https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

ELM-TAM: a structure-based, function-oriented land model embracing fine-root system complexity

Bin Wang[1], M. Luke McCormack [2], Daniel M. Ricciuto[1], Xiaojuan Yang[1], Min Xu[1], and Forrest M. Hoffman[1]

1 Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830 USA

2 Center for Tree Science, The Morton Arboretum, Lisle, Illinois 60532 USA

Correspondence: Bin Wang (wangb@ornl.gov)

Abstract

Land models typically represent the entire fine-root system as a single homogeneous pool, collapsing structural and functional complexity into one set of parameters. This simplification weakens the belowground feedbacks that balance source-driven carbon dynamics. Here, we integrate the TAM (Transport and Absorptive fine roots with Mycorrhizal fungi) framework into the E3SM Land Model (ELM), replacing the single fine-root pool with three pools — transport roots, absorptive roots, and mycorrhizal fungi — defined by their roles in resource conduction, nutrient acquisition, and symbiotic exchange, each governed by pool-specific C:N ratios, longevity, allocation coefficients, and litter chemistry (21 new parameters). Sobol sensitivity analysis at 13 FLUXNET sites spanning 13 natural PFTs in ELM reveals a two-strategy parameterization dichotomy: deciduous and grass plant functional types (PFTs) concentrate variance in 2–3 allocation parameters, while evergreen PFTs distribute sensitivity across 5–6 parameters spanning chemistry, longevity, and allocation. Surrogate-assisted Bayesian calibration

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

produces emergent fine-root properties that diverge from ELM defaults, with system-level nitrogen demand increasing at 10 of 14 PFTs (up to 4.3×) and absorptive-root longevity spanning 0.5–3.5 yr. A three-tier validation framework—monthly evaluation, temporal out-of-sample, and spatial out-of-sample—demonstrates that these changes yield transferable improvements: improved seasonal amplitude fidelity (mean GPP amplitude ratio 0.99 vs. 1.38 for the baseline), near-elimination of ER bias at a deciduous site (MBE from +515 to +33 gC m^{-2} yr^{-1}), and cross-continental parameter transfer for C4 grasses (81 % RMSE reduction). Degradation at evergreen sites where the baseline overestimates productivity identifies a scope boundary where a belowground-only improvement cannot remedy upstream source-side bias. The function-oriented design provides natural coupling points for future extensions, including mechanistic root–water interactions and depth-resolved root profiles. ELM-TAM demonstrates that resolving belowground structural heterogeneity strengthens sink-based controls on the carbon cycle without overparameterization, establishing a foundation for these extensions and for global benchmarking in Earth System land models.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

1. Introduction

Land models that predict terrestrial dynamics within the Earth System still carry substantial uncertainty. Achieving robust, reliable, and realistic predictions is a critical need, as land projections inform proactive environmental, energy, and socio-economic policy (Schimel and Miller, 2023; Rockström et al., 2021). Theory-driven model development that synthesizes physiochemical, biological, and ecological understanding into process representations has been ongoing since the 1970s and will continue with the complementary support from data-driven approaches. A fundamental challenge throughout this effort has been accurately specifying the external and internal controls of ecosystem dynamics to predict emergent carbon cycling. Model benchmarking shows performance improvements over time, yet current models still present large uncertainties and low confidence in these formulations (Hoffman et al. 2025). While land use change and ecological disturbances pose challenges for representing exogenous controls (Randerson et al., 2025; Liu et al., 2025), one fundamental endogenous control that remains poorly formulated is the balance between source-driven and sink-driven regulation of vegetation carbon turnover. Because this source–sink balance constitutes an ecological basis of ecosystem carbon dynamics, improving its representation is a high priority.

The conceptual importance of balancing source- and sink-driven controls has long been recognized, and increasing empirical evidence reveals fine-root systems as playing a central role in this balance. Over half a century ago, Mooney (1972) framed belowground allocation as a first-order control on the whole-plant carbon balance of gain, use, and loss. Roughly half of GPP is allocated belowground in many ecosystems, and the fine roots and mycorrhizal fungi that receive this carbon regulate nutrient and water uptake, respiration, and litter inputs to soil (Weigelt et al., 2021). At the ecosystem scale, the source–sink imbalance manifests as a widely observed

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

decoupling between photosynthesis and growth. CO_2 enrichment does not simply translate into proportional biomass growth (Norby et al., 2004); increases in gross primary production (GPP) often fail to yield commensurate increases in net primary production (NPP), and NPP itself can decouple from long-term carbon storage under global change (Körner et al., 2007; Leuzinger et al., 2013). Such decoupling is the signature of belowground sink limitation on the carbon cycle. Yet empirical advances over the past two decades have established that these fine-root systems are far more structurally complex. Different root orders display distinct anatomy, morphology, chemistry, longevity, and mycorrhizal colonization (Pregitzer et al., 2002; Guo et al., 2008; McCormack et al., 2015, 2017). This variation defines a clear functional differentiation: lower-order absorptive roots, in close association with mycorrhizal fungi, specialize in nutrient acquisition and turn over rapidly, while higher-order transport roots persist for years as structural conduits.

Despite this empirical progress, the structural representation of roots in land models has remained essentially static for five decades. Aboveground model development has progressed through multiple generations of increasing complexity: from big-leaf radiation schemes to two-stream canopy models, multilayer canopy representations (Bonan et al., 2021), and most recently vegetation demographic models (e.g., Moorcroft et al., 2001; Fisher et al., 2018) that resolve within-stand competition and across-stand heterogeneity arising from disturbances based on individual-based forest gap models (Shugart et al. 1984; Shugart et al., 2018; Zhang et al. 2026). Each generation introduced new structural degrees of freedom to better capture controls on photosynthesis and light interception. Belowground structural representation, by contrast, has scarcely moved since its origin. Nearly half a century ago, Shugart et al. (1977) introduced a single lumped compartment of "small roots"—the very tissues we now call fine roots—into one of the

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

earliest forest ecosystem models, and that same undifferentiated pool still underpins the belowground structure of today's land models. This asymmetry persists even in models that have substantially improved belowground process representation. Cost–benefit optimization frameworks (e.g., the FUN model; Fisher et al., 2010), dynamic nitrogen cycle models (Thornton et al., 2007; Zaehle and Friend, 2010), and root–soil hydraulic schemes (Dukes et al., 2026) have all improved individual processes including nutrient uptake, water acquisition, and respiration, but all still operate on a single, undifferentiated fine-root pool (Wang et al., 2023). The structural problem thus persists regardless of the process sophistication layered on top.

This structural misspecification has measurable consequences for model predictions. When models treat fine roots as a single pool with uniform chemistry, longevity, and respiration rates, they conflate tissues that in reality turn over on timescales ranging from months (absorptive roots) to nearly a decade (transport roots). The resulting homogeneous pool cannot correctly partition carbon among litter pools of different decomposability, cannot generate pool-specific nutrient demands, and cannot represent the differential respiratory costs of functionally distinct tissues. Model structure constrains performance by governing parameter interactions; consequently, structure and parameters are inseparable sources of uncertainty (Beven, 2006; Gupta et al., 2012; Foster et al., 2026). A practical signature of this misspecification is that current models predict near-constant carbon use efficiency under rising CO_2 (i.e., relative changes in NPP closely tracking relative changes in GPP) against rich empirical evidence of a more complex response (Friend et al., 2019; Graven et al., 2024). The resulting overestimation of CO_2 fertilization is not merely a tuning problem; it reflects a structural inability to represent the sink-side feedbacks that modulate how photosynthate is cycled belowground (Fatichi et al., 2019; Kou-Giesbrecht et al., 2025).

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

This structural gap has persisted despite decades of empirical progress for both observational and conceptual reasons. On the observational side, belowground processes have historically been far more difficult to measure than aboveground ones, and the hierarchical structure of fine-root systems was not systematically characterized until the 2000s (Guo et al. 2008; Weigelt et al., 2021). On the conceptual side, models are not animations of physical reality (Levins 1966). Incorporating the full anatomical complexity of branching root orders and mycorrhizal associations into an ecosystem model, though epistemically worthwhile (Edmonds, 2000), is not tractable at the scale of Earth system prediction. What is needed is a principled structural compression, that is, a representation that captures the essential functional differentiation of fine-root systems without replicating their full anatomical complexity.

The TAM framework provides such a compression. We proposed TAM to aggregate fine-root systems and their mycorrhizal associations into three functional pools: transport roots (T), absorptive roots (A), and mycorrhizal fungi (M) (Wang et al., 2023). Each pool is defined by its dominant function (structural support and resource conduction; nutrient acquisition; symbiotic nutrient exchange), and each carries distinct properties for C:N stoichiometry, longevity, allocation, and litter chemistry. Such a three-pool structure intends to capture the essential functional differentiation documented empirically including the contrast between persistent transport tissues and ephemeral absorptive tissues, the distinct chemical signature of mycorrhizal biomass, while remaining tractable for global-scale application. This aggregation follows a long tradition of structure-based, function-oriented lumping in ecology (Raunkiaer, 1934; Root, 1967; Grime, 1974; Smith et al. 1997). Nair (2023) characterized this three-pool representation as a step toward a more holistic view of belowground structure in vegetation models. Wang et al. (2023) demonstrated that TAM improved site-level simulations of ecosystem carbon fluxes in a stand-

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

alone ecosystem model. However, the framework has not yet been implemented in a land model capable of representing all major vegetation types and of coupling with other Earth system components.

Here we present ELM-TAM, the E3SM Land Model with the TAM framework fully integrated. We chose ELM for its representative structure and its role in energy and climate research (Leung et al. 2020), although the structural simplification that ELM-TAM addresses—a single homogeneous fine-root pool—is shared by all major land models. The current development establishes the TAM structure and its integration into the carbon–nutrient cycle; the function-oriented design provides natural coupling points for extensions such as mechanistic root–water interactions and depth-resolved root profiles in later phases. We evaluate ELM-TAM through three stages. First, we conduct a Sobol global sensitivity analysis across FLUXNET sites spanning all natural plant functional types (PFTs) in ELM to characterize how parameter importance varies across vegetation types. Second, we perform Bayesian calibration at these sites to estimate PFT-specific TAM parameters and examine the emergent fine-root system properties they imply. Third, we apply a three-tier validation framework of increasing independence (monthly evaluation at calibration sites, temporal out-of-sample testing, and spatial out-of-sample testing at new sites) to assess whether the added structural complexity yields transferable predictive improvements. Together, these analyses are designed to establish whether a principled structural decomposition of fine-root pools can serve as a foundation for improving sink-based controls in Earth system predictions.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

2. Model Description

2.1 Overview of ELM-TAM

ELM-TAM decomposes ELM's single fine-root pool into three functionally distinct TAM pools within the existing prognostic C–N–P framework (**Fig. 1**). The model inherits the spatial tiling and biogeochemical structure of ELM without modification to the model grid or soil discretization (e.g., Burrows et al. 2020). Each grid cell is subdivided into topographic units, land units, columns, and plant functional type (PFT) patches; the TAM development focuses on the natural vegetation land unit, in which a single soil column is shared by multiple PFT patches in fractional coverage. Globally, 14 PFTs (11 woody + 3 grass types) represent the natural vegetation of the terrestrial biosphere. Each PFT tracks 68 prognostic C/N/P state variables—24 for carbon, 22 each for nitrogen and phosphorus. Five organs (leaf, live stem, dead stem, live coarse root, dead coarse root) carry the standard active/storage/transfer triplet (15 pools per element), while fine root splits into three TAM pools sharing a single storage and transfer pool (5 pools per element), giving 20 organ-level pools per element. All three elements carry a temporary labile pool; carbon adds an excess maintenance-respiration pool and growth-respiration storage and transfer pools; nitrogen and phosphorus additionally carry a re-translocation pool. The soil column is discretized into layers with vertically resolved carbon, nutrient, and water pools. The T, A, and M states are tracked as column-integrated quantities per PFT patch, even though the root profile used for nutrient uptake and litter distribution is vertically distributed (Sect. 2.4). Root activity is constrained by bedrock depth and, in permafrost regions, by the active layer depth. By resolving belowground heterogeneity at the level of functionally distinct tissue pools, ELM-TAM introduces an explicit mechanism through which belowground sinks feed back on whole-ecosystem carbon dynamics, which is the structural basis for stronger sink-based control on carbon cycling.

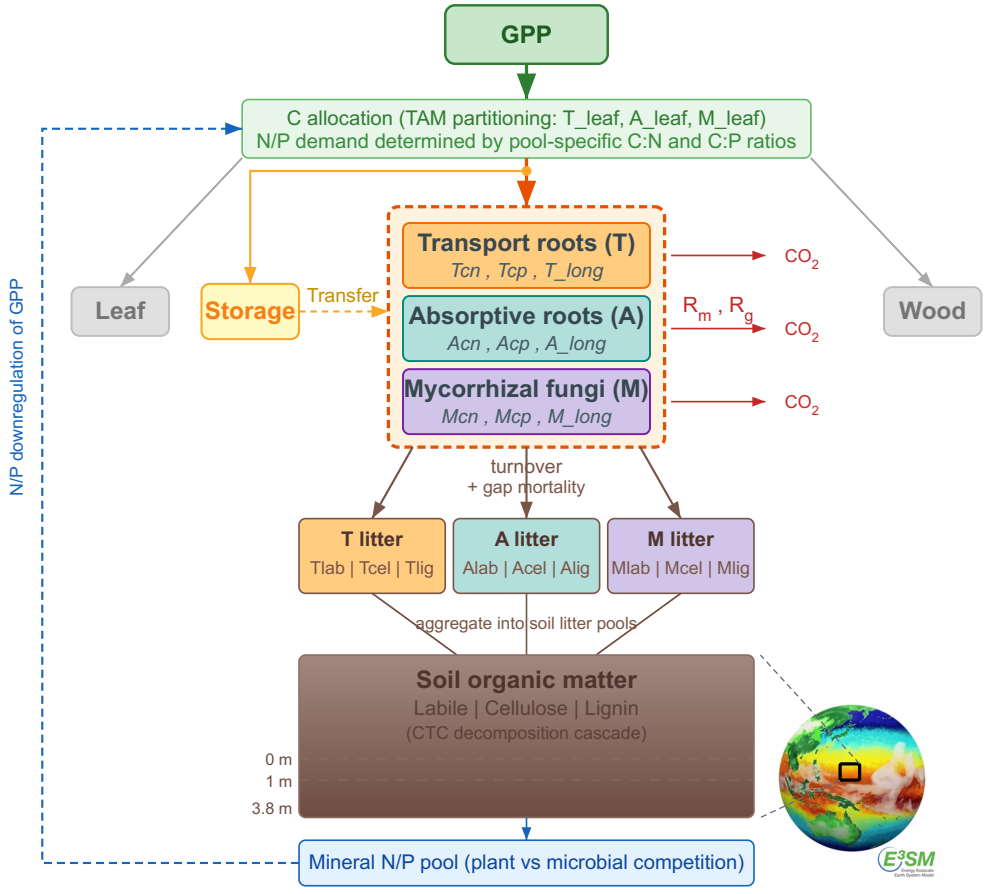
https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

The scope of the TAM modification reflects a deliberate parsimony–fidelity balance: we resolve the belowground structural heterogeneity that most directly governs carbon allocation, nutrient demand, and litter quality, while retaining standard ELM formulations for processes not yet coupled to the TAM structure. Photosynthesis uses the soil moisture stress factor rather than a mechanistic root-hydraulic approach, so the TAM pools do not feed back on transpiration or stomatal conductance in the current implementation. Nutrient re-translocation operates on leaves and live wood but not on fine roots in standard ELM, and the TAM pools inherit this treatment; all nitrogen and phosphorus in senescing TAM tissue enters the litter pathway. Soil organic matter decomposition follows the converging trophic cascade (CTC) framework with vertically resolved pools and relative-demand nutrient competition. Fire dynamics and land cover change are likewise unmodified. This design establishes the TAM structure and isolates the structural effect of fine-root pool differentiation on carbon and nutrient cycling, providing a clean baseline against which future extensions, e.g., mechanistic root–water coupling and depth-resolved TAM profiles, can be evaluated incrementally.

Concretely, the TAM modification introduces 21 new parameters governing the properties of the three fine-root pools (**Table 1**). These parameters fall into five functional categories including C:N and C:P ratios, carbon pool sizes, longevity, litter chemistry fractions, and root–leaf allocation scaling, and collectively determine how photosynthetic carbon is partitioned belowground, how quickly each pool turns over, and what litter chemistry enters the soil. The implementation modified 41 Fortran modules in the E3SM codebase; the complete source code is available at https://github.com/bioatmosphere/E3SM/tree/bwang. The subsections below describe the three principal process categories altered by the TAM structure: growth and allocation (Sect.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

204 2.2), respiration (Sect. 2.3), and turnover and decomposition (Sect. 2.4). Throughout, we use
205 subscript $i \in \{T, A, M\}$ to index the three TAM pools.
206



207

208 **Figure 1. Conceptual schematic of the ELM-TAM structure**. The single fine-root pool in
209 standard ELM is replaced by three structurally and functionally distinct pools—transport roots (T),
210 absorptive roots (A), and mycorrhizal fungi (M)—each with pool-specific C:N and C:P ratios,
211 longevity, allocation coefficients, and litter chemistry. Photosynthetic carbon is allocated to TAM
212 pools through direct allocation and indirect transfer via a shared storage pool; storage release is

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

phenology-dependent: continuous (evergreen), onset pulse (seasonal deciduous), and adaptive (stress-deciduous). TAM litter fluxes are split by biochemical composition (labile, cellulose, lignin), vertically distributed, and merged into the existing soil organic matter decomposition cascade.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Table 1. Definitions, units, and prior ranges of the 21 TAM parameters introduced in ELM-TAM. Seventeen parameters are free, which are constrained by observations and experiments (Wang et al. 2023) and sampled from uniform priors during Sobol sensitivity analysis and MCMC calibration. Four are directly constrained, among which three lignin litter fractions determined algebraically (lignin = 1 − labile − cellulose) and the mycorrhizal root–leaf ratio is derived from the total fine-root allocation once transport and absorptive allocations are set).

#	Symbol	Code name	Description	Unit	Prior range	Status
Carbon-to-nitrogen ratios						
1	Tcn	*froottcn*	Transport fine-root C:N ratio	g C g^{-1} N	20–184	Free
2	Acn	*frootacn*	Absorptive fine-root C:N ratio	g C g^{-1} N	11–119	Free
3	Mcn	*frootmcn*	Mycorrhizal C:N ratio	g C g^{-1} N	7–25	Free
Carbon-to-phosphorus ratios						
4	Tcp	*froottcp*	Transport fine-root C:P ratio	g C g^{-1} P	375–1125	Free
5	Acp	*frootacp*	Absorptive fine-root C:P ratio	g C g^{-1} P	250–750	Free
6	Mcp	*frootmcp*	Mycorrhizal C:P ratio	g C g^{-1} P	10–1600	Free
Root–leaf allocation scaling						
7	T_leaf	*froott_leaf*	Transport fine-root to leaf allocation ratio	–	0.05–0.40	Free
8	A_leaf	*froota_leaf*	Absorptive fine-root to leaf allocation ratio	–	0.20–0.60	Free
9	M_leaf	*frootm_leaf*	Mycorrhizal to leaf allocation ratio (derived)	–	—	Constrained
Longevity (turnover time)						
10	T_long	*froott_long*	Transport fine-root longevity	yr	3–10	Free
11	A_long	*froota_long*	Absorptive fine-root longevity	yr	0.5–4	Free
12	M_long	*frootm_long*	Mycorrhizal longevity	yr	0.13–1	Free
Litter chemistry — labile fractions						
13	Tlab	*frt_flab*	Transport fine-root labile litter fraction	–	0.10–0.35	Free
14	Alab	*fra_flab*	Absorptive fine-root labile litter fraction	–	0.10–0.35	Free
15	Mlab	*frm_flab*	Mycorrhizal labile litter fraction	–	0.10–0.35	Free
Litter chemistry — cellulose fractions						
16	Tcel	*frt_fcel*	Transport fine-root cellulose litter fraction	–	0.35–0.65	Free

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

17	Acel	*fra_fcel*	Absorptive fine-root cellulose litter fraction	–	0.35–0.65	Free
18	Mcel	*frm_fcel*	Mycorrhizal cellulose litter fraction	–	0.35–0.65	Free
Litter chemistry — lignin fractions (derived)						
19	Tlig	*frt_flig*	Transport fine-root lignin fraction	–	—	Constrained
20	Alig	*fra_flig*	Absorptive fine-root lignin fraction	–	—	Constrained
21	Mlig	*frm_flig*	Mycorrhizal lignin fraction	–	—	Constrained

223

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

2.2 TAM growth and allocation

Carbon allocation is the primary mechanism through which the TAM structure exerts sink-based control: by partitioning photosynthate among three pools with distinct stoichiometry, the model generates pool-specific nutrient demands that feed back on growth. TAM growth proceeds through two complementary pathways—direct allocation of current photosynthate and indirect transfer from a shared storage pool—both governed by PFT-specific allometric coefficients and stoichiometric ratios.

2.2.1 Carbon availability and allocation allometry

As in standard ELM, ELM-TAM first determines the carbon available for new growth by subtracting maintenance respiration and storage replenishment demands from gross primary production (GPP):

$$C_{\text{avail}}(p) = \text{GPP}(p) - R_{\text{maint}}(p)$$

where p denotes the PFT patch. The available carbon is then allocated to all tissue types—leaves, TAM pools, stems, and coarse roots—according to PFT-specific allometric ratios. In standard ELM, a single parameter a_1 (*froot_leaf*) governs the ratio of new fine-root carbon to new leaf carbon. ELM-TAM replaces this single parameter with three allocation parameters, one for each TAM pool:

$$a_{1,i} = a_1 \cdot f_{r,i}, \quad i \in \{T, A, M\}$$

where $f_{r,i}$ is the fractional partitioning of total fine-root allocation to pool i, determined by the root–leaf scaling parameters (*froott_leaf, froota_leaf, frootm_leaf*). The partitioning fractions satisfy:

$$f_{r,T} + f_{r,A} + f_{r,M} \leq 1$$

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

When the sum equals 1, all allocated carbon enters TAM pools; when the sum is less than 1, the residual implicitly represents root exudation.

The TAM partitioning and pool-specific C:N and C:P ratios jointly determine the PFT-specific allocation allometry. For woody PFTs (PFTs 1–11), the carbon, nitrogen, and phosphorus allometric ratios, that is, expressing the total new growth per unit new leaf growth, are:

$$C_{\text{allom}} = (1 + g_1)[1 + a_1 + a_3(1 + a_2)]$$

$$N_{\text{allom}} = \frac{1}{\text{CN}_{\text{leaf}}} + a_1 \sum_{i=1}^{3} \frac{f_{r,i}}{\text{CN}_{fr,i}} + \frac{a_3 a_4 (1 + a_2)}{\text{CN}_{\text{livewood}}} + \frac{a_3 (1 - a_4)(1 + a_2)}{\text{CN}_{\text{deadwood}}}$$

$$P_{\text{allom}} = \frac{1}{\text{CP}_{\text{leaf}}} + a_1 \sum_{i=1}^{3} \frac{f_{r,i}}{\text{CP}_{fr,i}} + \frac{a_3 a_4 (1 + a_2)}{\text{CP}_{\text{livewood}}} + \frac{a_3 (1 - a_4)(1 + a_2)}{\text{CP}_{\text{deadwood}}}$$

where a_1 through a_4 are the standard ELM allometric parameters (a_1: fine root to leaf ratio; a_2: coarse root to stem ratio; a_3: stem to leaf ratio; a_4: live wood to total wood ratio), g_1 is the growth respiration fraction, and $\text{CN}_{fr,i}$ and $\text{CP}_{fr,i}$ are the C:N and C:P ratios of TAM pool i (froottcn, frootacn, frootmcn and their phosphorus analogues). For non-woody PFTs (PFTs 12–14), the same expressions apply without the woody tissue terms (i.e., $a_2 = a_3 = a_4 = 0$):

$$C_{\text{allom}} = (1 + g_1)(1 + a_1)$$

$$N_{\text{allom}} = \frac{1}{\text{CN}_{\text{leaf}}} + a_1 \sum_{i=1}^{3} \frac{f_{r,i}}{\text{CN}_{fr,i}}$$

These equations reveal the key structural effect of the TAM framework on allocation: nitrogen and phosphorus demands of fine-root growth are now controlled by three distinct C:N (and C:P) ratios rather than a single ratio. Because absorptive roots and mycorrhizal fungi typically have lower C:N ratios than transport roots, consistent with the empirical pattern of decreasing C:N with decreasing root order (Pregitzer et al., 2002; McCormack et al., 2015), the TAM structure generally

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

increases the nitrogen demand per unit carbon allocated belowground. This is the mechanism through which fine-root pool differentiation strengthens the sink-based nutrient constraint on growth. Given the allometric ratios, the total nitrogen and phosphorus demands for new growth are:

$$N_{\text{demand}}(p) = C_{\text{avail}}(p) \cdot \frac{N_{\text{allom}}}{C_{\text{allom}}}$$

$$P_{\text{demand}}(p) = C_{\text{avail}}(p) \cdot \frac{P_{\text{allom}}}{C_{\text{allom}}}$$

2.2.2 Nutrient competition and downregulation

The elevated nutrient demands generated by pool-specific stoichiometry enter the existing ELM nutrient competition framework without modification to the competition algorithm itself. All PFTs sharing a soil column compete as a unified plant community against microbial decomposers, nitrifiers, and denitrifiers for available soil mineral nitrogen (sequentially: NH_4^+ first, then NO_3^- for unmet demand) and solution phosphorus. Combined plant demand is distributed vertically using availability-weighted nutrient uptake profiles. After competition is resolved independently for nitrogen and phosphorus, Liebig's Law of the Minimum is applied: whichever nutrient is more limiting controls the downregulation of GPP and, consequently, actual carbon allocation.

2.2.3 Direct allocation to TAM pools

Once nutrient limitation has downregulated GPP, the actual carbon allocation to each TAM pool is:

$$C_{\text{pool}\to i}(p) = C_{\text{nlc}}(p) \cdot a_1 \cdot f_{r,i} \cdot (1 - f_{\text{cur}})$$

where C_{nlc} is the nutrient-limited new leaf carbon, f_{cur} is the fraction allocated to current-year growth (versus storage), and $i \in \{T, A, M\}$. The corresponding nitrogen allocations are:

$$N_{\text{pool}\to i}(p) = C_{\text{nlc}}(p) \cdot a_1 \cdot \frac{f_{r,i}}{\text{CN}_{fr,i}} \cdot (1 - f_{\text{cur}})$$

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

and analogously for phosphorus with $CP_{fr,i}$. The allocation to the shared TAM storage pool aggregates the contributions from all three pools:

$$C_{\text{pool}\to\text{TAM,stor}}(p) = C_{\text{nlc}}(p) \cdot a_1 \cdot f_{\text{cur}}$$

$$N_{\text{pool}\to\text{TAM,stor}}(p) = C_{\text{nlc}}(p) \cdot a_1 \cdot \left(\sum_{i=1}^{3} \frac{f_{r,i}}{\text{CN}_{fr,i}}\right) \cdot f_{\text{cur}}$$

2.2.4 Indirect transfer from storage

Not all belowground growth comes from current photosynthate. TAM growth from stored carbon follows PFT-dependent phenological mechanisms through a single shared transfer pool: carbon moves from storage to a transfer pool, and from the transfer pool to the three displayed TAM pools proportional to their scaling relationships with leaf (*froott_leaf, froota_leaf, frootm_leaf*). The phenological strategy of each PFT determines the timing and rate of this transfer.

For evergreen PFTs, a constant background transfer rate operates year-round. A fixed fraction of the storage pool (0.5 yr^{-1}) is moved to the transfer pool, converted to a per-second rate by dividing by the number of seconds per year:

$$F_{\text{stor}\to\text{xfer}}(p) = \frac{0.5}{\tau_{\text{yr}}} \cdot C_{\text{TAM,stor}}(p)$$

where $\tau_{\text{yr}} = 365.25 \times 86400$ s. All mass in the transfer pool is then moved to displayed growth at each time step, partitioned evenly among the three TAM pools.

For seasonal deciduous PFTs (PFTs 3, 7, 8), TAM growth from storage occurs exclusively during a dedicated onset period (n_{onset} = 30 days). The transfer rate ramps linearly during this period: the fraction of current storage moved into the transfer pool increases at each time step until the final step, when all remaining storage is transferred. During onset, the rate coefficient at each time step is:

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

$$r_{\text{xfer}} = \frac{2 \cdot \Delta t}{\tau_{\text{onset}}^2}$$

where τ_{onset} is the remaining time in the onset period. On the final time step, all residual storage is flushed:

$$F_{\text{stor}\rightarrow\text{xfer}}(p) = \frac{1}{\Delta t} \cdot C_{\text{TAM,stor}}(p)$$

For stress-deciduous PFTs (PFTs 6, 10, 13, 14—grasses and tropical drought-deciduous trees), the phenological behavior depends on climate seasonality. In strongly seasonal climates, these PFTs behave like seasonal deciduous plants with discrete onset–offset cycles that may occur multiple times per year. In weakly seasonal environments (e.g., aseasonal tropics), they converge toward evergreen behavior: a long growing season factor (*lgsf*, ranging from 0 to 1 as the continuous active period exceeds one year) modulates a background transfer rate:

$$F_{\text{bgtr}}(p) = \frac{1}{\tau_{\text{yr}}} \cdot \text{lgsf}(p) \cdot C_{\text{TAM,stor}}(p)$$

This mechanism ensures that stress-deciduous plants with persistent canopies in the absence of a suitable cold or drought trigger transition to quasi-evergreen root maintenance.

2.3 TAM respiration

Respiration is the second pathway through which the TAM structure alters carbon cycling: because the three pools differ in nitrogen concentration, they generate distinct maintenance respiration costs even under a shared rate formulation.

2.3.1 Maintenance respiration

Maintenance respiration follows the standard ELM formulation, in which the respiration rate of each live tissue pool is proportional to its nitrogen content and modulated by a temperature response function:

$$R_{\text{maint},i}(p) = N_i(p) \cdot r_{\text{base}} \cdot f(T)$$

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

where $N_i(p)$ is the nitrogen content of TAM pool i, r_{base} is a base respiration rate per unit nitrogen (identical for all TAM pools and unchanged from ELM), and $f(T)$ is a Q_{10}-based temperature response. The TAM framework does not introduce pool-specific respiration rates or temperature sensitivities. Differentiation in maintenance respiration among the three pools arises entirely from their distinct nitrogen contents, which are governed by pool-specific C:N ratios. Because absorptive roots and mycorrhizal fungi have lower C:N ratios (higher nitrogen concentrations) than transport roots, they incur proportionally higher maintenance respiration costs per unit carbon, a parsimonious representation of the metabolic cost of maintaining nutrient-acquisition tissues.

2.3.2 Growth respiration

Growth respiration associated with TAM pool construction is calculated as a fixed fraction of carbon allocation fluxes:

$$R_{\text{growth},i}(p) = g_{\text{perc}} \cdot C_{\text{alloc},i}(p)$$

where $g_{\text{perc}} = 0.3$ is the growth respiration coefficient (ratio of growth respiration carbon to new growth carbon), unchanged from ELM. For TAM growth from the storage pool, the timing of growth respiration is governed by g_{now}, a PFT-specific parameter (default 0.5) that partitions growth respiration between the time of initial allocation to storage and the time of display growth:

$$R_{\text{growth,stor}}(p) = g_{\text{perc}} \cdot g_{\text{now}} \cdot C_{\text{alloc,stor}}(p)$$

with the remaining fraction $(1 - g_{\text{now}})$ deferred to the onset period when storage is converted to displayed TAM biomass.

2.4 TAM turnover and decomposition

Turnover is the third and final pathway through which the TAM structure reshapes carbon cycling: pool-specific longevity and litter chemistry determine not only how fast carbon leaves each pool but what biochemical form it takes when entering the soil.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

2.4.1 Endogenous turnover

TAM turnover to litter is controlled by pool-specific longevity parameters (*froott_long, froota_long, frootm_long*), with PFT-dependent phenological mechanisms distinguishing evergreen, seasonal deciduous, and stress-deciduous behavior.

For evergreen PFTs, TAM litterfall occurs year-round at a constant rate inversely proportional to pool longevity:

$$F_{i\to\text{litter}}(p) = \frac{1}{\lambda_i} \cdot C_i(p)$$

where λ_i is the longevity of TAM pool i (in seconds) and $C_i(p)$ is the current carbon content. The empirical basis for distinguishing longevity among the three pools derives from field observations of a dramatic increase in survivorship with root diameter even within narrow ranges ($\leq$ 0.5 mm) and from direct measurements of fungal tissue lifespan (e.g., Reid et al., 1993; McCormack et al., 2010; Allen and Kitajima, 2013). The longevity ordering ($\lambda_T > \lambda_A > \lambda_M$) reflects the empirical pattern that thicker, higher-order transport roots persist longer than ephemeral absorptive tips and their fungal partners.

For seasonal deciduous PFTs, TAM litterfall occurs during a discrete offset period (n_{off} = 15 days) triggered when day length drops below a critical threshold after the summer solstice. During the offset period, litter flux follows an accelerating ramp:

$$F_{i\to\text{litter}}(p,t) = F_{i\to\text{litter}}(p,t-\Delta t) + \frac{2\Delta t}{\tau_{\text{off}}^2} \cdot [C_i(p) - F_{i\to\text{litter}}(p,t-\Delta t) \cdot \tau_{\text{off}}]$$

where τ_{off} is the remaining time in the offset period. The rate starts near zero and accelerates, ensuring that on the final time step all remaining TAM carbon is transferred to litter:

$$F_{i\to\text{litter}}(p,t_{\text{final}}) = \frac{1}{\Delta t} \cdot C_i(p) + C_{\text{alloc}\to i}(p)$$

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

For stress-deciduous PFTs, the dual phenological behavior described in Sect. 2.2.4 extends to turnover. In strongly seasonal climates, offset can be triggered by sustained drought, freezing, or extreme short day length, producing complete shedding via the same accelerating ramp as seasonal deciduous PFTs, with potentially multiple onset–offset cycles per year. In weakly seasonal environments, stress-deciduous PFTs converge toward evergreen-like background turnover, with the long growing season factor modulating the litterfall rate.

2.4.2 Gap-phase mortality

In addition to endogenous turnover, all three TAM pools are subject to gap-phase mortality representing the aggregate effects of wind throw, insect attack, disease, drought stress, and age-related decline. The TAM framework uses the same PFT-specific annual mortality rate (m = 0.02 yr^{-1} by default) as other plant tissues, converted to a per-second rate and applied uniformly to each TAM pool and its corresponding storage and transfer pool:

$$F_{i,\mathrm{mort}}(p) = \frac{m}{\tau_{\mathrm{yr}}} \cdot C_i(p)$$

The mortality rate may also vary by soil order, reflecting the influence of soil fertility on stand turnover. Fire-induced and land-cover-change mortality are treated separately in ELM and are not modified by the TAM framework.

2.4.3 Litter chemistry and decomposition

Carbon, nitrogen, and phosphorus fluxes from TAM litter are split into three biochemical fractions — labile, cellulose, and lignin — using pool-specific litter chemistry parameters:

$$F_{i,\mathrm{labile}} = f_{i,\mathrm{lab}} \cdot F_{i\to\mathrm{litter}}$$

$$F_{i,\mathrm{cellulose}} = f_{i,\mathrm{cel}} \cdot F_{i\to\mathrm{litter}}$$

$$F_{i,\mathrm{lignin}} = f_{i,\mathrm{lig}} \cdot F_{i\to\mathrm{litter}}$$

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

where $f_{i,\mathrm{lab}}$, $f_{i,\mathrm{cel}}$, and $f_{i,\mathrm{lig}}$ are the labile, cellulose, and lignin fractions for TAM pool i (**Table 1**), with $f_{i,\mathrm{lig}} = 1 - f_{i,\mathrm{lab}} - f_{i,\mathrm{cel}}$. The same biochemical split applies to the nitrogen and phosphorus components of TAM litter, so the C:N and C:P ratios of each biochemical fraction at a given soil layer equal the weighted average of the contributing TAM pools at that layer.

TAM litter fluxes are vertically distributed within the soil profile following the PFT-specific cumulative root distribution of Jackson et al. (1996):

$$r_{\mathrm{root},j} = \beta^{100\,z_{j-1}} - \beta^{100\,z_j}$$

where β is a PFT-dependent root distribution parameter (dimensionless, typically 0.94–0.99; higher values correspond to deeper rooting profiles), z_{j-1} and z_j are the depths (m) of the upper and lower interfaces of soil layer j, and the factor of 100 converts depth from meters to centimeters to match the original parameterization. The profile is normalized over the hydrologically active soil layers above permafrost to ensure $\sum_{j=1}^{N_{\mathrm{active}}} r_{\mathrm{root},j} = 1$. In fully frozen columns, all fine root litter is assigned to the top soil layer. Because ELM-TAM does not maintain vertically resolved root biomass pools (Sect. 2.1), this profile serves exclusively to distribute litter inputs into the vertically resolved soil organic matter pools. The same profile is applied to all three TAM pools without differentiation, consistent with the assumption that the first-order control on litter distribution is the shared rooting depth of the PFT rather than pool-specific depth preferences.

The vertically distributed TAM litter fluxes from all three pools are aggregated and merged with leaf litter and coarse woody debris fragmentation into the three soil litter pools (labile, cellulose, and lignin) at each layer. These litter pools are subsequently decomposed following the converging trophic cascade (CTC) model with first-order kinetics subject to temperature, moisture, and stoichiometric constraints, as in standard ELM. The net effect of the TAM litter chemistry is that the same total fine-root carbon flux now enters decomposition with a composition that reflects

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

the relative contributions of three functionally distinct tissues — transport roots contributing more lignin-rich, slowly decomposable litter, while absorptive roots and mycorrhizal fungi contribute more labile, rapidly cycling substrates. This differentiated litter input is absent in standard ELM and represents a qualitatively new control on soil carbon dynamics.

3. Methods

The structural decomposition introduced above raises three questions: which of the 21 new TAM parameters most strongly control simulated carbon fluxes, what parameter values best reproduce observed fluxes, and whether the resulting parameterization transfers across temporal scales, time periods, and geographic locations. To address these questions, we designed a workflow comprising Sobol global sensitivity analysis (Sect. 3.2), Bayesian parameter calibration at 13 FLUXNET eddy covariance sites spanning all natural PFTs in ELM (Sect. 3.3), and a three-tier validation framework of increasing independence (Sect. 3.4). Because the thousands of forward evaluations required by sensitivity analysis and Markov chain Monte Carlo (MCMC) sampling are computationally prohibitive with full ELM-TAM runs, we trained neural-network surrogate models as efficient emulators (Sect. 3.1). These surrogates served as the forward model for sensitivity analysis and calibration, after which the calibrated parameter sets were applied in full ELM-TAM simulations for validation.

3.1 Simulation sites and experimental setup

We identified one representative FLUXNET eddy covariance site for each of the 14 natural PFTs in ELM, selecting for data availability, record length, and representativeness of the assigned PFT (**Fig. 2**; Table A1). Of these, 13 sites entered the surrogate-based sensitivity analysis and MCMC calibration workflow. PFT 9 (broadleaf evergreen shrub) was excluded because PFT 9 has

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

the smallest spatial coverage among the 14 natural PFTs in ELM and representative flux tower data are limited; its TAM parameter values were instead directly prescribed from published fine-root trait observations and expert judgment (**Table 2**). The dominant PFT at each site's grid cell was confirmed against the ELM surface dataset, and site marker colors in Fig. 2 indicate the dominant mycorrhizal association—ectomycorrhizal (EM), arbuscular (AM), or dual (EM/AM)—which governs the relative functional importance of the mycorrhizal pool in the TAM framework.

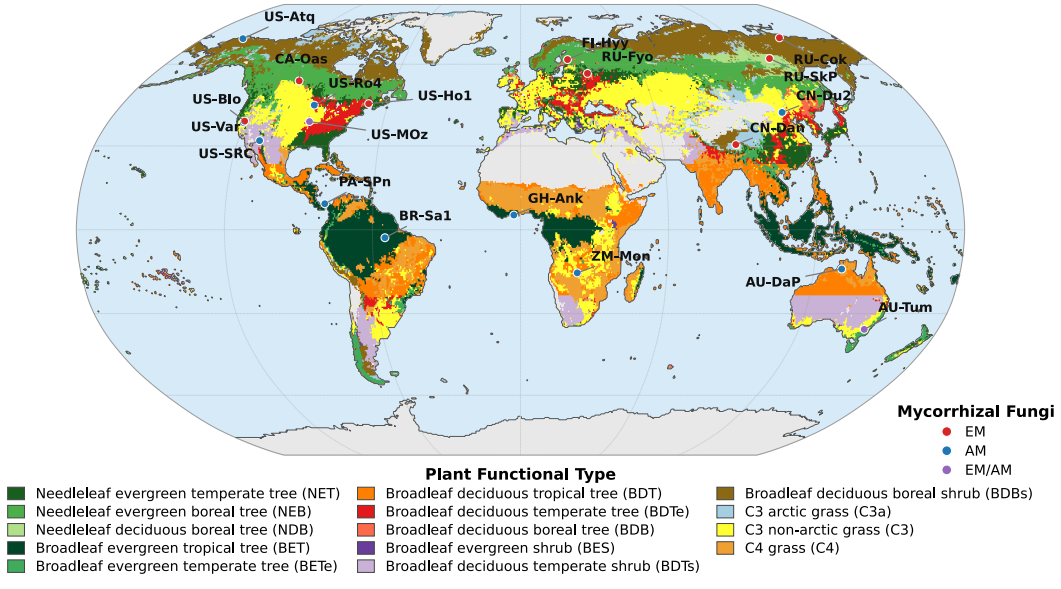


Figure 2. Global distribution of dominant plant functional types in the ELM surface dataset (0.5° resolution) with FLUXNET eddy covariance sites used for sensitivity analysis, calibration, and validation. Colors indicate the PFT with the highest fractional coverage at each grid cell; gray shading marks land areas where no single natural PFT exceeds 5 % coverage. Thirteen sites (one per PFT, excluding PFT 9; see text) entered the surrogate-based sensitivity analysis and MCMC calibration; 7 additional sites were used for spatial out-of-sample validation

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

(Tier 3; Sect. 3.4.3). Marker colors denote the dominant mycorrhizal association: ectomycorrhizal (EM, red), arbuscular mycorrhizal (AM, blue), or dual association (EM/AM, purple).

In addition to the 13 calibration sites, **Fig. 2** shows the seven independent validation sites used for spatial out-of-sample testing (Tier 3; Sect. 3.4.3), which span seven PFTs across four continents. The geographic separation between calibration and validation sites of the same PFT—for example, US-Ho1 (Maine, USA) versus US-Blo (California, USA) for PFT 1, or US-Atq (Alaska, USA) versus CN-Dan (Tibetan Plateau, China) for PFT 12—illustrates the stringency of the spatial transferability test.

All simulations followed a common three-stage protocol: (i) accelerated spinup (compset ICB1850CNRDCTCBC) to equilibrate carbon and nitrogen pools, (ii) final spinup (compset ICB1850CNPRDCTCBC) with coupled carbon–nitrogen–phosphorus cycling, and (iii) transient simulation (compset ICB20TRCNPRDCTCBC) driven by meteorological forcing. Each site was run as a single-point simulation in which domain, surface (except for reconfigured PFT composition), and land use data were extracted from global datasets at 0.5° resolution using a KDTree nearest-neighbor lookup, with coordinates subsequently overwritten to the exact site latitude and longitude. Meteorological forcing for the calibration period was derived from the GSWP3 reanalysis dataset, which covers 1901–2014 and serves as the default atmospheric driver in ELM; extended validation runs (Tier 2; Sect. 3.4.2) used the CRUJRA2024 reanalysis dataset covering 1901–2023.

At each of the 13 calibration sites, we ran the baseline ELM with default parameters and an ensemble of ELM-TAM simulations in which the 17 free TAM parameters (**Table 1**) were perturbed across their prior ranges. The four constrained parameters (*frt_flig, fra_flig, frm_flig*,

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

and *frootm_leaf*) were excluded from the ensemble because each pool's three litter chemistry fractions (labile + cellulose + lignin) must sum to 1.0, so the lignin (and one cellulose) fraction is determined by the other two. Ensemble outputs of annual GPP and ER were used to train site-specific surrogate models as described below.

3.1.1 Surrogate model construction

Because surrogate models underpin both sensitivity analysis and calibration, their fidelity is a prerequisite for the analyses that follow. Surrogate models were built by training multi-layer perceptron (MLP) neural networks (scikit-learn MLPRegressor; Pedregosa et al., 2011) to emulate ELM-TAM annual GPP and ER as functions of the 17 free TAM parameters. A separate surrogate was trained for each target variable at each site. Ensemble (n=1000) simulation outputs were filtered to remove invalid samples (if exist for reasons of fill values or physically implausible values), then split into training (80 %) and testing (20 %) sets with a fixed random seed for reproducibility. Both input parameters and target variables were independently standardized to zero mean and unit variance prior to training.

Optimal network configurations were identified through an exhaustive grid search with 5-fold cross-validation, evaluating 36 hyperparameter combinations. The search space included hidden-layer architectures of one layer (50 or 100 neurons) and two layers (100 and 50 neurons); tanh and ReLU activation functions; Adam and L-BFGS solvers; L2 regularization strengths of 0.0001, 0.01, and 0.05; and constant or adaptive learning-rate schedules. Training was capped at 2000 iterations with early stopping enabled: 20 % of the training data was reserved as an internal validation set, and training halted after 10 consecutive iterations without improvement in validation loss. Surrogate performance was evaluated using the coefficient of determination (R^2) between predicted and simulated values on both training and held-out test sets. Across the 13

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

calibration sites, all surrogates achieved $R^2 > 0.97$ for both GPP and ER on held-out data, confirming that the emulators faithfully reproduce ELM-TAM behavior across the full parameter space (Fig. A1). At inference time, new parameter vectors were standardized using the stored training scalers, passed through the best-performing model, and inverse-transformed to recover physical units.

3.2 Sensitivity analysis

With the surrogate models, the first analytical step was to identify which of the 17 free TAM parameters most strongly control simulated GPP and ER, and whether these controls vary across PFTs. We performed global sensitivity analysis using the Sobol variance-decomposition method (Sobol', 2001) implemented in the SALib Python package (Herman and Usher, 2017), with surrogates serving as the forward model. The Sobol method decomposes model output variance into contributions from individual parameters (first-order indices, S1), pairwise parameter interactions (second-order indices, S2), and all higher-order interactions. The total-effect index (ST) for a given parameter captures its first-order effect plus all interaction terms in which it participates.

For each of the 13 calibration PFTs (PFT 9 excluded), we generated 10,000 quasi-random parameter combinations from the parameter prior distributions using Saltelli's extension of the Sobol sequence (Saltelli, 2002). Each combination was evaluated through the site-specific surrogate model, and Sobol indices were computed separately for GPP and ER. For each simulation year we computed a combined sensitivity index as the mean of the GPP and ER indices [$S_{combined,y} = (S_{GPP,y} + S_{ER,y}) / 2$], then averaged these combined indices across all years to obtain the final per-parameter ranking to reflect parameter influence on the net ecosystem carbon cycle. To enable cross-PFT comparison, we applied two normalization schemes: (i) max-

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

normalization, dividing each PFT's indices by its maximum value to reveal the relative importance ranking within each PFT, and (ii) absolute ST values to compare magnitudes across PFTs. We separated the top eight most influential parameters from the remaining nine using an empirical gap in the ST distribution at 0.15 (**Fig. A2**).

3.3 Parameter calibration

Guided by the sensitivity analysis, we calibrated the 17 free TAM parameters at each of the 13 calibration sites using Bayesian inference via an enhanced Metropolis–Hastings Markov chain Monte Carlo (MCMC) algorithm with adaptive proposal distributions (Haario et al., 2001). The surrogate models served as the forward model in place of full ELM-TAM runs, reducing the computational cost of likelihood evaluations by several orders of magnitude while preserving the parameter–response relationships learned from the ELM-TAM ensemble. For PFT 9 (broadleaf evergreen shrub), which was excluded from the surrogate-based workflow, TAM parameter values were directly prescribed from published fine-root trait data and expert judgment rather than MCMC optimization; the resulting values are included in **Table 2** alongside the 13 MCMC-calibrated PFTs.

Uniform prior distributions were assigned to each parameter based on empirical constraints compiled from field observations of fine-root traits across ecosystems (Wang et al., 2023; T**able 1**). These priors bound the physically and biologically plausible range for each parameter while remaining uninformative within those bounds. The likelihood function assumes independent Gaussian observational errors:

$$p(\mathbf{y} \mid \boldsymbol{\theta}) = \prod_{v=1}^{V} \prod_{t=1}^{T_v} \frac{1}{\sqrt{2\pi\sigma_{v,t}^2}} \exp\left(-\frac{(y_{v,t} - f_v(\boldsymbol{\theta}, t))^2}{2\sigma_{v,t}^2}\right)$$

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

where V is the number of observed variables (GPP and ER), T_v is the number of annual time points for variable v, $y_{v,t}$ is the observation at time t, $\sigma_{v,t}$ is the observational uncertainty, and $f_v(\boldsymbol{\theta}, t)$ is the surrogate-model prediction for parameter vector $\boldsymbol{\theta}$.

Observational data were obtained from the FLUXNET2015 dataset (Pastorello et al., 2020), which provides standardized eddy covariance measurements with comprehensive quality control and uncertainty quantification. We used the night-time partitioning method with constant u* threshold (NT-CUT) reference values for GPP and ER along with their 5th- and 95th-percentile uncertainty bounds, which reflect the spread of plausible values arising from different gap-filling strategies, friction velocity threshold selection, and flux partitioning approaches. We used annual flux observations for calibration because (i) the surrogate models were trained on annual model output, (ii) annual aggregation reduces temporal autocorrelation, better satisfying the independence assumption in the Gaussian likelihood, and (iii) annual carbon budgets are more robust to structural model errors in sub-annual dynamics.

The 90 % confidence interval (5th to 95th percentile) was converted to a standard deviation assuming a Gaussian error distribution:

$$\sigma_{v,t} = \frac{P_{95}(v,t) - P_5(v,t)}{3.29}$$

where the conversion factor 3.29 derives from the normal-distribution relationship $2 \times 1.645\sigma \approx 3.29\sigma$ for the 5th-to-95th percentile span. Only observations with quality-control flags exceeding 0.8 (indicating > 80 % measured rather than gap-filled data) were retained for calibration; invalid or low-quality measurements were excluded from the likelihood calculation.

Given the 17-dimensional parameter space and the potential for correlated parameters, we implemented a three-phase adaptive burn-in strategy to accelerate convergence. Phase 1 (30 % of burn-in) emphasizes aggressive exploration using enlarged proposal steps to locate high-

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

probability regions. Phase 2 (50 % of burn-in) implements adaptive Metropolis sampling in which the proposal covariance is iteratively updated based on the empirical covariance of recent chain samples, with scaling factors adjusted to maintain acceptance rates near the optimal 23 % for multivariate normal proposals (Roberts et al., 1997). Phase 3 (20 % of burn-in) applies conservative fine-tuning to stabilize the proposal distribution. Prior to burn-in, a preconditioning phase of 50 samples provides rapid initial estimates of the covariance structure using eigenvalue decomposition for numerical stability. To improve mixing efficiency for correlated parameters, we implemented block sampling in which strongly correlated parameter groups, identified from empirical correlation analysis of the chain, are updated jointly. Global sensitivity analysis results informed the proposal distribution: smaller step sizes were applied to high-sensitivity parameters (requiring precise estimation) and larger steps to low-sensitivity parameters (enabling faster exploration). The algorithm employs eigenvalue decomposition for proposal generation to ensure numerical stability when the covariance matrix becomes ill-conditioned, with automatic detection and correction of conditioning issues. Convergence diagnostics include real-time monitoring of acceptance rates and effective sample size (ESS) calculated via FFT-based autocorrelation analysis. Multiple chains were initialized from overdispersed starting points to ensure adequate exploration of the posterior. For each of the 13 calibrated PFTs, we report the maximum a posteriori estimate for each parameter as the calibrated value used in subsequent validation runs (**Table 2**).

3.4 Validation

Calibration establishes that ELM-TAM can reproduce observed fluxes at the sites and years used for parameter estimation; the critical question is whether this skill transfers. We evaluated ELM-TAM predictive skill through a three-tier validation design that progressively increases the

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

degree of independence between calibration and evaluation data. Tier 1 tests whether the TAM structure produces realistic seasonal dynamics when only annual budgets were constrained, which is a necessary condition before interpreting performance at other sites or time periods. Tier 2 isolates the question of whether calibrated values are overfit to the training period by testing parameter robustness under novel climate forcing. Tier 3 provides the most stringent test: whether PFT-level parameters transfer to independent sites across diverse climates, continents, and ecosystems. Together, these tiers provide a systematic assessment of ELM-TAM transferability prior to comprehensive global benchmarking with ILAMB (Collier et al., 2018). **Figure 2** provides a spatial overview of all the sites entering the validation framework.

3.4.1 Tier 1: Monthly evaluation at calibration sites

To test whether ELM-TAM captures sub-annual dynamics not targeted during calibration, we evaluated model predictions against monthly eddy covariance observations at all 13 calibration sites over the same years used for calibration. Parameter calibration was performed against annual sums of GPP and ER at the 13 calibration sites. The annual objective function constrains the total carbon budget at each site but leaves the seasonal distribution of fluxes within each year entirely unconstrained. This tier is particularly informative for the TAM framework because the phenology-dependent allocation and turnover schemes (Sect. 2.2–2.3: year-round background transfer for evergreen PFTs versus discrete onset/offset periods for deciduous PFTs) produce distinct seasonal signatures that annual calibration does not directly constrain. Realistic seasonal dynamics under annual-only calibration would indicate that the TAM structure encodes mechanistically meaningful seasonal behavior rather than compensating errors that cancel at the annual scale. Monthly observations of GPP and ER were derived from the FLUXNET2015 dataset using the NT-CUT method, aggregated from half-hourly to monthly resolution, retaining only

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

months with quality-control flags exceeding 0.8. While mean seasonal cycles were analyzed at six representative sites spanning the three functional strategies (evergreen: FI-Hyy, AU-Tum; deciduous: PA-SPn, US-MOz; grass: US-Var, AU-DaP), a Taylor diagram was generated to summarize correlation and normalized variability across all 13 calibration sites.

3.4.2 Tier 2: Temporal out-of-sample validation

To test whether the optimized parameter sets generalize beyond the calibration period, we ran ELM-TAM with the posterior-mode parameter estimates at a subset of calibration sites for which extended eddy covariance observations are available and easily accessible from the AmeriFlux network: US-Ho1 (PFT 1, needleleaf evergreen temperate tree, 2015–2021), US-MOz (PFT 7, broadleaf deciduous temperate tree, 2015–2021), and US-Var (PFT 13, C3 non-Arctic grass, 2015–2021). These extended runs used the CRUJRA2024 atmospheric forcing dataset that covers 1901–2023. This change in meteorological driver introduces an additional degree of independence, as the model must reproduce observed fluxes under forcing data it was not calibrated against. For each site, we ran both the optimized ELM-TAM parameterization and the baseline ELM with default parameters, both driven by CRUJRA2024, to quantify the improvement attributable to the TAM structure and its calibrated parameters.

These three sites sample distinct sensitivity profiles across the PFT spectrum—an evergreen tree (US-Ho1), a deciduous tree (US-MOz), and a grass (US-Var)—providing a representative test of parameter robustness under novel forcing. We report annual time series of predicted and observed GPP and ER at these three sites (**Fig. 7**) to visualize how ELM-TAM tracks interannual variability, complementing the aggregate metrics in the validation scorecard (**Fig. 5**).

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

3.4.3 Tier 3: Spatial out-of-sample validation

We ran ELM-TAM with optimized parameters at seven sites not included in the calibration, spanning seven PFTs across four continents (**Fig. 2**; **Table A2**). The strongest test of model generality is whether PFT-level parameters calibrated at one site transfer to an independent site of the same PFT. The validation sites were selected to maximize ecological and climatic contrast with their corresponding calibration sites while sharing the same PFT classification: US-Blo (PFT 1, 2000–2006; parameters from US-Ho1), RU-Fyo (PFT 2, 1998–2014; from FI-Hyy), GH-Ank (PFT 4, 2011–2014; from BR-Sa1), ZM-Mon (PFT 6, 2000–2009; from PA-SPn), CN-Dan (PFT 12, 2004–2005; from US-Atq), CN-Du2 (PFT 13, 2006–2008; from US-Var), and US-Ro4 (PFT 14, 2014–2023; from AU-DaP).

These sites differ from their respective calibration sites in climate, species composition, and stand characteristics. For example, US-Blo (ponderosa pine, Mediterranean climate, California) contrasts with US-Ho1 (spruce–hemlock, humid continental, Maine) despite both being PFT 1; RU-Fyo (spruce–birch mixed forest, central European Russia) differs from FI-Hyy (Scots pine, southern Finland); and CN-Dan (alpine meadow at 4300 m on the Tibetan Plateau) presents a radically different environment from US-Atq (Arctic coastal plain tundra, Alaska) despite both being classified as C3 Arctic grass (Fig. 2). These contrasts test whether the calibrated parameters capture PFT-level functional strategies rather than site-specific conditions—the core premise underlying global application of ELM-TAM. For all spatial validation sites, we compared ELM-TAM predictions against eddy covariance observations from the FLUXNET2015 dataset or the AmeriFlux network.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

3.4.4 Performance metrics

Model performance was assessed using four complementary metrics applied to GPP and ER predictions. Root mean square error (RMSE; gC m^{-2} yr^{-1}) quantifies the magnitude of prediction errors:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_{\mathrm{obs},t} - y_{\mathrm{mod},t})^2}$$

Mean bias error (MBE; gC m^{-2} yr^{-1}) characterizes systematic over- or under-estimation:

$$\mathrm{MBE} = \frac{1}{n}\sum_{t=1}^{n}(y_{\mathrm{mod},t} - y_{\mathrm{obs},t})$$

For each metric, we report values for both the optimized ELM-TAM and baseline ELM to quantify the improvement attributable to the TAM framework. Performance was evaluated separately for GPP and ER to assess whether the model captures carbon uptake and respiratory fluxes with comparable skill. For the monthly evaluation (Tier 1), we additionally report the Pearson correlation coefficient (r) of monthly time series and the ratio of simulated to observed seasonal amplitude ($A_{\mathrm{mod}}/A_{\mathrm{obs}}$, where amplitude is the difference between maximum and minimum monthly flux) to separately characterize the model's ability to reproduce the phase and magnitude of the seasonal cycle. The primary summary metric used to compare ELM-TAM against baseline ELM across all sites and tiers is the relative RMSE change:

$$\Delta\mathrm{RMSE}\% = \frac{\mathrm{RMSE}_{\text{ELM-TAM}} - \mathrm{RMSE}_{\mathrm{ELM}}}{\mathrm{RMSE}_{\mathrm{ELM}}} \times 100$$

Negative values indicate that ELM-TAM reduces prediction error relative to the baseline. A validation scorecard offering an overview of ΔRMSE% across all sites and tiers is presented, in addition to detailed site-level results for each tier.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

4 Results

4.1 Sensitivity of ELM-TAM simulating GPP and ER

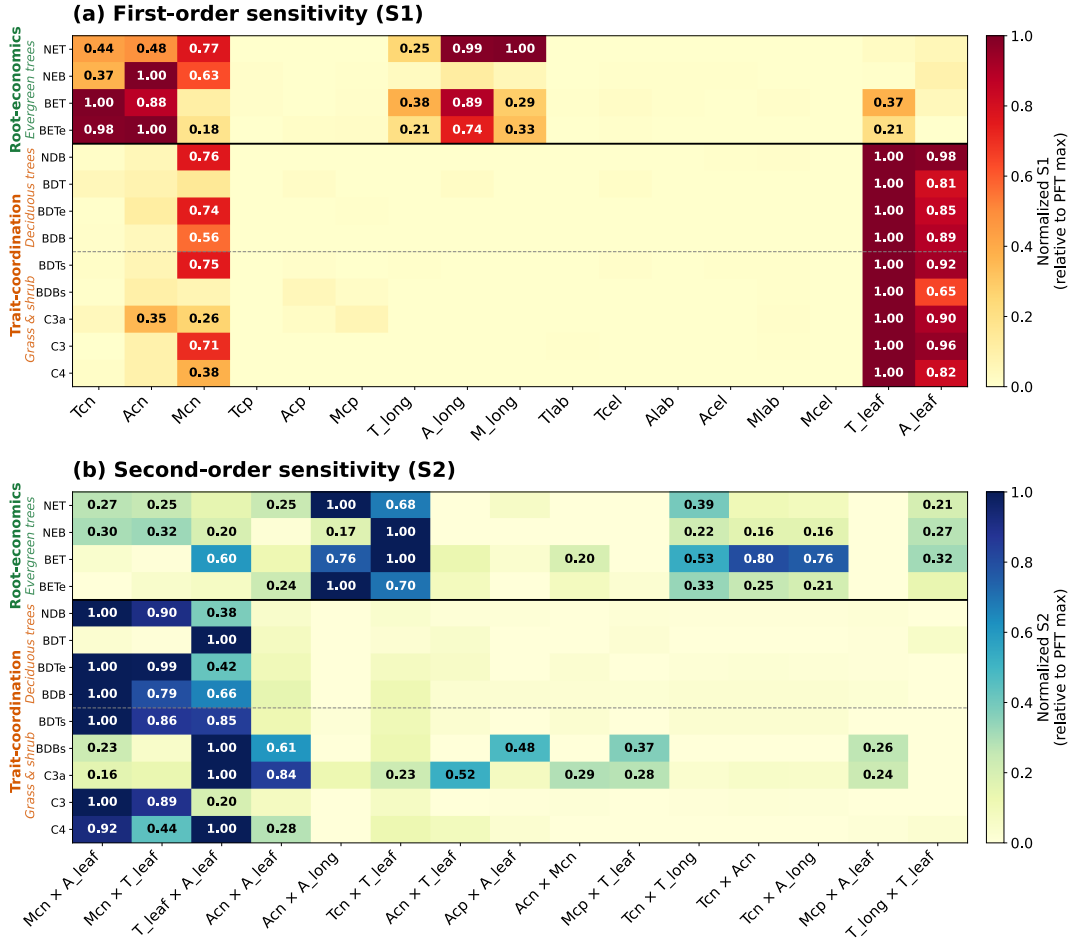


Fig. 3. Sobol sensitivity analysis reveals a two-strategy parameterization dichotomy across plant functional types. Cross-site sensitivity of 17 TAM root parameters for combined carbon fluxes (GPP + ER) across 13 PFTs. PFTs are grouped by parameterization strategy: root-economics (evergreen trees, PFTs 1, 2, 4, 5) and trait-coordination (deciduous trees and grass/shrub, PFTs 3, 6–8, 10–14). PFT 9 is excluded owing to its limited spatial coverage and flux tower data availability. **(a)** Normalized first-order sensitivity (S1): values within each PFT row are

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

scaled by that PFT's maximum S1 so that the darkest cell identifies the dominant parameter. **(b)** Normalized second-order sensitivity (S2) for the 15 strongest pairwise interactions (out of 136 possible pairs), scaled by each PFT's maximum S2. Values ≥ 0.15 are annotated. All indices are derived from Sobol variance decomposition of combined GPP and ER, averaged across multi-year simulations. Total-effect indices (ST) for the eight most influential parameters are shown in **Fig. A2**.

We performed Sobol global sensitivity analysis to quantify how 17 free TAM parameters of the 21 total (the remaining 4 are constrained by sum-to-one relationships) influence ecosystem carbon fluxes across 13 calibrated PFTs. The analysis decomposed output variance into first-order effects (S1), pairwise interactions (S2), and total effects (ST), revealing a clear two-strategy dichotomy in parameter control of carbon cycling (**Fig. 3**).

Evergreen tree PFTs (1, 2, 4, 5) distributed sensitivity across 3–6 parameters encompassing root chemistry, longevity, and allocation (**Fig. 3a**). These PFTs showed normalized S1 = 0.8–1.0 for multiple parameters, primarily C:N ratios (*frootacn, froottcn, frootmcn*) and root longevity (*froota_long, frootm_long*). For example, PFT 2 (boreal needleleaf evergreen, FI-Hyy) showed *frootacn* = 1.00, *frootmcn* = 0.63, and *froottcn* = 0.37, indicating that root chemical quality, rather than allocation, controls carbon cycling in this system. Tropical and temperate evergreen broadleaf forests (PFTs 4–5) showed similar multi-parameter patterns but with additional emphasis on absorptive-root longevity (*froota_long*: S1 = 0.74–0.89). The interaction analysis reinforced this pattern: PFT 4 (Amazon tropical forest, BR-Sa1) exhibited strong pairwise interactions between chemistry and longevity parameters (*froottcn* × *froota_long*: S2 = 0.80; *frootacn* × *froota_long*: S2 = 0.76), indicating that carbon flux sensitivity in tropical evergreen systems emerges from

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

trade-offs across the root economics spectrum rather than simple parameter additivity (**Fig. 3b**). The total-effect indices (ST) confirm this multi-parameter pattern: evergreen PFTs show ST ≥ 0.15 for five or more parameters spanning all three functional categories (**Fig. A2**). Because dominant parameters (C:N ratios and root longevities) span the acquisition–conservation trade-off (short-lived/cheap/N-rich tissue vs. long-lived/expensive/N-poor tissue), which is precisely the axis defining the root economics spectrum in trait databases (Reich, 2014), we refer to these PFTs as root-economics strategy.

In contrast, deciduous and grass PFTs (3, 6–8, 10–14; 69 % of the 13 analyzed systems) exhibited remarkably convergent sensitivity patterns dominated by just two root-to-leaf allocation parameters: *froott_leaf* and *froota_leaf* (normalized S1 = 1.00), with S1 < 0.3 for most other parameters, though *frootmcn* (mycorrhizal C:N) reaches moderate sensitivity (S1 = 0.3–0.8) at several deciduous and shrub PFTs (**Fig. 3a**). This convergence persisted across dramatic environmental and taxonomic gradients spanning 63° of latitude and encompassing both woody and herbaceous growth forms: tropical deciduous forests (PFT 6, PA-SPn), temperate deciduous forests (PFT 7, US-MOz), boreal deciduous stands (PFT 8, CA-Oas), Arctic grasses (PFT 12, US-Atq), Mediterranean grasslands (PFT 13, US-Var), and tropical savannas (PFT 14, AU-DaP). The interaction analysis confirmed that second-order effects in these PFTs were concentrated among the same allocation parameters, reinforcing rather than complicating the main-effect hierarchy (**Fig. 3b**). Because of the dominance by parameters *froott_leaf, froota_leaf*, and *frootm_leaf*, which are about the below- and aboveground allometric coupling (Poorter et al. 2012), we refer to these PFTs as trait-coordination strategy.

The total sensitivity indices (ST), integrating main effects, interactions, and higher-order terms, further delineate the two strategies (**Fig. A2**). The eight parameters with the highest ST in

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

at least one PFT—*froott_leaf, froota_leaf, froottcn, frootacn, frootmcn, froota_long, froott_long, and frootm_long*—are cleanly separated from the remaining nine by an empirical gap (the 8th-ranked parameter reaches ST = 0.155 while the 9th drops to 0.081). ST values for these top eight range from 0.12 to 0.70 across PFTs. In deciduous and grass PFTs the two allocation parameters dominate total sensitivity, whereas in evergreen PFTs the ST budget is spread across C:N ratios, longevity, and allocation—consistent with the first-order patterns in **Fig. 3a**. Litter chemistry parameters (*frt_flab, fra_fcel, frm_fcel*) contributed negligible sensitivity across all PFTs for the combined GPP + ER metric. Longevity parameters exhibited both high main-effect and high interaction sensitivity with C:N ratios exclusively in evergreen PFTs, providing a mechanistic basis for the observed strategy separation.

This two-strategy dichotomy, trait-coordination in deciduous and grass systems versus root-economics in evergreen systems, has direct implications for parameter estimation. It predicts that deciduous and grass PFTs, controlled by only 2–3 parameters, should be tractable targets for calibration, while evergreen PFTs, with sensitivity distributed across 5–6 interacting parameters, face greater equifinality risk. The calibration results below test this prediction.

4.2 Calibration results

We calibrated the 17 free TAM parameters independently for each of 13 PFTs using Bayesian MCMC optimization against annual GPP and ER observations at one FLUXNET site per PFT (**Table 1**). For PFT 9 (broadleaf evergreen shrub), TAM parameter values were prescribed from published fine-root trait data owing to its limited spatial coverage and flux tower data availability (**Fig.2**). The remaining 4 parameters (*frootm_leaf, frt_flig, fra_flig, frm_flig*) are determined by sum-to-one constraints. **Table 2** reports the full set of 21 parameter values,

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

including13 sets of MCMC posterior modes, 1 set of literature-prescribed values, and 4 derived for all 14 PFTs.

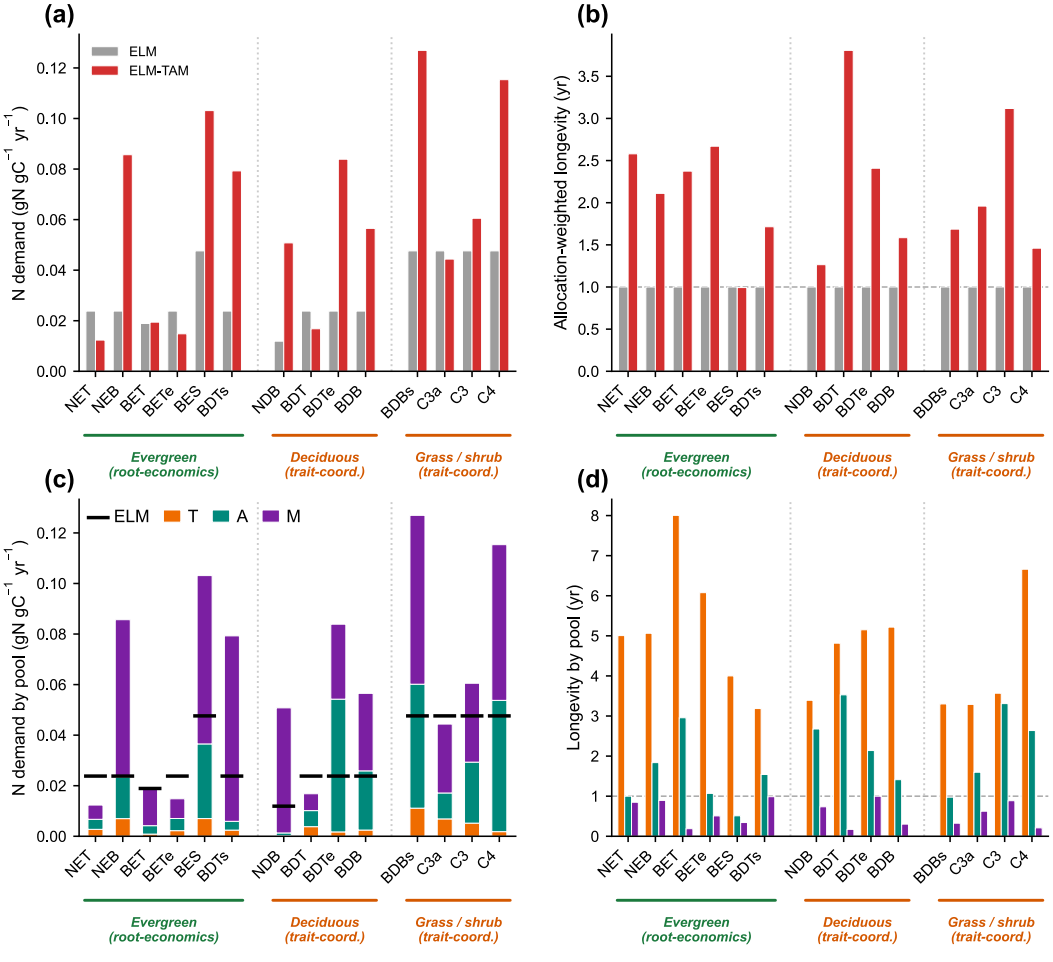
https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Table 2. TAM parameter values for 14 plant functional types: MCMC posterior modes for 13 calibrated PFTs; literature-prescribed values for PFT 9.

PFT	Code	*froott_leaf*	*froota_leaf*	*frootm_leaf*	*frootlcn*	*frootacn*	*frootmcn*	*froott_long*	*froota_long*	*frootm_long*	*frootlcp*	*frootacp*	*frootmcp*	*frt_flab*	*fra_flab*	*frm_flab*	*frt_fcel*	*fra_fcel*	*frm_fcel*	*frt_flig*	*fra_flig*	*frm_flig*
1	*NET*	0.40	0.51	0.09	145	129	16	5.01	1.00	0.85	901	658	1264	0.23	0.34	0.35	0.65	0.65	0.37	0.12	0.01	0.28
2	*NEB*	0.25	0.20	0.55	35	12	9	5.06	1.84	0.90	764	494	604	0.11	0.25	0.26	0.64	0.44	0.65	0.25	0.31	0.10
3	*NDB*	0.05	0.20	0.75	113	93	8	3.39	2.68	0.74	538	375	693	0.13	0.28	0.34	0.37	0.40	0.36	0.50	0.32	0.30
4	*BET*	0.12	0.45	0.43	112	107	22	8.01	2.96	0.19	1015	477	1588	0.26	0.28	0.24	0.61	0.64	0.39	0.13	0.08	0.37
5	*BETemp*	0.34	0.52	0.14	151	108	18	6.08	1.07	0.51	756	597	784	0.22	0.22	0.30	0.61	0.57	0.62	0.17	0.21	0.08
6	*BDTrop*	0.40	0.54	0.07	104	84	10	4.82	3.53	0.17	635	643	938	0.11	0.34	0.11	0.36	0.47	0.65	0.53	0.19	0.24
7	*BDTemp*	0.18	0.59	0.24	104	11	8	5.15	2.14	1.00	719	691	676	0.30	0.35	0.30	0.58	0.65	0.60	0.12	0.00	0.10
8	*BDB*	0.14	0.56	0.31	55	24	10	5.22	1.41	0.30	552	268	314	0.11	0.14	0.35	0.35	0.35	0.37	0.54	0.51	0.28
9	*BES*	0.16	0.44	0.40	45	30	12	4.00	0.51	0.34	600	400	600	0.30	0.28	0.32	0.41	0.55	0.44	0.30	0.17	0.24
10	*BDTemps*	0.28	0.20	0.52	115	58	7	3.19	1.54	0.99	478	289	304	0.29	0.22	0.10	0.36	0.37	0.39	0.36	0.41	0.51
11	*BDBS*	0.40	0.28	0.33	71	11	10	3.30	0.98	0.33	509	306	561	0.11	0.14	0.31	0.57	0.35	0.65	0.32	0.51	0.04
12	*C3Arc*	0.29	0.59	0.13	83	115	9	3.29	1.60	0.63	567	452	965	0.31	0.15	0.10	0.35	0.65	0.62	0.33	0.20	0.28
13	*C3*	0.35	0.53	0.12	134	44	7	3.57	3.31	0.89	578	330	206	0.31	0.34	0.31	0.38	0.38	0.50	0.31	0.28	0.20
14	*C4*	0.07	0.33	0.60	76	13	19	6.66	2.64	0.21	763	500	782	0.23	0.27	0.23	0.60	0.60	0.55	0.17	0.13	0.22

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

760 The calibrated parameters produce system-level fine-root properties that diverge markedly
761 from ELM's defaults, even though both models simulate the same ecosystem carbon fluxes. ELM
762 prescribes uniform fine-root properties for all PFTs: a single root pool with fixed longevity (1 yr),
763 a single C:N ratio, a single C:P ratio, and fixed litter chemistry fractions. ELM-TAM replaces this
764 single pool with three functionally distinct pools whose calibrated properties aggregate to PFT-
765 specific system-level values for nitrogen demand and effective turnover time (**Fig. 4**).
766



767

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Fig. 4. Emergent fine-root properties from calibration: ELM defaults versus ELM-TAM calibrated values across 14 PFTs, grouped by parameterization strategy. (a) System-level fine-root nitrogen demand, computed as the allocation-weighted sum of pool-specific N demands (gN gC^{-1} yr^{-1}). Gray bars: ELM default (single pool); red bars: ELM-TAM (three-pool aggregate). **(b)** Allocation-weighted fine-root longevity (yr). The dashed line marks ELM's uniform default of 1 yr. **(c)** ELM-TAM nitrogen demand decomposed by pool: transport roots (orange), absorptive roots (teal), and mycorrhizal fungi (purple). Black dashes indicate the ELM default for reference. **(d)** ELM-TAM longevity decomposed by pool. Dotted vertical lines separate the three PFT groups; colored brackets and labels follow the strategy grouping of Fig. 3.

Although ELM-TAM also introduces pool-specific C:P ratios (*froottcp, frootacp, frootmcp*; **Table 2**), the emergent-property analysis here focuses on nitrogen rather than phosphorus for three reasons. First, the sensitivity analysis showed that C:P parameters contributed negligible first-order sensitivity to the combined GPP + ER metric across all 13 PFTs, ranking consistently below the empirical ST gap at 0.15. Second, the calibration targeted carbon fluxes at sites where nitrogen is the dominant nutrient limitation; phosphorus limitation is most acute in highly weathered tropical soils, which are represented by only one PFT (PFT 4) in the current calibration set. Third, ELM's phosphorus cycle operates on longer timescales through mineral weathering and sorption dynamics that are less directly coupled to the annual carbon fluxes used as calibration targets (Zaehle and Friend, 2010). The C:P parameters are nonetheless retained because they govern litter P content entering the soil decomposition module and will become critical when ELM-TAM is evaluated against phosphorus-limited ecosystems or when P-cycle outputs are used as calibration targets in future work.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

ELM-TAM's calibrated N demand exceeds ELM's default at 10 of 14 PFTs (**Fig. 4a**); four PFTs show decreased N demand (PFTs 1, 5, 6, and 12), where the calibrated C:N ratios are higher than ELM's defaults. The increases are largest for evergreen PFTs; NEB (PFT 2) reaches 0.086 gN gC^{-1} yr^{-1}, roughly 3.6× the ELM default, reflecting the low C:N ratios of absorptive roots and mycorrhizal fungi in these systems (**Table 2**). Among the deciduous group, three of four PFTs show increased N demand (PFTs 3, 7, 8; up to 4.3×), driven primarily by the absorptive-root pool which accounts for the largest share of total N demand (**Fig. 4c**), while PFT 6 (tropical deciduous) shows a decrease (0.71×) owing to its high calibrated C:N ratios. Grass and shrub PFTs show intermediate increases, with the boreal deciduous shrub (PFT 11) displaying the highest N demand in this group (0.127 gN gC^{-1} yr^{-1}), followed closely by the C4 grass (PFT 14, 0.115 gN gC^{-1} yr^{-1}), both driven by high absorptive-root allocation. The pool decomposition (**Fig. 4c**) reveals that across all PFTs, the absorptive-root and mycorrhizal pools together contribute 77–100% of total system N demand despite comprising a smaller fraction of total root biomass, a pattern consistent with empirical observations of fine-root nutrient economics (McCormack et al., 2015; See et al., 2019).

The calibrated parameters also restructure fine-root turnover. ELM assigns a uniform 1-yr fine-root longevity to all PFTs (**Fig. 4b**). ELM-TAM's allocation-weighted longevity exceeds this default at 13 of 14 PFTs (PFT 9, the prescribed Mediterranean shrub at 0.99 yr, is the sole exception), with the largest increases in the deciduous group: BDT (PFT 6) reaches 3.8 yr and BDTe (PFT 7) reaches 2.4 yr. The pool decomposition (**Fig. 4d**) shows that this divergence arises from the long-lived transport root pool, whose calibrated longevity ranges from 3 to 8 yr across PFTs—substantially longer than the absorptive roots (0.5–3.5 yr) and mycorrhizal fungi (0.17–1.0 yr). Transport roots dominate the weighted average because they receive a substantial allocation

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

fraction (*froott_leaf* = 0.05–0.40; Table 2) while turning over slowly. Absorptive roots turn over faster than transport roots across all PFTs but span a broad range themselves, from 0.5 yr (PFT 9, broadleaf evergreen shrub) to 3.5 yr (PFT 6, tropical deciduous), consistent with minirhizotron observations showing order-of-magnitude variation in fine-root lifespan across biomes (McCormack et al., 2012; Iversen et al., 2017). Mycorrhizal fungi are the shortest-lived pool at every PFT (0.17–1.0 yr), consistent with the rapid turnover of external hyphae documented in isotope-tracer studies. Among the evergreen PFTs, allocation-weighted longevity ranges from 1.7 yr (PFT 10) to 2.7 yr (PFT 5), a narrower spread than in the deciduous group, reflecting the more balanced allocation across the three pools characteristic of the root-economics strategy.

The three-group ordering in **Fig. 4** reveals systematic contrasts that echo the sensitivity dichotomy. Within the evergreen group (root-economics strategy), the PFTs with the largest N demand increases (PFTs 2, 10; 3.3–3.6×) are those where calibrated C:N ratios diverge most from ELM's defaults, while others (PFTs 1, 5) show decreased N demand, reflecting the distributed, multi-parameter control identified in the sensitivity analysis. Several deciduous PFTs (trait-coordination strategy) show the largest longevity increases (PFT 6: 3.8 yr; PFT 7: 2.4 yr), consistent with their allocation-dominated sensitivity where *froott_leaf* , *froota_leaf* , and *froota_leaf* control the partitioning among long-lived transport root and short-lived absorptive root and mycorrhizal fungi pools. Grass and shrub PFTs, also trait-coordination, show a distinct pattern: moderate longevity increases with relatively high absorptive-root N demand, reflecting the rapid-turnover growth strategy characteristic of herbaceous systems.

These emergent differences in N demand and turnover constitute the mechanism through which ELM-TAM alters carbon cycling relative to ELM. Higher N demand drives stronger nutrient limitation feedbacks; differentiated turnover rates produce heterogeneous litter inputs with

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

pool-specific chemistry fractions (**Table 2**), which in turn alter soil decomposition dynamics and nutrient mineralization timing. The magnitude of these changes (up to 4× for N demand where increases occur, up to 3.8× for effective longevity) indicates that the TAM structural decomposition into three pools is not a minor perturbation but a fundamental restructuring of the belowground carbon and nutrient economy. Whether these calibration-derived changes translate into improved flux predictions under independent conditions is tested in the three-tier validation below.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

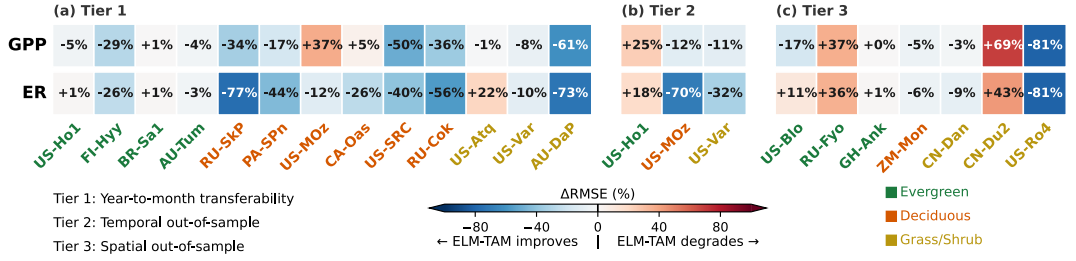


Fig. 5. Validation scorecard of ΔRMSE (%) for ELM-TAM relative to baseline ELM across three evaluation tiers. Each cell shows the percentage change in RMSE when replacing the baseline with ELM-TAM; negative values (blue) indicate improvement, positive values (red) indicate degradation. Site labels are colored by groups: evergreen/root-economics (green), deciduous/trait-coordination (orange), and grass and shrub (gold). **(a)** Tier 1—monthly RMSE at 13 calibration sites, computed from FLUXNET MM observations and monthly model output over the calibration period. **(b)** Tier 2—annual RMSE at 3 calibration sites during the post-calibration period (2015–2021). **(c)** Tier 3—annual RMSE at 7 independent validation sites not used in calibration.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

4.3 Validation results

The scorecard reveals a central pattern that cuts across all three tiers: RMSE reductions concentrate at grass and deciduous sites, while degradations concentrate at evergreen sites (**Fig. 5**). This functional-strategy dependence mirrors the sensitivity dichotomy and suggests that PFTs controlled by fewer allocation parameters (deciduous/grass) are more tractable calibration targets than PFTs with distributed parameter sensitivity (evergreen), which require fundamentally different parameterization approaches. The three tiers form a hierarchy of increasing independence that together bracket ELM-TAM's current credibility: emergent seasonal dynamics not targeted during calibration (Tier 1) establish mechanistic plausibility, temporal generalization at deciduous and grassland sites (Tier 2) demonstrates predictive skill, and the mixed spatial transferability results (Tier 3) define the next methodological frontier—transitioning from single-site to multi-site and even tighter trait-constrained calibration. The subsections below present the tier-specific evidence underlying these patterns.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

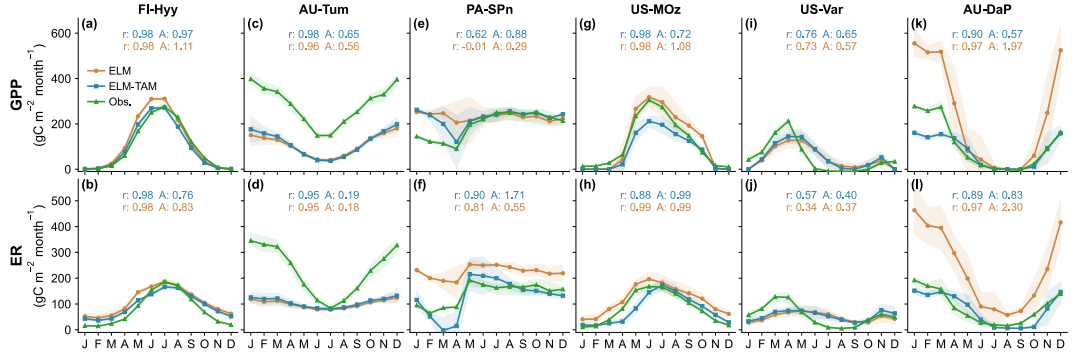


Fig. 6. Composite mean seasonal cycles at six representative calibration sites. Each column shows one site with GPP (top) and ER (bottom); two sites per functional strategy (evergreen: FI-Hyy, AU-Tum; deciduous: PA-SPn, US-MOz; grass: US-Var, AU-DaP). Solid lines show multi-year mean monthly fluxes (gC m^{-2} $month^{-1}$) for ELM (orange), ELM-TAM (blue), and FLUXNET observations (green triangles); shading indicates ±1 standard deviation across years. Pearson correlation (r) and amplitude ratio (A = [model max − min] / [obs max − min]) between each model's mean seasonal cycle and observations are reported. Observations are converted from gC m^{-2} d^{-1} (FLUXNET MM) to gC m^{-2} $month^{-1}$ by multiplying by days per month.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

4.3.1 Tier 1—Year-to-month transferability

Tier 1 tests whether parameters optimized against annual sums capture seasonal patterns that were not explicitly constrained by comparing monthly model predictions against FLUXNET eddy covariance observations (MM datasets) at all 13 calibration sites using correlation coefficient (r), peak month agreement, and seasonal amplitude ratio (**Fig. 6**) and then disaggregating by functional strategy to connect the results to the sensitivity dichotomy (**Fig. A3**). Overall, ELM-TAM reproduces monthly patterns not targeted during calibration, with systematic amplitude improvements across most sites—evidence that the differentiated root pools capture mechanistically meaningful seasonal processes rather than achieving annual agreement through compensating errors.

In detail, across the 13 evaluated sites, ELM-TAM and baseline ELM achieved comparable mean monthly correlations for GPP (r = 0.70 vs. 0.70) and ER (r = 0.84 vs. 0.86). Site-level differences are nonetheless substantial. At PA-SPn (tropical deciduous, PFT 6), ELM-TAM improves the seasonal correlation over the baseline for both GPP (r = 0.62 vs. −0.01) and ER (r = 0.90 vs. 0.81; Fig. 6), indicating that the TAM stress-deciduous phenology scheme captures the basic seasonal rhythm even where the absolute phase remains shifted. Peak month agreement reinforces this picture: ELM-TAM correctly identifies the peak GPP month at 7 of 13 sites and the peak ER month at 6 of 13, with the strongest agreement at high-latitude deciduous and boreal sites (US-Atq, US-Var, AU-DaP, and RU-Cok) and the weakest in the tropics. At BR-Sa1 (PFT 4) both models place peak GPP in July rather than the observed December, and at PA-SPn both produce peak GPP in January instead of October, which are phase errors that likely reflect limitations in prescribed phenology and moisture-driven seasonality in the forcing data rather than TAM-specific deficiencies.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

ELM-TAM also produces more realistic seasonal amplitudes than the baseline across the majority of sites. The mean GPP amplitude ratio is 0.99 for ELM-TAM versus 1.38 for the baseline; for ER, it is 1.05 versus 1.34. Individual sites illustrate the magnitude of this correction: at RU-SkP (deciduous boreal, PFT 3), ELM-TAM achieves near-perfect ER amplitude (A = 1.01) compared to the baseline's three-fold overestimate (A = 3.03), and at AU-DaP (C4 grass, PFT 14), ELM-TAM reduces the GPP amplitude ratio from 1.97 to 0.57 (**Fig. 6**).

Disaggregating by functional strategy clarifies where these improvements originate. The Taylor diagram (**Fig. A3**) confirms that ELM-TAM's primary monthly-scale improvement is in variability rather than correlation: ELM-TAM markers cluster closer to unit normalized standard deviation, while baseline markers scatter out to nσ = 2.0–2.6. Deciduous and grass PFTs show higher mean monthly correlations (GPP r = 0.77; ER r = 0.81) than evergreen PFTs (GPP r = 0.55; ER r = 0.89), with the largest individual ER correlation gain at US-Var (Δr = +0.23). Amplitude improvements also concentrate in the deciduous and grass group (mean GPP amplitude ratio 0.85 vs. baseline 1.40), while evergreen PFTs show little change as a group, though individual sites such as FI-Hyy achieve strong seasonal reproduction (GPP r = 0.98, amplitude ratio 0.97). This strategy dependence connects directly to the sensitivity analysis: deciduous PFTs, controlled by allocation parameters (*froott_leaf*, *froota_leaf*, and *frootm_leaf*), produce seasonal dynamics through discrete onset/offset periods that generate clear seasonal signals even under annual-only calibration, while evergreen PFTs governed by C:N ratios and longevity generate weaker seasonal contrasts.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

4.3.2 Tier 2—Temporal out-of-sample validation

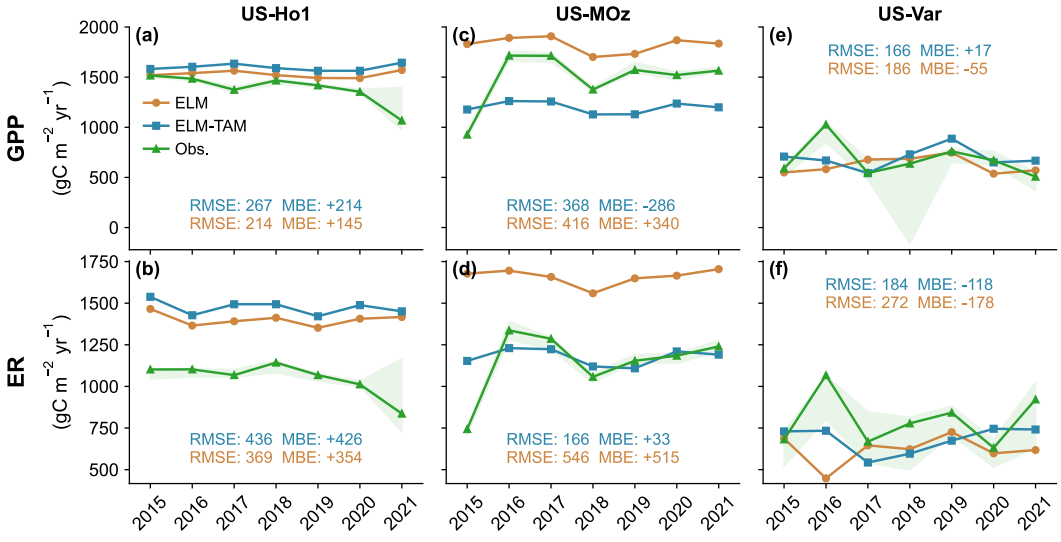


Fig. 7. Tier 2 temporal out-of-sample validation: annual GPP and ER at three calibration sites during 2015–2021. Each column shows one site with GPP (top) and ER (bottom): US-Ho1 (PFT 1, needleleaf evergreen temperate), US-MOz (PFT 7, broadleaf deciduous temperate), US-Var (PFT 13, C3 non-Arctic grass). Lines connect annual values (gC m^{-2} yr^{-1}) for ELM (orange), ELM-TAM (blue), and AmeriFlux observations (green triangles); green shading shows the 5th–95th percentile uncertainty range of the observations. RMSE and MBE (gC m^{-2} yr^{-1}) are reported for each model.

Tier 2 tests whether calibrated parameters generalize to years outside the calibration window. We evaluated ELM-TAM at three sites (US-Ho1, US-MOz, US-Var) for 2015–2021 using CRUJRA2024 forcing and AmeriFlux observations (**Fig. 7**). Beyond aggregate error metrics, the direction of MBE provides mechanistic insight into how the TAM framework alters carbon flow through the plant–soil system.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

At US-MOz, ELM-TAM substantially outperformed the baseline, most dramatically for ER (RMSE = 165 vs. 545 gC/m²/yr). The baseline overestimated both fluxes (GPP MBE = +340, ER MBE = +515 gC/m²/yr), consistent with known positive GPP bias in standard ELM at productive deciduous sites. ELM-TAM corrected the ER bias almost entirely (MBE = +33 gC/m²/yr) but reversed the sign of the GPP bias (MBE = −286 gC/m²/yr). This sign reversal has a clear mechanistic origin: the TAM allocation scheme redirects carbon from leaves to the three root pools via *froott_leaf* and *froota_leaf*—the dominant parameters for deciduous PFTs (Sect. 4.1)—correcting the baseline's GPP overestimate but overcorrecting to produce an underestimate of comparable magnitude. The near-elimination of ER bias, however, indicates that the TAM framework's partitioning of belowground carbon across pools with distinct turnover rates produces a more realistic respiratory flux than the baseline's single fine-root pool. The year-by-year trajectory (**Fig. 7**) confirms this: ELM-TAM's ER predictions track observations throughout 2015–2021, while the baseline consistently overshoots by 360–934 gC/m²/yr.

At US-Var, ELM-TAM improved both fluxes relative to the baseline (GPP RMSE = 166 vs. 186 gC/m²/yr; ER RMSE = 184 vs. 272 gC/m²/yr). The baseline underestimated both GPP (MBE = −55 gC/m²/yr) and ER (MBE = −178 gC/m²/yr) at this California annual grassland. The negative GPP bias reflects a structural limitation of standard ELM: its single fine-root pool with fixed turnover does not capture the rapid root proliferation that C3 grasses use to exploit ephemeral soil moisture pulses. ELM-TAM's separate absorptive root pool—with its own C:N ratio (*frootacn*) and fast turnover (*froota_long*)—represents this strategy more faithfully, nearly eliminating the GPP bias (MBE = +17 gC/m²/yr). The ER bias also improved (MBE = −118 vs. −178 gC/m²/yr), though a negative residual persists. This remaining underestimate likely reflects soil decomposition limitations outside the TAM scope: the Mediterranean wet-season rewetting of

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

dry soils triggers microbial respiration pulses (the Birch effect) that ELM's soil carbon module does not fully capture (Wang and Allison 2021). The year-by-year data (**Fig. 7**) support this interpretation, as the largest ER underestimates occur in high-rainfall years (e.g., 2016) when the Birch effect would be strongest.

At US-Ho1, ELM-TAM did not improve over the baseline during the validation period (GPP RMSE = 267 vs. 214 gC/m²/yr; ER RMSE = 436 vs. 369 gC/m²/yr). Both models overestimated fluxes throughout the record (Fig. 7), and ELM-TAM amplified the positive bias for both GPP (MBE = +214 vs. +145 gC/m²/yr) and ER (MBE = +426 vs. +354 gC/m²/yr). At this site, ELM-TAM's calibrated root C:N ratios are higher than ELM's defaults (**Fig. 4a**), reducing system-level N demand. In ELM's N-limitation framework, lower N demand weakens the down-regulation of potential GPP, allowing more carbon to be assimilated. The additional carbon allocated to longer-lived transport roots (**Fig. 4d**) in turn raises maintenance respiration and litter inputs, amplifying ER. Because the baseline already overestimates potential photosynthesis at this site, both effects, relaxed N constraint on GPP and elevated root respiration, compound rather than correct the positive bias. Even in years of anomalously low observed productivity (e.g., 2021: observed GPP = 1067 gC/m²/yr, both models >1570), the positive bias persists. The degradation at US-Ho1 illustrates a scope boundary rather than a structural deficiency in ELM-TAM: the dominant error source is upstream canopy photosynthesis, which neither model corrects. A holistic recalibration that jointly optimizes photosynthesis and root parameters would allow the improved belowground representation to operate on a more realistic carbon input, potentially converting the degradation into an improvement.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

4.3.3 Tier 3—Spatial out-of-sample validation

Spatial transferability was tested at seven independent sites spanning seven PFTs across four continents (**Fig. 5c; Fig. A4a–b**). Site-level MBE values for both models are shown in Fig. **A4c–d**. Of the 14 site–variable pairs evaluated, ELM-TAM improved RMSE relative to the baseline in 8 cases (57 %), with one site (GH-Ank, n = 4) showing negligible change (<1 % ΔRMSE).

The most successful parameter transfer occurred at US-Ro4 (Rosemount, Minnesota, 2014–2023), where ELM-TAM reduced both GPP and ER RMSE by 81 % relative to the baseline (GPP: 164 vs. 845 gC/m²/yr; ER: 217 vs. 1136 gC/m²/yr). The baseline massively overestimated both fluxes (GPP MBE = +830, ER MBE = +1083 gC/m²/yr; **Fig. A4c–d**), while ELM-TAM brought predictions close to observations (GPP MBE = −62, ER MBE = +194 gC/m²/yr). The parameters calibrated at AU-DaP (tropical C4 savanna, Darwin, Australia) thus transferred remarkably well to a temperate C4 grassland in a continental climate, suggesting that the TAM parameterization captures fundamental C4 grass physiology rather than site-specific conditions.

Moderate improvements emerged at two additional sites. At US-Blo (Blodgett Forest, California, 2000–2006), ELM-TAM improved GPP relative to the baseline (RMSE = 441 vs. 533 gC/m²/yr, −17 %), reducing the negative bias (MBE = −324 vs. −444 gC/m²/yr), though ER was slightly degraded (+11 %). At ZM-Mon (Mongu, Zambia, 2000–2009), ELM-TAM modestly improved both GPP (−5 %) and ER (−6 %), with the TAM parameters from PA-SPn (tropical deciduous, Panama) partially correcting the baseline's ER underestimate.

Three sites showed degradation that illuminates the limits of single-site calibration for PFTs spanning wide environmental gradients. At CN-Du2 (Inner Mongolia, PFT 13, 2006–2008), the US-Var grassland parameters produced excessive productivity when transferred to this

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

semiarid steppe (GPP RMSE +69 %), because allocation parameters calibrated under the Mediterranean climate of coastal California do not generalize to a continental semiarid regime. At RU-Fyo (central Russia, PFT 2, 1998–2014), the FI-Hyy boreal parameters underestimated fluxes (GPP RMSE +37 %, ER +36 %; n = 17 years), reflecting the climatic and species-composition differences between a Scots pine stand in southern Finland and a mixed spruce–birch forest in central European Russia. At CN-Dan (Tibetan Plateau, 4300 m, PFT 12, 2004–2005), both models overestimated GPP by a factor of four, as the US-Atq Arctic tundra parameters cannot represent an alpine meadow despite sharing the same PFT classification.

Across the seven sites, parameter transfer success is partly associated with the magnitude of baseline model bias. The strongest improvement (US-Ro4, −81 % RMSE) occurred where the baseline had the largest systematic error, while degradation occurred where calibration and validation sites differ substantially in productivity or climate despite sharing a PFT class. These degradations are not failures of the TAM framework itself but of the single-site calibration strategy applied to PFTs that span wide environmental gradients. The TAM structure provides the mechanistic scaffolding for differentiated root representation; realizing its potential at these PFTs requires multi-site calibration, sub-PFT stratification, and/or trait-based priors.

5 Discussion

The three-tier validation, together with the sensitivity analysis and calibration results, establishes ELM-TAM as a structurally viable alternative to the homogeneous fine-root representation in ELM and other land models. The validation also delineates where the TAM framework succeeds, where it falls short, and what this means for the path toward global application. We organize the discussion around five themes: the mechanistic evidence for sink-

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

based control (5.1), the parameterization dichotomy and its implications (5.2), the roadmap from single-site calibration to global benchmarking (5.3), limitations of the current implementation (5.4), and broader implications for land model development strategy (5.5).

5.1 TAM strengthens sink-based control through belowground structural realism

The central motivation of ELM-TAM is to restore the balance between source-driven and sink-based controls on carbon cycling. In conventional land models, photosynthesis determines growth and downstream fluxes follow passively, leaving belowground processes underrepresented (Fatichi et al., 2014; Cabon et al., 2022; Friend et al., 2019). ELM-TAM strengthens the sink side, so that it operates on a more equal footing with the source side. Three lines of validation evidence indicate that this rebalancing improves model performance.

First, the Tier 2 results at US-MOz deliver the most direct demonstration. The baseline ELM overestimated ER by +515 gC/m²/yr. This overestimate is a consequence of the source-driven cascade: overestimated GPP inflates all downstream carbon pools and their litter inputs to soil respiration. ELM-TAM nearly eliminates this bias (MBE = +33 gC/m²/yr; **Fig. 7**) by partitioning belowground carbon into three functionally distinct pools with different turnover rates. Fast-cycling absorptive roots contribute contemporaneous litter, long-lived transport roots provide a slow background decomposition flux, and mycorrhizal tissue mediates an intermediate pathway. This heterogeneous litter input profile produces a respiratory flux that tracks observed seasonal and interannual dynamics far more closely than the baseline's single-pool turnover. That this improvement emerges in a post-calibration period (2015–2021) under different meteorological forcing provides evidence that the differential turnover rates encode real process relationships rather than compensating for training-period forcing characteristics.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Second, the Tier 1 seasonal amplitude correction demonstrates that sink-based control operates at the sub-annual scale. ELM-TAM's mean amplitude ratio is 0.99 for GPP and 1.05 for ER, compared to 1.38 and 1.34 for the baseline (**Fig. 6; Fig. A3**). The baseline's amplitude overestimate arises because a single root pool with one turnover rate either releases too much carbon during the growing season or too little during the dormant season; the TAM framework's differentiated pools smooth out this temporal mismatch. This amplitude correction was not explicitly constrained during calibration (annual sums were the only objective), yet it emerges consistently across deciduous, grass, and several evergreen PFTs, a hallmark of structural correctness in which the right mechanisms produce the right dynamics at scales they were not trained on.

Third, the evergreen degradation at US-Ho1 (Tier 2) delineates the scope boundary of a belowground-only structural improvement. Both the baseline and ELM-TAM overestimate GPP at this site, and ELM-TAM's MBE worsens from +145 to +214 gC/m²/yr (**Fig. 7; Fig. A4c**). Two specific pathways connect the TAM structure to this amplification. First, the calibration assigns PFT 1 higher pool-specific C:N ratios than ELM's default (**Fig. 4**), which reduces system-level N demand and thereby weakens the N-based down-regulation of potential GPP; more carbon is assimilated, compounding the upstream photosynthesis bias rather than correcting it. Second, longer-lived transport roots accumulate more standing biomass, inflating maintenance respiration and litter inputs that further amplify ER. These pathways do not indicate a failure of the TAM framework itself. They indicate that a belowground structural improvement cannot remedy a source-side bias when TAM parameters are calibrated in isolation from the photosynthesis module. The constructive implication is that evergreen PFTs require holistic source–sink coupled

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

recalibration, in which canopy and belowground parameters are optimized jointly rather than sequentially.

Underpinning these flux improvements is a fundamentally restructured belowground nutrient economy. The calibration shows that TAM amplifies system-level N demand by up to 4.3× and replaces ELM's uniform 1-yr default with absorptive-root longevity ranging from 0.5 to 3.5 yr across PFTs (**Fig. 4**). These changes intensify the nutrient limitation feedback and produce heterogeneous litter inputs with pool-specific chemistry, altering soil decomposition dynamics and nutrient mineralization timing. This represents one key mechanism through which ELM-TAM strengthens sink-based control. The generally equal or better ER performance relative to GPP across deciduous and grass PFTs (**Fig. 5**) supports this interpretation. Yet the strength and direction of this belowground regulation vary sharply across vegetation types: the US-MOz correction and the US-Ho1 degradation are comparable in magnitude but opposite in sign. What determines whether TAM's structural realism helps or hinders at a given PFT?

5.2 The two-strategy parameterization dichotomy and its validation consequences

The answer lies in the two-strategy parameterization dichotomy identified in the sensitivity analysis (**Fig. 3**): trait-coordination PFTs (69% of the 13 analyzed systems) concentrate variance in 2–3 allocation parameters, while root-economics PFTs distribute sensitivity across 5–6 parameters spanning chemistry, longevity, and allocation. The validation independently confirms that these two strategies have fundamentally different transferability properties.

For deciduous and grass PFTs governed by the trait-coordination strategy, the validation results support robust transferability. These PFTs transfer well temporally: US-MOz and US-Var both improve over the baseline during 2015–2021 (Tier 2; **Fig. 7**). The concentrated sensitivity means that once *froott_leaf*, *froota_leaf*, and *frootm_leaf* are correctly calibrated, the model's

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

seasonal and interannual behavior is largely determined. However, spatial transfer is more fragile for PFTs spanning wide productivity ranges. At CN-Du2 (Tier 3), the US-Var parameters degrade GPP by +69% when applied to a semiarid Inner Mongolian steppe, because the allocation ratios calibrated at a productive Californian grassland overcorrect at a site where water and temperature constrain productivity far more severely. The implication is that for allocation-dominated PFTs, multi-site calibration or productivity-dependent scaling of *froott_leaf*, *froota_leaf*, and *frootm_leaf* would likely resolve the spatial transfer problem without requiring additional structural complexity. The 2–3 parameter sensitivity means that even with multiple calibration sites, the optimization problem remains tractable.

For evergreen PFTs governed by the root-economics strategy, the outcome diverges sharply. These PFTs degrade even temporally: US-Ho1 worsens during 2015–2021 despite strong calibration-period performance. Spatially, RU-Fyo (PFT 2) degrades by +37% despite a 17-year comparison. The distributed sensitivity across 5–6 parameters makes calibration susceptible to equifinality: multiple parameter combinations can produce similar annual fluxes during calibration but diverge under novel conditions. The validation thus prescribes a fundamentally different strategy for evergreen PFTs: tighter trait-based priors to reduce the effective parameter space before flux-based optimization. Our calibration already draws on trait measurements compiled from the published literature and trait databases to set prior ranges for root C:N, longevity, and mycorrhizal parameters (Sect. 3.3; Wang et al., 2023). The limitation is that current trait coverage for evergreen root systems is relatively sparse compared with deciduous counterparts (Iversen et al., 2017), so the priors are necessarily broad. Expanded measurements, e.g., direct tissue C:N, minirhizotron longevity estimates, and mycorrhizal colonization surveys, would narrow these priors sufficiently to transform the calibration from a 6-parameter optimization against flux data

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

alone into a 2–3 parameter optimization, potentially restoring the tractability that deciduous PFTs enjoy.

The two-strategy dichotomy thus prescribes distinct paths toward global application. The two strategies map onto phenological type rather than geography or taxonomy: trait-coordination governs the deciduous and grass PFTs (PFTs 3, 6, 7, 8, 10, 11, 12, 13, and 14), where short leaf and fine-root lifespans tie belowground allocation tightly to seasonal source activity, whereas root-economics governs the evergreen PFTs (PFTs 1, 2, 4, and 5), where multi-year tissue persistence makes chemistry and longevity the primary levers on carbon return. Because this mapping is phenological, paired leaf–root measurements from functional trait databases [e.g., TRY (Kattge et al., 2020) and FRED (Iversen et al., 2017)] could be synthesized into predictive scaling relationships that cross taxonomic and geographic boundaries. Such relationships would provide globally applicable parameter estimates for the 9 trait-coordination PFTs while focusing targeted observational campaigns on the evergreen systems where they are most needed. The validation results indicate where these two different strategies are most urgently required.

5.3 From single-site calibration to global application: a roadmap

The Tier 3 validation provides quantitative evidence for translating these prescriptions into a concrete path from the current calibration to global application. As a starting point, we report the complete set of calibrated TAM parameter values for all 14 PFTs in Table 2, providing the community with a ready-to-use baseline parameterization that can be refined as additional observations become available. The US-Ro4 result (−81% RMSE for both GPP and ER) demonstrates that PFT-level calibration can transfer across continents when the TAM parameterization captures fundamental physiology. In this case, the relevant physiology is C4 grass carbon cycling. Parameters calibrated at a tropical C4 savanna (AU-DaP, Darwin, Australia)

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

transferred remarkably well to a temperate C4 grassland in a continental climate (Rosemount, Minnesota), providing an encouraging proof of concept that TAM parameters can encode PFT-level functional strategies rather than site-specific conditions.

The degradation cases indicate what is needed to achieve this more broadly. Three strategies emerge from the validation: First, multi-site calibration would address the spatial transfer failures observed within PFTs that span wide productivity ranges. Rather than calibrating *froott_leaf* , *froota_leaf* , and *frootm_leaf* at a single site, the MCMC objective function should aggregate misfit across multiple sites per PFT. The Tier 3 results indicate which PFTs need this most urgently (**Fig. A4**): PFT 13, where CN-Du2 degrades by +69% under US-Var parameters, and PFT 2, where RU-Fyo degrades by +37% under FI-Hyy parameters. For PFTs where the sensitivity analysis shows only 2–3 dominant parameters, multi-site optimization remains computationally feasible even with the surrogate model approach.

Second, sub-PFT stratification is needed for PFTs that span qualitatively different environments where no single parameter set can be adequate. PFT 12 (C3 Arctic grass), for example, encompasses Arctic coastal tundra (US-Atq) and high-altitude alpine meadows (CN-Dan, 4300 m). These environments are so climatologically distinct that shared parameters inevitably compromise at both sites. Splitting PFT 12 into Arctic and alpine sub-PFTs, or conditioning parameters on environmental covariates such as mean annual temperature and growing-season length, would address this without proliferating the total number of PFTs.

Third, tighter trait-informed priors offer the most promising path forward for evergreen PFTs, where equifinality undermines the current calibration. The present study already incorporates published measurements and trait database values to set prior ranges for root and mycorrhizal parameters (Sect. 3.3). The priors nevertheless remain broad because direct root and

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

mycorrhizal observations for evergreen systems are relatively sparse. The TRY database (Kattge et al., 2020) and regional trait networks (e.g., Wang et al., 2022) offer extensive leaf-economic measurements and increasingly include corresponding fine-root and mycorrhizal fungi traits, yet root longevity and chemistry data for evergreen species remain comparatively thin (Iversen et al., 2017). Expanding this coverage through targeted field campaigns, guided by the clear mechanistic interpretation of sensitivity patterns, would narrow the trait-informed priors and, in turn, collapse the 5–6 parameter optimization toward a more tractable 2–3 parameter problem akin to that enjoyed by deciduous PFTs.

With these strategies in place, regional and global benchmarking with ILAMB (Collier et al., 2018) represents the natural next step. The demonstration of robust performance across an order of magnitude in productivity (from ~100 to >3000 gC/m²/yr) and across all major biome types establishes the framework's readiness for larger-scale evaluation. Most current land models share the simplification of homogenizing diverse belowground processes into a single pool (Wang et al., 2023). ELM-TAM would provide a platform to test whether this homogenization propagates systematic biases into regional and global carbon balance estimates.

5.4 Limitations of the current implementation

The roadmap outlined above assumes that the current limitations are surmountable. Here we distinguish a methodological limitation of the calibration strategy from structural simplifications that define the boundaries of what ELM-TAM can represent.

On the methodological side, the calibration targets only two flux variables, GPP and ER, leaving the 17-dimensional parameter space underconstrained. For evergreen PFTs, where sensitivity is distributed across 5–6 interacting parameters (Sect. 4.1), this underconstraint manifests as equifinality: multiple parameter combinations yield similar annual fluxes during

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

calibration but diverge under novel conditions. Even for the best-performing deciduous sites, compensating errors reveal the cost of limited constraints: at US-MOz, ELM-TAM nearly eliminates the baseline ER bias but flips the GPP bias from +340 to −286 gC/m²/yr, achieving correct ER partly at the expense of canopy productivity. Incorporating additional observational targets such as soil respiration partitioning, root biomass, mycorrhizal colonization rates, or litter chemistry could break the parameter degeneracy that permits such trade-offs. A complementary extension, motivated by the US-Ho1 scope boundary (Sect. 5.1), is holistic source–sink coupled recalibration, in which canopy and belowground parameters are optimized jointly rather than treating TAM parameters as a downstream correction layer applied atop fixed photosynthesis settings. Even so, TAM's gains would remain bounded by ELM's unchanged process representations. The persistent ER residual at US-Var (MBE = −118 gC/m²/yr), traceable to soil decomposition limitations, and the tropical phase errors at BR-Sa1 and PA-SPn, traceable to the phenology scheme, illustrate that TAM is a necessary but not sufficient improvement.

Part of this insufficiency traces not to calibration but to the TAM framework itself, which incorporates structural simplifications that bound its representational scope. Two are particularly critical, although it is imperative to eventually relax the assumption of big-leaf approach in ELM-TAM to integrate with demographic vegetation representation (Wang et al., 2023). First, the TAM pools do not feed back on ecosystem hydrology. ELM-TAM retains the empirical soil moisture stress factor rather than coupling root traits to a mechanistic hydraulic scheme. Absorptive root density and mycorrhizal networks govern water uptake and drought responses in reality (e.g., Kakouridis et al., 2022), but they are not coupled to the water cycle in the current implementation. This decoupling likely contributes to the spatial transfer failures at water-limited sites such as CN-Du2 (semiarid steppe), where root-driven drought responses may be as important as carbon

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

allocation in controlling productivity. Mechanistic root–water coupling, together with mechanistic root-nutrient coupling, would allow TAM to influence ecosystem water balance alongside carbon and nutrient cycling.

Another related issue is applying the same vertical root distribution to all three TAM pools. In reality, absorptive roots and mycorrhizal hyphae concentrate in organic-rich topsoil while transport roots extend deeper (e.g., McElrone et al., 2004; Robin et al., 2019). Because litter chemistry differentiation and exudation are the primary mechanism through which TAM alters soil carbon dynamics, its impact is muted when all three litter streams enter the same depth profile. Depth-resolved TAM pools would direct labile absorptive and mycorrhizal litter and exudates into shallow, fast-cycling soil layers and lignin-rich transport root litter into deeper, slower pools. This vertical separation could amplify TAM's influence on both the magnitude and distribution of soil carbon storage (e.g., Billings et al., 2025).

Decoupled root hydraulics and uniform vertical distribution thus define two concrete extensions that would broaden TAM's mechanistic scope. Each issue requires new process representation in coordination with photosynthetic allocation rather than additional parameter tuning (Wang et al., 2023). This distinction of process representation vs. parameter tuning matters for how the community should prioritize development effort.

5.5 Implications for land model development strategy

Our results suggest that the marginal return on development investment may be higher for improving root representation than for further refining aboveground processes. Belowground structural representation has lagged aboveground refinement for decades (Wang et al., 2023). Our results demonstrate that belowground structural complexity yields measurable, transferable gains. Evidence from ELM-TAM supports this view across multiple lines of evidence: ecologically

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

coherent parameter differentiation (**Fig. 4**), consistent ER improvement at deciduous and grass PFTs (**Fig. 5**), emergent seasonal amplitude correction at scales not targeted during calibration (**Fig. 6**), and cross-continental parameter transfer at US-Ro4 (−81% RMSE; **Fig. 5**). The evergreen degradation at US-Ho1 and RU-Fyo reinforces rather than undermines this conclusion: gains are largest where observational constraints match the added structural complexity, pointing to a clear path for extending these benefits to currently underperforming PFTs. This argues for a targeted strategy — prioritize TAM implementation for PFTs where it demonstrably helps (deciduous and grass) while pursuing tighter trait-informed priors, supported by expanded root and mycorrhizal trait measurements, for PFTs where performance remains bounded by sparse data (evergreen).

This selective success also challenges the prevailing assumption that increased model complexity necessarily degrades performance through equifinality and overparameterization (Luo and Schuur 2020; Famiglietti et al., 2021). Although TAM introduces 21 new parameters, only 8 of 17 free parameters exceed the empirical ST threshold at any PFT, and the calibrated values produce emergent properties—PFT-specific N demand and turnover time (**Fig. 4**)—that are consistent with independent root trait observations. Model complexity should therefore be evaluated not by parameter count alone but by the alignment between model structure and the dominant controls on ecosystem function.

These findings also carry a cautionary message for the growing trend toward replacing mechanistic process development with machine learning emulators. Such tools are valuable for model analysis, as our own surrogate-based sensitivity analysis and calibration demonstrate, but they are fundamentally constrained by the physics embedded in the model they emulate. A surrogate of a structurally incomplete model inherits and potentially amplifies its structural biases (Gyllingberg et al., 2023; Nartallo-Kaluarachchi et al., 2025). ELM-TAM's improvements arise

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

precisely because the TAM framework introduces new mechanistic pathways that no emulator of the original ELM could have represented. Continued investment in mechanistic process representation therefore remains essential for the structural advances from which predictive gains in Earth system land models ultimately arise.

6. Conclusions

This study presented ELM-TAM, a new version of the E3SM Land Model that replaces the single homogeneous fine-root pool with three functionally distinct pools representing transport roots, absorptive roots, and mycorrhizal fungi. This structural decomposition introduces 21 new parameters governing root allocation, chemistry, longevity, and litter quality. Through Sobol global sensitivity analysis, Bayesian calibration at 13 FLUXNET sites spanning all natural PFTs in ELM, and a three-tier validation framework of increasing independence, we evaluated whether this added belowground complexity yields measurable and transferable improvements in simulated carbon fluxes.

The sensitivity analysis revealed a two-strategy parameterization dichotomy: deciduous and grass PFTs (69% of analyzed systems) follow a trait-coordination strategy controlled by 2–3 allocation parameters, while evergreen PFTs follow a root-economics strategy that distributes sensitivity across 5–6 parameters spanning C:N ratios, longevity, and allocation. This dichotomy persists across more than 100° of latitude and across woody to herbaceous growth forms, suggesting that phenological strategy determines the dominant controls on belowground carbon cycling. An empirical gap in total-effect indices separates 8 influential parameters from the remaining 9, indicating that the effective dimensionality is far smaller than the nominal parameter count.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Bayesian calibration produced emergent fine-root properties that diverge from ELM's defaults: system-level nitrogen demand increased at 10 of 14 PFTs (up to 4.3×), with absorptive roots and mycorrhizal fungi contributing 77–100% of total N demand; absorptive-root longevity ranged from 0.5 to 3.5 yr across PFTs, replacing ELM's uniform 1-yr default with PFT-specific turnover shaped by pool composition. The three-tier validation established that these changes improve predictive skill in transferable ways. Seasonally (Tier 1), ELM-TAM improved amplitude fidelity not targeted during calibration (mean GPP amplitude ratio 0.99 vs. 1.38 for the baseline). Temporally (Tier 2), the differentiated pools nearly eliminated the baseline's ER overestimate at US-MOz (MBE from +515 to +33 gC m^{-2} yr^{-1}). Spatially (Tier 3), C4 grass parameters transferred across continents with 81% RMSE reduction, while PFTs spanning wide environmental gradients (PFTs 2, 12, 13) require multi-site calibration or sub-PFT stratification. The validation also defined a boundary: at evergreen sites where the baseline already overestimates productivity, a belowground-only structural improvement cannot remedy the upstream source-side bias and instead amplifies it.

The two-strategy dichotomy prescribes distinct development paths: multi-site calibration of 2–3 allocation parameters for trait-coordination PFTs, sub-PFT stratification for PFTs that span qualitatively different environments (e.g., PFT 12), and tighter trait-informed priors, supported by expanded root and mycorrhizal trait measurements, for root-economics PFTs. Because the current development focuses on the TAM structure and its integration into the carbon–nutrient cycle, the function-oriented design provides natural coupling points for mechanistic root–water interactions and depth-resolved root profiles in future phases. Global benchmarking with ILAMB represents the natural next step. More broadly, ELM-TAM demonstrates that mechanistic belowground complexity, when grounded in observable traits and guided by sensitivity analysis, can improve

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

predictive skill without creating an overparameterization burden. The consistent pattern across all validation tiers—improvement where TAM addresses genuine structural deficiencies, interpretable degradation where errors originate upstream—establishes the framework's mechanistic credibility and identifies the steps needed to extend it from site-level demonstration to global application in Earth system land models.

Appendix A

A1 Site metadata

Table A1. Calibration and sensitivity analysis sites. Thirteen FLUXNET eddy covariance sites used for ELM-TAM development, one per ELM plant functional type (PFT). Thirteen sites entered the surrogate-based sensitivity analysis and MCMC calibration; PFT 9 was prescribed from literature. Coordinates are site-level values used in single-point ELM simulations. The calibration period denotes the years over which annual GPP and ER were compared to FLUXNET2015 observations. The mycorrhizal association (EM = ectomycorrhizal, AM = arbuscular mycorrhizal) governs the relative functional importance of the mycorrhizal pool in the TAM framework.

PFT	Site	PFT name	Lat	Lon	IGBP	Myc.	Calibration period	Calibration
1	US-Ho1	Needleleaf evergreen temperate tree	45.20	−68.74	ENF	EM	2000–2006	MCMC

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

2	FI-Hyy	Needleleaf evergreen boreal tree	61.85	24.29	ENF	EM	1997–2014	MCMC
3	RU-SkP	Needleleaf deciduous boreal tree	62.26	129.17	ENF	EM	2004–2014	MCMC
4	BR-Sa1	Broadleaf evergreen tropical tree	−2.86	−54.96	EBF	AM	2002–2011	MCMC
5	AU-Tum	Broadleaf evergreen temperate tree	−35.66	148.15	EBF	EM/AM	2002–2014	MCMC
6	PA-SPn	Broadleaf deciduous tropical tree	9.32	−79.63	DBF	AM	2008–2012	MCMC
7	US-MOz	Broadleaf deciduous temperate tree	38.74	−92.20	DBF	EM/AM	2005–2014	MCMC
8	CA-Oas	Broadleaf deciduous boreal tree	53.63	−106.20	DBF	EM	1997–2010	MCMC
10	US-SRC	Broadleaf deciduous	31.91	−110.84	OSH	AM	2008–2014	MCMC

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

		temperate shrub						
11	RU-Cok	Broadleaf deciduous boreal shrub	70.83	147.49	OSH	EM	2003–2014	MCMC
12	US-Atq	C3 Arctic grass	70.47	−157.41	WET	AM	2004–2008	MCMC
13	US-Var	C3 non-Arctic grass	38.41	−120.95	GRA	AM	2001–2014	MCMC
14	AU-DaP	C4 grass	−14.06	131.32	GRA	AM	2008–2014	MCMC

Table A2. Tier 3 spatial validation sites. Seven independent FLUXNET sites used for spatial out-of-sample validation (Tier 3; Sect. 3.4.3). Each site shares the same ELM PFT classification as its corresponding calibration site but differs in climate, species composition, and geographic location. Parameters were transferred from the calibration site without re-tuning.

Validation site	PFT	Calibration donor	Validation period	Lat	Lon	Key contrast
US-Blo	1	US-Ho1	2000–2006	38.90	−120.63	Ponderosa pine, Mediterranean vs. spruce–hemlock, humid continental

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

RU-Fyo	2	FI-Hyy	1998–2014	56.46	32.92	Spruce–birch mixed forest, central Russia vs. Scots pine, Finland
GH-Ank	4	BR-Sa1	2011–2014	5.27	−2.69	West African semi-deciduous forest vs. Amazonian rainforest
ZM-Mon	6	PA-SPn	2000–2009	−15.44	23.25	Miombo woodland, Zambia vs. tropical dry forest, Panama
CN-Dan	12	US-Atq	2004–2005	30.50	91.07	Alpine meadow, Tibetan Plateau (4300 m) vs. Arctic tundra, Alaska
CN-Du2	13	US-Var	2006–2008	42.05	116.28	Temperate steppe, Inner Mongolia vs. Mediterranean grassland, CA
US-Ro4	14	AU-DaP	2014–2023	44.68	−93.07	Tallgrass prairie, Minnesota vs. tropical savanna, N. Australia

1325

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

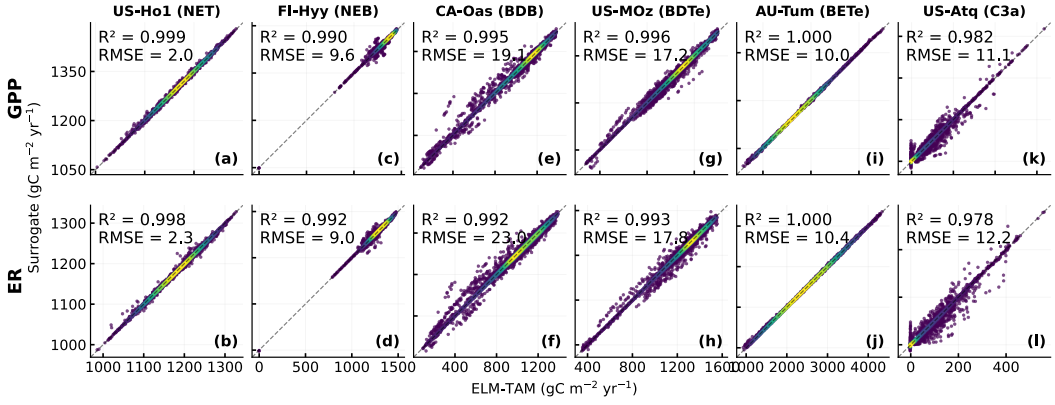


Figure A1. Surrogate model performance for six representative calibration sites. Each panel shows a 1:1 scatter plot of surrogate-predicted versus ELM-TAM-simulated annual flux values (gC m^{-2} yr^{-1}) across 1000 ensemble members and all calibration years, with points colored by kernel density estimation. Left column: gross primary production (GPP); right column: ecosystem respiration (ER). Coefficient of determination (R^2) and root mean square error (RMSE) are annotated in each panel. The dashed line indicates the 1:1 reference.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

A2 Surrogate model performance

Neural-network surrogate models (multi-layer perceptron regressors; Sect. 3.1) served as computationally efficient emulators of ELM-TAM for both sensitivity analysis and MCMC calibration. Because all downstream inferences depend on surrogate fidelity, we assessed emulator accuracy by comparing surrogate predictions against the actual ELM-TAM ensemble outputs on held-out test data (20 % of ensemble members withheld during training).

Figure A1 presents density-coloured 1:1 scatter plots of surrogate-predicted versus ELM-TAM-simulated annual GPP and ER for six representative sites spanning the PFT diversity: US-Ho1 (PFT 1, needleleaf evergreen temperate), FI-Hyy (PFT 2, needleleaf evergreen boreal), CA-Oas (PFT 8, broadleaf deciduous boreal), US-MOz (PFT 7, broadleaf deciduous temperate), AU-Tum (PFT 5, broadleaf evergreen temperate), and US-Atq (PFT 12, C3 Arctic grass). These sites sample both parameterization strategies (root-economics and trait-coordination) and span a wide range of flux magnitudes, from approximately 200 gC m^{-2} yr^{-1} at the Arctic tundra site (US-Atq) to over 5000 gC m^{-2} yr^{-1} at the temperate broadleaf evergreen site (AU-Tum).

All 12 panels exhibit tight clustering along the 1:1 line with $R^2 \geq 0.978$, and the highest-fidelity surrogates (AU-Tum, US-Ho1) achieve $R^2 = 1.000$ and 0.999 respectively. Root mean square errors are small relative to the flux magnitudes at each site (RMSE / mean flux < 2 % in all cases shown). Point density, depicted by kernel density estimation coloring, confirms that the surrogate fidelity is uniformly high across the full parameter space rather than being concentrated in a narrow region. These results hold across all 13 calibration sites (not shown), confirming that the surrogate models are sufficiently accurate to support the sensitivity analysis and Bayesian calibration presented in the main text.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

A3 Sensitivity analysis: total-effect indices

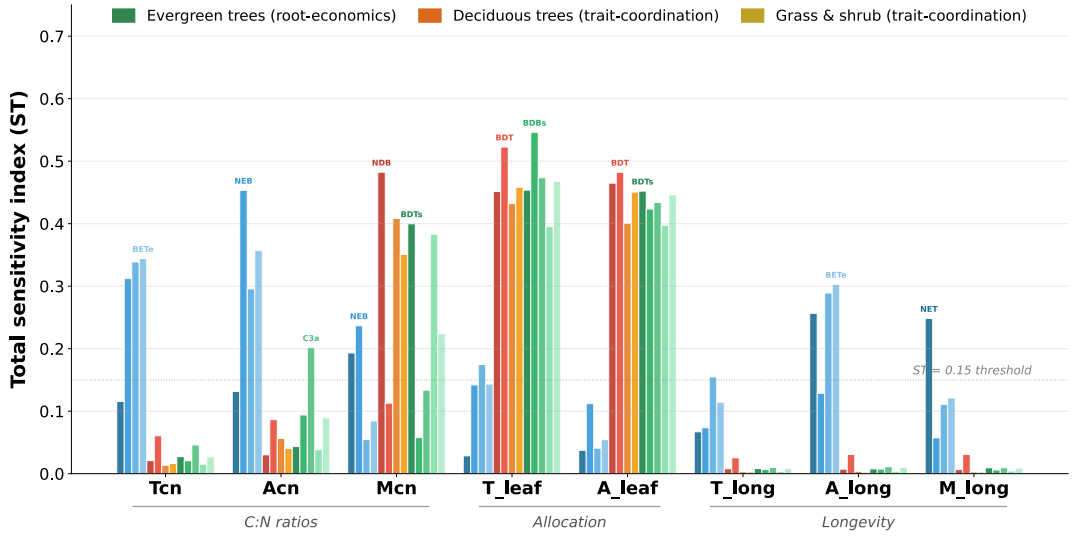


Figure A2. Total-effect sensitivity indices (ST) for the eight most influential TAM parameters across 13 calibrated PFTs, grouped by functional category. Bar colors denote parameterization strategy: blue for root-economics (evergreen PFTs) and orange for trait-coordination (deciduous and grass PFTs). ST values represent the combined GPP + ER sensitivity. The eight parameters were identified by an empirical gap at ST = 0.15 in the cross-PFT distribution; the remaining nine parameters contribute negligible total sensitivity.

Figure A2 complements the first-order and second-order Sobol heatmap (**Fig.3**) by showing the absolute magnitudes of total-effect sensitivity indices (ST) for the eight most influential TAM parameters. Total-effect indices integrate direct effects, pairwise interactions, and all higher-order terms, providing an upper bound on each parameter's contribution to output variance. The eight parameters were separated from the remaining nine by an empirical gap in the cross-PFT ST distribution at 0.15. Bar heights represent the combined GPP + ER sensitivity

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

(S_combined = (S_GPP + S_ER) / 2) for each of the 13 calibrated PFTs, grouped by functional category (chemistry, allocation, longevity, litter chemistry) and colored by parameterization strategy (blue: root-economics, evergreen PFTs; orange: trait-coordination, deciduous and grass PFTs).

For trait-coordination PFTs, the allocation parameters froott_leaf and froota_leaf dominate the ST budget, consistent with their S1 dominance in **Fig. 3a**. For root-economics PFTs, ST is more evenly distributed across C:N ratios (froottcn, frootacn, frootmcn), longevity (froott_long, froota_long, frootm_long), and allocation, reflecting the broader parameter sensitivity landscape characteristic of evergreen vegetation. The clear separation between the two strategy groups — visible as distinct color clustering within each functional category — reinforces the dichotomy identified in the main-text sensitivity analysis and provides a mechanistic basis for the parameterization strategies discussed in Sect. 5.2.

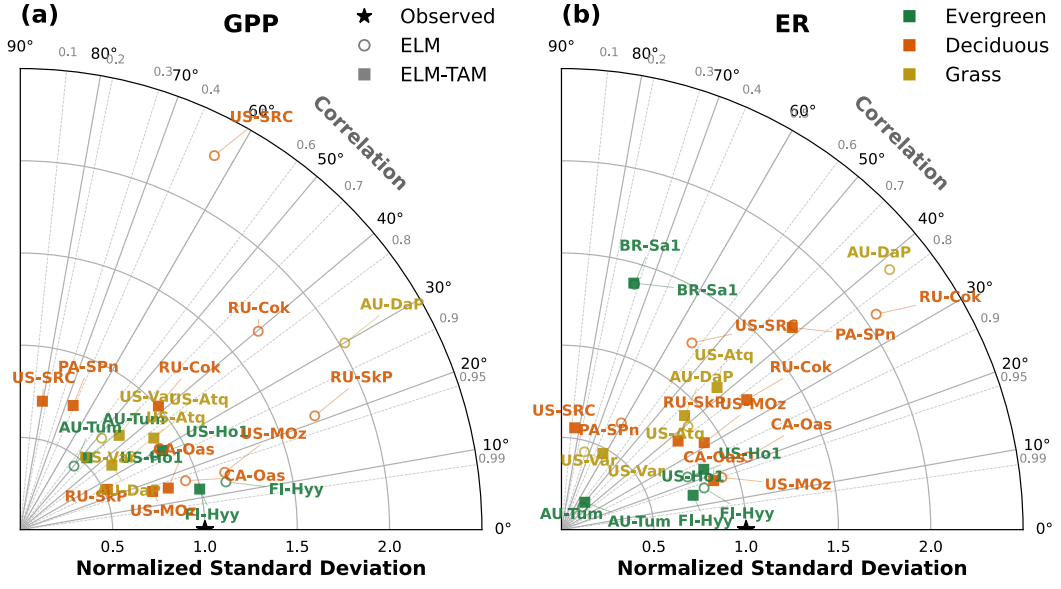
A4 Validation: Taylor diagram for Tier 1 monthly evaluation

The Taylor diagram provides a compact summary of Tier 1 monthly evaluation performance across all 13 calibration sites evaluated (ES-LJu excluded; Sect. 4.3). Each point represents the monthly time series of GPP or ER at one site, positioned according to the Pearson correlation coefficient (angular axis), normalized standard deviation (radial axis), and centered root mean square difference (contours) relative to observations. Filled symbols denote ELM-TAM; open symbols denote baseline ELM.

The diagram reveals two key patterns. First, ELM-TAM shifts most site–variable combinations toward higher correlation and lower centered RMSD relative to baseline ELM, with the largest improvements at deciduous and grass sites where the trait-coordination strategy

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

1394 concentrates sensitivity in allocation parameters that govern seasonal dynamics. Second, the
1395 normalized standard deviation ratios cluster near 1.0 for ELM-TAM, indicating that the optimized
1396 model reproduces the observed amplitude of the seasonal cycle more faithfully than baseline ELM,
1397 whose markers scatter out to nσ = 2.0–2.6, reflecting systematic overestimation of seasonal
1398 variability. These patterns are consistent with the composite seasonal cycles shown for six
1399 representative sites in **Fig. 6**.
1400



1401

1402 **Figure A3. Taylor diagram summarizing Tier 1 monthly evaluation of GPP and ER at 13**
1403 **calibration sites**. Each point represents the monthly time series at one site for one variable. Filled
1404 markers: ELM-TAM with calibrated TAM parameters; open markers: baseline ELM. Angular
1405 position indicates Pearson correlation with observations; radial distance indicates normalized
1406 standard deviation (model / observed); gray contours show centered root mean square difference.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Points closer to the reference point (correlation = 1, normalized SD = 1) indicate better agreement with observations.

A5 Validation: Tier 3 spatial transferability

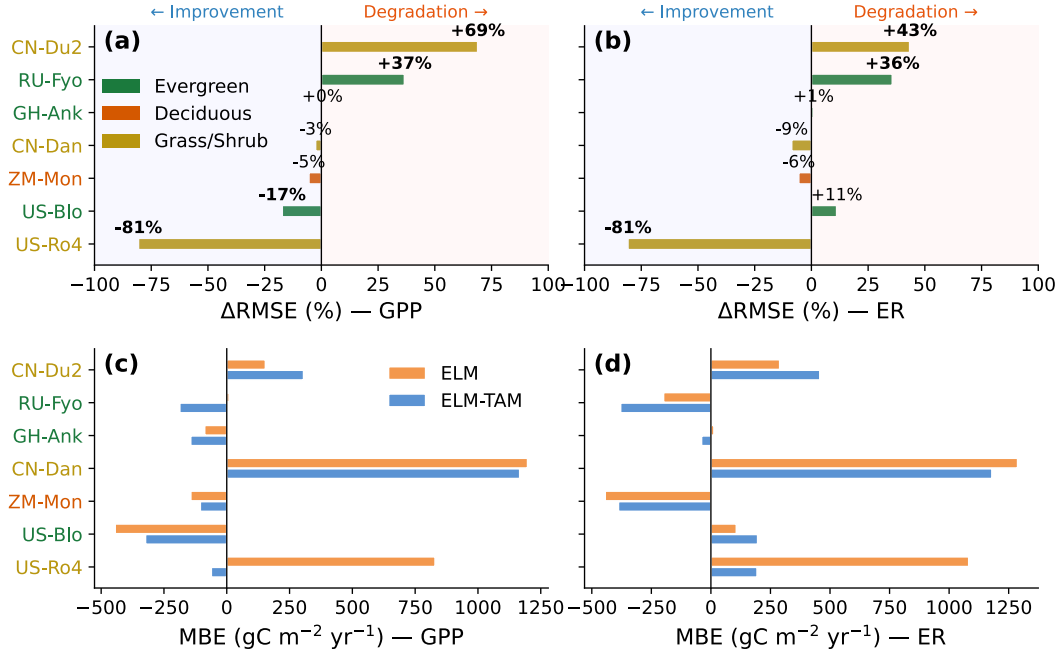


Figure A4. Tier 3 spatial validation at seven independent sites, sorted by GPP ΔRMSE% magnitude. **(a–b)** Relative RMSE change (ΔRMSE%) for GPP and ER; negative values indicate that ELM-TAM reduces prediction error. Bars are colored by parameterization strategy: evergreen/root-economics (green), deciduous/trait-coordination (orange), and grass/shrub (gold). **(c–d)** Mean bias error (MBE; gC m^{-2} yr^{-1}) for GPP and ER, showing paired bars for baseline ELM (orange) and ELM-TAM (blue) at each site. Bars closer to zero indicate smaller systematic bias.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Fig. A4 complements the validation scorecard (**Fig. 5c**) by presenting the relative RMSE change (ΔRMSE%) and mean bias error (MBE) at the seven spatial validation sites, sorted by GPP ΔRMSE% magnitude and color-coded by functional strategy. Of 14 site–variable pairs, ELM-TAM improves RMSE in 8 cases (57 %). The ΔRMSE% panels (a–b) show that the strongest improvements concentrate at grass and deciduous sites (e.g., US-Ro4, −81 %), while degradations occur where calibration and validation sites differ substantially in climate despite sharing a PFT class (e.g., CN-Du2, +69 %). The MBE panels (c–d) reveal the underlying bias structure: at US-Ro4, ELM-TAM nearly eliminates a baseline overestimate exceeding 800 gC m^{-2} yr^{-1} for both GPP and ER, whereas at CN-Dan both models show large positive biases that the TAM parameters do not correct. These patterns indicate that the degradations reflect limits of the single-site calibration strategy rather than the TAM framework itself.

Code and data availability

Data and code used in this manuscript were archived at https://doi.org/10.5281/zenodo.20310624 (Wang 2026). Living analysis code available at https://github.com/bioatmosphere/TAM, and the ELM-TAM codebase at https://github.com/bioatmosphere/ELM-TAM.

Author contributions

B.W. conceived the study, performed simulations and analyses, and wrote the manuscript with contributions from all co-authors.

Competing interests: The authors declare that they have no conflict of interest.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Acknowledgements

We thank the ORNL Data-Model Integration Group for consistent conversations and inspirations. This research used resources of the Compute and Data Environment for Science (CADES) at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

Financial support

This work was supported by the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation (RUBISCO) Science Focus Area, which is sponsored by the Regional and Global Model Analysis (RGMA) activity of the Earth & Environmental Systems Modeling (EESM) Program in the Earth and Environmental Systems Sciences Division (EESSD) of the Office of Biological and Environmental Research (BER) in the US Department of Energy Office of Science. This work was also supported by the Terrestrial Ecosystem Science Scientific Focus Area, which is sponsored by the Environmental System Science Program in the EESSD of BER.

References

Allen, M. F., & Kitajima, K. (2013). In situ high-frequency observations of mycorrhizas. New Phytologist, 200, 222-228.

Bonan, G. B., Patton, E. G., Finnigan, J. J., Baldocchi, D. D., & Harman, I. N. (2021). Moving beyond the incorrect but useful paradigm: Reevaluating big-leaf and multilayer plant canopies to model biosphere-atmosphere fluxes –A review. Agricultural and Forest Meteorology, 306, 108435.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Burrows, S. M., Maltrud, M., Yang, X., Zhu, Q., Jeffery, N., Shi, X., Ricciuto, D., Wang, S., Bisht, G., Tang, J., Wolfe, J., Harrop, B. E., Singh, B., Brent, L., Baldwin, S., Zhou, T., Cameron-Smith, P., Keen, N., Collier, N., … Leung, L. R. (2020). The DOE E3SM v1.1 biogeochemistry configuration: Description and simulated ecosystem-climate responses to historical changes in forcing. Journal of Advances in Modeling Earth Systems, 12, e2019MS001766.

Beven, K.: A manifesto for the equifinality thesis, J. Hydrol., 320, 18–36, https://doi.org/10.1016/j.jhydrol.2005.07.007, 2006.

Billings, S. A., Sullivan, P. L., Li, L., Hirmas, D. R., Nippert, J. B., Ajami, H., Flores, A. N., Singha, K., Keen, R. M., Markewitz, D., Chorover, J., Ederer, A., Silver, W. L., Unruh, M., Gerson, J., and Hart, S. C.: Contrasting depth dependencies of plant root presence and mass across biomes underscore prolific root–regolith interactions, AGU Adv., 6, e2025AV002072, https://doi.org/10.1029/2025AV002072, 2025.

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., Mu, M., and Randerson, J. T.: The International Land Model Benchmarking (ILAMB) System: Design, theory, and implementation, J. Adv. Model. Earth Syst., 10, 2731–2754, https://doi.org/10.1029/2018MS001354, 2018.

Cabon, A., Kannenberg, S. A., Arain, A., Babst, F., Baldocchi, D., Belmecheri, S., ... & Anderegg, W. R. (2022). Cross-biome synthesis of source versus sink limits to tree growth. Science, 376, 758-761.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Dale, M. B. (1970). Systems analysis and ecology. Ecology, 51(1), 2–16. https://doi.org/10.2307/1933596

Dukes, J. S., Xu, C., Liao, C., Novick, K. A., Phillips, R. P., Beverly, D. P., Fang, Y., Jacobs, E. M., McAdam, S. A. M., Paudel, I., Rimer, I. M., and Robbins, Z. J.: Improving the representation of plant water stress and water use in Earth System Models, New Phytol., 249, 39–55, https://doi.org/10.1111/nph.70687, 2026.

Edmonds, B. Complexity and scientific modelling, Foundations of Science, 5, 379–390, 2000

Famiglietti, C. A., Smallman, T. L., Levine, P. A., Flack-Prain, S., Quetin, G. R., Meyer, V., Parazoo, N. C., Stettz, S. G., Yang, Y., Bonal, D., Bloom, A. A., Williams, M., & Konings, A. G. (2021). Optimal model complexity for terrestrial carbon cycle prediction. Biogeosciences, 18(8), 2727–2754. https://doi.org/10.5194/bg-18-2727-2021

Fatichi S, Pappas C, Zscheischler J, Leuzinger S (2019) Modelling carbon sources and sinks in terrestrial vegetation. New Phytologist 221, 652–668. https://doi.org/10.1111/nph.15451

Fatichi, S., Leuzinger, S., & Körner, C. (2014). Moving beyond photosynthesis: from carbon source to sink-driven vegetation modeling. New Phytologist, 201, 1086-1095. https://doi.org/10.1111/nph.12614

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Fisher, J. B., Sitch, S., Malhi, Y., Fisher, R. A., Huntingford, C., and Tan, S.-Y.: Carbon cost of plant nitrogen acquisition: A mechanistic, globally applicable model of plant nitrogen uptake, retranslocation, and fixation, Global Biogeochem. Cycles, 24, GB1014, https://doi.org/10.1029/2009GB003621, 2010.

Fisher, R. A., Koven, C. D., Anderegg, W. R. L., Christoffersen, B. O., Dietze, M. C., Farrior, C. E., Holm, J. A., Hurtt, G. C., Knox, R. G., Lawrence, P. J., Lichstein, J. W., Longo, M., Matheny, A. M., Medvigy, D., Muller-Landau, H. C., Powell, T. L., Serbin, S. P., Sato, H., Shuman, J. K., Smith, B., Trugman, A. T., Viskari, T., Verbeeck, H., Weng, E., Xu, C., Xu, X., Zhang, T., and Moorcroft, P. R.: Vegetation demographics in Earth System Models: A review of progress and priorities, Glob. Change Biol., 24, 35–54, https://doi.org/10.1111/gcb.13910, 2018.

Foster, A. C., Hawkins, L. R., Kennedy, D., Bonan, G. B., Fisher, R. A., Needham, J. F., Knox, R. G., Koven, C. D., Wieder, W. R., Dagon, K., and Lawrence, D. M.: Contrasting parametric sensitivities in two global vegetation models using parameter perturbation ensembles, J. Adv. Model. Earth Syst., 18, e2025MS005590, https://doi.org/10.1029/2025MS005590, 2026.

Friend, A.D., Eckes-Shephard, A.H., Fonti, P. et al. On the need to consider wood formation processes in global vegetation models and a suggested approach. Annals of Forest Science 76, 49 (2019). https://doi.org/10.1007/s13595-019-0819-x

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Friend, A. D., et al.: Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO_2, P. Natl. Acad. Sci. USA, 116, 24563–24573, https://doi.org/10.1073/pnas.1903740116, 2019.

Gaskin, T., Pavliotis, G. A., & Girolami, M. (2023). Neural parameter calibration for large-scale multiagent models. Proceedings of the National Academy of Sciences, 120(7), e2216415120. https://doi.org/10.1073/pnas.2216415120

Graven, H. D., Warren, H., Gibbs, H. K., Khatiwala, S., Koven, C., Landschützer, P., Levin, I., Shi, M., & Wieder, W. (2024). Bomb radiocarbon evidence for strong global carbon uptake and turnover in terrestrial vegetation. Science, 384(6702), 1335–1339. https://doi.org/10.1126/science.adl4443

Green, J. K., & Keenan, T. F. (2022). The limits of forest carbon sequestration. Science, 376(6594), 692-693. https://www.science.org/doi/10.1126/science.abo6547

Guo, D., Xia, M., Wei, X., Chang, W., Liu, Y., & Wang, Z. (2008). Anatomical traits associated with absorption and mycorrhizal colonization are linked to root branch order in twenty-three Chinese temperate tree species. New Phytologist, 180, 673–683.

Gupta, H. V., Clark, M. P., Vrugt, J. A., Abramowitz, G., & Ye, M. (2012). Towards a comprehensive assessment of model structural adequacy. Water Resources Research, 48(8).

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Grime, J. P.: Vegetation classification by reference to strategies, Nature, 250, 26–31, https://doi.org/10.1038/250026a0, 1974.

Gupta, H. V., Clark, M. P., Vrugt, J. A., Abramowitz, G., and Ye, M.: Towards a comprehensive assessment of model structural adequacy, Water Resour. Res., 48, W08301, https://doi.org/10.1029/2011WR011044, 2012.

Gyllingberg, L., Birhane, A., and Sumpter, D. J. T.: The lost art of mathematical modelling, Math. Biosci., 362, 109033, https://doi.org/10.1016/j.mbs.2023.109033, 2023.

Haario, H., Saksman, E., and Tamminen, J.: An adaptive Metropolis algorithm, Bernoulli, 7, 223–242, https://doi.org/10.2307/3318737, 2001.

Hendrick, R. L. and Pregitzer, K. S.: The dynamics of fine root length, biomass, and nitrogen content in two northern hardwood ecosystems, Can. J. Forest Res., 23, 2507–2520, https://doi.org/10.1139/x93-312, 1993.

Herman, J. and Usher, W.: SALib: An open-source Python library for sensitivity analysis, J. Open Source Softw., 2, 97, https://doi.org/10.21105/joss.00097, 2017.

He, Y., Ding, J., Dorji, T., Wang, T., Li, J., & Smith, P. 2022. Observation-based global soil heterotrophic respiration indicates underestimated turnover and sequestration of soil carbon by terrestrial ecosystem models. Global Change Biology. https://doi.org/10.1111/gcb.16286

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Haario, H., Saksman, E., & Tamminen, J. (2001). An adaptive Metropolis algorithm.

Hoffman, F. M., Hassler, B., Swaminathan, R., Lewis, J., Andela, B., Collier, N., Hegedűs, D., Lee, J., Pascoe, C., Pflüger, M., Stockhause, M., Ullrich, P., Xu, M., Bock, L., Chun, F., Gier, B. K., Kelley, D. I., Lauer, A., Lenhardt, J., . . . Dunne, J. P. (2025). Rapid evaluation framework for the CMIP7 assessment fast track. EGUsphere. https://doi.org/10.5194/egusphere-2025-2685

Iversen, C. M., McCormack, M. L., Powell, A. S., Blackwood, C. B., Freschet, G. T., Kattge, J., Roumet, C., Stover, D. B., Soudzilovskaia, N. A., Valverde-Barrantes, O. J., van Bodegom, P. M., & Violle, C. (2017). A global Fine-Root Ecology Database to address below-ground challenges in plant ecology. New Phytologist, 215, 15–26.

Jackson, R. B., Canadell, J., Ehleringer, J. R., Mooney, H. A., Sala, O. E., and Schulze, E. D.: A global analysis of root distributions for terrestrial biomes, Oecologia, 108, 389–411, https://doi.org/10.1007/BF00333714, 1996.

Kakouridis, A., Hagen, J. A., Kan, M. P., Mambelli, S., Feldman, L. J., Herman, D. J., Weber, P. K., Pett-Ridge, J., and Firestone, M. K.: Routes to roots: direct evidence of water transport by arbuscular mycorrhizal fungi to host plants, New Phytol., 236, 210–221, https://doi.org/10.1111/nph.18281, 2022.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Kattge, J., et al.: TRY plant trait database — enhanced coverage and open access, Glob. Change Biol., 26, 119–188, https://doi.org/10.1111/gcb.14904, 2020.

Körner, C., Morgan, J., & Norby, R. (2007). CO2 fertilization: when, where, how much?. In Terrestrial ecosystems in a changing world (pp. 9-21). Berlin, Heidelberg: Springer Berlin Heidelberg.

Kou-Giesbrecht, S., Reis Ely, C. R., Perakis, S. S., et al. (2025). Overestimated natural biological nitrogen fixation translates to an exaggerated CO2 fertilization effect in Earth system models. Proceedings of the National Academy of Sciences, 122(48), e2514628122. https://doi.org/10.1073/pnas.2514628122

Leung, L. R., Bader, D. C., Taylor, M. A., & McCoy, R. B. (2020). An introduction to the E3SM special collection: Goals, science drivers, development, and analysis. Journal of Advances in Modeling Earth Systems, 12(11), e2019MS001821. https://doi.org/10.1029/2019MS001821

Leuzinger, S., Manusch, C., Bugmann, H., & Wolf, A. (2013). A sink-limited growth model improves biomass estimation along boreal and alpine tree lines. Global Ecology and Biogeography, 22, 924-932. https://doi.org/10.1111/geb.12047

Levins, R. (1966). The Strategy of Model Building in Population Biology. American Scientist, 54(4), 421–431. https://www.jstor.org/stable/27836590

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Luo, Y., & Schuur, E. A. G. (2020). Model parameterization to represent processes at unresolved scales and changing properties of evolving systems. Global Change Biology, 26(3), 1109–1117. https://doi.org/10.1111/gcb.14939

Liu, J., Sleeter, B. M., Zhu, Z., Cochrane, M. A., et al. (2025). Land change, fire, and climate weaken carbon sink in the conterminous United States. Science Advances, 11(46). https://doi.org/10.1126/sciadv.adx7823

McCormack, M. L., Pritchard, S. G., Breland, S., Davis, M. A., Prior, S. A., Brett Runion, G., ... & Rogers, H. H. (2010). Soil fungi respond more strongly than fine roots to elevated CO_2 in a model regenerating longleaf pine-wiregrass ecosystem. Ecosystems, 13, 901-916.

McCormack, M. L., Adams, T. S., Smithwick, E. A., & Eissenstat, D. M. (2012). Predicting fine root lifespan from plant functional traits in temperate trees. New Phytologist, 195, 823–831.

McCormack, M. L., Dickie, I. A., Eissenstat, D. M., Fahey, T. J., Fernandez, C. W., Guo, D., Helmisaari, H.-S., Hobbie, E. A., Iversen, C. M., Jackson, R. B., Leppälammi-Kujansuu, J., Norby, R. J., Phillips, R. P., Pregitzer, K. S., Pritchard, S. G., Rewald, B., and Zadworny, M.: Redefining fine roots improves understanding of below-ground contributions to terrestrial biosphere processes, New Phytol., 207, 505–518, https://doi.org/10.1111/nph.13363, 2015.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

McElrone, A. J., Pockman, W. T., Martínez-Vilalta, J., and Jackson, R. B.: Variation in xylem structure and function in stems and roots of trees to 20 m depth, New Phytol., 163, 507–517, https://doi.org/10.1111/j.1469-8137.2004.01127.x, 2004.

Mooney, H. A.: The carbon balance of plants, Annu. Rev. Ecol. Syst., 3, 315–346, https://doi.org/10.1146/annurev.es.03.110172.001531, 1972.

Moorcroft, P. R., Hurtt, G. C., and Pacala, S. W.: A method for scaling vegetation dynamics: the Ecosystem Demography model (ED), Ecol. Monogr., 71, 557–586, https://doi.org/10.1890/0012-9615(2001)071[0557:AMFSVD]2.0.CO;2, 2001.

Nair, R.: Rooting vegetation models in realism, Glob. Change Biol., 29, 2868–2870, https://doi.org/10.1111/gcb.16699, 2023.

Nartallo-Kaluarachchi, R., Lambiotte, R., and Goriely, A.: From reductionism to realism: holistic mathematical modelling for complex biological systems, J. R. Soc. Interface, 22, 20250468, https://doi.org/10.1098/rsif.2025.0468, 2025.

Norby, R. J., Ledford, J., Reilly, C. D., Miller, N. E., & O'Neill, E. G. (2004). Fine-root production dominates response of a deciduous forest to atmospheric CO2 enrichment. Proceedings of the National Academy of Sciences (PNAS), 101(26), 9689-9693. https://doi.org/10.1073/pnas.0403491101

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Oleson, K. W., et al.: Technical Description of version 4.5 of the Community Land Model (CLM), NCAR Tech. Note NCAR/TN-503+STR, National Center for Atmospheric Research, Boulder, CO, https://doi.org/10.5065/D6RR1W7M, 2013.

Pastorello, G., Trotta, C., Canfora, E. et al. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. Sci Data 7, 225 (2020). https://doi.org/10.1038/s41597-020-0534-3

Paul MJ, Foyer CH. 2001. Sink regulation of photosynthesis. Journal of Experimental Botany 52: 1383–1400.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É.: Scikit-learn: machine learning in Python, J. Mach. Learn. Res., 12, 2825–2830, 2011.

Poorter, H., Niklas, K. J., Reich, P. B., Oleksyn, J., Poot, P., & Mommer, L. (2012). Biomass allocation to leaves, stems and roots: Meta-analyses of interspecific variation and environmental control. New Phytologist, 193(1), 30–50. https://doi.org/10.1111/j.1469-8137.2011.03952.x

Pregitzer, K. S., DeForest, J. L., Burton, A. J., Allen, M. F., Ruess, R. W., & Hendrick, R. L. (2002). Fine root architecture of nine North American trees. Ecological Monographs, 72, 293–309.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Randerson, J. T., Li, Y., Fu, W., Primeau, F., ... & Goulden, M. L. (2025). The weak land carbon sink hypothesis. Science Advances, 11(37). https://doi.org/10.1126/sciadv.adr5489

Raunkiaer, C.: The Life Forms of Plants and Statistical Plant Geography, Clarendon Press, Oxford, 1934.

Reich, P. B. (2014). The world-wide 'fast–slow' plant economics spectrum: a traits manifesto. Journal of Ecology, 102(2), 275-301. https://doi.org/10.1111/1365-2745.12211

Reid, J. B., Sorensen, I., & Petrie, R. A. (1993). Root demography in kiwifruit (*Actinidia deliciosa*). Plant, Cell & Environment, 16, 949-957.

Roberts, G. O., Gelman, A., and Gilks, W. R.: Weak convergence and optimal scaling of random walk Metropolis algorithms, Ann. Appl. Probab., 7, 110–120, https://doi.org/10.1214/aoap/1034625254, 1997.

Robin, A., Pradier, C., Sanguin, H., Mahé, F., Lambais, G. R., de Araujo Pereira, A. P., Germon, A., Santana, M. C., Tisseyre, P., Pablo, A. L., Heuillard, P., Sauvadet, M., Bouillet, J. P., Andreote, F. D., Plassard, C., de Moraes Gonçalves, J. L., Cardoso, E. J. B. N., Laclau, J. P., Hinsinger, P., & Jourdan, C. (2019). How deep can ectomycorrhizas go? A case study on Pisolithus down to 4 meters in a Brazilian eucalypt plantation. Mycorrhiza, 29, 637–648.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Rockström, J., Beringer, T., Hole, D., Griscom, B., Mascia, M. B., Folke, C., & Creutzig, F. (2021). We need biosphere stewardship that protects carbon sinks and builds resilience. Proceedings of the National Academy of Sciences, 118(38), e2115218118.

Root, R. B.: The niche exploitation pattern of the blue-gray gnatcatcher, Ecol. Monogr., 37, 317–350, https://doi.org/10.2307/1942327, 1967.

Saltelli, A.: Making best use of model evaluations to compute sensitivity indices, Comput. Phys. Commun., 145, 280–297, https://doi.org/10.1016/S0010-4655(02)00280-1, 2002.

Sainsbury-Dale, M., Zammit-Mangion, A., & Huser, R. (2024). Likelihood-free parameter estimation with neural Bayes estimators. The American Statistician, 78(1), 1–14. https://doi.org/10.1080/00031305.2023.2249522

Schimel, D., & Miller, C. (2023). Do two climate wrongs make a right? AGU Advances, 4, e2023AV001020.

Smith, T. M., Shugart, H. H., & Woodward, F. I. (1997). Plant functional types: Their relevance to ecosystem properties and global change. Cambridge University Press.

See, C. R., McCormack, M. L., Hobbie, S. E., Flores-Moreno, H., Silver, W. L., & Kennedy, P. G. (2019). Global patterns in fine root decomposition: climate, chemistry, mycorrhizal association and woodiness. Ecology Letters, 22(6), 946–953. https://doi.org/10.1111/ele.13248

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Shugart, H. H. (1984). A theory of forest dynamics. Springer.

Shugart, H. H., Harris, W. F., Edwards, N. T., and Ausmus, B. S.: Modelling of belowground processes associated with roots in deciduous forest ecosystems, Pedobiologia, 17, 382–388, 1977.

Shugart, H. H., B. Wang, R. Fischer, J. Ma, J. Fang, X. Yan, A. Huth, and A. H. Armstrong, "Gap models and their individual-based relatives in the assessment of the consequences of global change," Environmental Research Letters, 2018. https://doi.org/10.1088/1748-9326/aaaacc

Smith, S. E. and Read, D. J.: Mycorrhizal Symbiosis, 2nd edn., Academic Press, San Diego, California, 1997.

Sobol', I. M.: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Math. Comput. Simul., 55, 271–280, https://doi.org/10.1016/S0378-4754(00)00270-6, 2001.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res.-Atmos., 106, 7183–7192, https://doi.org/10.1029/2000JD900719, 2001.

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

Thornton, P. E., Lamarque, J.-F., Rosenbloom, N. A., and Mahowald, N. M.: Influence of carbon–nitrogen cycle coupling on land model response to CO_2 fertilization and climate variability, Global Biogeochem. Cycles, 21, GB4018, https://doi.org/10.1029/2006GB002868, 2007.

Walker, W. S., Gorelik, S. R., Cook-Patton, S. C., Baccini, A., Farina, M. K., Solvik, K. K., ... & Griscom, B. W. (2022). The global potential for increased storage of carbon on land. Proceedings of the National Academy of Sciences, 119, e2111312119.

Wang, B. (2026). ELM-TAM: a structure-based, function-oriented land model embracing fine-root system complexity [Data set]. Zenodo. https://doi.org/10.5281/zenodo.20310624

Wang, B., & Allison, S. D. (2021). Drought legacies mediated by trait trade-offs in soil microbiomes. Ecosphere, 12(6), e03562.

Wang, B., et al.: Embracing fine-root system complexity in terrestrial ecosystem modelling, Glob. Change Biol., 29, 2494–2508, https://doi.org/10.1111/gcb.16659, 2023.

Wang, H., Harrison, S. P., Li, M., Prentice, I. C., Qiao, S., Wang, R., Xu, H., Mengoli, G., Peng, Y., and Yang, Y.: The China plant trait database version 2, Sci. Data, 9, 769, https://doi.org/10.1038/s41597-022-01884-4, 2022.

Weigelt, A., Mommer, L., Andraczek, K., Iversen, C. M., Bergmann, J., Bruelheide, H., Fan, Y., Freschet, G. T., Guerrero-Ramírez, N. R., Kattge, J., Kuyper, T. W., Laughlin, D. C., Meier, I. C.,

https://doi.org/10.5194/egusphere-2026-2759
Preprint. Discussion started: 24 July 2026

van der Plas, F., Poorter, H., Roumet, C., van Ruijven, J., Sabatini, F. M., Semchenko, M., … McCormack, M. L. (2021). An integrated framework of plant form and function: The belowground perspective. New Phytologist, 232, 42–59.

Zaehle, S. and Friend, A. D.: Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates, Global Biogeochem. Cycles, 24, GB1005, https://doi.org/10.1029/2009GB003521, 2010.

Zhang, X., Bin Wang, Herman Shugart, Forrest Hoffman, Robert Patton. Toward Continental Forest Dynamics Modeling at Individual-tree Resolution on Exascale Systems. The international conference for high performance computing networking storage, and analysis 2026. (*in review*)