada, the Canada First Research Excellence Fund's Global Water Futures program (Northern Water Futures), and the Polar
761 Continental Shelf Program. This research made use of eddy covariance data from the AmeriFlux network (<https://ameriflux.lbl.gov>). Ad-
762 ditional eddy covariance data were directly provided by the principal investigators O. Sonnentag and S. N. Williamson. We thank all the
763 principal investigators and site teams of the eddy covariance tower sites (listed in Table 1) for their sustained efforts in site maintenance,
764 data processing, and data sharing. The SMAP SPL4CMDL Version 8 product can be downloaded from the National Snow and Ice Data
765 Center (NSIDC) website at: <https://nsidc.org/data/SPL4CMDL/Versions/8>. The SMAP SPL4SMGP Version 8 product can be downloaded
766 at: <https://nsidc.org/data/SPL4SMGP/Versions/8>. The MERRA-2 M2T1NXSLV Version 5.12.4 product is available at: [goldsmr4.gesdisc.](https://goldsmr4.gesdisc.eosdis.nasa.gov/data/MERRA2/M2T1NXSLV.5.12.4)
767 [/eosdis.nasa.gov/data/MERRA2/M2T1NXSLV.5.12.4](https://goldsmr4.gesdisc.eosdis.nasa.gov/data/MERRA2/M2T1NXSLV.5.12.4). All the authors thank the indigenous and northern communities for agreeing to the
768 installation of eddy covariance towers on their territory. As a non-native English speaker, R. Madelon acknowledges the use of artificial in-
769 telligence tools to improve the writing quality of the manuscript. Finally, the authors acknowledge the constructive reviews from the referees,
770 which have improved both the clarity of the manuscript and its content.

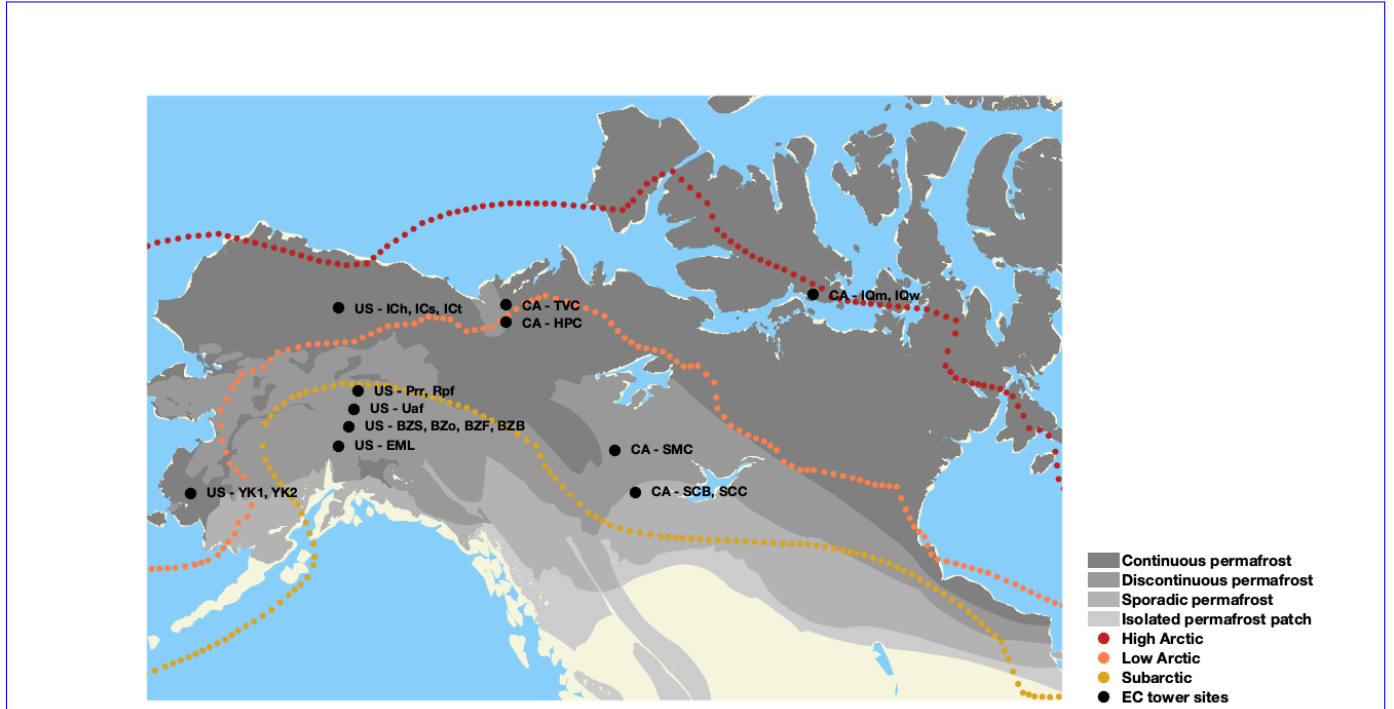


Figure 1. Locations of the 20 eddy covariance (EC) towers providing measurements of net ecosystem CO₂ exchange (NEE_{EC}), NEE-derived gross primary production (GPP_{EC}) and ecosystem respiration (ER_{EC}) from April 2015 to December 2022. The High Arctic, Low Arctic and Subarctic zones were delineated following the Conservation of Arctic Flora and Fauna (CAFF) working group of the Arctic Council, using Moderate Resolution Imaging Spectroradiometer (MODIS) and Landsat imagery (Potapov et al., 2008). The permafrost extent is estimated in percent areal coverage (Brown et al., 2002): continuous (>90-100% areal extent), discontinuous (>50-90%), sporadic (10-50%) and isolated patches (<10%). Due to overlapping, a single color dot may represent up to four EC towers on the map. Information for each EC tower is listed in Table 1. The figure was inspired by Madelon et al. (2025) and Mavrovic et al. (2023a).

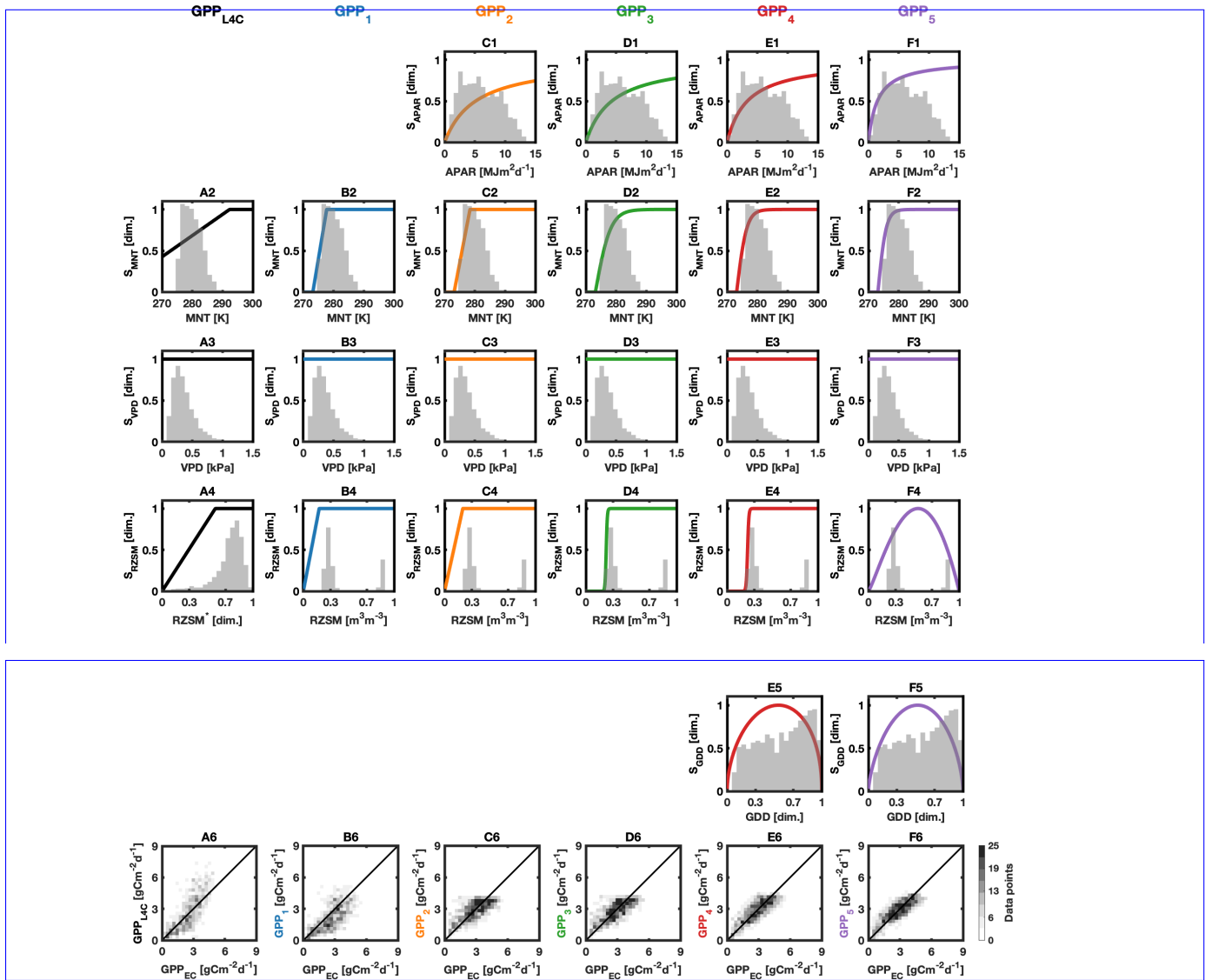


Figure 2. Ecosystem responses used to compute modeled **gross primary production (GPP)** from the L4C model (GPP_{L4C} , column A) and from the Arctic-Subarctic (AS) adapted formulations (GPP_1 through GPP_5 , columns B–F) in **upland tundra** during the growing season. The ecosystem responses are expressed as stress scalars, S_x , for each environmental variable x , where x represents absorbed photosynthetically active radiation (APAR), minimum air temperature (MNT), vapor pressure deficit (VPD), root zone soil moisture (RZSM), and normalized growing degree days (GDD). In the background of each subplot, the histogram of the corresponding environmental variable is shown in light grey. **All ecosystem responses of a given model formulation are calibrated jointly using eddy covariance GPP (GPP_{EC}) as reference, and not based on the histogram values.** In subplot A4, RZSM* refers to the normalized RZSM used as input in the original L4C model. Subplots A1 and B1 are omitted because APAR is used as a direct input, rather than through a stress scalar, in GPP_{L4C} and GPP_1 formulations. Similarly, subplots A5–D5 are omitted because GDD is not used as input in the GPP_{L4C} , GPP_1 , GPP_2 , and GPP_3 formulations. Refer to Sections 3 and 4.1 for a detailed description of each model formulation. The last row displays scatter plots between modeled GPP against GPP_{EC} . It is a spatiotemporal comparison, accounting for both temporal and spatial variability. It includes all available GPP_{EC} data points for upland tundra after filtering, including those used for calibration (1,628 data points; Section 4.3).

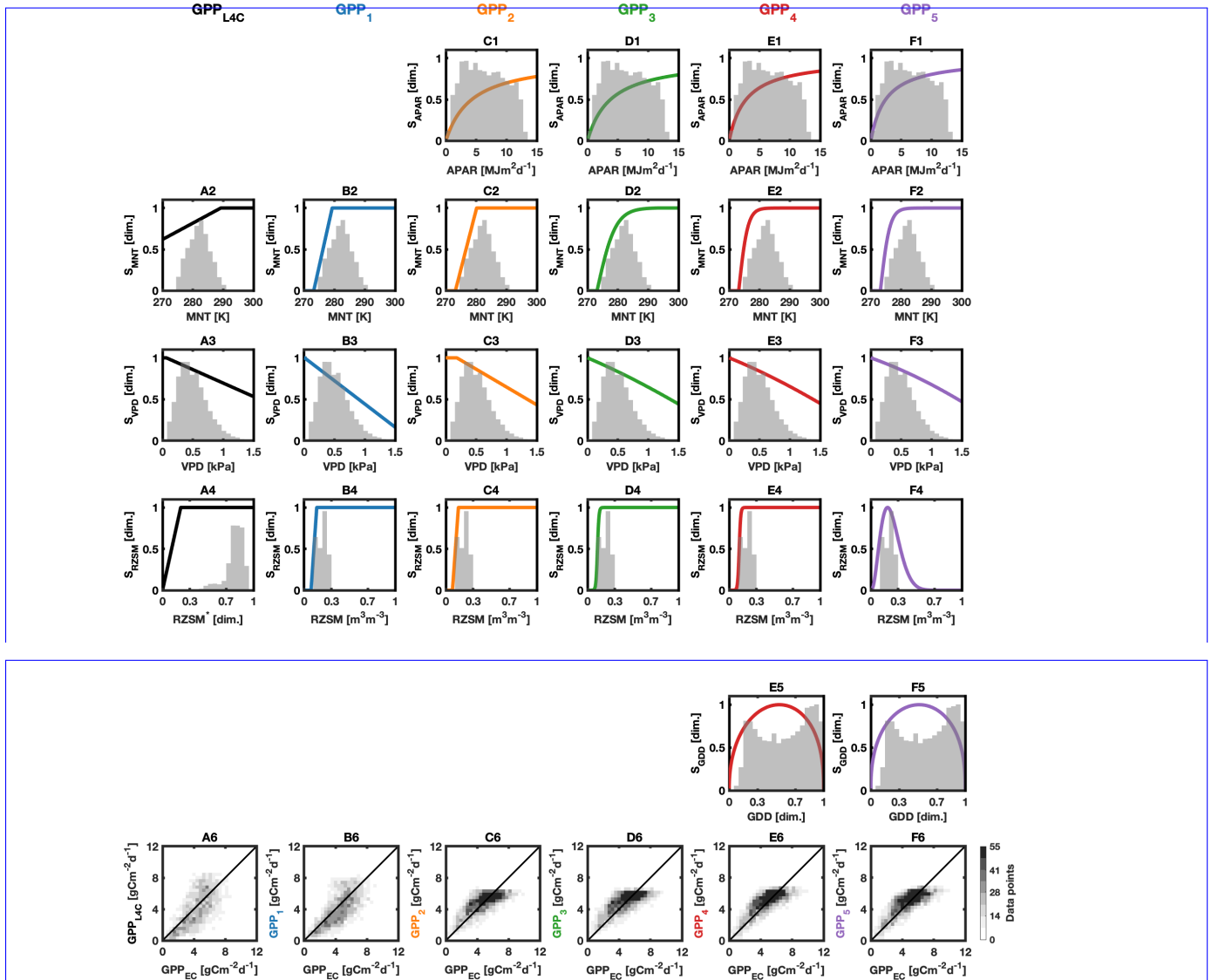


Figure 3. Same as Figure 2, but for **taiga forests**. The spatiotemporal comparison includes all available GPP_{EC} data points for taiga forests after filtering, including those used for calibration (4,632 data points; Section 4.3).

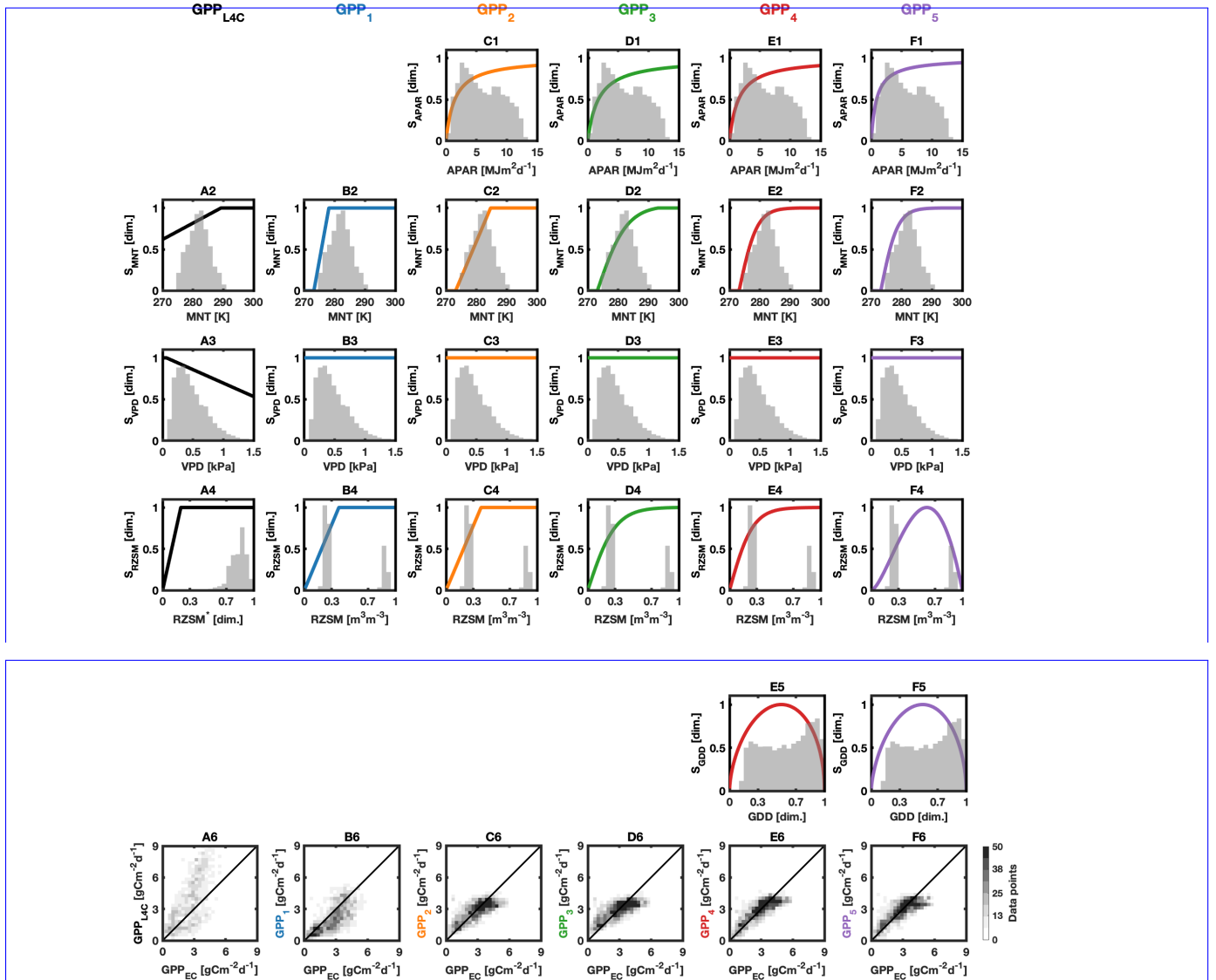


Figure 4. Same as Figure 2, but for **wetlands**. The spatiotemporal comparison includes all available GPP_{EC} data points for wetlands after filtering, including those used for calibration (3,653 data points; Section 4.3).

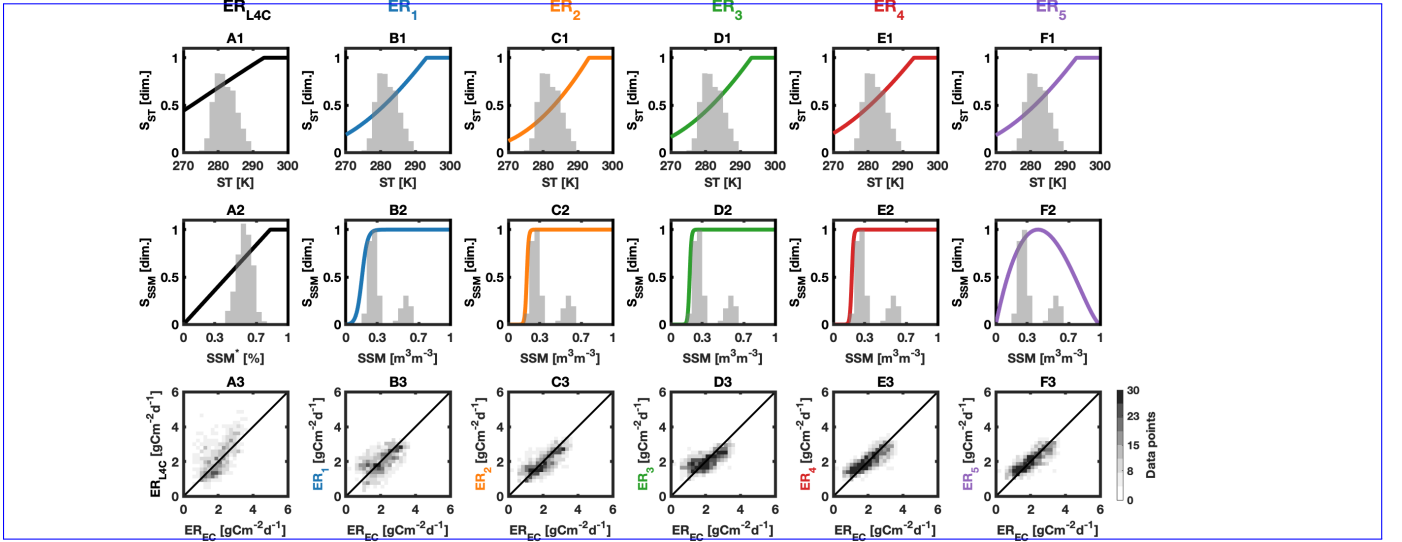


Figure 5. Ecosystem responses used to compute modeled **ecosystem respiration (ER)** from the L4C model (ER_{L4C} , column A) and from the Arctic-Subarctic (AS) adapted formulations (ER_1 through ER_5 , columns B–F) in **upland tundra** during the growing season. The ecosystem responses are expressed as stress scalars, S_x , for each environmental variable x , where x represents soil temperature (ST) and surface soil moisture (SSM). In the background of each subplot, the histogram of the corresponding environmental variable is shown in light grey. **All ecosystem responses of a given model formulation are calibrated jointly using daily-averaged eddy covariance ER (ER_{EC}) as reference, and not based on the histogram values.** In subplot A2, SSM* refers to the SSM in relative wetness unit used as input in the original L4C model. Refer to Sections 3 and 4.1 for a detailed description of each model formulation. The last row displays scatter plots between modeled ER against ER_{EC} . It is a spatiotemporal comparison, accounting for both temporal and spatial variability. It includes all available ER_{EC} data points for upland tundra after filtering, including those used for calibration (1,650 data points; Section 4.3). GPP_5 was used as input for ER_1 through ER_5 (Section 5.2.1).

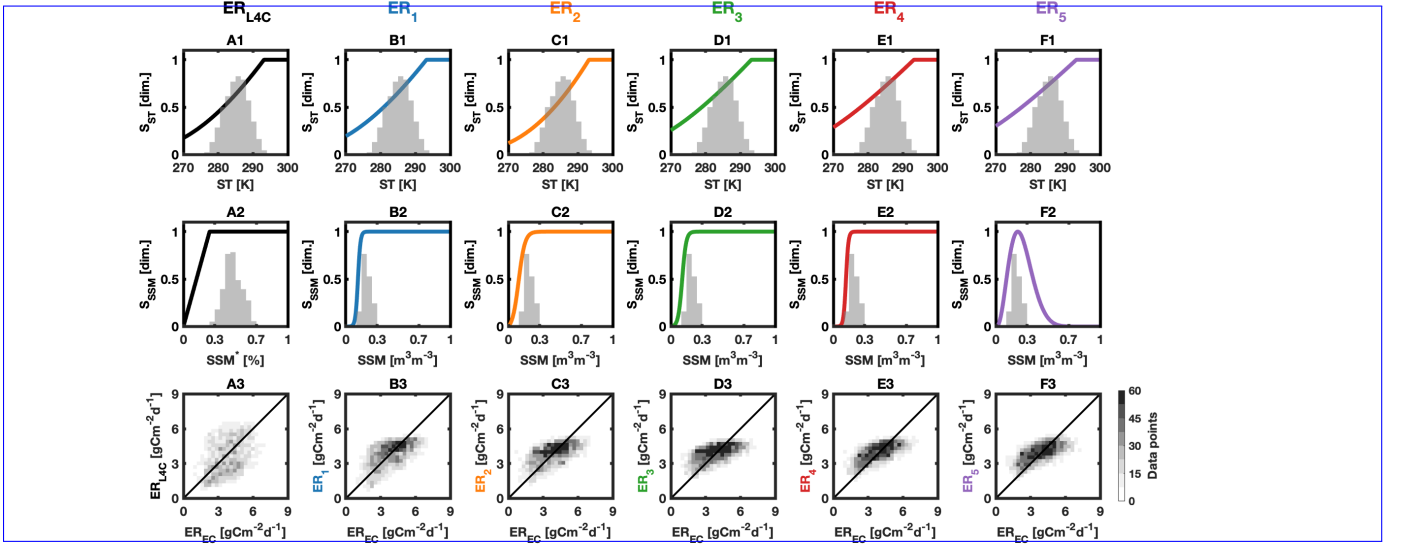


Figure 6. Same as Figure 5, but for **taiga forests**. The spatiotemporal comparison includes all available ER_{EC} data points for taiga forests after filtering, including those used for calibration (4,632 data points; Section 4.3). GPP_5 was used as input for ER_1 through ER_5 (Section 5.2.2).

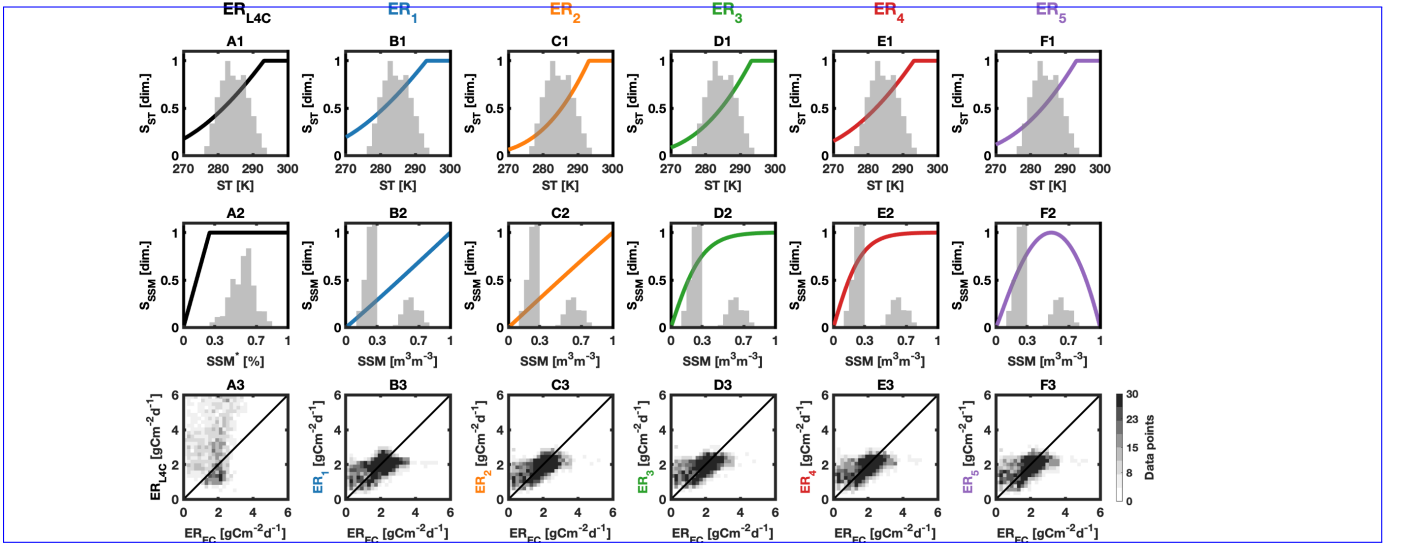


Figure 7. Same as Figure 5, but for **wetlands**. The spatiotemporal comparison includes all available ER_{EC} data points for wetlands after filtering, including those used for calibration (3,653 data points; Section 4.3). GPP_4 was used as input for ER_1 through ER_5 (Section 5.2.3).

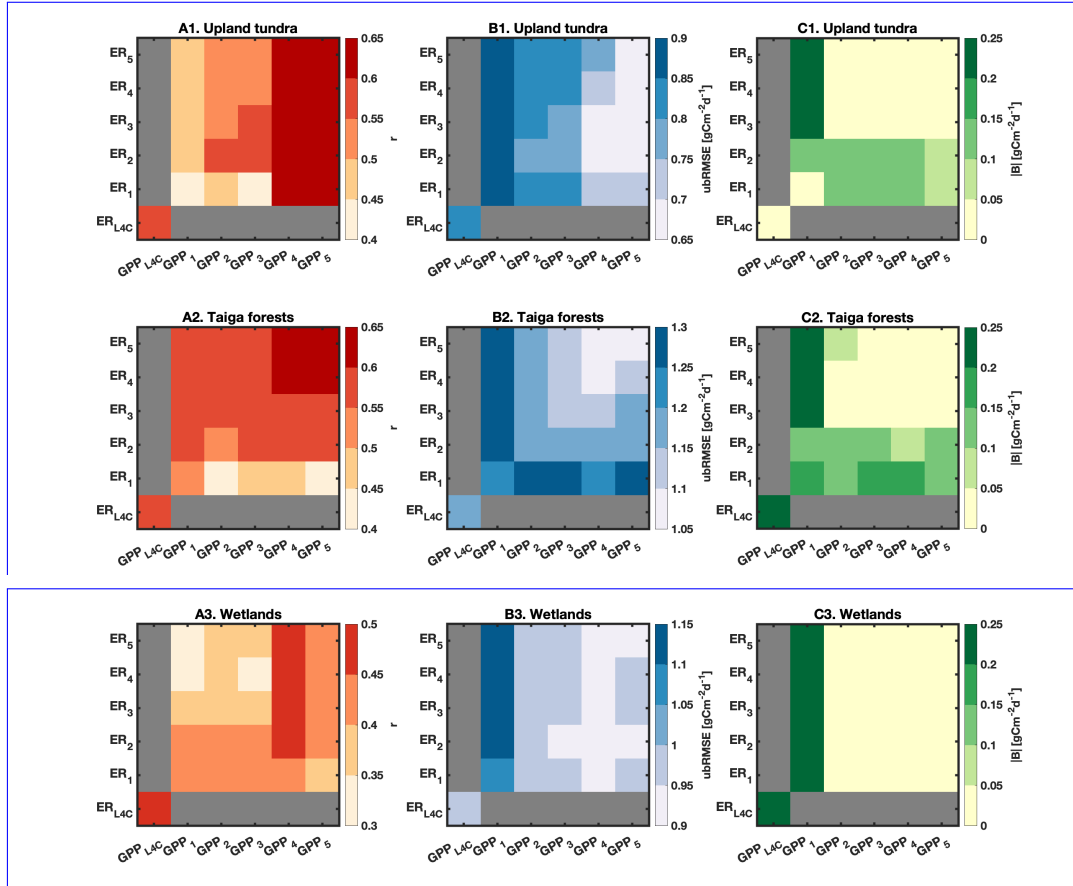


Figure 8. Spatiotemporal performance of modeled **net ecosystem CO₂ exchange (NEE)** against daily-averaged eddy covariance NEE (NEE_{EC}) for **upland tundra, taiga forests, and wetlands** during the growing season. NEE is computed as the difference between ecosystem respiration (ER) and gross primary production (GPP). GPP_{L4C} and ER_{L4C} refer to outputs from the original L4C model, while GPP_1 through GPP_5 and ER_1 through ER_5 correspond to the Arctic–Subarctic (AS) adapted formulations (Sections 3 and 4.1). **The (ER_i , GPP_j) grid cell corresponds to the $NEE_{i,j}$ configuration.** The Pearson correlation coefficient is denoted by r (dimensionless), and $uRMSE$ and $|B|$ denote the unbiased root mean square error and absolute bias, respectively (in $gCm^{-2}d^{-1}$). Spatiotemporal metrics were derived from 100 random splits (70 % training, 30 % testing), accounting for both spatial and temporal variability. The total number of data points used for training (testing) is 1,155 (495) for upland tundra, 3,243 (1,389) for taiga forests, and 2,557 (1,096) for wetlands (Section 4.3). Reported values correspond to median metrics computed over the testing splits (Tables E1-E5).

Site name	Site ID	Coordinates	Years	Ecosystem	PFT	NDP	Reference
Iqaluktuuttiaq Mesic	CA-IQm	69.08°N, -104.58°E	[2022, 2023[	Upland Tundra	GRA	68	Madelon et al. (2025)
Iqaluktuuttiaq Wetland	CA-IQw	69.08°N, -104.58°E	[2022, 2023[	Wetland	GRA	22	Madelon et al. (2025)
Havikpak Creek	CA-HPC	68.32°N, -133.52°E	[2016, 2023[	Taiga Forest	SHR	371	Sonnentag and Marsh (2021a)
Scotty Creek Bog	CA-SCB	61.31°N, -121.30°E	[2015, 2023[	Wetland	ENF	633	Sonnentag and Quinton (2021)
Scotty Creek Landscape	CA-SCC	61.31°N, -121.30°E	[2015, 2023[	Taiga Forest	ENF	733	Sonnentag and Quinton (2018)
Smith Creek	CA-SMC	63.15°N, -123.25°E	[2017, 2023[	Taiga Forest	ENF	312	Sonnentag (2021)
Trail Valley Creek	CA-TVC	68.75°N, -133.50°E	[2015, 2023[	Upland Tundra	SHR	620	Sonnentag and Marsh (2021b)
Bonanza Creek Black Spruce	US-BZS	64.70°N, -148.32°E	[2017, 2023[	Taiga Forest	ENF	652	Euskirchen (2022d)
Bonanza Creek Old Thermokarst Bog	US-BZo	64.69°N, -148.33°E	[2018, 2023[	Wetland	ENF	548	Euskirchen (2022c)
Bonanza Creek Rich Fen	US-BZF	64.70°N, -148.31°E	[2017, 2023[	Wetland	ENF	689	Euskirchen (2022b)
Bonanza Creek Thermokarst Bog	US-BZB	64.70°N, -148.32°E	[2017, 2023[	Wetland	ENF	680	Euskirchen (2022a)
Eight Mile Lake	US-EML	63.88°N, -149.25°E	[2015, 2021[	Upland Tundra	SHR	571	Bracho et al. (2021)
Imnavait Creek Heath Tundra	US-IC _h	68.61°N, -149.30°E	[2015, 2023[	Upland Tundra	SHR	314	Euskirchen et al. (2022a)
Imnavait Creek Sedge Tundra	US-IC _s	68.61°N, -149.31°E	[2015, 2023[	Wetland	SHR	336	Euskirchen et al. (2022b)
Imnavait Creek Tussock Tundra	US-IC _t	68.61°N, -149.30°E	[2021, 2023[	Upland Tundra	SHR	77	Euskirchen et al. (2022c)
Poker Flats Black Spruce	US-Prr	65.12°N, -147.49°E	[2015, 2023[	Taiga Forest	ENF	827	Iwahana et al. (2023)
Poker Flats Fire Scar	US-Rpf	65.12°N, -147.43°E	[2015, 2023[	Taiga Forest	ENF	852	Ueyama et al. (2023b)
University Of Fairbanks	US-Uaf	64.87°N, -147.86°E	[2015, 2023[	Taiga Forest	ENF	885	Ueyama et al. (2023a)
Yukon-Kuskokwim Delta Burned	US-YK1	61.27°N, -163.22°E	[2019, 2023[	Wetland	SHR	330	Natali (2024)
Yukon-Kuskokwim Delta Unburned	US-YK2	61.26°N, -163.26°E	[2019, 2023[	Wetland	SHR	415	Natali (2025)

Table 1. List of the 20 eddy covariance (EC) tower sites used in this study. Ecosystem types were assigned based on site descriptions while plant functional type (PFT) classes were assigned using the Moderate Resolution Imaging Spectroradiometer (MODIS) MCD12Q1 type 5 product (Friedl and Sulla-Menashe, 2019). GRA, SHR, and ENF stand for grassland, shrubland, and evergreen needleleaf forest, respectively. Column 7 shows the number of data points (NDP) within the EC net ecosystem CO₂ exchange (NEE_{EC}), gross primary production (GPP_{EC}), and ecosystem respiration (ER_{EC}) time series after daily averaging and data filtering (see Sections 2 and 4.2).

Abbreviation	Definition
AS	Arctic-Subarctic
L4C model	Soil Moisture Active Passive Level-4 Terrestrial Carbon Flux model
EC	Eddy covariance
PFT	Plant functional type
NEE	Net ecosystem CO ₂ exchange [gCm ⁻² d ⁻¹]
GPP	Gross primary production [gCm ⁻² d ⁻¹]
ER	Ecosystem respiration [gCm ⁻² d ⁻¹]
AR	Autotrophic respiration [gCm ⁻² d ⁻¹]
HR	Heterotrophic respiration [gCm ⁻² d ⁻¹]
NPP	Net primary production [gCm ⁻² d ⁻¹]
PAR	Photosynthetically active radiation [MJm ⁻² d ⁻¹]
FPAR	Canopy-intercepted fraction of absorbed photosynthetically active radiation [dim.]
APAR	Canopy-absorbed photosynthetically active radiation [MJm ⁻² d ⁻¹]
MNT	Minimum air temperature [K]
AT	Mean air temperature [K]
VPD	Vapor pressure deficit [kPa]
RZSM	Rootzone soil moisture [m ³ m ⁻³]
SSM	Surface soil moisture [m ³ m ⁻³]
ST	Soil temperature [K]
SOC	Soil organic carbon [gCm ⁻²]
SOC ₁	Labile soil organic carbon pool [gCm ⁻²]
SOC ₂	Structural soil organic carbon pool [gCm ⁻²]
SOC ₃	Recalcitrant soil organic carbon pool [gCm ⁻²]
L _{fall}	Litterfall [gCm ⁻²]
GDD	Normalized growing degree days [dim.]
S _x	Stress scalar corresponding to the environmental variable x [dim.]
LAI	Leaf area index [dim.]
NDVI	Normalized difference vegetation index [dim.]

Table 2. Summary of frequently used abbreviations.

AS-adapted formulation	Free parameters	Model specificity
GPP ₁	$\epsilon_{\max}, \text{MNT}_{\max}, \text{VPD}_{\min}, \text{VPD}_{\max}, \text{RZSM}_{\min}, \text{RZSM}_{\max}$	$\text{MNT}_{\min}$ is set to 273.15 K
GPP ₂	$\text{GPP}_{\max}, \text{APAR}_{\text{crit}}, \text{MNT}_{\max}, \text{VPD}_{\min}, \text{VPD}_{\max}, \text{RZSM}_{\min}, \text{RZSM}_{\max}$	Nonlinear light response to APAR
GPP ₃	$\text{GPP}_{\max}, \text{APAR}_{\text{crit}}, \gamma_{\text{MNT}}, \gamma_{\text{VPD}}, \text{VPD}_{\text{crit}}, \gamma_{\text{RZSM}}, \text{RZSM}_{\text{crit}}$	Nonlinear responses to MNT, VPD and RZSM
GPP ₄	$\text{GPP}_{\max}, \text{APAR}_{\text{crit}}, \gamma_{\text{MNT}}, \gamma_{\text{VPD}}, \text{VPD}_{\text{crit}}, \gamma_{\text{RZSM}}, \text{RZSM}_{\text{crit}}, \text{a}_{\text{GDD}}, \text{b}_{\text{GDD}}$	Incorporation of GDD
GPP ₅	$\text{GPP}_{\max}, \text{APAR}_{\text{crit}}, \gamma_{\text{MNT}}, \gamma_{\text{VPD}}, \text{VPD}_{\text{crit}}, \text{a}_{\text{RZSM}}, \text{b}_{\text{RZSM}}, \text{a}_{\text{GDD}}, \text{b}_{\text{GDD}}$	Bell-shaped response to RZSM
ER ₁	$\alpha, k_1, \lambda, \beta_0, \gamma_{\text{SSM}}, \text{SSM}_{\text{crit}}$	Constant daily allocation of annual mean NPP to L_{fall}
ER ₂	$\alpha, k_1, \lambda, \beta_0, \gamma_{\text{SSM}}, \text{SSM}_{\text{crit}}$	Dynamic daily allocation of annual mean NPP to L_{fall}
ER ₃	$\alpha, R_{\text{base}}, \beta_0, \gamma_{\text{SSM}}, \text{SSM}_{\text{crit}}$	Constant R_{base} instead of SOC dynamic
ER ₄	$\alpha, R_0, \omega, \beta_0, \gamma_{\text{SSM}}, \text{SSM}_{\text{crit}}$	Dynamic $R_{\text{base}}(t)$
ER ₅	$\alpha, R_0, \omega, \beta_0, \text{a}_{\text{SSM}}, \text{b}_{\text{SSM}}$	Bell-shaped response to SSM

Table 3. Summary of the specificity of each Arctic-Subarctic (AS) adapted model formulation of the L4C model. Column 1 lists the gross primary production (GPP) and ecosystem respiration (ER) formulations, labeled GPP₁ through GPP₅ and ER₁ through ER₅, respectively. The free parameters estimated during the optimization process are provided in column 2. Each GPP and ER formulation is incrementally built on the previous one by incorporating its modifications along with an additional change summarized in column 3. APAR, MNT, VPD, RZSM, SSM, GDD, NPP, L_{fall} , SOC, R_{base} , denote absorbed photosynthetically active radiation, minimum air temperature, vapor pressure deficit, rootzone soil moisture, surface soil moisture, normalized growing degree days, net primary productivity, litterfall, soil organic carbon, and baseline heterotrophic respiration, respectively. Refer to Sections 3 and 4.1 for a description of each model formulation and parameter.

Spatiotemporal evaluation										
GPP	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
GPP_{L4C}	training	0.64	1.47	0.41	0.58	1.93	-0.37	0.48	2.19	1.31
	testing	0.64	1.45	0.42	0.58	1.93	-0.38	0.48	2.19	1.31
GPP₁	training	0.56	1.19	-0.32	0.62	1.75	-0.44	0.53	1.33	-0.40
	testing	0.56	1.20	-0.32	0.61	1.76	-0.45	0.53	1.33	-0.40
GPP₂	training	0.62	0.98	-0.01	0.67	1.43	-0.04	0.63	1.04	-0.05
	testing	0.62	0.99	-0.01	0.66	1.44	-0.04	0.63	1.04	-0.05
GPP₃	training	0.65	0.95	-0.00	0.67	1.41	-0.01	0.65	1.01	-0.01
	testing	0.65	0.96	-0.00	0.67	1.42	-0.02	0.65	1.01	-0.01
GPP₄	training	0.75	0.84	-0.01	0.73	1.30	-0.02	0.72	0.92	-0.01
	testing	0.74	0.84	-0.01	0.73	1.30	-0.02	0.72	0.92	-0.01
GPP₅	training	0.77	0.80	-0.02	0.74	1.29	-0.02	0.68	0.99	-0.05
	testing	0.77	0.80	-0.02	0.74	1.29	-0.02	0.67	0.99	-0.05

Site-level evaluation									
GPP	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
GPP_{L4C}	0.62	1.38	0.51	0.63	1.57	-0.36	0.66	1.62	1.27
GPP₁	0.58	1.11	-0.22	0.61	1.72	-0.57	0.65	1.15	-0.62
GPP₂	0.63	0.91	0.05	0.66	1.33	-0.20	0.71	0.85	-0.22
GPP₃	0.65	0.86	0.08	0.70	1.31	-0.14	0.74	0.79	-0.22
GPP₄	0.77	0.71	0.18	0.79	1.15	-0.13	0.79	0.76	-0.23
GPP₅	0.76	0.65	0.04	0.76	1.12	-0.33	0.76	0.82	-0.19

Table 4. Performance of modeled **gross primary production (GPP)** against daily averaged eddy covariance (EC) GPP (GPP_{EC}) for upland tundra, taiga forests, and wetlands during the growing season. GPP_{L4C} refers to outputs from the original L4C model, while GPP_1 through GPP_5 correspond to the five Arctic–Subarctic (AS) adapted formulations (Sections 3 and 4.1). The Pearson correlation coefficient is denoted by r (dimensionless), and ubRMSE and B denote the unbiased root mean square error and bias, respectively (in $gCm^{-2}d^{-1}$). A positive (negative) B indicates overestimation (underestimation) of GPP_{EC} . **Upper table:** Spatiotemporal performance metrics derived from 100 random splits (70 % training, 30 % testing), accounting for both spatial and temporal variability. The total number of data points used for training (testing) is 1,155 (495) for upland tundra, 3,243 (1,389) for taiga forests, and 2,557 (1,096) for wetlands (Section 4.3). Median values across the 100 splits are reported. **Lower table:** Site-level (temporal-only) performance metrics computed for each EC tower using model outputs obtained with the median parameter values across the 100 splits. Metric values are reported as the median across towers. For this analysis, training and testing data are pooled.

Spatiotemporal evaluation										
ER	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	training	0.43	0.99	0.39	0.33	1.71	-0.11	0.23	1.81	1.76
	testing	0.44	0.99	0.37	0.34	1.71	-0.12	0.24	1.80	1.77
ER₁	training	0.52	0.72	-0.10	0.51	1.28	-0.15	0.50	0.80	-0.04
	testing	0.52	0.72	-0.12	0.52	1.28	-0.17	0.49	0.81	-0.05
ER₂	training	0.61	0.63	-0.11	0.53	1.25	-0.12	0.53	0.78	-0.02
	testing	0.61	0.64	-0.12	0.54	1.25	-0.12	0.53	0.78	-0.03
ER₃	training	0.60	0.62	-0.00	0.52	1.23	0.00	0.51	0.79	-0.02
	testing	0.60	0.62	-0.01	0.54	1.23	0.02	0.50	0.80	-0.03
ER₄	training	0.73	0.53	-0.00	0.58	1.17	0.00	0.51	0.79	-0.02
	testing	0.72	0.53	-0.03	0.59	1.17	0.02	0.51	0.80	-0.04
ER₅	training	0.72	0.54	-0.01	0.59	1.17	0.01	0.50	0.80	-0.04
	testing	0.70	0.55	-0.02	0.60	1.17	0.03	0.49	0.81	-0.05

Site-level evaluation									
ER	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	0.45	0.94	0.35	0.56	1.18	-0.28	0.43	1.10	1.24
ER₁	0.58	0.51	0.06	0.61	0.96	-0.15	0.65	0.56	-0.12
ER₂	0.66	0.46	0.03	0.65	0.95	-0.09	0.66	0.53	-0.09
ER₃	0.57	0.47	0.08	0.65	0.96	0.00	0.67	0.51	-0.18
ER₄	0.58	0.41	-0.01	0.65	1.00	-0.05	0.64	0.48	-0.15
ER₅	0.60	0.42	0.00	0.63	1.01	-0.13	0.63	0.51	-0.16

Table 5. Same as Table 4, but for **ecosystem respiration (ER)**. **GPP₅** was used as the GPP input for ER₁ through ER₅ for upland tundra and taiga forests. For wetlands, **GPP₄** was used instead.

771 **Appendix A: EC tower measurements of NEE and derived GPP and ER**

772 EC tower measurements of NEE are subject to systematic errors, which mostly arise from unmet assumptions, instrument
773 design and calibration, physical phenomena (e.g. storage terms), and terrain-specific conditions. These errors are generally well
774 characterized and are typically corrected using software, such as EddyPro, as part of the standard flux processing workflow
775 (Aubinet et al., 2012; Burba, 2022). NEE measurements are also affected by random errors, notably turbulence sampling error,
776 which arises when large eddies are not adequately captured within a 30-minute window (Finkelstein and Sims, 2001). The
777 standard deviation of this error tends to follow a consistent pattern across ecosystem types and increases linearly with the flux
778 magnitude (Aubinet et al., 2012). Overall, random errors in NEE are difficult to quantify, but using simultaneous measurements
779 from two collocated EC towers, they have been estimated at 15 % for a 30-minute interval (Eugster et al., 1997; Dragoni et al.,
780 2007).

781 Gaps in NEE are usually filled using the Marginal Distribution Sampling (MDS) method (Falge et al., 2001; Reichstein
782 et al., 2005; Aubinet et al., 2012). In this approach, a missing measurement is replaced by the mean of valid values observed
783 under similar meteorological conditions within a ± 7 -day time window. Meteorological conditions are considered similar when
784 variables such as solar radiation, temperature, and vapor pressure deficit do not deviate beyond predefined thresholds within
785 the time window. When no suitable meteorological analogues are available, missing measurements are replaced by the mean
786 of valid values from adjacent days at the same time of day (± 1 hour).

787 There is no universal flux-partitioning algorithm to separate NEE into its ER and GPP components. The most established
788 method assumes that nighttime NEE consists solely of the ER component, since photosynthesis, and therefore GPP, is con-
789 sidered negligible in the absence of light (Reichstein et al., 2005; Aubinet et al., 2012). Nighttime ER is modeled using an
790 exponential function, with air or soil temperature as the primary driver (Lloyd and Taylor, 1994). Air temperature is gener-
791 ally preferred because it better represents the landscape surrounding the EC tower, whereas soil temperature varies spatially
792 and with depth across heterogeneous terrain (Helbig et al., 2017a, b). Daytime ER is then extrapolated to isolate the GPP
793 contribution from the NEE measurements. Alternative approaches fit a light-response curve combined with a Q_{10} equation to
794 NEE measurements, accounting for the effects of photosynthetically active radiation (PAR) on GPP and air temperature on
795 ER (Falge et al., 2001; Gilmanov et al., 2003; Lasslop et al., 2010b; Runkle et al., 2013; Helbig et al., 2017a). ER and GPP
796 respond to more than temperature and light, respectively, and depend on other drivers such as SOC availability or moisture
797 limitations. However, establishing a model formulation with more variables is complex and such additional measurements are
798 not systematically available (Aubinet et al., 2012). To overcome this problem, the model regression is generally performed
799 over short time intervals of 4 through 15 days (Reichstein et al., 2005; Lasslop et al., 2010b; Runkle et al., 2013). This ap-
800 proach enables the effects of additional drivers to be implicitly incorporated into the estimation of free parameters and accounts
801 for seasonal parameter variability, reflecting changes in ecosystem state that are not represented in the model (Lasslop et al.,
802 2010b; Aubinet et al., 2012). As GPP and ER are modeled using additional data and rely on various assumptions, they have
803 greater uncertainties than tower measurements of NEE (Lasslop et al., 2010a).

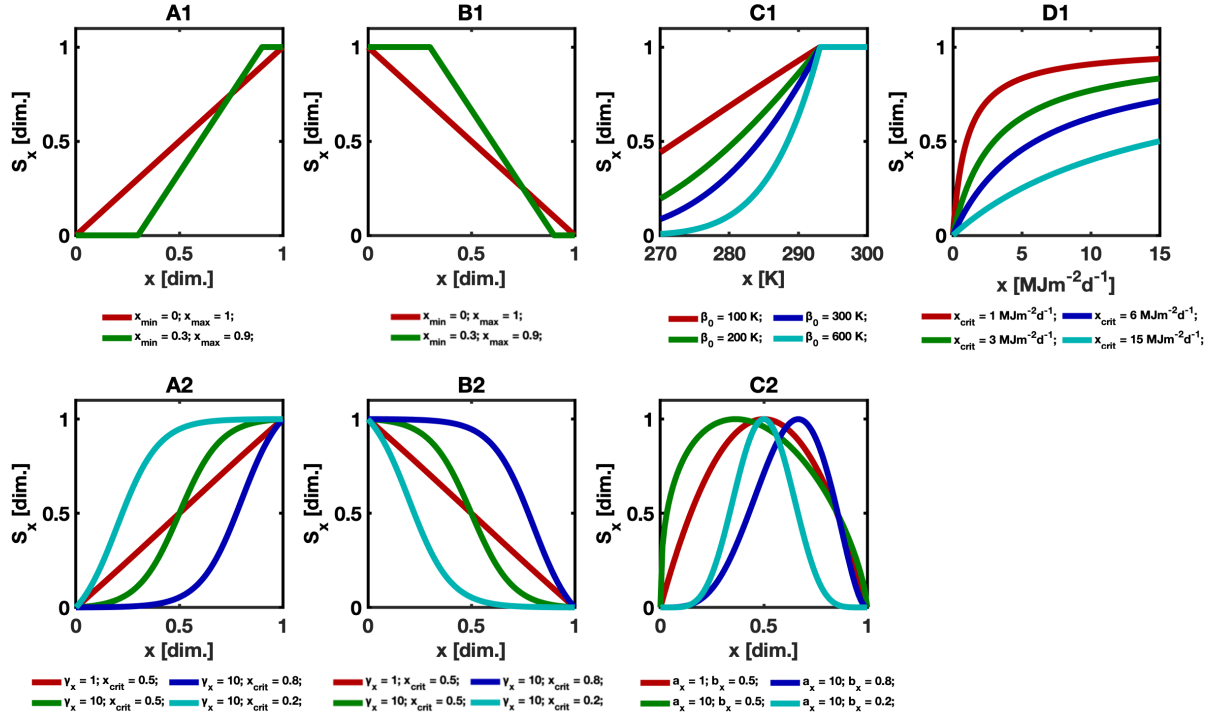


Figure B1. Behavior of the ecosystem response functions used in the L4C model and the Arctic-Subarctic adapted formulations. Each panel shows the same response function evaluated across a range of parameter values to illustrate how parameterization affects function shape. Refer to Sections 3 and 4.1 for a detailed description of each model formulation. Panels A1, B1, and C1 correspond to Equations 7a,c,d 7b, and 7e, respectively; panel D1 corresponds to Equation 8b. Panels A2 and B2 correspond to Equations 9a,c and 9b, respectively; panel C2 corresponds to Equations 10b and 11.

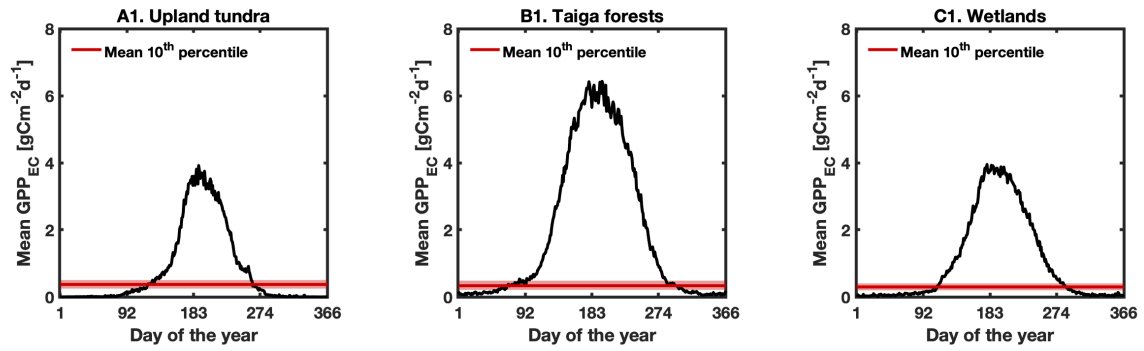


Figure B2. Mean seasonal cycle of daily averaged eddy covariance (EC) gross primary production (GPP_{EC}) for upland tundra, taiga forests and wetlands. The horizontal red line indicates the estimated separation between dormant and transitional periods with active photosynthetic periods (i.e., the growing season) using a percentile-based GPP_{EC} threshold (Section 4.2). This threshold corresponds to the mean 10th percentile computed across EC towers, with the 5th and 15th percentiles shown as variability bounds.

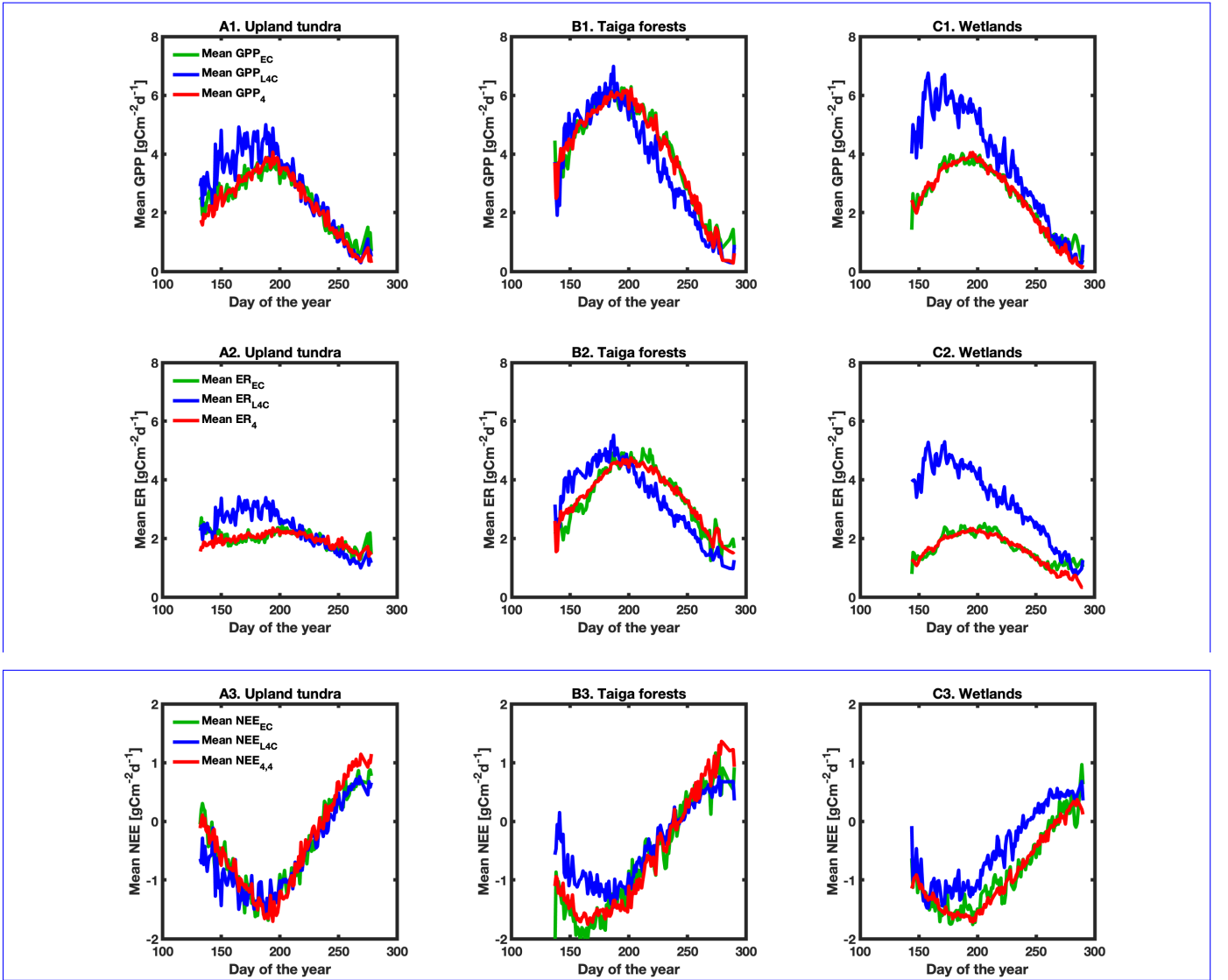


Figure B3. Mean seasonal cycle of daily averaged gross primary production (GPP), ecosystem respiration (ER), and net ecosystem CO₂ exchange (NEE) across eddy covariance tower sites for upland tundra, taiga forests, and wetlands. GPP_{EC}, ER_{EC}, and NEE_{EC} denote eddy covariance data. GPP_{L4C}, ER_{L4C}, and NEE_{L4C} represent outputs from the original L4C model, whereas GPP₄, ER₄, and NEE_{4,4} (i.e., ER₄ – GPP₄) correspond to outputs from the Arctic–Subarctic adapted formulations using test case 4 for both GPP and ER (Sections 3 and 4.1). The data shown represent the growing season after applying the filtering criteria described in Section 4.2.

GPP	Upland Tundra		Taiga Forests		Wetlands	
	Δ AIC	Δ BIC	Δ AIC	Δ BIC	Δ AIC	Δ BIC
GPP₁	-482	-487	-574	-580	-3116	-3122
GPP₂	-1010	-1010	-2036	-2036	-4592	-4592
GPP₃	-1081	-1081	-2134	-2134	-4734	-4734
GPP₄	-1380	-1369	-2682	-2669	-5200	-5188
GPP₅	-1486	-1476	-2737	-2725	-4847	-4836

Table C1. Median differences in Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) between GPP₁ through GPP₅ and GPP_{L4C}. GPP_{L4C} refers to GPP outputs from the original L4C model, while GPP₁ through GPP₅ represent the five Arctic–Subarctic (AS) adapted formulations (Sections 3 and 4.1).

Spatiotemporal evaluation										
GPP	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
GPP₄	training	0.73	0.86	-0.01	0.73	1.32	-0.01	0.72	0.92	-0.00
	testing	0.73	0.86	0.00	0.72	1.33	-0.02	0.72	0.92	-0.00
GPP₅	training	0.74	0.83	-0.03	0.73	1.30	-0.01	0.67	0.99	-0.04
	testing	0.74	0.84	-0.01	0.73	1.30	-0.02	0.67	0.99	-0.05

Site-level evaluation									
GPP	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
GPP₄	0.77	0.76	0.16	0.78	1.19	-0.12	0.76	0.77	-0.22
GPP₅	0.76	0.73	0.02	0.76	1.17	-0.34	0.74	0.82	-0.19

Table C2. Same as Table 4, but here GPP₄ and GPP₅ use growing degree days (GDD) as inputs, which were normalized by the long-term minimum and maximum across all years rather than by annual values (Section 6.1).

Spatiotemporal evaluation										
GPP	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
GPP_{L4C}	training	0.64	1.46	0.43	0.58	1.92	-0.37	0.46	2.21	1.31
	testing	0.65	1.45	0.36	0.60	1.89	-0.40	0.54	2.10	1.29
GPP₁	training	0.56	1.20	-0.33	0.62	1.75	-0.44	0.52	1.34	-0.41
	testing	0.56	1.20	-0.30	0.61	1.82	-0.52	0.57	1.25	-0.42
GPP₂	training	0.63	0.98	-0.01	0.67	1.43	-0.04	0.62	1.05	-0.05
	testing	0.62	0.98	0.01	0.66	1.44	-0.11	0.66	0.97	-0.10
GPP₃	training	0.65	0.96	-0.00	0.67	1.41	-0.02	0.64	1.02	-0.01
	testing	0.65	0.97	0.04	0.66	1.42	-0.15	0.69	0.96	-0.04
GPP₄	training	0.75	0.83	-0.01	0.73	1.30	-0.02	0.71	0.94	-0.01
	testing	0.74	0.84	-0.02	0.72	1.30	-0.07	0.77	0.84	-0.05
GPP₅	training	0.77	0.79	-0.02	0.74	1.28	-0.02	0.66	1.00	-0.05
	testing	0.76	0.81	-0.03	0.73	1.29	-0.01	0.72	0.92	-0.17

Site-level evaluation									
GPP	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
GPP_{L4C}	0.62	1.38	0.51	0.63	1.57	-0.36	0.66	1.62	1.27
GPP₁	0.58	1.11	-0.22	0.60	1.73	-0.55	0.65	1.16	-0.63
GPP₂	0.63	0.91	0.06	0.67	1.34	-0.20	0.71	0.85	-0.28
GPP₃	0.65	0.86	0.08	0.70	1.31	-0.17	0.74	0.78	-0.24
GPP₄	0.77	0.71	0.19	0.79	1.15	-0.11	0.79	0.76	-0.22
GPP₅	0.76	0.66	0.08	0.76	1.12	-0.35	0.76	0.81	-0.20

Table C3. Same as Table 4, but instead of 100 runs with 70—30 % training–testing splits, the data were split into 6-year training and 2-year testing periods. The period of study spans a total of 8 years, yielding 28 possible runs instead of 100.

Spatiotemporal evaluation										
GPP	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
GPP_{L4C}	training	0.63	1.48	0.40	0.59	1.90	-0.34	0.48	2.16	1.30
	testing	0.62	1.38	0.51	0.64	1.81	-0.45	0.56	1.89	1.34
GPP₁	training	0.56	1.19	-0.32	0.62	1.74	-0.44	0.55	1.32	-0.41
	testing	0.57	1.08	-0.16	0.60	1.79	-0.56	0.57	1.28	-0.33
GPP₂	training	0.64	0.97	-0.01	0.67	1.44	-0.04	0.63	1.04	-0.05
	testing	0.62	0.94	0.10	0.62	1.43	-0.19	0.63	0.94	0.02
GPP₃	training	0.66	0.94	-0.00	0.69	1.39	-0.01	0.65	1.01	-0.01
	testing	0.64	0.90	0.14	0.66	1.40	-0.06	0.64	0.93	0.05
GPP₄	training	0.76	0.82	-0.01	0.74	1.29	-0.01	0.73	0.92	-0.01
	testing	0.76	0.77	0.19	0.73	1.25	-0.10	0.71	0.85	0.00
GPP₅	training	0.77	0.79	-0.02	0.75	1.27	-0.01	0.69	0.97	-0.04
	testing	0.74	0.73	0.19	0.73	1.29	-0.13	0.65	0.92	-0.04

Site-level evaluation									
GPP	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
GPP_{L4C}	0.62	1.38	0.51	0.63	1.57	-0.36	0.66	1.62	1.27
GPP₁	0.58	1.12	-0.19	0.61	1.71	-0.65	0.65	1.16	-0.53
GPP₂	0.62	0.89	0.08	0.67	1.37	-0.04	0.71	0.85	-0.32
GPP₃	0.65	0.86	0.11	0.70	1.31	-0.27	0.74	0.78	-0.27
GPP₄	0.78	0.71	0.18	0.79	1.16	-0.16	0.78	0.77	-0.26
GPP₅	0.76	0.65	0.03	0.77	1.13	-0.31	0.77	0.82	-0.30

Table C4. Same as Table 4, but the 70–30 % training-testing splits were applied to the set of eddy covariance (EC) tower sites rather than to the individual data points, such that all data from a given site were assigned to either the training or the testing split. Because the number of EC tower sites differs among ecosystems, this yielded 5, 21, and 28 possible runs for upland tundra, taiga forests, and wetlands, respectively, instead of 100.

Spatiotemporal evaluation										
ER	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER _{L4C}	training	0.43	0.99	0.39	0.33	1.71	-0.11	0.23	1.81	1.76
	testing	0.44	0.99	0.37	0.34	1.71	-0.12	0.24	1.80	1.77
ER ₁	training	0.34	0.87	-0.32	0.35	1.58	-0.62	0.37	0.90	-0.20
	testing	0.36	0.86	-0.34	0.35	1.57	-0.63	0.38	0.90	-0.21
ER ₂	training	0.55	0.67	-0.19	0.46	1.35	-0.33	0.51	0.79	-0.03
	testing	0.55	0.67	-0.22	0.47	1.33	-0.34	0.51	0.79	-0.03
ER ₃	training	0.56	0.64	-0.00	0.47	1.28	0.00	0.46	0.82	-0.02
	testing	0.57	0.64	-0.01	0.49	1.27	0.01	0.46	0.82	-0.03
ER ₄	training	0.70	0.56	0.00	0.55	1.21	0.00	0.46	0.82	-0.03
	testing	0.68	0.57	-0.01	0.55	1.21	0.02	0.45	0.82	-0.04
ER ₅	training	0.69	0.57	-0.00	0.56	1.20	0.01	0.42	0.84	-0.05
	testing	0.67	0.58	-0.01	0.56	1.20	0.03	0.42	0.85	-0.06

Site-level evaluation									
ER	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER _{L4C}	0.45	0.94	0.35	0.56	1.18	-0.28	0.43	1.10	1.24
ER ₁	0.33	0.66	-0.01	0.35	1.31	-0.67	0.46	0.69	-0.39
ER ₂	0.55	0.51	-0.11	0.57	1.07	-0.34	0.63	0.51	-0.19
ER ₃	0.53	0.50	0.17	0.53	1.02	-0.01	0.60	0.54	-0.17
ER ₄	0.48	0.44	-0.02	0.68	1.05	-0.05	0.63	0.49	-0.21
ER ₅	0.44	0.45	0.08	0.68	1.06	-0.12	0.60	0.54	-0.22

Table D1. Same as Table 4, but for **ecosystem respiration (ER)**. **GPP₁** was used as the GPP input for ER₁ through ER₅.

Spatiotemporal evaluation										
ER	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	training	0.43	0.99	0.39	0.33	1.71	-0.11	0.23	1.81	1.76
	testing	0.44	0.99	0.37	0.34	1.71	-0.12	0.24	1.80	1.77
ER₁	training	0.40	0.77	-0.12	0.43	1.37	-0.17	0.46	0.83	-0.06
	testing	0.43	0.75	-0.12	0.44	1.37	-0.17	0.46	0.83	-0.07
ER₂	training	0.53	0.67	-0.13	0.47	1.30	-0.15	0.53	0.78	-0.03
	testing	0.56	0.66	-0.14	0.48	1.31	-0.14	0.53	0.78	-0.04
ER₃	training	0.57	0.64	-0.00	0.49	1.26	0.00	0.48	0.81	-0.02
	testing	0.58	0.63	-0.03	0.50	1.25	0.02	0.48	0.81	-0.03
ER₄	training	0.70	0.56	0.00	0.55	1.20	0.00	0.49	0.80	-0.02
	testing	0.68	0.57	-0.01	0.56	1.20	0.02	0.49	0.81	-0.03
ER₅	training	0.70	0.56	-0.00	0.57	1.19	0.01	0.47	0.81	-0.04
	testing	0.69	0.57	-0.01	0.57	1.19	0.03	0.48	0.82	-0.05

Site-level evaluation									
ER	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	0.45	0.94	0.35	0.56	1.18	-0.28	0.43	1.10	1.24
ER₁	0.41	0.57	0.19	0.52	1.04	-0.27	0.61	0.60	-0.17
ER₂	0.52	0.52	-0.01	0.60	1.00	-0.19	0.68	0.51	-0.19
ER₃	0.54	0.47	0.13	0.56	1.02	0.01	0.62	0.53	-0.16
ER₄	0.51	0.42	-0.05	0.68	1.04	-0.09	0.66	0.49	-0.22
ER₅	0.46	0.44	0.02	0.67	1.05	-0.15	0.64	0.52	-0.22

Table D2. Same as Table 4, but for **ecosystem respiration (ER)**. **GPP₂** was used as the GPP input for ER₁ through ER₅.

Spatiotemporal evaluation										
ER	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	training	0.43	0.99	0.39	0.33	1.71	-0.11	0.23	1.81	1.76
	testing	0.44	0.99	0.37	0.34	1.71	-0.12	0.24	1.80	1.77
ER₁	training	0.44	0.74	-0.09	0.42	1.37	-0.16	0.46	0.82	-0.04
	testing	0.47	0.73	-0.11	0.43	1.36	-0.16	0.45	0.82	-0.04
ER₂	training	0.55	0.66	-0.12	0.46	1.30	-0.11	0.49	0.80	-0.03
	testing	0.57	0.65	-0.12	0.47	1.30	-0.10	0.50	0.80	-0.04
ER₃	training	0.58	0.63	-0.00	0.48	1.26	0.00	0.48	0.81	-0.02
	testing	0.59	0.63	-0.01	0.50	1.26	0.02	0.48	0.81	-0.03
ER₄	training	0.71	0.55	-0.00	0.55	1.21	0.00	0.50	0.80	-0.03
	testing	0.70	0.55	-0.01	0.56	1.20	0.03	0.50	0.80	-0.04
ER₅	training	0.71	0.55	-0.00	0.57	1.19	0.01	0.48	0.81	-0.03
	testing	0.70	0.55	-0.01	0.57	1.19	0.03	0.49	0.81	-0.04

Site-level evaluation									
ER	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	0.45	0.94	0.35	0.56	1.18	-0.28	0.43	1.10	1.24
ER₁	0.44	0.55	0.20	0.51	1.01	-0.24	0.60	0.58	-0.13
ER₂	0.53	0.51	0.01	0.60	1.00	-0.25	0.63	0.53	-0.19
ER₃	0.54	0.49	0.16	0.56	1.03	0.00	0.63	0.54	-0.12
ER₄	0.59	0.42	-0.06	0.69	1.05	-0.08	0.66	0.50	-0.18
ER₅	0.63	0.43	0.06	0.67	1.06	-0.13	0.65	0.52	-0.21

Table D3. Same as Table 4, but for **ecosystem respiration (ER)**. **GPP₃** was used as the GPP input for ER₁ through ER₅.

Spatiotemporal evaluation										
ER	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	training	0.43	0.99	0.39	0.33	1.71	-0.11	0.23	1.81	1.76
	testing	0.44	0.99	0.37	0.34	1.71	-0.12	0.24	1.80	1.77
ER₁	training	0.44	0.77	-0.13	0.48	1.33	-0.17	0.50	0.80	-0.04
	testing	0.45	0.77	-0.15	0.49	1.33	-0.17	0.49	0.81	-0.05
ER₂	training	0.56	0.65	-0.15	0.51	1.27	-0.10	0.53	0.78	-0.02
	testing	0.56	0.65	-0.17	0.52	1.26	-0.09	0.53	0.78	-0.03
ER₃	training	0.58	0.63	-0.00	0.52	1.24	0.00	0.51	0.79	-0.02
	testing	0.58	0.63	-0.02	0.53	1.23	0.02	0.50	0.80	-0.03
ER₄	training	0.73	0.53	-0.00	0.57	1.18	0.00	0.51	0.79	-0.02
	testing	0.72	0.54	-0.02	0.58	1.19	0.02	0.51	0.80	-0.04
ER₅	training	0.72	0.54	-0.01	0.59	1.17	0.00	0.50	0.80	-0.04
	testing	0.71	0.54	-0.01	0.59	1.17	0.03	0.49	0.81	-0.05

Site-level evaluation									
ER	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	0.45	0.94	0.35	0.56	1.18	-0.28	0.43	1.10	1.24
ER₁	0.57	0.59	0.35	0.59	1.01	-0.27	0.65	0.56	-0.12
ER₂	0.66	0.51	0.09	0.66	0.94	-0.17	0.66	0.53	-0.09
ER₃	0.56	0.48	0.18	0.64	0.96	-0.01	0.67	0.51	-0.18
ER₄	0.60	0.43	0.01	0.67	1.01	-0.06	0.64	0.48	-0.15
ER₅	0.63	0.45	0.01	0.65	1.02	-0.13	0.63	0.51	-0.16

Table D4. Same as Table 4, but for **ecosystem respiration (ER)**. **GPP₄** was used as the GPP input for ER₁ through ER₅.

Spatiotemporal evaluation										
ER	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	training	0.43	0.99	0.39	0.33	1.71	-0.11	0.23	1.81	1.76
	testing	0.44	0.99	0.37	0.34	1.71	-0.12	0.24	1.80	1.77
ER₁	training	0.52	0.72	-0.10	0.51	1.28	-0.15	0.47	0.82	-0.04
	testing	0.52	0.72	-0.12	0.52	1.28	-0.17	0.46	0.82	-0.06
ER₂	training	0.61	0.63	-0.11	0.53	1.25	-0.12	0.51	0.78	-0.02
	testing	0.61	0.64	-0.12	0.54	1.25	-0.12	0.51	0.79	-0.03
ER₃	training	0.60	0.62	-0.00	0.52	1.23	0.00	0.50	0.80	-0.02
	testing	0.60	0.62	-0.01	0.54	1.23	0.02	0.49	0.80	-0.05
ER₄	training	0.73	0.53	-0.00	0.58	1.17	0.00	0.49	0.81	-0.03
	testing	0.72	0.53	-0.03	0.59	1.17	0.02	0.48	0.81	-0.05
ER₅	training	0.72	0.54	-0.01	0.59	1.17	0.01	0.47	0.82	-0.04
	testing	0.70	0.55	-0.02	0.60	1.17	0.03	0.46	0.83	-0.06

Site-level evaluation									
ER	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	0.45	0.94	0.35	0.56	1.18	-0.28	0.43	1.10	1.24
ER₁	0.58	0.51	0.06	0.61	0.96	-0.15	0.62	0.58	-0.06
ER₂	0.66	0.46	0.03	0.65	0.95	-0.09	0.64	0.52	-0.18
ER₃	0.57	0.47	0.08	0.65	0.96	0.00	0.65	0.52	-0.14
ER₄	0.58	0.41	-0.01	0.65	1.00	-0.05	0.61	0.48	-0.14
ER₅	0.60	0.42	0.00	0.63	1.01	-0.13	0.59	0.51	-0.19

Table D5. Same as Table 4, but for **ecosystem respiration (ER)**. **GPP₅** was used as the GPP input for ER₁ through ER₅.

ER	Upland Tundra		Taiga Forests		Wetlands	
	ΔAIC	ΔBIC	ΔAIC	ΔBIC	ΔAIC	ΔBIC
GPP₁ as input						
ER₁	-331	-336	-25	-31	-5138	-5144
ER₂	-977	-982	-1363	-1369	-5973	-5978
ER₃	-1174	-1184	-1916	-1929	-5786	-5797
ER₄	-1511	-1516	-2246	-2252	-5775	-5781
ER₅	-1484	-1489	-2337	-2344	-5600	-5606
GPP₂ as input						
ER₁	-742	-747	-1393	-1400	-5700	-5706
ER₂	-1015	-1020	-1732	-1738	-6032	-6038
ER₃	-1187	-1197	-1990	-2002	-5849	-5861
ER₄	-1509	-1514	-2290	-2296	-5873	-5879
ER₅	-1510	-1515	-2380	-2386	-5783	-5789
GPP₃ as input						
ER₁	-811	-816	-1433	-1439	-5740	-5746
ER₂	-1056	-1061	-1744	-1750	-5922	-5928
ER₃	-1210	-1221	-1977	-1990	-5835	-5846
ER₄	-1540	-1546	-2277	-2283	-5899	-5905
ER₅	-1553	-1558	-2367	-2373	-5830	-5836
GPP₄ as input						
ER₁	-729	-734	-1621	-1627	-5853	-5859
ER₂	-1062	-1067	-1931	-1937	-6007	-6013
ER₃	-1210	-1220	-2122	-2135	-5944	-5955
ER₄	-1602	-1607	-2395	-2401	-5936	-5942
ER₅	-1606	-1611	-2498	-2504	-5868	-5873
GPP₅ as input						
ER₁	-881	-886	-1844	-1850	-5774	-5780
ER₂	-1185	-1190	-2030	-2036	-5976	-5982
ER₃	-1269	-1279	-2146	-2158	-5863	-5874
ER₄	-1613	-1619	-2455	-2461	-5840	-5846
ER₅	-1577	-1582	-2482	-2488	-5742	-5748

Table D6. Same as Table C1, but for ecosystem respiration (ER).

Spatiotemporal evaluation										
ER	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	training	0.42	0.99	0.39	0.33	1.71	-0.11	0.23	1.80	1.79
	testing	0.49	0.96	0.38	0.35	1.71	-0.12	0.27	1.79	1.68
ER₁	training	0.52	0.71	-0.10	0.52	1.27	-0.15	0.49	0.80	-0.04
	testing	0.54	0.70	-0.13	0.51	1.28	-0.15	0.50	0.83	-0.02
ER₂	training	0.62	0.63	-0.11	0.53	1.24	-0.11	0.53	0.78	-0.02
	testing	0.65	0.57	-0.15	0.52	1.25	-0.10	0.54	0.78	0.01
ER₃	training	0.60	0.62	-0.00	0.53	1.23	0.00	0.50	0.79	-0.02
	testing	0.60	0.61	-0.04	0.52	1.23	0.02	0.53	0.80	0.00
ER₄	training	0.71	0.55	-0.01	0.59	1.17	0.00	0.50	0.79	-0.02
	testing	0.76	0.48	0.00	0.58	1.17	0.04	0.55	0.81	-0.01
ER₅	training	0.70	0.56	-0.01	0.59	1.17	0.00	0.49	0.80	-0.04
	testing	0.73	0.51	-0.02	0.59	1.16	0.02	0.55	0.81	-0.02

Site-level evaluation									
ER	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	0.45	0.94	0.35	0.56	1.18	-0.28	0.43	1.10	1.24
ER₁	0.57	0.51	0.11	0.60	0.96	0.02	0.64	0.58	0.02
ER₂	0.66	0.46	0.03	0.64	0.96	0.00	0.66	0.52	-0.10
ER₃	0.58	0.45	0.13	0.64	0.97	0.05	0.67	0.53	-0.04
ER₄	0.58	0.41	-0.08	0.65	1.01	-0.11	0.64	0.48	-0.13
ER₅	0.62	0.41	-0.03	0.63	1.02	-0.13	0.63	0.50	-0.16

Table D7. Same as Table 5, but instead of 100 runs with 70—30 % training–testing splits, the data were split into 6-year training and 2-year testing periods. The period of study spans a total of 8 years, yielding 28 possible runs instead of 100.

Spatiotemporal evaluation										
ER	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	training	0.42	1.00	0.40	0.34	1.71	-0.10	0.24	1.78	1.74
	testing	0.45	0.94	0.35	0.34	1.65	-0.16	0.35	1.51	1.83
ER₁	training	0.53	0.70	-0.11	0.52	1.28	-0.17	0.51	0.80	-0.04
	testing	0.49	0.53	0.25	0.49	1.29	-0.18	0.55	0.74	-0.16
ER₂	training	0.63	0.61	-0.10	0.54	1.22	-0.11	0.54	0.79	-0.01
	testing	0.65	0.44	-0.02	0.51	1.25	-0.20	0.57	0.73	-0.03
ER₃	training	0.58	0.62	-0.00	0.53	1.23	0.00	0.53	0.79	-0.01
	testing	0.64	0.43	0.14	0.51	1.23	-0.06	0.55	0.76	-0.05
ER₄	training	0.70	0.56	-0.03	0.58	1.17	0.00	0.53	0.78	-0.01
	testing	0.71	0.52	0.03	0.57	1.21	-0.08	0.52	0.77	-0.22
ER₅	training	0.71	0.55	-0.01	0.60	1.15	0.00	0.53	0.80	-0.04
	testing	0.69	0.46	0.04	0.57	1.21	-0.09	0.53	0.78	-0.17

Site-level evaluation									
ER	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
ER_{L4C}	0.45	0.94	0.35	0.56	1.18	-0.28	0.43	1.10	1.24
ER₁	0.58	0.51	0.08	0.61	0.97	-0.22	0.65	0.57	-0.06
ER₂	0.66	0.44	0.04	0.65	0.95	-0.12	0.66	0.52	-0.10
ER₃	0.61	0.44	0.10	0.64	0.96	0.02	0.66	0.53	0.10
ER₄	0.58	0.40	0.09	0.66	1.01	0.03	0.64	0.51	0.12
ER₅	0.61	0.42	-0.00	0.66	1.04	0.06	0.63	0.51	-0.12

Table D8. Same as Table 5, but the 70–30 % training-testing splits were applied to the set of eddy covariance (EC) tower sites rather than to the individual data points, such that all data from a given site were assigned to either the training or the testing split. Because the number of EC tower sites differs among ecosystems, this yielded 5, 21, and 28 possible runs for upland tundra, taiga forests, and wetlands, respectively, instead of 100.

Spatiotemporal evaluation										
NEE	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
NEE _{L4C}	training	0.54	0.87	-0.03	0.55	1.17	0.26	0.47	0.99	0.46
	testing	0.57	0.84	-0.03	0.55	1.17	0.26	0.48	0.99	0.45
NEE _{1,1}	training	0.45	0.87	0.00	0.53	1.20	-0.18	0.40	1.05	0.21
	testing	0.44	0.87	-0.01	0.52	1.22	-0.18	0.40	1.06	0.20
NEE _{2,1}	training	0.50	0.92	0.14	0.55	1.33	0.13	0.41	1.17	0.38
	testing	0.50	0.92	0.12	0.55	1.35	0.13	0.42	1.16	0.38
NEE _{3,1}	training	0.49	0.99	0.32	0.59	1.42	0.44	0.36	1.22	0.38
	testing	0.49	0.97	0.30	0.58	1.44	0.46	0.36	1.22	0.38
NEE _{4,1}	training	0.47	1.11	0.33	0.59	1.51	0.44	0.32	1.31	0.37
	testing	0.48	1.09	0.32	0.59	1.52	0.47	0.32	1.30	0.37
NEE _{5,1}	training	0.49	1.09	0.32	0.60	1.50	0.44	0.32	1.29	0.35
	testing	0.49	1.07	0.31	0.60	1.52	0.47	0.32	1.28	0.34

Site-level evaluation									
NEE	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
NEE _{L4C}	0.44	0.90	0.06	0.54	1.04	0.40	0.50	0.79	0.73
NEE _{1,1}	0.39	0.78	-0.04	0.54	1.09	-0.17	0.55	0.80	0.58
NEE _{2,1}	0.38	0.85	0.12	0.53	1.36	-0.02	0.54	0.97	0.74
NEE _{3,1}	0.38	0.92	0.36	0.55	1.40	0.34	0.53	0.97	0.74
NEE _{4,1}	0.43	1.04	0.21	0.57	1.50	0.35	0.48	1.04	0.72
NEE _{5,1}	0.41	1.02	0.10	0.58	1.48	0.34	0.48	1.02	0.68

Table E1. Same as Table 4, but for **net ecosystem CO₂ exchange (NEE)**. Modeled NEE is computed as the difference between modeled ecosystem respiration (ER) and gross primary production (GPP). NEE_{L4C} denotes NEE from the original L4C model, while NEE_{i,1} refers to outputs from the Arctic–Subarctic (AS) adapted ER_i and **GPP₁** formulations.

Spatiotemporal evaluation										
NEE	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
NEE _{L4C}	training	0.54	0.87	-0.03	0.55	1.17	0.26	0.47	0.99	0.46
	testing	0.57	0.84	-0.03	0.55	1.17	0.26	0.48	0.99	0.45
NEE _{1,2}	training	0.45	0.83	-0.10	0.44	1.25	-0.12	0.39	0.97	-0.01
	testing	0.46	0.82	-0.10	0.44	1.26	-0.13	0.40	0.97	-0.02
NEE _{2,2}	training	0.54	0.78	-0.12	0.53	1.18	-0.11	0.42	0.97	0.02
	testing	0.55	0.77	-0.11	0.53	1.19	-0.11	0.42	0.96	0.03
NEE _{3,2}	training	0.52	0.81	0.01	0.56	1.17	0.04	0.38	0.99	0.03
	testing	0.53	0.80	-0.01	0.56	1.18	0.04	0.39	0.98	0.03
NEE _{4,2}	training	0.50	0.87	0.02	0.59	1.15	0.04	0.36	1.00	0.03
	testing	0.51	0.85	-0.00	0.59	1.16	0.04	0.36	0.99	0.03
NEE _{5,2}	training	0.50	0.87	0.01	0.60	1.14	0.04	0.36	0.99	0.01
	testing	0.52	0.85	-0.01	0.59	1.16	0.05	0.37	0.98	0.01

Site-level evaluation									
NEE	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
NEE _{L4C}	0.44	0.90	0.06	0.54	1.04	0.40	0.50	0.79	0.73
NEE _{1,2}	0.42	0.79	-0.15	0.49	1.06	-0.08	0.42	0.81	0.24
NEE _{2,2}	0.43	0.74	-0.13	0.50	1.00	-0.19	0.42	0.83	0.24
NEE _{3,2}	0.40	0.75	-0.05	0.53	1.08	-0.07	0.41	0.83	0.22
NEE _{4,2}	0.47	0.78	-0.10	0.53	1.10	-0.06	0.38	0.83	0.21
NEE _{5,2}	0.44	0.76	-0.20	0.55	1.10	0.05	0.39	0.81	0.19

Table E2. Same as Table 4, but for **net ecosystem CO₂ exchange (NEE)**. Modeled NEE is computed as the difference between modeled ecosystem respiration (ER) and gross primary production (GPP). NEE_{L4C} denotes NEE from the original L4C model, while NEE_{i,2} refers to outputs from the Arctic–Subarctic (AS) adapted ER_i and **GPP₂** formulations.

Spatiotemporal evaluation										
NEE	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
NEE _{L4C}	training	0.54	0.87	-0.03	0.55	1.17	0.26	0.47	0.99	0.46
	testing	0.57	0.84	-0.03	0.55	1.17	0.26	0.48	0.99	0.45
NEE _{1,3}	training	0.44	0.84	-0.09	0.47	1.26	-0.14	0.40	0.97	-0.03
	testing	0.45	0.82	-0.10	0.46	1.26	-0.15	0.41	0.96	-0.03
NEE _{2,3}	training	0.55	0.78	-0.11	0.57	1.17	-0.09	0.43	0.95	-0.02
	testing	0.56	0.77	-0.12	0.56	1.18	-0.10	0.44	0.95	-0.01
NEE _{3,3}	training	0.54	0.80	0.00	0.58	1.14	0.01	0.39	0.97	-0.01
	testing	0.55	0.78	-0.00	0.58	1.15	0.01	0.39	0.97	-0.01
NEE _{4,3}	training	0.52	0.84	0.00	0.60	1.12	0.02	0.35	0.99	-0.02
	testing	0.54	0.81	-0.01	0.60	1.13	0.01	0.34	0.99	-0.02
NEE _{5,3}	training	0.52	0.84	0.00	0.61	1.12	0.02	0.37	0.98	-0.03
	testing	0.53	0.82	-0.01	0.60	1.13	0.02	0.37	0.98	-0.03

Site-level evaluation									
NEE	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
NEE _{L4C}	0.44	0.90	0.06	0.54	1.04	0.40	0.50	0.79	0.73
NEE _{1,3}	0.51	0.79	-0.01	0.48	1.06	-0.13	0.47	0.77	0.28
NEE _{2,3}	0.49	0.75	-0.08	0.54	0.98	-0.24	0.44	0.78	0.22
NEE _{3,3}	0.46	0.75	0.09	0.54	1.01	-0.07	0.43	0.78	0.27
NEE _{4,3}	0.47	0.76	-0.11	0.55	1.07	-0.10	0.41	0.78	0.27
NEE _{5,3}	0.44	0.78	-0.05	0.56	1.07	0.06	0.42	0.78	0.22

Table E3. Same as Table 4, but for **net ecosystem CO₂ exchange (NEE)**. Modeled NEE is computed as the difference between modeled ecosystem respiration (ER) and gross primary production (GPP). NEE_{L4C} denotes NEE from the original L4C model, while NEE_{i,3} refers to outputs from the Arctic–Subarctic (AS) adapted ER_i and **GPP₃** formulations.

Spatiotemporal evaluation										
NEE	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
NEE _{L4C}	training	0.54	0.87	-0.03	0.55	1.17	0.26	0.47	0.99	0.46
	testing	0.57	0.84	-0.03	0.55	1.17	0.26	0.48	0.99	0.45
NEE _{1,4}	training	0.60	0.75	-0.12	0.49	1.24	-0.15	0.46	0.94	-0.03
	testing	0.61	0.74	-0.12	0.49	1.24	-0.16	0.44	0.93	-0.04
NEE _{2,4}	training	0.69	0.67	-0.14	0.57	1.17	-0.09	0.48	0.93	-0.01
	testing	0.70	0.66	-0.14	0.57	1.17	-0.09	0.48	0.92	-0.01
NEE _{3,4}	training	0.70	0.68	0.01	0.58	1.14	0.02	0.48	0.93	-0.01
	testing	0.71	0.66	0.00	0.58	1.15	0.01	0.48	0.92	-0.01
NEE _{4,4}	training	0.64	0.76	0.00	0.64	1.07	0.02	0.46	0.94	-0.01
	testing	0.66	0.74	0.01	0.64	1.08	0.02	0.46	0.94	-0.02
NEE _{5,4}	training	0.62	0.78	-0.00	0.65	1.07	0.02	0.47	0.93	-0.03
	testing	0.64	0.76	0.01	0.65	1.08	0.02	0.47	0.93	-0.03

Site-level evaluation									
NEE	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
NEE _{L4C}	0.44	0.90	0.06	0.54	1.04	0.40	0.50	0.79	0.73
NEE _{1,4}	0.67	0.69	-0.13	0.54	1.02	-0.22	0.58	0.71	0.31
NEE _{2,4}	0.66	0.67	-0.09	0.57	1.00	-0.27	0.59	0.71	0.14
NEE _{3,4}	0.66	0.68	-0.21	0.59	0.97	-0.08	0.59	0.72	0.17
NEE _{4,4}	0.67	0.67	0.00	0.56	1.04	-0.15	0.60	0.70	0.17
NEE _{5,4}	0.62	0.68	-0.16	0.60	0.95	-0.14	0.60	0.72	0.17

Table E4. Same as Table 4, but for **net ecosystem CO₂ exchange (NEE)**. Modeled NEE is computed as the difference between modeled ecosystem respiration (ER) and gross primary production (GPP). NEE_{L4C} denotes NEE from the original L4C model, while NEE_{i,4} refers to outputs from the Arctic–Subarctic (AS) adapted ER_i and **GPP₄** formulations.

Spatiotemporal evaluation										
NEE	Splits	Upland Tundra			Taiga Forests			Wetlands		
		r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
NEE _{L4C}	training	0.54	0.87	-0.03	0.55	1.17	0.26	0.47	0.99	0.46
	testing	0.57	0.84	-0.03	0.55	1.17	0.26	0.48	0.99	0.45
NEE _{1,5}	training	0.60	0.77	-0.08	0.42	1.27	-0.13	0.39	0.98	0.01
	testing	0.61	0.75	-0.08	0.43	1.26	-0.14	0.39	0.97	-0.00
NEE _{2,5}	training	0.70	0.68	-0.09	0.55	1.17	-0.11	0.44	0.96	0.03
	testing	0.71	0.66	-0.08	0.55	1.17	-0.12	0.44	0.95	0.03
NEE _{3,5}	training	0.65	0.71	0.02	0.57	1.16	0.02	0.41	0.97	0.02
	testing	0.66	0.69	0.02	0.57	1.16	0.01	0.41	0.97	0.01
NEE _{4,5}	training	0.64	0.72	0.02	0.63	1.09	0.02	0.42	0.97	0.02
	testing	0.66	0.70	0.02	0.63	1.10	0.02	0.41	0.97	0.01
NEE _{5,5}	training	0.64	0.72	0.01	0.65	1.06	0.02	0.43	0.96	0.00
	testing	0.66	0.70	0.00	0.65	1.08	0.02	0.42	0.95	-0.00

Site-level evaluation									
NEE	Upland Tundra			Taiga Forests			Wetlands		
	r	ubRMSE	B	r	ubRMSE	B	r	ubRMSE	B
NEE _{L4C}	0.44	0.90	0.06	0.54	1.04	0.40	0.50	0.79	0.73
NEE _{1,5}	0.63	0.73	-0.10	0.45	1.09	-0.18	0.51	0.76	0.24
NEE _{2,5}	0.65	0.71	-0.05	0.55	0.96	-0.28	0.55	0.77	0.09
NEE _{3,5}	0.65	0.70	0.01	0.53	1.01	-0.04	0.56	0.74	0.25
NEE _{4,5}	0.64	0.68	0.06	0.59	0.96	0.04	0.56	0.71	0.25
NEE _{5,5}	0.65	0.68	0.00	0.60	0.94	-0.03	0.56	0.70	0.23

Table E5. Same as Table 4, but for **net ecosystem CO₂ exchange (NEE)**. Modeled NEE is computed as the difference between modeled ecosystem respiration (ER) and gross primary production (GPP). NEE_{L4C} denotes NEE from the original L4C model, while NEE_{i,5} refers to outputs from the Arctic–Subarctic (AS) adapted ER_i and **GPP₅** formulations.

- 809 Aubinet, M., Vesala, T., and Papale, D.: Eddy covariance: a practical guide to measurement and data analysis, Springer Science & Business
810 Media, 2012.
- 811 Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes,
812 J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw U, K., Pilegaard, K., Schmid,
813 H., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: A New Tool to Study the Temporal and Spatial
814 Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities., *Am. Meteorol. Soc.*, 82(11), 2415–2434,
815 [https://doi.org/https://doi.org/10.1175/1520-0477\(2001\)082<2415:FANTTS>2.3.CO;2](https://doi.org/https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2), 2007.
- 816 Bao, S., Wutzler, T., Koirala, S., Cuntz, M., Ibrom, A., Besnard, S., Walther, S., Šigut, L., Moreno, A., Weber, U., Wohlfahrt, G., Cleverly,
817 J., Migliavacca, M., Woodgate, W., Merbold, L., Elmar, V., and Carvalhais, N.: Environment-sensitivity functions for gross primary
818 productivity in light use efficiency models, *Agricultural and Forest Meteorology*, 312, 108 708, 2022.
- 819 Bell, J., Palecki, M., Baker, C., Collins, W., Lawrimore, J., Leeper, R., Hall, M., Kochendorfer, J., Meyers, T., Wilson, T., and Diamond, H.:
820 U.S. Climate Reference Network soil moisture and temperature observations., *J. Hydrometeorol.*, 14, 977–988, 2013.
- 821 Blume-Werry, G., Milbau, A., Teuber, L. M., Johansson, M., and Dorrepaal, E.: Dwelling in the deep—strongly increased root growth and
822 rooting depth enhance plant interactions with thawing permafrost soil, *New Phytologist*, 223, 1328–1339, 2019.
- 823 Bowling, D. R., Logan, B. A., Hufkens, K., Aubrecht, D. M., Richardson, A. D., Burns, S. P., Anderegg, W. R. L., Blanken, P. D., and Eiriks-
824 son, D. P.: Limitations to winter and spring photosynthesis of a Rocky Mountain subalpine forest, *Agricultural and Forest Meteorology*,
825 252, 241–255, 2018.
- 826 Bracho, R., Celis, G., Rodenhizer, H., See, C., and Schuur, E. A.: AmeriFlux BASE US-EML Eight Mile Lake Permafrost thaw gradient,
827 Healy Alaska, Ver. 4-5, AmeriFlux AMP, (Dataset)., <https://doi.org/https://doi.org/10.17190/AMF/1418678>, 2021.
- 828 Brown, J., Ferrians, O., Heginbottom, J., and Melnikov, E.: Circum-Arctic Map of Permafrost and Ground-Ice Conditions, Version 2.,
829 Boulder, Colorado USA., 2002.
- 830 Burba, G.: Eddy covariance method for scientific, industrial, agricultural and regulatory applications: A field book on measuring ecosystem
831 gas exchange and areal emission rates, LI-Cor Biosciences, 2013.
- 832 Burba, G.: Eddy covariance method for scientific, regulatory, and commercial applications, LI-COR Biosciences, 2022.
- 833 Chapin, F., Woodwell, G., Randerson, J., Rastetter, E., Lovett, G., Baldocchi, D., Clark, D., Harmon, M., Schimel, D., Valentini, R., Wirth,
834 C., Aber, J., Cole, J., Goulden, M., Harden, J., Heimann, M., Howarth, R., Matson, P., McGuire, A., Melillo, J., Mooney, H., Neff,
835 J., Houghton, R., Pace, M., Ryan, M., Running, S., Sala, O., Schlesinger, W., and Schulze, E.-D.: Reconciling Carbon-cycle Concepts,
836 Terminology, and Methods., *Ecosystems*, 9, 1041–1050, 2006.
- 837 Chen, N., Zhang, Y., Yuan, F., Song, C., Xu, M., Wang, Q., Hao, G., Bao, T., Zuo, Y., Liu, J., Song, Y., Sun, L., Guo, Y., Zhang, H., Ma, G.,
838 Du, Y., Xu, X., and Wang, X.: Warming-induced vapor pressure deficit suppression of vegetation growth diminished in northern peatlands,
839 *Nature Communications*, 14, 7885, 2023.
- 840 Churchill, A. C., Turetsky, M. R., McGuire, A. D., and Hollingsworth, T. N.: Response of plant community structure and primary productivity
841 to experimental drought and flooding in an Alaskan fen, *Canadian Journal of Forest Research*, 45, 185–193, 2015.
- 842 Colliander, A., Jackson, T., Bindlish, R., Chan, S., Das, N., Kim, S., Cosh, M., Dunbar, R., Dang, L., Pashaian, L., Asanuma, J., Aida,
843 K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K., Goodrich, D., al Jassar, H., Lopez-Baeza, E., Martínez Fernández, J.,
844 González-Zamora, A., Livingston, S., McNairn, H., Pacheco, A., Moghaddam, M., Montzka, C., Notarnicola, C., Niedrist, G., Pellarin,

845 T., Prueger, J., Pulliainen, J., Rautiainen, K., Ramos, J., Seyfried, M., Starks, P., Su, Z., Zeng, Y., van der Velde, R., Thibeault, M., Dorigo,
846 W., Vreugdenhil, M., Walker, J. P., Wu, X., Monerris, A., O'Neill, P. E., Entekhabi, D., Njoku, E., and Yueha, S.: Validation of SMAP
847 surface soil moisture products with core validation sites, *Remote sensing of environment*, 191, 215–231, 2017.

848 Crawford, R. M. M.: *Tundra-taiga biology*, Oxford University Press, 2013.

849 Davis, T. W., Prentice, I. C., Evans, B. J., Wang, H., and Gilbert, X.: The Global ecosystem Production in Space and Time (GePiSaT) Model
850 of the Terrestrial Biosphere, 2014, H53J–04, 2014.

851 Derksen, C., Xu, X., Scott Dunbar, R., Colliander, A., Kim, Y., Kimball, J. S., Black, T. A., Euskirchen, E., Langlois, A., Loranty, M. M.,
852 Marsh, P., and Rautiainen, K., Roy, A., Royer, A., and Stephens, J.: Retrieving landscape freeze/thaw state from Soil Moisture Active
853 Passive (SMAP) radar and radiometer measurements, *Remote Sensing of Environment*, 194, 48–62, 2017.

854 Dragoni, D., Schmid, H. P., Grimmond, C. S. B., and Loescher, H. W.: Uncertainty of annual net ecosystem productivity estimated using
855 eddy covariance flux measurements, *Journal of Geophysical Research: Atmospheres*, 112, 2007.

856 Elmendorf, S. C. and Hollister, R. D.: Limits on phenological response to high temperature in the Arctic, *Scientific Reports*, 13, 208, 2023.

857 Endsley, K., Kimball, J., and Reichle, R.: Soil respiration phenology improves modeled phase of terrestrial net ecosystem exchange in
858 northern hemisphere, *Journal of Advances in Modeling Earth Systems*, 14, e2021MS002804, 2022.

859 Ensminger, I., Svishnikov, D., Campbell, D. A., Funk, C., Jansson, S., Lloyd, J., Shibistova, O., and Öquist, G.: Intermittent low temperatures
860 constrain spring recovery of photosynthesis in boreal Scots pine forests, *Global Change Biology*, 10, 995–1008, 2004.

861 Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J.,
862 Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C.,
863 Spencer, M. W., Thurman, S. W., Tsang, L., and Van Zyl, J.: The soil moisture active passive (SMAP) mission, *Proceedings of the IEEE*,
864 98, 704–716, 2010.

865 Entekhabi, D., Yueh, S., O'Neill, P. E., and Kellogg, K. H.: *SMAP Handbook*, Tech. rep., Jet Propulsion Laboratory, NASA, 2014.

866 Eugster, W., McFadden, J. P., and Chapin, F. S.: A comparative approach to regional variation in surface fluxes using mobile eddy correlation
867 towers, *Boundary-Layer Meteorology*, 85, 293–307, 1997.

868 Euskirchen, E.: AmeriFlux BASE US-BZB Bonanza Creek Thermokarst Bog, Ver. 4-5, AmeriFlux AMP, (Dataset).,
869 <https://doi.org/https://doi.org/10.17190/AMF/1773401>, 2022a.

870 Euskirchen, E.: AmeriFlux BASE US-BZF Bonanza Creek Rich Fen, Ver. 4-5, AmeriFlux AMP, (Dataset).,
871 <https://doi.org/https://doi.org/10.17190/AMF/1756433>, 2022b.

872 Euskirchen, E.: AmeriFlux BASE US-BZo Bonanza Creek Old Thermokarst Bog, Ver. 3-5, AmeriFlux AMP, (Dataset).,
873 <https://doi.org/https://doi.org/10.17190/AMF/1846662>, 2022c.

874 Euskirchen, E.: AmeriFlux BASE US-BZS Bonanza Creek Black Spruce, Ver. 3-5, AmeriFlux AMP, (Dataset).,
875 <https://doi.org/https://doi.org/10.17190/AMF/1756434>, 2022d.

876 Euskirchen, E., Shaver, G., and Bret-Harte, S.: AmeriFlux BASE US-IC_h Imnavait Creek Watershed Heath Tundra, Ver. 4-5, AmeriFlux
877 AMP, (Dataset)., <https://doi.org/https://doi.org/10.17190/AMF/1246133>, 2022a.

878 Euskirchen, E., Shaver, G., and Bret-Harte, S.: AmeriFlux BASE US-IC_s Imnavait Creek Watershed Wet Sedge Tundra, Ver. 7-5, AmeriFlux
879 AMP, (Dataset)., <https://doi.org/https://doi.org/10.17190/AMF/1246130>, 2022b.

880 Euskirchen, E., Shaver, G., and Bret-Harte, S.: AmeriFlux BASE US-IC_t Imnavait Creek Watershed Tussock Tundra, Ver. 5-5, AmeriFlux
881 AMP, (Dataset)., <https://doi.org/https://doi.org/10.17190/AMF/1246131>, 2022c.

882 Euskirchen, E. S., McGuire, A. D., Chapin III, F. S., Yi, S., and Thompson, C. C.: Changes in vegetation in northern Alaska under scenarios
883 of climate change, 2003–2100: implications for climate feedbacks, *Ecological applications*, 19, 1022–1043, 2009.

884 Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., and Granier,
885 A.: Gap filling strategies for defensible annual sums of net ecosystem exchange, *Agricultural and forest meteorology*, 107, 43–69, 2001.

886 Finkelstein, P. L. and Sims, P. F.: Sampling error in eddy correlation flux measurements, *Journal of Geophysical Research: Atmospheres*,
887 106, 3503–3509, 2001.

888 Fisher, J. B., Hayes, D. J., Schwalm, C. R., Huntzinger, D. N., Stofferahn, E., Schaefer, K. and Luo, Y., Wullschleger, S. D., Goetz, S., Miller,
889 C. E., Griffith, P., Chadburn, S., Chatterjee, A., Ciais, P., Douglas, T., Genet, H., Ito, A., Neigh, C., Poulter, B., Rogers, B., Sonnentag, O.,
890 Tian, H., Wang, W., Xue, Y., Yang, Z.-L., Zeng, N., , and Zhang, Z.: Missing pieces to modeling the Arctic-Boreal puzzle, *Environmental*
891 *Research Letters*, 13, 020 202, 2018.

892 Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R., Heimann, M., Thonicke, K., Zaehle, S., and Reichstein, M.: Enhanced seasonal CO₂
893 exchange caused by amplified plant productivity in northern ecosystems, *Science*, 351, 696–699, 2016.

894 Foster, K. T., Sun, W., Merder, J., Kurz, W., Sinha, E., Nesdoly, A., Metsaranta, J., Hararuk, O., Bond-Lamberty, B. P., Schwalm, C., Natali,
895 S., Huntzinger, D. N., and Michalak, A.: Permafrost, Peatland and Agricultural Regions Key to Reconciling Top-Down and Bottom-Up
896 Estimates of North American Carbon Uptake, in: *AGU Fall Meeting Abstracts*, vol. 2024, pp. A43T-08, 2024.

897 Fotouo Makouate, H. and Zude-Sasse, M.: Advances in Growing Degree Days Models for Flowering to Harvest: Optimizing Crop Manage-
898 ment with Methods of Precision Horticulture—A Review, *Horticultrae*, 11, 1415, 2025.

899 Friedl, M. and Sulla-Menashe, D.: MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006,
900 <https://doi.org/10.5067/MODIS/MCD12Q1.006>, 2019.

901 Fu, Y. H., Piao, S., Zhao, H., Jeong, S.-J., Wang, X., Vitasse, Y., Ciais, P., and Janssens, I. A.: Unexpected role of winter precipitation in
902 determining heat requirement for spring vegetation green-up at northern middle and high latitudes, *Global change biology*, 20, 3743–3755,
903 2014.

904 Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R.,
905 Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R.,
906 Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and
907 Zhao, B.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), *Journal of climate*, 30, 5419–5454,
908 2017.

909 Gilmanov, T. G., Verma, S. B., Sims, P. L., Meyers, T. P., Bradford, J. A., Burba, G. G., and Suyker, A. E.: Gross primary production and light
910 response parameters of four Southern Plains ecosystems estimated using long-term CO₂-flux tower measurements, *Global Biogeochemical*
911 *Cycles*, 17, 2003.

912 Gonsamo, A., Chen, J. M., and D’Odorico, P.: Deriving land surface phenology indicators from CO₂ eddy covariance measurements, *Eco-*
913 *logical Indicators*, 29, 203–207, 2013.

914 Hagedorn, F., Imboden, J., Moiseev, P. A., Gao, D., Frossard, E., Schleppe, P., Christen, D., Gavazov, K., and Fetzer, J.: Distinct changes in
915 carbon, nitrogen, and phosphorus cycling in the litter layer across two contrasting forest–tundra ecotones, *Biogeosciences*, 22, 2959–2977,
916 2025.

917 Hayes, D. J., McGuire, A. D., Kicklighter, D. W., Gurney, K. R., Burnside, T., and Melillo, J. M.: Is the northern high-latitude land-based
918 CO₂ sink weakening?, *Global Biogeochemical Cycles*, 25, 2011.

919 He, H., Moore, T., Lafleur, P., Sonnentag, O., Humphreys, E., Wu, M., and Roulet, N.: Spring phenology in photosynthesis control and
920 modeling for a temperate bog, *Frontiers in Environmental Science*, 13, 1548 578, 2025.

921 Heinzelmann, V., Marinissen, J., Aerts, R., Cornelissen, J. H. C., and Bokhorst, S.: Stronger Drought Response of CO₂ Fluxes in Tundra
922 Heath Compared to Sphagnum Peatland in the Sub-Arctic, *Global Change Biology*, 31, e70 210, 2025.

923 Helbig, M., Chasmer, L. E., Desai, A. R., Kljun, N., Quinton, W. L., and Sonnentag, O.: Direct and indirect climate change effects on carbon
924 dioxide fluxes in a thawing boreal forest–wetland landscape, *Global Change Biology*, 23, 3231–3248, 2017a.

925 Helbig, M., Chasmer, L. E., Kljun, N., Quinton, W. L., Treat, C. C., and Sonnentag, O.: The positive net radiative greenhouse gas forcing of
926 increasing methane emissions from a thawing boreal forest-wetland landscape, *Global change biology*, 23, 2413–2427, 2017b.

927 Hu, F. and Bliss, L.: Tundra, *Encyclopedia Britannica*, <https://www.britannica.com/science/tundra>, [Date accessed: 2025-06-11], 2025.

928 Huang, X., Xiao, J., and Ma, M.: Evaluating the performance of satellite-derived vegetation indices for estimating gross primary productivity
929 using FLUXNET observations across the globe, *Remote Sensing*, 11, 1823, 2019.

930 Hugelius, G., Strauss, J., Zubrzycki, S., Harden, J., Schuur, E., Ping, C.-L., Schirrmeister, L., Grosse, G., Michaelson, G., Koven, C., and
931 O'Donnell, J.: Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps, *Biogeo-*
932 *sciences*, 11, 6573–6593, 2014.

933 Ise, T. and Moorcroft, P. R.: The global-scale temperature and moisture dependencies of soil organic carbon decomposition: an analysis using
934 a mechanistic decomposition model, *Biogeochemistry*, 80, 217–231, 2006.

935 Iwahana, G., Kobayashi, H., Ikawa, H., and Suzuki, R.: AmeriFlux BASE US-Prr Poker Flat Research Range Black Spruce Forest, Ver. 4-5,
936 AmeriFlux AMP, (Dataset)., <https://doi.org/https://doi.org/10.17190/AMF/1246153>, 2023.

937 Jia, G. J., Epstein, H. E., and Walker, D. A.: Greening of arctic Alaska, 1981–2001, *Geophysical Research Letters*, 30, 2003.

938 Jones, L., Kimball, J., Reichle, R., Madani, N., Glassy, J., Ardizzone, J., Colliander, A., Cleverly, J., Desai, A., Eamus, D., Euskirchen,
939 E., Hutley, L., Macfarlane, C., and Scott, R.: The SMAP Level 4 Carbon Product for Monitoring Ecosystem Land–Atmosphere CO₂
940 Exchange., *IEEE T. Geosci. Remote*, 55(11), 6517–6532, <https://doi.org/960.10.1109/TGRS.2017.2729343>, 2017.

941 Juday, G.: Taiga, Plants, Animals, Climate, Location, & Facts, *Encyclopedia Britannica*, <https://www.britannica.com/science/taiga>, [Date
942 accessed: 2025-06-11], 2025.

943 Kerr, Y., Waldteufel, P., Richaume, P., Wigneron, J., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S., Leroux, D.,
944 Mialon, A., and Delwart, S.: The SMOS Soil Moisture Retrieval Algorithm, *IEEE Transactions on Geoscience and Remote Sensing*, 50,
945 1384–1403, <https://doi.org/10.1109/TGRS.2012.2184548>, 2012.

946 Kim, Y., Kim, S.-D., Enomoto, H., Kushida, K., Kondoh, M., and Uchida, M.: Latitudinal distribution of soil CO₂ efflux and temperature
947 along the Dalton Highway, Alaska, *Polar science*, 7, 162–173, 2013.

948 Kimball, J. S., Jones, L. A., Zhang, K., Heinsch, F. A., McDonald, K. C., and Oechel, W. C.: A Satellite Approach to Estimate Land–
949 Atmosphere CO₂ Exchange for Boreal and Arctic Biomes Using MODIS and AMSR-E, *IEEE Transactions on Geoscience and Remote*
950 *Sensing*, 47, 569–587, 2008.

951 Kimball, J. S., Endsley, A., Jones, L. A., Kundig, T., and Reichle, R.: SMAP L4 Global Daily 9 km EASE-Grid Carbon Net Ecosystem
952 Exchange, Version 8, <https://doi.org/10.5067/U7SN8JDZL0UC>, 2025.

953 Kreuzwieser, J., Papadopolou, E., and Rennenberg, H.: Interaction of flooding with carbon metabolism of forest trees, *Plant Biology*, 6,
954 299–306, 2004.

955 Lasslop, G., Reichstein, M., Detto, M., Richardson, A. D., and Baldocchi, D. D.: Comment on Vickers et al.: Self-correlation between
 956 assimilation and respiration resulting from flux partitioning of eddy-covariance CO₂ fluxes, *Agricultural and Forest Meteorology*, 150,
 957 312–314, 2010a.

958 Lasslop, G., Reichstein, M., Papale, D., Richardson, A. D., Arneeth, A., Barr, A., Stoy, P., and Wohlfahrt, G.: Separation of net ecosystem
 959 exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation, *Global change*
 960 *biology*, 16, 187–208, 2010b.

961 Leclerc, M. and Thurtell, G.: Footprint prediction of scalar fluxes using a Markovian analysis, *Boundary-Layer Meteorology*, 52, 247–258,
 962 1990.

963 Lees, K., Quaife, T., Artz, R., Khomik, M., and Clark, J.: Potential for using remote sensing to estimate carbon fluxes across northern
 964 peatlands—A review, *Science of the Total Environment*, 615, 857–874, 2018.

965 Lievens, H., Demuzere, M., Marshall, H., Reichle, R. H., Brucker, L., Brangers, I., de Rosnay, P., Dumont, M., Giroto, M., Immerzeel,
 966 W. W., Jonas, T., Kim, E., Koch, I., Marty, C., Saloranta, T., Schöber, J., , and De Lannoy, G.: Snow depth variability in the Northern
 967 Hemisphere mountains observed from space, *Nature communications*, 10, 4629, 2019.

968 Lloyd, J. and Taylor, J. A.: On the Temperature Dependence of Soil Respiration, *Functional Ecology*, 8, 315–323, [http://www.jstor.org/stable/](http://www.jstor.org/stable/2389824)
 969 2389824, 1994.

970 López, J., Way, D. A., and Sadok, W.: Systemic effects of rising atmospheric vapor pressure deficit on plant physiology and productivity,
 971 *Global Change Biology*, 27, 1704–1720, 2021.

972 López-Blanco, E., Exbrayat, J., Lund, M., Christensen, T. R., Tamstorf, M. P., Slevin, D., Hugelius, G., Bloom, A. A., and Williams, M.:
 973 Evaluation of terrestrial pan-Arctic carbon cycling using a data-assimilation system, *Earth System Dynamics*, 10, 233–255, 2019.

974 Lucchesi, R.: File Specification for GEOS FP, http://gmao.gsfc.nasa.gov/pubs/office_notes, 2018.

975 Luo, X., Zhao, R., Chu, H., Collalti, A., Faticchi, S., Keenan, T. F., Lu, X., Nguyen, N., Prentice, I. C., Sun, W., Yu, K., and Yu, L.: Global
 976 variation in vegetation carbon use efficiency inferred from eddy covariance observations, *Nature Ecology & Evolution*, 9, 1414–1425,
 977 2025.

978 Madelon, R., Kimball, J. S., Endsley, K. A., De Lannoy, G. J. M., Sonnentag, O., Alcock, H., Detto, M., Virkkala, A. M., Rogers, B. M., Watts,
 979 J. D., Mavrovic, A., Williamson, S. N., Humphreys, E., Colliander, A., Mialon, A., and Roy, A.: Assessing the SMAP Level-4 Carbon
 980 Product over the Arctic and Subarctic Zones, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2025.

981 Maire, V., Martre, P., Kattge, J., Gastal, F., Esser, G., Fontaine, S., and Soussana, J.-F.: The coordination of leaf photosynthesis links C and
 982 N fluxes in C₃ plant species, *PloS one*, 7, e38 345, 2012.

983 Maltby, E. and Immirzi, P.: Carbon dynamics in peatlands and other wetland soils: Regional and global perspectives, *Chemosphere*, 27,
 984 999–1023, 1993.

985 Martinez Molera, L.: Machine Learning Q&A: All About Model Validation, MathWorks Online Documentation, [https://www.mathworks.](https://www.mathworks.com/campaigns/offers/next/all-about-model-validation.html)
 986 [com/campaigns/offers/next/all-about-model-validation.html](https://www.mathworks.com/campaigns/offers/next/all-about-model-validation.html), accessed: 2026-01-29, 2025.

987 MathWorks, Inc.: MATLAB R2023b, MathWorks, Natick, Massachusetts, USA, <https://www.mathworks.com/products/matlab.html>, 2023.

988 Mavrovic, A., Sonnentag, O., Lemmetyinen, J., Baltzer, J. L., Kinnard, C., and Roy, A.: Reviews and syntheses: Recent advances in mi-
 989 crowave remote sensing in support of terrestrial carbon cycle science in Arctic–boreal regions, *Biogeosciences*, 20, 2941–2970, 2023a.

990 Mavrovic, A., Sonnentag, O., Lemmetyinen, J., Voigt, C., Rutter, N., Mann, P., Sylvain, J.-D., and Roy, A.: Environmental controls of winter
 991 soil carbon dioxide fluxes in boreal and tundra environments, *Biogeosciences*, 20, 5087–5108, <https://doi.org/10.5194/bg-20-5087-2023>,
 992 2023b.

993 McCallum, I., Franklin, O., Moltchanova, E., Merbold, L., Schmullius, C., Shvidenko, A., Schepaschenko, D., and Fritz, S.: Improved light
 994 and temperature responses for light-use-efficiency-based GPP models, *Biogeosciences*, 10, 6577–6590, [https://doi.org/10.5194/bg-10-](https://doi.org/10.5194/bg-10-6577-2013)
 995 6577-2013, 2013.

996 McGuire, A. D., Christensen, T., Hayes, D., Herault, A., Euskirchen, E., Kimball, J., Koven, C., Laffleur, P., Miller, P., Oechel, W., Peylin, P.,
 997 Williams, M., and Yi, Y.: An assessment of the carbon balance of Arctic tundra: comparisons among observations, process models, and
 998 atmospheric inversions, *Biogeosciences*, 9, 3185–3204, 2012.

999 McPartland, M. Y., Falkowski, M. J., Reinhardt, J. R., Kane, E. S., Kolka, R., Turetsky, M. R., Douglas, T. A., Anderson, J., Edwards,
 1000 J. D., Palik, B., and Montgomery, R. A.: Characterizing boreal peatland plant composition and species diversity with hyperspectral remote
 1001 sensing, *Remote Sensing*, 11, 1685, 2019.

1002 Mialon, A., Rodríguez-Fernández, N. J., Santoro, M., Saatchi, S., Mermoz, S., Bousquet, E., and Kerr, Y. H.: Evaluation of the sensitivity of
 1003 SMOS L-VOD to forest above-ground biomass at global scale, *Remote Sensing*, 12, 1450, 2020.

1004 Mirabel, A., Girardin, M. P., Metsaranta, J., Way, D., and Reich, P. B.: Increasing atmospheric dryness reduces boreal forest tree growth,
 1005 *Nature Communications*, 14, 6901, 2023.

1006 Mishra, U., Hugelius, G., Shelef, E., Yang, Y., Strauss, J., Lupachev, A., Harden, J., Jastrow, J., Ping, C., Riley, W., and Schuur, E.: Spatial
 1007 heterogeneity and environmental predictors of permafrost region soil organic carbon stocks, *Sci. Adv.*, 7, eaaz5236, 2021.

1008 Myneni, R. and Knyazikhin, Y.: VIIRS/NPP Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V001,
 1009 <https://doi.org/10.5067/VIIRS/VNP15A2H.001>, 2018.

1010 Myneni, R. B., Keeling, C., Tucker, C. J., Asrar, G., and Nemani, R. R.: Increased plant growth in the northern high latitudes from 1981 to
 1011 1991, *Nature*, 386, 698–702, 1997.

1012 Natali, S.: AmeriFlux BASE US-YK1 Yukon-Kuskokwim Delta, Izaviknek-Kingaglia uplands, Burned 2015, Ver. 1-5, AmeriFlux AMP,
 1013 (Dataset), <https://doi.org/https://doi.org/10.17190/AMF/2331384>, 2024.

1014 Natali, S.: AmeriFlux BASE US-YK2 Yukon-Kuskokwim Delta, Izaviknek-Kingaglia uplands, Unburned, Ver. 2-5, AmeriFlux AMP,
 1015 (Dataset), <https://doi.org/https://doi.org/10.17190/AMF/2331385>, 2025.

1016 Natali, S., Watts, J., Rogers, B., Potter, S., Ludwig, S., Selbmann, A.-K., Sullivan, P., Abbott, B., Arndt, K., Birch, L., Björkman, M., Bloom,
 1017 A., Celis, G., Christensen, T., Christiansen, C., Commane, R., Cooper, E., Crill, P., Czimczik, C., Davydov, S., Du, J., Egan, J., Elberling,
 1018 B., Euskirchen, E., Friborg, T., Genet, H., Göckede, M., Goodrich, J., Grogan, P., Helbig, M., Jafarov, E., Jastrow, J., Kalhori, A., Kim,
 1019 Y., Kimball, J., Kutzbach, L., Lara, M., Larsen, K., Lee, B.-Y., Liu, Z., Loranty, M., Lund, M., Lupascu, M., Madani, N., Malhotra, A.,
 1020 Matamala, R., McFarland, J., McGuire, A., Michelsen, A., Minions, C., Oechel, W., Olefeldt, D., Parmentier, F.-J., Pirk, N., Poulter, B.,
 1021 Quinton, W., Rezanezhad, F., Risk, D., Sachs, T., Schaefer, K., Schmidt, N., Schuur, E., Semenchuk, P., Shaver, G., Sonnentag, O., Starr,
 1022 G., Treat, C., Waldrop, M., Wang, Y., Welker, J., Wille, C., Xu, Zhang, Z., Zhuang, Q., and Zona, D.: Large loss of CO₂ in winter observed
 1023 across the northern permafrost region., *Nature Climate Change*, 9, 852–857, 2019.

1024 Natali, S. M., Schuur, E. A., and Rubin, R. L.: Increased plant productivity in Alaskan tundra as a result of experimental warming of soil and
 1025 permafrost, *Journal of ecology*, 100, 488–498, 2012.

1026 Nawaz, A. F., Gargiulo, S., Pichierri, A., and Casolo, V.: Exploring the Role of Non-Structural Carbohydrates (NSCs) Under Abiotic Stresses
 1027 on Woody Plants: A Comprehensive Review, *Plants*, 14, 328, 2025.

1028 Oechel, W. C., Hastings, S. J., Vourltis, G., Jenkins, M., Riechers, G., and Grulke, N.: Recent change of Arctic tundra ecosystems from a
 1029 net carbon dioxide sink to a source, *Nature*, 361, 520–523, 1993.

1030 Olefeldt, D., Euskirchen, E. S., Harden, J., Kane, E., McGuire, A. D., Waldrop, M. P., and Turetsky, M. R.: A decade of boreal rich fen
1031 greenhouse gas fluxes in response to natural and experimental water table variability, *Global change biology*, 23, 2428–2440, 2017.

1032 Pallandt, M., Kumar, J., Mauritz, M., Schuur, E., Virkkala, A.-M., Celis, G., Hoffman, F., and Göckede, M.: Representativeness as-
1033 sessment of the pan-Arctic eddy covariance site network and optimized future enhancements., *Biogeosciences*, 19(3), 559–583,
1034 <https://doi.org/10.5194/bg-19-559-2022>, 2022.

1035 Panwar, A., Migliavacca, M., Nelson, J. A., Cortés, J., Bastos, A., Forkel, M., and Winkler, A. J.: Methodological challenges and new
1036 perspectives of shifting vegetation phenology in eddy covariance data, *Scientific Reports*, 13, 13 885, 2023.

1037 Parazoo, N. C., Arneth, A., Pugh, T. A. M., Smith, B., Steiner, N., Luus, K., Commene, R., Benmergui, J., Stofferahn, E., Liu, J., Rödenbeck,
1038 C., Kawa, R., Euskirchen, E., Zona, D., Arndt, K., Oechel, W., and Miller, C.: Spring photosynthetic onset and net CO₂ uptake in Alaska
1039 triggered by landscape thawing, *Global Change Biology*, 24, 3416–3435, 2018.

1040 Peng, J., Tang, J., Xie, S., Wang, Y., Liao, J., Chen, C., Sun, C., Mao, J., Zhou, Q., and Niu, S.: Evidence for the acclimation of ecosystem
1041 photosynthesis to soil moisture, *Nature Communications*, 15, 9795, 2024.

1042 Potapov, P., Hansen, M. C., Stehman, S. V., Loveland, T. R., and Pittman, K.: Combining MODIS and Landsat imagery to estimate and map
1043 boreal forest cover loss, *Remote sensing of environment*, 112, 3708–3719, 2008.

1044 Prince, M., Roy, A., Royer, A., and Langlois, A.: Timing and spatial variability of fall soil freezing in boreal forest and its effect on SMAP
1045 L-band radiometer measurements, *Remote sensing of environment*, 231, 111 230, 2019.

1046 Pu, J., Yan, K., Roy, S., Zhu, Z., Rautiainen, M., Knyazikhin, Y., and Myneni, R. B.: Sensor-independent LAI/FPAR CDR: reconstruct-
1047 ing a global sensor-independent climate data record of MODIS and VIIRS LAI/FPAR from 2000 to 2022, *Earth System Science Data*
1048 *Discussions*, 2023, 1–29, 2023.

1049 Pulliainen, J., Aurela, M., Aalto, T., Böttcher, K., Cohen, J., Derksen, C., Heimann, M., Helbig, M., Kolari, P., Kontu, A., Krasnova, A.,
1050 Launiainen, S., Lemmetyinen, J., Lindqvista, H., Lindroth, A., Lohila, A., Luojusa, K., Mammarella, I., Markkanen, T., Nevala, E., Noe,
1051 S., Peichl, M., Pumpanen, J., Rautiainen, K., Salminen, M., Sonnentag, O., Takala, M., Thum, T., Vesala, T., and Vestin, P.: Increase
1052 in gross primary production of boreal forests balanced out by increase in ecosystem respiration, *Remote Sensing of Environment*, 313,
1053 114 376, 2024.

1054 Péwé, T.: Permafrost, *Encyclopedia Britannica*, <https://www.britannica.com/science/permafrost>, [Date accessed: 2025-06-11], 2025.

1055 Ramage, J., Kuhn, M., Virkkala, A., Voigt, C., Marushchak, M. E., Bastos, A., Biasi, C., Canadell, J. G., Ciais, P., López-Blanco, E., Natali,
1056 S. M., Olefeldt, D., Potter, S., Poulter, B., Rogers, B. M., Schuur, E. A. G., Treat, C., Turetsky, M. R., Watts, J., and Hugelius, G.: The
1057 net GHG balance and budget of the permafrost region (2000–2020) from ecosystem flux upscaling, *Global Biogeochemical Cycles*, 38,
1058 e2023GB007 953, 2024.

1059 Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., and Laaksonen, A.: The Arctic has
1060 warmed nearly four times faster than the globe since 1979, *Communications Earth & Environment*, 3, 168, 2022.

1061 Rautiainen, K., Parkkinen, T., Lemmetyinen, J., Schwank, M., Wiesmann, A., Ikonen, J., Derksen, C., Davydov, S., Davydova, A., Boike, J.,
1062 Langer, M., Drusch, M., and Pulliainen, J.: SMOS prototype algorithm for detecting autumn soil freezing, *Remote Sensing of Environment*,
1063 180, 346–360, 2016.

1064 Reichle, R., De Lannoy, G., Koster, R., Crow, W., Kimball, J., Liu, Q., and Bechtold, M.: SMAP L4 Global 9 km EASE-Grid Surface and
1065 Root Zone Soil Moisture Land Model Constants, Version 8, <https://doi.org/10.5067/PXQIBL2ALDZD>, 2025.

1066 Reichle, R. H., Liu, Q., Koster, R. D., Crow, W. T., De Lannoy, G. J., Kimball, J. S., Ardizzone, J. V., Bosch, D., Colliander, A., Cosh, M.,
 1067 Kolassa, J., Mahanama, S. P., Prueger, J., Starks, P., and Walker, J. P.: Version 4 of the SMAP level-4 soil moisture algorithm and data
 1068 product, *Journal of Advances in Modeling Earth Systems*, 11, 3106–3130, 2019.

1069 Reichle, R. H., Liu, Q., Ardizzone, J. V., Bechtold, M., Crow, W. T., De Lannoy, G., Kimball, J. S., and Koster, R. D.: Soil moisture active
 1070 passive (smap) project assessment report for version 7 of the l4_sm data product, 2023.

1071 Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., and Granier, A.:
 1072 On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm, *Global change*
 1073 *biology*, 11, 1424–1439, 2005.

1074 Robinson, S. D. and Moore, T. R.: The influence of permafrost and fire upon carbon accumulation in high boreal peatlands, *Northwest*
 1075 *Territories, Canada, Arctic, Antarctic, and Alpine Research*, 32, 155–166, 2000.

1076 Rouse, W. R., Douglas, M. S. V., Hecky, R. E., Hershey, A. E., Kling, G. W., Lesack, L., Marsh, P., McDonald, M., Nicholson, B. J., Roulet,
 1077 N. T., and SMOL, J. P.: Effects of climate change on the freshwaters of arctic and subarctic North America, *Hydrological processes*, 11,
 1078 873–902, 1997.

1079 Runkle, B. R. K., Sachs, T., Wille, C., Pfeiffer, E.-M., and Kutzbach, L.: Bulk partitioning the growing season net ecosystem exchange of CO₂
 1080 in Siberian tundra reveals the seasonality of its carbon sequestration strength, *Biogeosciences*, 10, 1337–1349, [https://doi.org/10.5194/bg-](https://doi.org/10.5194/bg-10-1337-2013)
 1081 [10-1337-2013](https://doi.org/10.5194/bg-10-1337-2013), 2013.

1082 Salmabadi, H., Pardo Lara, R., Berg, A., Mavrovic, A., Hanes, C., Montpetit, B., and Roy, A.: Improving Seasonally Frozen
 1083 Ground Monitoring Using Soil Freezing Characteristic Curve in Permittivity–Temperature Space, *EGUsphere*, 2025, 1–36,
 1084 <https://doi.org/10.5194/egusphere-2025-620>, 2025.

1085 Schaefer, K., Schwalm, C. R., Williams, C., Arain, M. A., Barr, A., Chen, J. M., Davis, K. J., Dimitrov, D., Hilton, T. W., Hollinger, D. Y.,
 1086 Humphreys, E., Poulter, B., Raczka, B. M., Richardson, A. D., Sahoo, A., Thornton, P., Vargas, R., Verbeeck, H., Anderson, R., Baker, I.,
 1087 Black, T. A., Bolstad, P., Chen, J., Curtis, P. S., Desai, A. R., Dietze, M., Dragoni, D., Gough, C., Grant, R. F., Gu, L., Jain, A., Kucharik,
 1088 C., Law, B., Liu, S., Lokipitiya, E., Margolis, H. A., Matamala, R., McCaughey, J. H., Monson, R., Munger, J. W., Oechel, W., Peng, C.,
 1089 Price, D. T., Ricciuto, D., Riley, W. J., Roulet, N., Tian, H., Tonitto, C., Torn, M., Weng, E., and X., Z.: A model-data comparison of gross
 1090 primary productivity: Results from the North American Carbon Program site synthesis, *Journal of Geophysical Research: Biogeosciences*,
 1091 117, 2012.

1092 Schaefer, K., Lantuit, H., Romanovsky, V. E., G., S. E. A., and Witt, R.: The impact of the permafrost carbon feedback on global climate,
 1093 *Environmental Research Letters*, 9(8), 085–003, <https://doi.org/10.1088/1748-9326/9/8/085003>, 2014.

1094 Schuepp, P., Leclerc, M., MacPherson, J., and Desjardins, R.: Footprint prediction of scalar fluxes from analytical solutions of the diffusion
 1095 equation, *Boundary-Layer Meteorology*, 50, 355–373, 1990.

1096 Schuur, E., Abbott, B., Bowden, W., Brovkin, V., Camill, P., Canadell, J., Chanton, J., Chapin, F., Christensen, T., Ciais, P., Crosby, B. T.,
 1097 Czimczik, C. I., Grosse, G., Harden, J., Hayes, D. J., Hugelius, G., Jastrow, J. D., Jones, J. B., Kleinen, T., Koven, C. D., Krinner, G.,
 1098 Kuhry, P., Lawrence, D. M., McGuire, A. D., Natali, S. M., O'Donnell, J. A., Ping, C. L., Riley, W. J., Rinke, A., Romanovsky, V. E.,
 1099 Sannel, A. B. K., Schädel, C., Schaefer, K., Sky, J., Subin, Z. M., Tarnocai, C., Turetsky, M. R., Waldrop, M. P., Walter Anthony, K. M.,
 1100 Wickland, K. P., Wilson, C. J., and Zimov, S. A.: Expert assessment of vulnerability of permafrost carbon to climate change, *Climatic*
 1101 *Change*, 119, 359–374, 2013.

1102 Sierra, C. A., Ceballos-Núñez, V., Hartmann, H., Herrera-Ramírez, D., and Metzler, H.: Ideas and perspectives: Allocation of carbon from
 1103 net primary production in models is inconsistent with observations of the age of respired carbon, *Biogeosciences*, 19, 3727–3738, 2022.

1104 Sonnentag, O.: AmeriFlux BASE CA-SMC Smith Creek, Ver. 1-5, AmeriFlux AMP, (Dataset).,
 1105 <https://doi.org/https://doi.org/10.17190/AMF/1767830>, 2021.

1106 Sonnentag, O. and Marsh, P.: AmeriFlux BASE CA-HPC Havikpak Creek, Ver. 1-5, AmeriFlux AMP, (Dataset).,
 1107 <https://doi.org/https://doi.org/10.17190/AMF/1773392>, 2021a.

1108 Sonnentag, O. and Marsh, P.: AmeriFlux BASE CA-TVC Trail Valley Creek, Ver. 1-5, AmeriFlux AMP, (Dataset).,
 1109 <https://doi.org/https://doi.org/10.17190/AMF/1767831>, 2021b.

1110 Sonnentag, O. and Quinton, W. L.: AmeriFlux BASE CA-SCC Scotty Creek Landscape, Ver. 1-5, AmeriFlux AMP, (Dataset).,
 1111 <https://doi.org/https://doi.org/10.17190/AMF/1480303>, 2018.

1112 Sonnentag, O. and Quinton, W. L.: AmeriFlux BASE CA-SCB Scotty Creek Bog, Ver. 2-5, AmeriFlux AMP, (Dataset).,
 1113 <https://doi.org/https://doi.org/10.17190/AMF/1498754>, 2021.

1114 Tarnocai, C., Canadell, J., Schuur, E., Kuhry, P., Mazhitova, G., and Zimov, S.: Soil organic carbon pools in the northern circumpolar
 1115 permafrost region, *Global biogeochemical cycles*, 23, 2009.

1116 Treat, C. C., Jones, M. C., Alder, J., and Frolking, S.: Hydrologic controls on peat permafrost and carbon processes: New insights from past
 1117 and future modeling, *Frontiers in Environmental Science*, 10, 892 925, 2022.

1118 Turetsky, M. R., Kane, E. S., Harden, J. W., Ottmar, R. D., Manies, K. L., Hoy, E., and Kasischke, E. S.: Recent acceleration of biomass
 1119 burning and carbon losses in Alaskan forests and peatlands, *Nature Geoscience*, 4, 27–31, 2011.

1120 Turetsky, M. R., Abbott, B. W., Jones, M. C., Anthony, K. W., Olefeldt, D., Schuur, E. A., Grosse, G., Kuhry, P., Hugelius, G., Koven, C.,
 1121 Lawrence, D. M., Carolyn, G., Sannel, A. B. K., and McGuire, A. D.: Carbon release through abrupt permafrost thaw, *Nature Geoscience*,
 1122 13, 138–143, 2020.

1123 Ueyama, M., Iwata, H., and Harazono, Y.: AmeriFlux BASE US-Uaf University of Alaska, Fairbanks, Ver. 11-5, AmeriFlux AMP, (Dataset).,
 1124 <https://doi.org/https://doi.org/10.17190/AMF/1480322>, 2023a.

1125 Ueyama, M., Iwata, H., and Harazono, Y.: AmeriFlux BASE US-Rpf Poker Flat Research Range: Succession from fire scar to deciduous
 1126 forest, Ver. 9-5, AmeriFlux AMP, (Dataset)., <https://doi.org/https://doi.org/10.17190/AMF/1579540>, 2023b.

1127 Valkenborg, B., De Lannoy, G. J. M., Gruber, A., Miralles, D. G., Köhler, P., Frankenberg, C., Desai, A. R., Humphreys, E., Klatt, J., Lohila,
 1128 A., Nilsson, M. B., Tuittila, E.-S., and Bechtold, M.: Drought and Waterlogging Stress Regimes in Northern Peatlands Detected Through
 1129 Satellite Retrieved Solar-Induced Chlorophyll Fluorescence, *Geophysical Research Letters*, 50, e2023GL105 205, 2023.

1130 Virkkala, A., Aalto, J., Rogers, B. M., Tagesson, T., Treat, C. C., Natali, S. M., Watts, J. D., Potter, S., Lehtonen, A., Mauritz, M., Schuur, E.
 1131 A. G., Kochendorfer, J., Zona, D., Oechel, W., Kobayashi, K., Humphreys, E., Goeckede, M., Iwata, H., Lafleur, P. M., Euskirchen, E. S.,
 1132 Bokhorst, S., Marushchak, M., Martikainen, P. J., Elberling, B., Voigt, C., Biasi, C., Sonnentag, O., Parmentier, F. W., Ueyama, M., Celis,
 1133 G., St.Louis, V. L., Emmerton, C. A., Peichl, M., Chi, J., Järveoja, J., Nilsson, M. B., Oberbauer, S. F., Torn, M. S., Park, S., Dolman, H.,
 1134 Mammarella, I., Chae, N., Poyatos, R., López-Blanco, E., Christensen, T. R., Kwon, M. J., Sachs, T., Holl, D., and Luoto, M.: Statistical
 1135 upscaling of ecosystem CO₂ fluxes across the terrestrial tundra and boreal domain: Regional patterns and uncertainties, *Global Change*
 1136 *Biology*, 27, 4040–4059, 2021.

1137 Virkkala, A., Rogers, B. M., Watts, J. D., Arndt, K. A., Potter, S., Wargowsky, I., Schuur, E. A. G., See, C., Mauritz, M., Boike, J., Bret-
 1138 Harte, S. M., Burke, E. J., Burrell, A., Chae, N., Chatterjee, A., Chevallier, F., Christensen, T. R., Commene, R., Dolman, H., Elberling,
 1139 B., Emmerton, C. A., Euskirchen, E. S., Feng, L., Goeckede, M., Grelle, A., Helbig, M., Holl, D., Järveoja, J., Kobayashi, H., Kutzbach,
 1140 L., Liu, J., Liujkx, I., López-Blanco, E., Lunneberg, K., Mammarella, I., Marushchak, M. E., Mastepanov, M., Matsuura, Y., Maximov, T.,
 1141 Merbold, L., Meyer, G., Nilsson, M. B., Niwa, Y., Oechel, W., Park, S., Parmentier, F. W., Peichl, M., Peters, W., Petrov, R., Quinton, W.,

1142 Rödenbeck, C., Sachs, T., Schulze, C., Sonnentag, O., St.Louis, V., Tuittila, E., Ueyama, M., Varlagin, A., Zona, D., and Natali, S. M.: An
 1143 increasing Arctic-Boreal CO₂ sink despite strong regional sources, *bioRxiv*, <https://doi.org/10.1101/2024.02.09.579581>, 2024.

1144 Wang, F., Chen, J. M., Gonsamo, A., Zhou, B., Cao, F., and Yi, Q.: A two-leaf rectangular hyperbolic model for estimat-
 1145 ing GPP across vegetation types and climate conditions, *Journal of Geophysical Research: Biogeosciences*, 119, 1385–1398,
 1146 <https://doi.org/https://doi.org/10.1002/2013JG002596>, 2014a.

1147 Wang, H., Prentice, I. C., and Davis, T. W.: Biophysical constraints on gross primary production by the terrestrial biosphere, *Biogeosciences*,
 1148 11, 5987–6001, 2014b.

1149 Wang, H., Prentice, I. C., Keenan, T. F., Davis, T. W., Wright, I. J., Cornwell, W. K., Evans, B. J., and Peng, C.: Towards a universal model
 1150 for carbon dioxide uptake by plants, *Nature plants*, 3, 734–741, 2017.

1151 Wania, R., Ross, I., and Prentice, I. C.: Integrating peatlands and permafrost into a dynamic global vegetation model: 1. Evaluation and
 1152 sensitivity of physical land surface processes, *Global Biogeochemical Cycles*, 23, 2009.

1153 Webb, E. E., Schuur, E. A., Natali, S. M., Oken, K. L., Bracho, R., Krapek, J. P., Risk, D., and Nickerson, N. R.: Increased wintertime CO₂
 1154 loss as a result of sustained tundra warming, *Journal of Geophysical Research: Biogeosciences*, 121, 249–265, 2016.

1155 Xiao, X., Jin, C., and Dong, J.: Gross primary production of terrestrial vegetation, in: *Biophysical applications of satellite remote sensing*,
 1156 pp. 127–148, Springer, 2013.

1157 Xu, B., Park, T., Yan, K., Chen, C., Zeng, Y., Song, W., Yin, G., Li, J., Liu, Q., Knyazikhin, Y., and Myneni, R. B.: Analysis of global
 1158 LAI/FPAR products from VIIRS and MODIS sensors for spatio-temporal consistency and uncertainty from 2012–2016, *Forests*, 9, 73,
 1159 2018.

1160 Yi, Y., Kimball, J. S., Watts, J. D., Natali, S. M., Zona, D., Liu, J., Ueyama, M., Kobayashi, H., Oechel, W., and Miller, C. E.: Investigating
 1161 the sensitivity of soil heterotrophic respiration to recent snow cover changes in Alaska using a satellite-based permafrost carbon model,
 1162 *Biogeosciences*, 17, 5861–5882, 2020.

1163 Zhang, M., Fu, L., Ma, D., Wang, X., and Liu, A.: Effects of Microtopography on Soil Microbial Community Structure and Abundance in
 1164 Permafrost Peatlands, *Microorganisms*, 12, 867, 2024.

1165 Zona, D., Gioli, B., Commane, R., Lindaas, J., Wofsy, S. C., Miller, C. E., Dinardo, S. J., Dengel, S., Sweeney, C., Karion, A., Chang, R.
 1166 Y. W., Henderson, J. M., Murphy, P. C., Goodrich, J. P., Moreaux, V., Liljedahl, A., Watts, J. D., Kimball, J. S., Lipson, D. A., and Oechel,
 1167 W. C.: Cold season emissions dominate the Arctic tundra methane budget, *Proceedings of the National Academy of Sciences*, 113, 40–45,
 1168 2016.

1169 Zona, D., Lafleur, P. M., Hufkens, K., Gioli, B., Bailey, B., Burba, G., Euskirchen, E. S., Watts, J. D., Arndt, K. A., Farina, M., Kimball,
 1170 J. S., Heimann, M., Göckede, M., Pallandt, M., Christensen, T. R., Mastepanov, M., López-Blanco, E., Dolman, A. J., Commane, R.,
 1171 Miller, C. E., Hashemi, J., Kutzbach, L., Holl, D., Boike, J., Wille, C., Sachs, T., Kalhori, A., Humphreys, E. R., Sonnentag, O., Meyer, G.,
 1172 Gosselin, G. H., Marsh, P., and C., O. W.: Pan-Arctic soil moisture control on tundra carbon sequestration and plant productivity, *Global
 1173 Change Biology*, 29, 1267–1281, 2023.