Received: 22 July 2024 | Accepted: 21 May 2025

DOI: 10.1111/2041-210X.70076

BRITISH ECOLOGICAL SOCIETY Methods in Ecology and Evolution

APPLICATION

pnetr: An R package for the PnET family of forest ecosystem models

Xiaojie Gao[1] | Zaixing Zhou[2] | Scott V. Ollinger[2] | Jaclyn Hatala Matthes[1] | Wenzhe Jiao[3] | Jonathan R. Thompson[1]

[1]Harvard Forest, Harvard University, Petersham, Massachusetts, USA

[2]Earth Systems Research Center, University of New Hampshire, Durham, New Hampshire, USA

[3]Department of Ecology and Conservation Biology, Texas A&M University, College Station, Texas, USA

Correspondence
Xiaojie Gao
Email: xiaojiegao@fas.harvard.edu

Funding information
National Science Foundation > Division of Environmental Biology, Grant/Award Number: 18-32210, 22-31681, 19-26454 and 22-05705; New Hampshire Agricultural Experiment Station, Grant/Award Number: NH00711; National Science Foundation > Office of Integrative Activities, Grant/Award Number: 19-20908

Handling Editor: Nicholas Clark

Abstract

1. Ecosystem models offer a rigorous way to formalize scientific theories and are critical to evaluating complex interactions among ecological and biogeochemical processes. In addition to simulation and prediction, ecosystem models are a valuable tool for testing hypotheses about mechanisms and empirical findings because they reveal critical internal processes that are difficult to observe directly.
2. However, many ecosystem models are difficult to manage and apply by scientists because of complex model structures, lack of consistent documentation, and low-level programming implementation.
3. Here, we present the '*pnetr*' R package, which is designed to provide an easy-to-manage ecosystem modelling framework and detailed documentation in both model structure and programming. The framework implements a family of widely used PnET (net photosynthesis, evapotranspiration) ecosystem models, which are relatively parsimonious but capture essential biogeochemical cycles of water, carbon and nitrogen. We chose the R programming language because it is familiar to many ecologists and has abundant statistical modelling resources. We showcase examples of model simulations and test the effects of phenology on carbon assimilation and wood production using data measured by the Environmental Measurement Station (EMS) eddy-covariance flux tower at Harvard Forest, MA.
4. We hope '*pnetr*' can facilitate further development of ecological theory and increase the accessibility of ecosystem modelling and ecological forecasting.

KEYWORDS
biogeochemistry, community ecology, ecological process, ecosystem modelling, PnET, Software

1 | INTRODUCTION

Ecosystem models integrate ecological processes to elucidate how ecosystems function through interactions between biology, chemistry and physics (Bacmeister et al., 2020; Bonan et al., 2019; Geary et al., 2020; Middendorp et al., 2016; Scheller et al., 2010, 2012). Ecosystem models play a critical role in evaluating the interactions among multiple ecological processes (Liang et al., 2023;

Thompson et al., 2011), investigating the causal effects of certain phenomena (Gustafson et al., 2017; Liang et al., 2018) and simulating potential future scenarios (De Bruijn et al., 2014; Prinn, 2013; Shifley et al., 2017). Complementary to statistical correlations, ecosystem models enable scientists to test new hypotheses by directly comparing multiple alternatives in ecological processes and investigating complex feedbacks, interactions and consequences. The integration of ecological forecasting and model-data fusion techniques provides a powerful tool for hypothesis testing in ecosystem models by predicting short- and long-term changes and aligning them with multi-source new observations (Dietze, 2017; Gettelman et al., 2022; Lewis et al., 2023). In addition, since not all variables are easily observable, conducting ecological forecasting using ecosystem models also provides an opportunity to link unobservable to observable processes and evaluate hypotheses in an integrated systematic view (Lewis et al., 2023; Liu et al., 2021). More importantly, these practices help update existing ecosystem model structures as new knowledge becomes available, which is crucial to advancing ecological theory.

However, despite recent technological developments, testing hypotheses in most ecosystem models and updating model structures are not trivial tasks. Existing model frameworks, such as the LANDIS-II Forest landscape model (Scheller & Mladenoff, 2004), the Functionally Assembled Terrestrial Ecosystem Simulator (FATES; Massoud et al., 2019) and the Lund-Potsdam-Jena General Ecosystem Simulator (LPJ-GUESS; Smith, 2001), have several limitations in the applications of hypothesis testing. First, for computational efficiency, existing ecosystem modelling frameworks are normally implemented in programming languages that are more widely used in software engineering, such as C++, C# and FORTRAN (Huang et al., 2020; Hurrell et al., 2013; Koven et al., 2020; Ma et al., 2022; Rastetter et al., 2022; Scheller et al., 2010). These implementations increase computing speed but can also pose technical barriers for ecologists to modify model structures. Collaborating with the corresponding software developers is one solution but it slows the iteration process due to communication and transferring materials. Consequently, while many new scientific findings have led to suggestions for new or altered representations of certain ecological processes (e.g. Richardson et al., 2012; Stocker et al., 2019), implementing these suggestions is often far from straightforward. For instance, although more realistic process-based models have been developed to simulate vegetation phenology over the past two decades (Basler, 2016; Chuine & Régnière, 2017; Gao et al., 2024; Teets et al., 2023), phenological processes in most ecosystem models have not been updated. This delay in updates may hinder the investigation of how current phenological shifts induced by climate change interact with other ecological processes and influence future ecosystem functions. Second, many existing models lack clear and consistent documentation in both model structures and programming. Commonly, a process in the model has been updated but the documentation has not, which increases the difficulty for users to catch up and can generate misinterpretation of the model results. Although the complexity of ecosystem models typically involves some preliminary knowledge, consistent documentation with version control can help simplify the difficulty and increase accessibility (Shifley et al., 2017). Third, due to their complex structures, many ecosystem modelling frameworks are difficult to manage by non-programmers in terms of preparing input data sets, understanding model details that are beyond their research domain, and modifying relevant processes for their use cases. Ecosystem models are meant to be complex; however, aggregation and simplification are helpful in many scientific explorations. Parsimonious models can facilitate scientific hypothesis testing by simulating the essential ecological processes but removing the ones that are too complex and not directly related to the scientific questions. Ecologists can be more confident in modifying parsimonious models without worrying about making mistakes in other processes.

To facilitate hypothesis testing using ecosystem models and updating model structures, we developed the '*pnetr*' ecosystem modelling framework using the statistical R language to implement the family of the PhotosyNthesis and EvapoTranspiration (PnET) ecosystem models (Gao, 2025). The PnET model family is a collection of forest ecosystem models that have been actively developed since the 1990s (Aber et al., 1995, 1993, 1997; Aber & Federer, 1992). Compared to many ecosystem models, PnET models are relatively parsimonious because they only focus on essential processes of water, carbon and nitrogen cycling and simplify processes in other aspects. The original PnET model operated at a monthly time step; it was developed in Visual Basic and later updated to C++. Another version based on C# was also developed for integration into the LANDIS-II landscape simulation platform (De Bruijn et al., 2014; Gustafson et al., 2023). Source code for some old versions or sub-models of the PnET family was also adapted to MATLAB and R; however, limited documentation is available. Compared to previous implementations of PnET, here we provide a comprehensive but easy-to-manage modelling framework, as well as detailed algorithm documentation for major sub-models in the PnET family. Our implementation supports both daily and monthly time steps in a consistent manner for all models. We chose the R statistical programming language to make the source code easy to modify and to better assist ecologists with the abundant statistical tools in the community for conducting scientific practices such as hypothesis testing, ecological forecasting and model data fusion.

2 | PnET AND '*PNETR*'

The PnET model family consists of a nested series of models simulating carbon, water and nitrogen dynamics in forest ecosystems. These simple and lumped-parameter models are built on two principal relationships: (1) maximum photosynthetic rate is a function of foliar nitrogen concentration and (2) transpiration is a function of realized photosynthetic rate (Aber et al., 1995, 1993, 1997; Aber & Federer, 1992). The PnET models have been successfully used to predict gross and net primary productivity (GPP and NPP), carbon and water balances and nitrogen dynamics in forest ecosystems responding to changes in climate, nitrogen deposition, land use and species composition at both site and grid levels (De Bruijn et al., 2014; Liang et al., 2023; Ollinger et al., 1998; Zhou et al., 2018).

Our '*pnetr*' R package provides three major PnET sub-models, including PnET-Day (Aber et al., 1996), PnET-II (Aber et al., 1995; Aber & Federer, 1992) and PnET-CN (Aber et al., 1997). The sub-models differ in structures and perspectives but share common routines (Figure 1). PnET-Day is the simplest model in the family; it only simulates the photosynthetic processes at leaf and canopy scales. PnET-II includes additional routines concerning water and carbon cycles, and PnET-CN further includes nitrogen cycles with complete feedback between carbon, water and nitrogen. The functions 'PnET_Day()', 'PnET_II()' and 'PnET-CN()' in the package correspond to each model simulation. These functions take input data describing climate conditions, site and vegetation characteristics (Table 1) and output various variables describing the water, carbon and nitrogen cycles. Both input and output data vary by the sub-models. The required climate data generally include maximum (*T*max) and minimum temperature (*T*min), photosynthetic active radiation (PAR), precipitation, CO_2 and O_3 concentration in the air and NO_3 and NH_4 deposition. The climate data can be daily, weekly or monthly, and its resolution determines the time-step of the model simulation. The site information includes the latitudinal location, soil water holding capacity and initial snowpack. The vegetation characteristics include a series of empirical values indicating information such as foliar nitrogen concentration, maximum foliar mass, phenological requirements and respiration rate. Default values for most vegetation characteristics are provided. The output variables generally include gross primary productivity (GPP), net primary productivity (NPP), net ecosystem productivity (NEP), evapotranspiration (ET) and biomass at each time step of the simulation as well as their monthly and annual summaries. The full list of the inputs and outputs, as well as example data sets, can be found in the package documentation. We also provide tools for sensitivity analysis so that users can investigate the relationships between multiple variables when they modify the model structures.

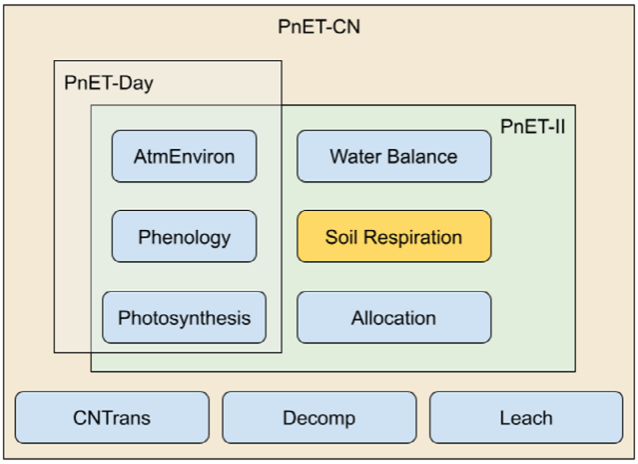

FIGURE 1 Major PnET sub-models and their components. Note that 'Soil Respiration' is included in the 'Decomp' routine when using PnET-CN.

There are nine major routines in the PnET models (Figure 1). '*AtmEnviron*' deals with the atmospheric environment and calculates meteorological variables such as average temperatures and vapour pressure deficit needed to run the simulations. '*Phenology*' controls the timing of plants' critical life-cycle events such as leaf development, woody growth, leaf senescence and dormancy. '*Photosynthesis*' concerns the processes in carbon assimilation through photosynthesis, which are driven by foliar nitrogen concentration and affected by phenology and environmental factors such as light radiation, temperature and water stress. '*Water Balance*' simulates the water cycle in the field including precipitation, snow and evapotranspiration. '*Soil Respiration*' quantifies the amount of carbon

TABLE 1 Some important input parameters for PnET model simulations (The full list and the corresponding default values can be found in the 'pnetr' package documentation).

Variable name	Description	Unit
*T*max, *T*min	Maximum/minimum daily/monthly temperature	°C
PAR	Photosynthetic active radiation	micromoles $m^{-2}s^{-1}$
Prec	Total amount of precipitation	mm day^{-1} ($month^{-1}$)
CO_2	CO_2 concentration in air (PnET-CN)	parts per million (ppm)
O_3	Cumulative daytime O_3 exposure (PnET-CN)	Dose >40 parts per billion (ppb)
NO3dep	Nitrogen deposition through NO_3 (PnET-CN)	g N m^{-2}
NH4dep	Nitrogen deposition through NH_4 (PnET-CN)	g N m^{-2}
Lat	Site latitude	degrees
WHC	Site water hold capacity	cm
SnowPack	Site initial snowpack on the first month (day)	cm equivalent water
GDDFolStart, GDDFolEnd	Growing degree days at which foliage production onsets/ends	degree days
SenescStart	Day of year after which leaf drop can occur	day of year
FolNCon	Foliar nitrogen concentration (% by weight)	g N 100 g^{-1} dry mass

emitted from respiration in soil. '*Allocation*' is the process that allocates the assimilated carbon and/or nitrogen into different pools in the plant including the whole plant pool, the bud pool, the wood pool and the root pool. Allocation only occurs monthly and annually. '*CNTrans*', '*Decomp*' and '*Leach*' determine the nitrogen cycle, which dynamically computes foliar nitrogen concentration over the years. The nitrogen cycle includes processes such as nitrogen accumulation, mineralization, immobilization, nitrification and leaching losses. The effects of CO_2 fertilization (Ollinger et al., 2002) and ozone concentration (Ollinger et al., 1997) on photosynthesis are also included in the PnET-CN model.

A simple schematic diagram of the PnET-CN model illustrates the major processes included in the PnET models (Figure 2). During the growing season, carbon is fixed through leaf photosynthesis, quantified by GPP and released by respiration at the same time. The assimilated carbon is then transported to the plant carbon mobile pool and allocated to grow leaves, wood and roots after deduction of autotrophic respiration. This net carbon storage is represented by net primary productivity (NPP). Carbon emissions from the decay of plant detritus and decomposition of soil organic matter are also quantified. The rate of photosynthesis of all PnET models depends on the nitrogen concentration in the leaves and is affected by environmental factors including light radiation, temperature, vapour pressure deficit and water stress. Foliar nitrogen in PnET-CN is estimated as part of the nitrogen cycle, whereas it is a user input in PnET-Day and PnET-II. The only source of water is precipitation. Depending on the temperature and site conditions, water entering the system can be in the form of rain and/or snow. The amount of water available for plant uptake is affected by processes such as fast flow and drainage. Plants lose water through evapotranspiration. The nitrogen cycle is coupled with the carbon cycle. The foliar nitrogen concentration determines the maximum net photosynthesis rate without water stress, which increases the internal non-structure plant carbon pool and thus NPP. Conversely, the increased plant NPP will increase the nitrogen demand, which reduces the non-structural plant nitrogen pool. The degree to which nitrification occurs is negatively correlated with the strength of plant demand for nitrogen in competition with nitrifiers. The amount of $NO3^-$ and $NH4^+$ in the soil solution affects plant nitrogen uptake and the nitrate leaching rate.

3 | WORKING EXAMPLES

Here we use three working examples including carbon dynamics simulation, sensitivity analysis and hypothesis testing to demonstrate

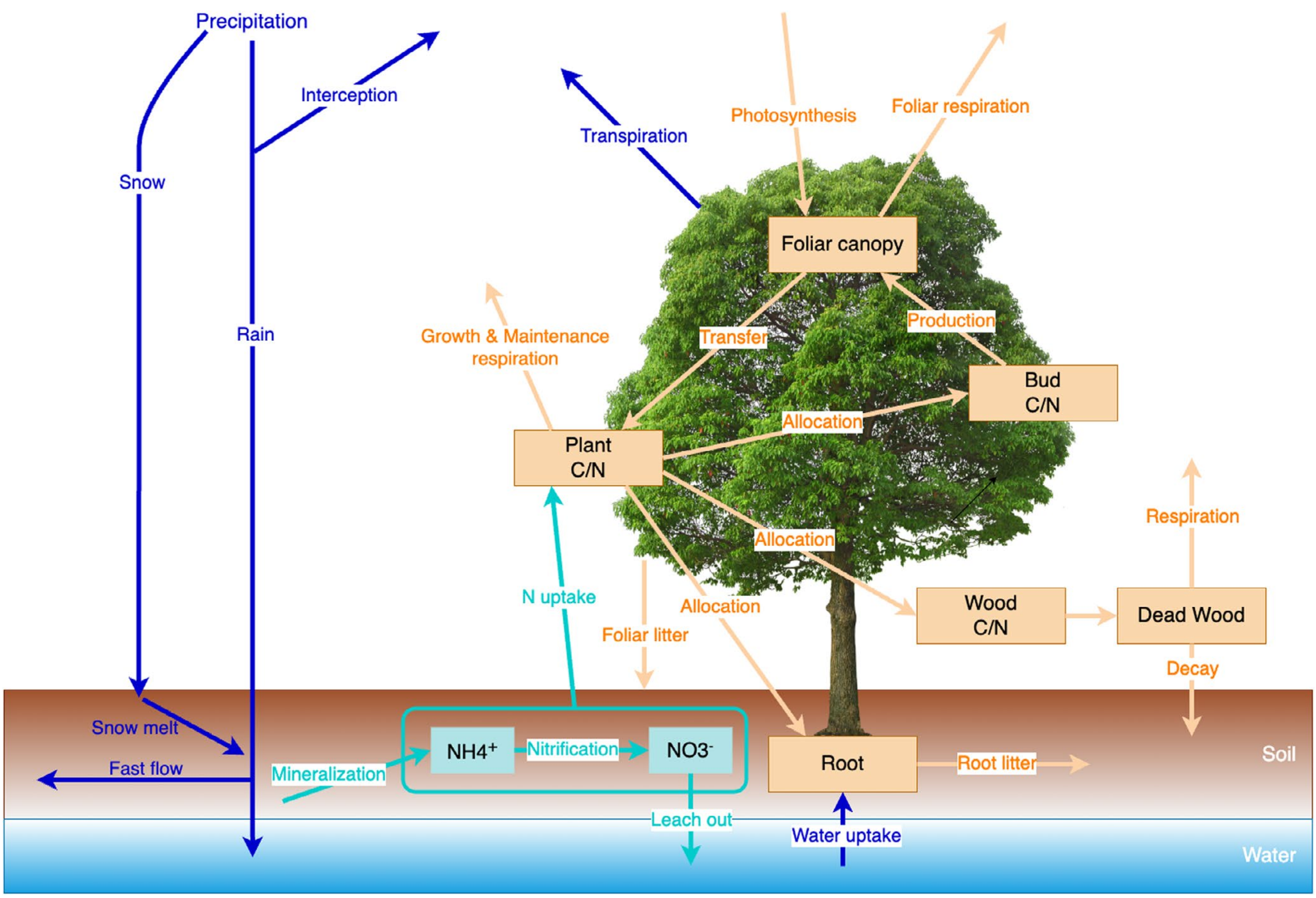

FIGURE 2 PnET-CN schematic diagram. Arrows in colour represent the major processes in the biogeochemical cycles including carbon (yellow), water (blue) and nitrogen (cyan).

the general usage of the 'pnetr' R package and how it can facilitate scientific exploration.

3.1 | Simulating ecosystem photosynthesis

As an example, we simulated GPP, NEP and plant internal carbon pools and evaluated model performance using data collected by the Environmental Measurement Station (EMS) eddy covariance (EC) flux tower at the Harvard Forest, MA (Munger & Wofsy, 2024). The example climate data set is included in the package. Air temperature input variables (*T*max, *T*min) and incoming photosynthetic active radiation (PAR) were aggregated to daily values from hourly measurements at the Harvard Forest long-term meteorological station (Boose et al., 2024). The EMS tower measured hourly net ecosystem CO_2 exchange (NEE) over the 1992–2023 time period. The EMS NEE data were filtered to remove periods of low turbulence using a seasonally variable friction velocity threshold (Pastorello et al., 2020) where the growing and non-growing seasons were defined by changes in Landsat-derived leaf greenness seasonality within the tower footprint (Gao et al., 2021). EMS tower GPP was partitioned from NEE by fitting a temperature-light response function to NEE data (Munger & Wofsy, 2024). Tower GPP and NEE were aggregated to daily and monthly sums for comparison to the PnET model output. We converted tower NEE into NEP by taking its negative and compared it with PnET-modelled NEP.

We ran all PnET sub-models at both monthly (Figures 3 and 4) and daily steps (Figure 5). In general, model simulations are reasonably consistent with EC measurements. The simulated GPP values in PnET-Day and PnET-II are similar due to limited water stress at this site. PnET-CN simulated GPP values align better with the EC measurements, suggesting the importance of accounting for the nitrogen cycle effect on photosynthesis. Although all models underestimate the higher GPP and NEP values during 2002–2012, their general variability is fitted

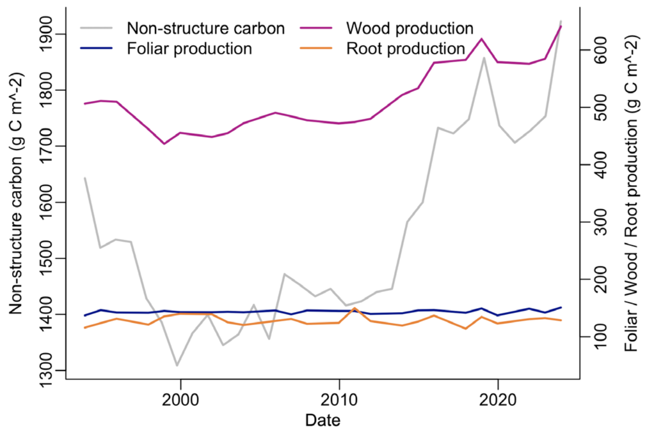

FIGURE 4 Annual non-structural carbon dynamics and net primary production simulated by the PnET-CN model. Annual net primary production is allocated to store carbon in foliar, wood and root production.

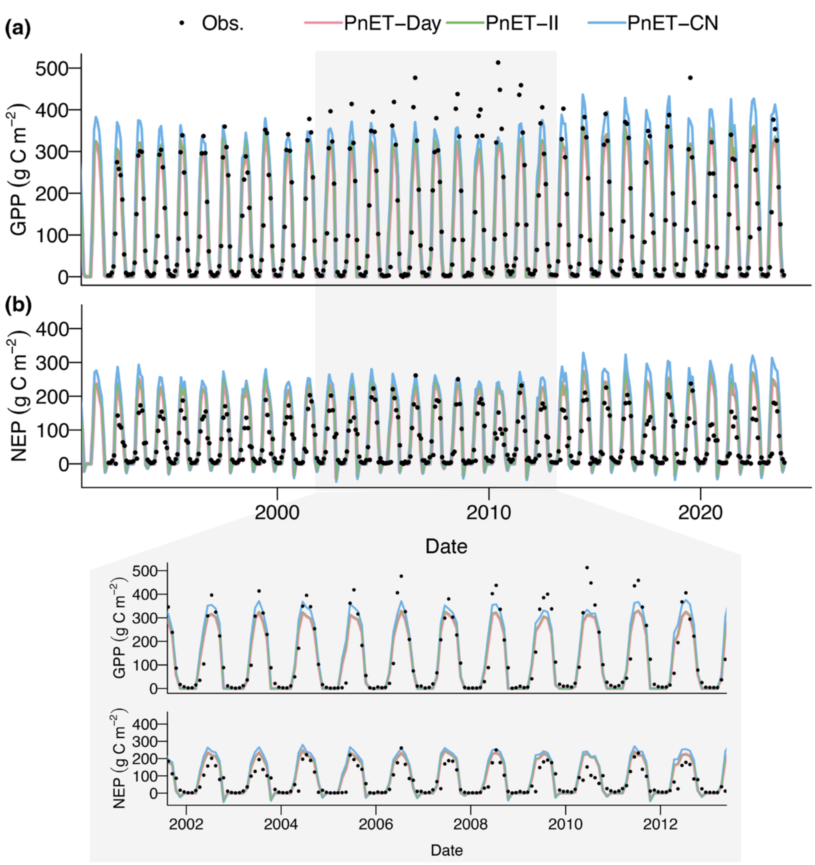

FIGURE 3 PnET models simulated gross primary productivity (a) and net ecosystem productivity (NEP) (b) compared with the measured data collected by the Environmental Measurement Station (EMS) eddy-covariance flux tower installed at Harvard Forest, MA. The EMS-measured net ecosystem CO_2 exchange (NEE) was converted to NEP by taking its negative. Models were simulated in monthly steps. The bottom subfigure shows the period during 2002–2012. Note that the results of PnET-Day and PnET-II are very similar in this case.

BRITISH ECOLOGICAL SOCIETY Methods in Ecology and Evolution

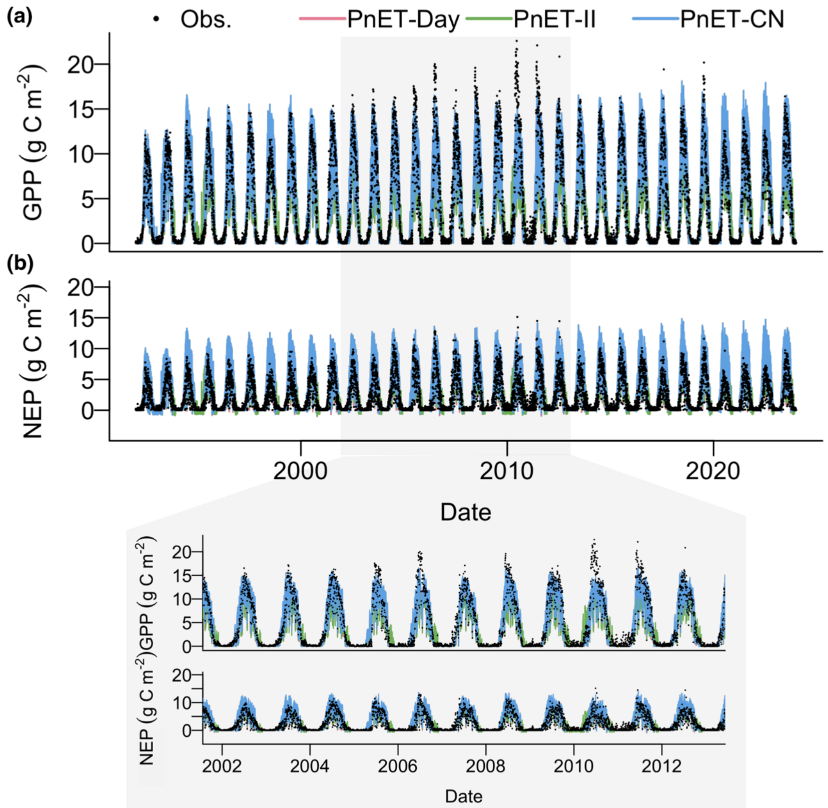

FIGURE 5 PnET models simulated gross primary productivity (a) and net ecosystem productivity (NEP) (b) compared with the measured data collected by the Environmental Measurement Station (EMS) eddy-covariance flux tower installed at Harvard Forest, MA. Note that the EMS-measured net ecosystem CO_2 exchange (NEE) was converted to NEP by taking its negative. Models were simulated in daily steps. The bottom subfigure shows the period during 2002–2012.

reasonably well, especially for GPP. The dynamics of carbon in the plant components show the carbon storage and allocation strategies for above- and below-ground growth (Figure 4).

The daily simulations reveal similar patterns to the monthly simulations but with more pronounced temporal variations (Figure 5). The daily versions of models enable investigating processes operating on shorter temporal scales (e.g. the daily variation of photosynthesis and water balances). Similar to the monthly results, although PnET models reasonably simulated the cyclic variations of carbon exchanges at this site, the consistent underestimation in the summer suggests that future studies are needed to further improve model algorithms and/or parameter values in processes. Indeed, this underestimation of the summer GPP problem at some sites has been previously noted (Zhou et al., 2018).

3.2 | Sensitivity analysis

Sensitivity analysis is often conducted to investigate the sensitivity of model output variables to input and/or internal parameter variations (McKenzie et al., 2019). We encourage users to modify the processes implemented in the '*pnetr*' package for their purposes, so we included a routine for users to easily conduct sensitivity analysis. There are two options of sensitivity analysis available. The first option is a simple automatic sensitivity analysis on the default parameter sets, which can be performed by calling an integrated 'DoSens()' function. With this option, parameter samples are prescribed by varying their default values by 10% using a Markov chain Monte Carlo sampling procedure (for more information, see documentation for 'GenerateSamples()' function). Note that caution needs to be taken using this method with user-defined parameter values, especially for monthly simulations, because slight changes in some parameters may be exaggerated in monthly scales. For example, when running models at a monthly time step, the variables controlling phenology (e.g. start-of-season, start-of-senescence) are measured in units of months. However, since phenological changes often occur within a span of several days, they might fall into a different month. This discrepancy can lead to significant impacts on model simulations because the temporal resolution fails to capture finer-scale phenological shifts. The automatic sample generation can be used as a template for users to modify as needed. The second option is to manually sample parameters from the distributions defined by the users. This option is a bit tedious but gives full control to the users. Once parameter samples are generated, either through the automatic or manual procedure, the model runs with each parameter set and the user-specified output variables (e.g., GPP) are recorded for each run. To reduce computing time, parallel processing implemented in the 'pnetr' package can be used to run multiple sampled parameter sets simultaneously across multiple computer processors. Finally, a random forest regression is performed to evaluate and visualize the importance of each parameter in explaining the target output variables.

As an example, here we show annual GPP and NPP sensitivity to 10% parameter variations in the PnET-CN model running at a monthly step (Figure 6), that is, run the model for 1000 iterations with each iteration randomly changing the default variable values by 10% of the magnitude. The meanings of the parameters can be found

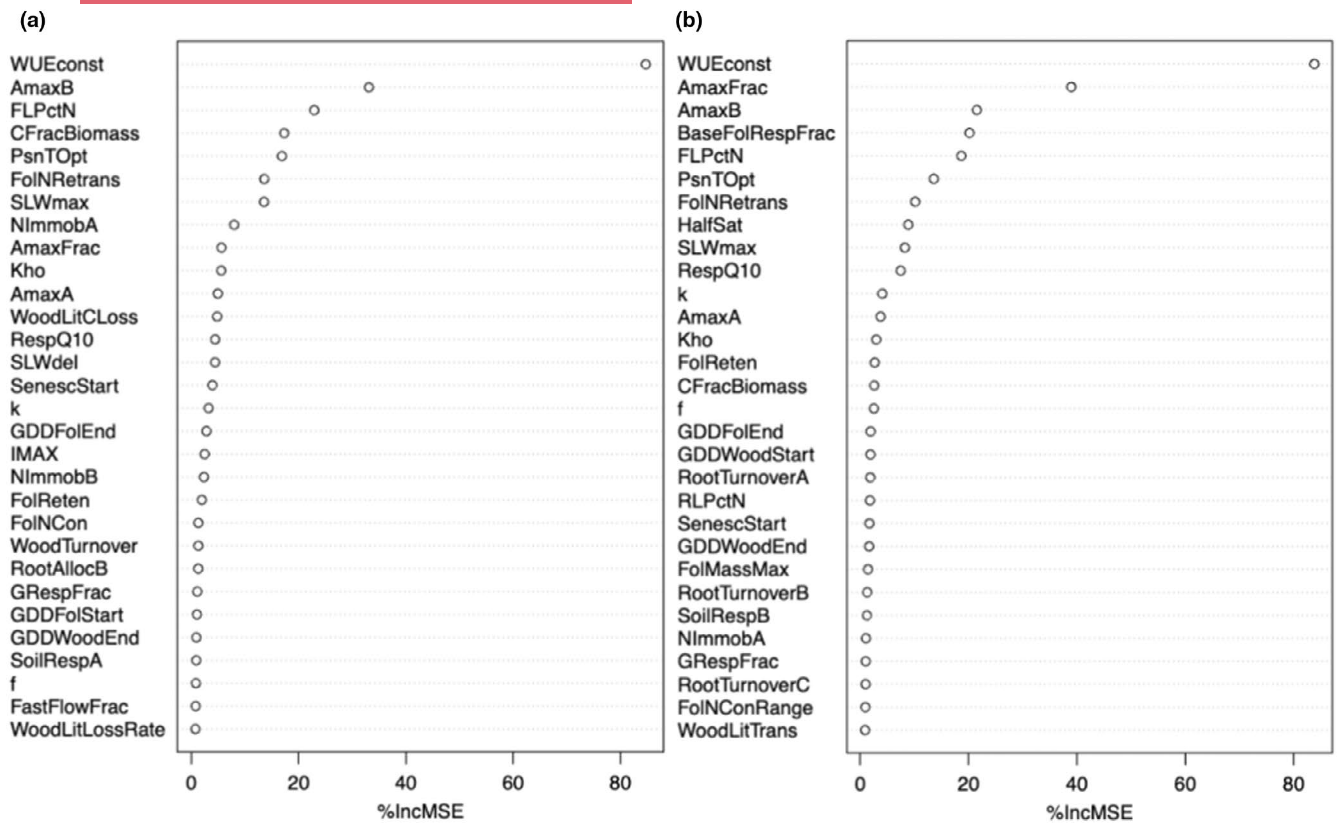

FIGURE 6 Variable importance of the monthly PnET-CN model in simulating annual gross primary productivity (a) and net primary productivity values (b). Variable names and descriptions can be found in the 'pnetr' package documentation.

in the package documentation. More importantly, the result of this sensitivity analysis may not be universal as the variation of individual parameters likely changes among study regions. For example, some parameters showing large sensitivity, such as the constant used to calculate water-use efficiency (WUEconst) from vapour pressure deficit, the slope of the relationship between foliar N and max photosynthetic rate (AmaxB), and the fraction of carbon in foliage mass (CFracBiomass) showed large control to GPP and/or NPP in the example. However, these parameters are usually prescribed using empirical values from observational analysis, so they typically produce less variability in model predictions than in an unconstrained sensitivity analysis. Therefore, this example is to show the sensitivity analysis feature of the package; to acquire useful insights about model processes, users should carefully determine which parameters are of interest and which parameters are viewed as constants.

3.3 | Hypothesis testing with 'pnetr'

Plant phenology is the first-order control of carbon exchange in vegetated ecosystems. Previous studies show that earlier plant spring leaf-out can significantly increase annual GPP, which measures the total amount of carbon absorbed (Gao et al., 2023; Keenan et al., 2014). However, earlier springs may not always increase wood production depending on many complex processes in carbon sequestration and allocation (Dow et al., 2022). Here, as a simple example, we use the daily PnET-CN model to test the hypothesis that spring phenology is not significantly correlated with wood production.

Specifically, we modified the phenology routine of the PnET-CN model to include historical phenological observations and quantified the relationship between the model simulated annual GPP and spring phenology. The original timing of the start of foliar development in the PnET-CN model is defined as the day-of-year (DOY) when the accumulated growing degree days (GDDs; base temperature 0°C) reach a threshold ('GDDFolStart'; Table 1); and the timing of the start of senescence is defined when the current DOY exceeds a threshold ('SenescStart'; Table 1). We replaced these phenological DOYs with satellite-derived phenology observations obtained from Landsat imagery (1984–2023) at the Harvard Forest EMS site, utilizing a Bayesian land surface phenology model (Gao et al., 2021). To match with the PnET-CN parameterization, we used the time when the EVI2 trajectory increases by 15% of its amplitude as the threshold to determine the 'GDDFolStart' DOY and decreases to 95% of its amplitude to determine the 'SenescStart' DOY (Figure 7a). We then investigated the impacts of historical changes in spring phenology on annual GPP, plant carbon pool (PlantC) and wood production (WoodProdCYr), respectively. The R implementation of the 'pnetr' package makes model modification and visualization simple.

The simulation result shows that earlier spring phenology is significantly correlated with higher annual GPP (Figure 7a; $R^2 = 0.24$;

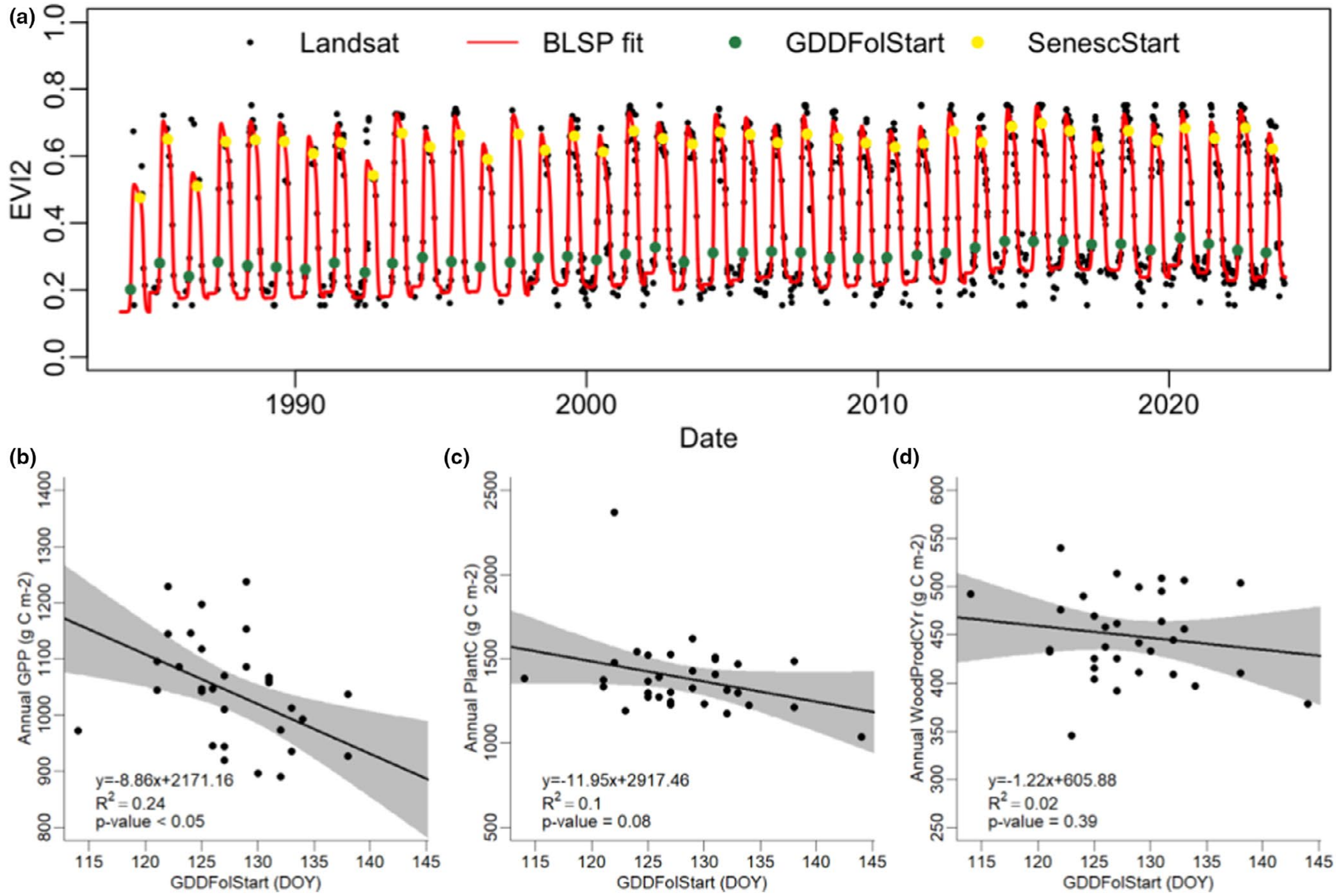

FIGURE 7 Testing the effects of spring and autumn phenology on carbon exchange. Phenology dates in the original PnET-CN model, that is, the start-of-foliar-development ('GDDFolStart') and the start-of-senescence ('SenescStart'), were replaced by phenology retrieved from fitting the two-band enhanced vegetation index (EVI2) time series of Landsat observations during 1984–2023 using a Bayesian land surface phenology (BLSP) model (a). The relationships between spring phenology and annual gross primary productivity (GPP; b), plant carbon pool (PlantC; b) and wood production (WoodProdCYr; d) are shown in scatter plots, in which the linear regression lines, confidence intervals (grey polygons) and statistical summary of the model fits are also shown.

p-value <0.05) and weakly correlated with higher annual PlantC (Figure 7b; $R^2=0.10$; p-value$=0.08$) but not with WoodProdCYr (Figure 7c; $R^2=0.02$; p-value$=0.39$), which suggests that spring phenology significantly affects carbon sequestration but may not directly affect wood production due to complex internal carbon processes. This result is consistent with previous studies (Dow et al., 2022; Keenan et al., 2014), although the simulation is a relatively simple example and the extension of the conclusion needs more investigation. More importantly, however, this simulation example shows that testing hypotheses in ecosystem models can reveal complex interactions between multiple variables and suggest how the apparent relationships propagate internally. Thus, the result of the model simulation can be used as either proof of observational phenomena or as clues for further investigations.

4 | APPLICATIONS

We envision four major applications of the 'pnetr' R package. First, simulating biogeochemical cycles. This is the most direct application. Once the parameters are carefully prescribed at a site, the models can simulate carbon, water and nitrogen cycles to better understand the biogeochemical processes. Using projected future environmental variables, model simulation provides plausible future scenarios, which can be a useful tool to understand future climate change impacts. Second, testing hypotheses. As we have shown, hypothesis testing is straightforward, and it allows investigations of interactions of multiple ecological processes. In addition to our example, the daily scale PnET-CN model has been used in evaluating how different process-based spring phenology models representing different hypotheses affect the simulations of photosynthetic productivity (Teets et al., 2023). The model's embedded processes represent our current ecological theory, and if a new scientific finding can consistently increase the accuracy of model simulations, it is a direct sign of theory advancement. Third, conducting model-data fusion analyses. Since neither the model nor the data are perfect, model-data fusion using Bayesian methods is a useful technique to integrate various big data into model simulations and quantify uncertainty. Model-data fusion can also be used to retrieve internal variables that are difficult to measure (Liu et al., 2021) or to develop near-real-time ecological monitoring systems (Dietze, 2017; Fer et al., 2018). Our R implementation benefits ecologists with abundant statistical resources, especially Bayesian methods, which are critical to model-data fusion. Fourth, providing hands-on examples for teaching. Although the package was developed mainly for scientific exploration, it can also be used as a tool for students to learn biogeochemical processes and get hands-on experience in ecological modelling. The elaborate documentation we provide can help students understand the variables and algorithms in ecology.

5 | CONCLUSION

With the aim of providing an easy-to-manage ecosystem modelling framework that captures essential carbon, water and nutrient processes for ecologists, we developed the 'pnetr' package to implement a family of PnET ecosystem models using the R programming language. We provide detailed documentation about the implemented algorithms to help users better understand the models and be confident in modifying the processes to suit their needs. Compared to more complex ecosystem models, the 'pnetr' package helps scientists focus more on ecology without the burden of acquiring savvy skills in computer science. Additionally, the R implementation gives users easy access to abundant statistical resources for modelling and visualization. Here, we present the package, provide working examples and propose applications. We hope the 'pnetr' R package can facilitate ecological theory development and scientific hypothesis testing and increase the accessibility of ecosystem modelling.

AUTHOR CONTRIBUTIONS

Xiaojie Gao and Jonathan R. Thompson conceptualized the idea. Xiaojie Gao developed the package, documented the algorithms and conducted the analyses with important feedback from Jonathan R. Thompson, Zaixing Zhou, Scott V. Ollinger, Jaclyn H. Matthes and Wenzhe Jiao; Jaclyn H. Matthes formatted the long-term eddy-covariance measurements for the model tests. Xiaojie Gao drafted the manuscript, and all authors contributed to editing and revising. All authors approved the final manuscript.

ACKNOWLEDGEMENTS

Xiaojie Gao and Jonathan R. Thompson were funded by NSF-DEB-LTER (Grant 18-32210 and NSF-DISES Grant 22-05705). Jaclyn H. Matthes was funded by NSF-DEB (Grant 22-31681 and NSF-EF Grant 19-26454). Zaixing Zhou and Scott V. Ollinger were funded by NSF-EPSCoR (Grant 19-20908, NSF-DEB-LTER Grants 18-32210 and the New Hampshire Agricultural Experiment Station, Hatch project NH00711). We thank the two anonymous reviewers for offering constructive comments and suggestions, which have substantially improved our manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.70076.

DATA AVAILABILITY STATEMENT

The source code of the R package, as well as an example data set, is available via https://zenodo.org/records/15486341 (Gao, 2025). The development version of the R package is also available on GitHub (https://github.com/hf-thompson-lab/pnetr). Either GitHub or Email can be used for communicating any questions or software bugs. The example data set was derived from the data measured by the Environmental Measurement Station (EMS) eddy-covariance flux tower at Harvard Forest, MA (https://harvardforest.fas.harvard.edu/research/towers) and the FLUXNET2015 data set (https://fluxnet.org/data/fluxnet2015-dataset, site ID: US-Ha1).

ORCID

Xiaojie Gao https://orcid.org/0000-0002-0607-7768
Jonathan R. Thompson https://orcid.org/0000-0003-0512-1226

REFERENCES

Aber, J., Ollinger, S., Federer, C., Reich, P., Goulden, M., Kicklighter, D., Melillo, J., & Lathrop, R. (1995). Predicting the effects of climate change on water yield and forest production in the northeastern United States. *Climate Research*, *5*, 207–222. https://doi.org/10.3354/cr005207

Aber, J. D., Driscoll, C., Federer, C. A., Lathrop, R., Lovett, G., Melillo, J. M., Steudler, P., & Vogelmann, J. (1993). A strategy for the regional analysis of the effects of physical and chemical climate change on biogeochemical cycles in northeastern (U.S.) forests. *Ecological Modelling*, *67*(1), 37–47. https://doi.org/10.1016/0304-3800(93)90098-D

Aber, J. D., & Federer, C. A. (1992). A generalized, lumped-parameter model of photosynthesis, evapotranspiration and net primary production in temperate and boreal forest ecosystems. *Oecologia*, *92*(4), 463–474. https://doi.org/10.1007/BF00317837

Aber, J. D., Ollinger, S. V., & Driscoll, C. T. (1997). Modeling nitrogen saturation in forest ecosystems in response to land use and atmospheric deposition. *Ecological Modelling*, *101*(1), 61–78. https://doi.org/10.1016/S0304-3800(97)01953-4

Aber, J. D., Reich, P. B., & Goulden, M. L. (1996). Extrapolating leaf CO_2 exchange to the canopy: A generalized model of Forest photosynthesis compared with measurements by Eddy correlation. *Oecologia*, *106*(2), 257–265.

Bacmeister, J. T., Hannay, C., Medeiros, B., Gettelman, A., Neale, R., Fredriksen, H. B., Lipscomb, W. H., Simpson, I., Bailey, D. A., Holland, M., Lindsay, K., & Otto-Bliesner, B. (2020). CO_2 increase experiments using the CESM: Relationship to climate sensitivity and comparison of CESM1 to CESM2. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2020MS002120. https://doi.org/10.1029/2020MS002120

Basler, D. (2016). Evaluating phenological models for the prediction of leaf-out dates in six temperate tree species across central Europe. *Agricultural and Forest Meteorology*, *217*, 10–21. https://doi.org/10.1016/j.agrformet.2015.11.007

Bonan, G. B., Lombardozzi, D. L., Wieder, W. R., Oleson, K. W., Lawrence, D. M., Hoffman, F. M., & Collier, N. (2019). Model structure and climate data uncertainty in historical simulations of the terrestrial carbon cycle (1850–2014). *Global Biogeochemical Cycles*, *33*(10), 1310–1326. https://doi.org/10.1029/2019GB006175

Boose, E., Gould, E., & VanScoy, M. (2024). Harvard forest climate data since 1964 [dataset]. *Environmental Data Initiative*, https://doi.org/10.6073/PASTA/4FB183A9B4480CE9748570C3716C3813

Chuine, I., & Régnière, J. (2017). Process-based models of phenology for plants and animals. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 159–182. https://doi.org/10.1146/annurev-ecolsys-110316-022706

De Bruijn, A., Gustafson, E. J., Sturtevant, B. R., Foster, J. R., Miranda, B. R., Lichti, N. I., & Jacobs, D. F. (2014). Toward more robust projections of forest landscape dynamics under novel environmental conditions: Embedding PnET within LANDIS-II. *Ecological Modelling*, *287*, 44–57. https://doi.org/10.1016/j.ecolmodel.2014.05.004

Dietze, M. (2017). *Ecological Forecasting*. Princeton University Press. https://doi.org/10.1515/9781400885459

Dow, C., Kim, A. Y., D'Orangeville, L., Gonzalez-Akre, E. B., Helcoski, R., Herrmann, V., Harley, G. L., Maxwell, J. T., McGregor, I. R., McShea, W. J., McMahon, S. M., Pederson, N., Tepley, A. J., & Anderson-Teixeira, K. J. (2022). Warm springs alter timing but not total growth of temperate deciduous trees. *Nature*, *608*, 552–557. https://doi.org/10.1038/s41586-022-05092-3

Fer, I., Kelly, R., Moorcroft, P. R., Richardson, A. D., Cowdery, E. M., & Dietze, M. C. (2018). Linking big models to big data: Efficient ecosystem model calibration through Bayesian model emulation. *Biogeosciences*, *15*(19), 5801–5830. https://doi.org/10.5194/bg-15-5801-2018

Gao, X. (2025). Source code of the 'pnetr' R package (v1.0). *Zenodo*, https://doi.org/10.5281/zenodo.15486341

Gao, X., Gray, J. M., & Reich, B. J. (2021). Long-term, medium spatial resolution annual land surface phenology with a Bayesian hierarchical model. *Remote Sensing of Environment*, *261*, 112484. https://doi.org/10.1016/j.rse.2021.112484

Gao, X., McGregor, I. R., Gray, J. M., Friedl, M. A., & Moon, M. (2023). Observations of satellite land surface phenology indicate that maximum leaf greenness is more associated with global vegetation productivity than growing season length. *Global Biogeochemical Cycles*, *37*(3), e2022GB007462. https://doi.org/10.1029/2022GB007462

Gao, X., Richardson, A. D., Friedl, M. A., Moon, M., & Gray, J. M. (2024). Thermal forcing versus chilling? Misspecification of temperature controls in spring phenology models. *Global Ecology and Biogeography*, *33*(12), e13932. https://doi.org/10.1111/geb.13932

Geary, W. L., Bode, M., Doherty, T. S., Fulton, E. A., Nimmo, D. G., Tulloch, A. I. T., Tulloch, V. J. D., & Ritchie, E. G. (2020). A guide to ecosystem models and their environmental applications. *Nature Ecology & Evolution*, *4*(11), 1459–1471. https://doi.org/10.1038/s41559-020-01298-8

Gettelman, A., Geer, A. J., Forbes, R. M., Carmichael, G. R., Feingold, G., Posselt, D. J., Stephens, G. L., van den Heever, S. C., Varble, A. C., & Zuidema, P. (2022). The future of earth system prediction: Advances in model-data fusion. *Science Advances*, *8*(14), eabn3488. https://doi.org/10.1126/sciadv.abn3488

Gustafson, E. J., Brian, M. R., Brian, S. R., & Zaixing, Z. (2023). *PnET-succession v 5.1: Comprehensive description of an ecophysiological succession extension within the LANDIS-II forest landscape model* [computer software]. The LANDIS-II Foundation. https://github.com/LANDIS-II-Foundation/Foundation-Publications/blob/main/Description%20of%20PnET-Succession%20v5.1.pdf

Gustafson, E. J., Miranda, B. R., De Bruijn, A. M. G., Sturtevant, B. R., & Kubiske, M. E. (2017). Do rising temperatures always increase forest productivity? Interacting effects of temperature, precipitation, cloudiness and soil texture on tree species growth and competition. *Environmental Modelling & Software*, *97*, 171–183. https://doi.org/10.1016/j.envsoft.2017.08.001

Huang, M., Xu, Y., Longo, M., Keller, M., Knox, R. G., Koven, C. D., & Fisher, R. A. (2020). Assessing impacts of selective logging on water, energy, and carbon budgets and ecosystem dynamics in Amazon forests using the functionally assembled terrestrial ecosystem simulator. *Biogeosciences*, *17*(20), 4999–5023. https://doi.org/10.5194/bg-17-4999-2020

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J.-F., Large, W. G., Lawrence, D., Lindsay, K., Lipscomb, W. H., Long, M. C., Mahowald, N., Marsh, D. R., Neale, R. B., Rasch, P., Vavrus, S., Vertenstein, M., Bader, D., … Marshall, S. (2013). The community earth system model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, *94*(9), 1339–1360. https://doi.org/10.1175/BAMS-D-12-00121.1

Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., Munger, J. W., O'Keefe, J., Schmid, H. P., Wing, I. S., Yang, B., & Richardson, A. D. (2014). Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Climate Change*, *4*(7), 598–604. https://doi.org/10.1038/nclimate2253

Koven, C. D., Knox, R. G., Fisher, R. A., Chambers, J. Q., Christoffersen, B. O., Davies, S. J., Detto, M., Dietze, M. C., Faybishenko, B., Holm, J., Huang, M., Kovenock, M., Kueppers, L. M., Lemieux, G., Massoud, E., McDowell, N. G., Muller-Landau, H. C., Needham, J. F., Norby, R. J., … Xu, C. (2020). Benchmarking and parameter sensitivity of physiological and vegetation dynamics using the functionally assembled terrestrial ecosystem simulator (FATES) at Barro Colorado Island, Panama. *Biogeosciences*, *17*(11), 3017–3044. https://doi.org/10.5194/bg-17-3017-2020

Lewis, A. S. L., Rollinson, C. R., Allyn, A. J., Ashander, J., Brodie, S., Brookson, C. B., Collins, E., Dietze, M. C., Gallinat, A. S., Juvigny-Khenafou, N., Koren, G., McGlinn, D. J., Moustahfid, H., Peters, J. A., Record, N. R., Robbins, C. J., Tonkin, J., & Wardle, G. M. (2023). The power of forecasts to advance ecological theory. *Methods in Ecology and Evolution*, *14*(3), 746–756. https://doi.org/10.1111/2041-210X.13955

Liang, Y., Duveneck, M. J., Gustafson, E. J., Serra-Diaz, J. M., & Thompson, J. R. (2018). How disturbance, competition, and dispersal interact to prevent tree range boundaries from keeping pace with climate change. *Global Change Biology*, *24*(1), e335–e351. https://doi.org/10.1111/gcb.13847

Liang, Y., Gustafson, E. J., He, H. S., Serra-Diaz, J. M., Duveneck, M. J., & Thompson, J. R. (2023). What is the role of disturbance in catalyzing spatial shifts in forest composition and tree species biomass under climate change? *Global Change Biology*, *29*(4), 1160–1177. https://doi.org/10.1111/gcb.16517

Liu, Y., Holtzman, N. M., & Konings, A. G. (2021). Global ecosystem-scale plant hydraulic traits retrieved using model–data fusion. *Hydrology and Earth System Sciences*, *25*(5), 2399–2417. https://doi.org/10.5194/hess-25-2399-2021

Ma, L., Hurtt, G., Ott, L., Sahajpal, R., Fisk, J., Lamb, R., Tang, H., Flanagan, S., Chini, L., Chatterjee, A., & Sullivan, J. (2022). Global evaluation of the ecosystem demography model (ED v3.0). *Geoscientific Model Development*, *15*(5), 1971–1994. https://doi.org/10.5194/gmd-15-1971-2022

Massoud, E. C., Xu, C., Fisher, R. A., Knox, R. G., Walker, A. P., Serbin, S. P., Christoffersen, B. O., Holm, J. A., Kueppers, L. M., Ricciuto, D. M., Wei, L., Johnson, D. J., Chambers, J. Q., Koven, C. D., McDowell, N. G., & Vrugt, J. A. (2019). Identification of key parameters controlling demographically structured vegetation dynamics in a land surface model: CLM4. 5 (FATES). *Geoscientific Model Development*, *12*(9), 4133–4164.

McKenzie, P. F., Duveneck, M. J., Morreale, L. L., & Thompson, J. R. (2019). Local and global parameter sensitivity within an ecophysiologically based forest landscape model. *Environmental Modelling & Software*, *117*, 1–13. https://doi.org/10.1016/j.envsoft.2019.03.002

Middendorp, R. S., Pérez, A. J., Molina, A., & Lambin, E. F. (2016). The potential to restore native woody plant richness and composition in a reforesting landscape: A modeling approach in the Ecuadorian

Andes. *Landscape Ecology*, *31*(7), 1581–1599. https://doi.org/10.1007/s10980-016-0340-7

Munger, W., & Wofsy, S. (2024). Canopy-atmosphere exchange of carbon, water and energy at Harvard Forest EMS tower since 1991 [dataset]. *Environmental Data Initiative*, https://doi.org/10.6073/PASTA/56C6FE02A07E8A8AAFF44A43A9D9A6A5

Ollinger, S. V., Aber, J. D., & Federer, A. (1998). Estimating regional forest productivity and water yield using an ecosystem model linked to a GIS. *Landscape Ecology*, *13*, 323–334.

Ollinger, S. V., Aber, J. D., & Reich, P. B. (1997). simulating ozone effects on forest productivity: Interactions among leaf-, canopy-, and stand-level processes. *Ecological Applications*, *7*(4), 1237–1251. https://doi.org/10.1890/1051-0761(1997)007[1237:SOEOFP]2.0.CO;2

Ollinger, S. V., Aber, J. D., Reich, P. B., & Freuder, R. J. (2002). Interactive effects of nitrogen deposition, tropospheric ozone, elevated CO_2 and land use history on the carbon dynamics of northern hardwood forests. *Global Change Biology*, *8*(6), 545–562. https://doi.org/10.1046/j.1365-2486.2002.00482.x

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Reichstein, M., Ribeca, A., van Ingen, C., Vuichard, N., Zhang, L., Amiro, B., Ammann, C., … Papale, D. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data*, *7*(1), 225. https://doi.org/10.1038/s41597-020-0534-3

Prinn, R. G. (2013). Development and application of earth system models. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(supplement_1), 3673–3680. https://doi.org/10.1073/pnas.1107470109

Rastetter, E. B., Kwiatkowski, B. L., Kicklighter, D. W., Barker Plotkin, A., Genet, H., Nippert, J. B., O'Keefe, K., Perakis, S. S., Porder, S., Roley, S. S., Ruess, R. W., Thompson, J. R., Wieder, W. R., Wilcox, K., & Yanai, R. D. (2022). N and P constrain C in ecosystems under climate change: Role of nutrient redistribution, accumulation, and stoichiometry. *Ecological Applications*, *32*(8), e2684. https://doi.org/10.1002/eap.2684

Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., Chen, J. M., Ciais, P., Davis, K. J., Desai, A. R., Dietze, M. C., Dragoni, D., Garrity, S. R., Gough, C. M., Grant, R., Hollinger, D. Y., Margolis, H. A., McCaughey, H., Migliavacca, M., … Xue, Y. (2012). Terrestrial biosphere models need better representation of vegetation phenology: Results from the North American carbon program site synthesis. *Global Change Biology*, *18*(2), 566–584. https://doi.org/10.1111/j.1365-2486.2011.02562.x

Scheller, R. M., Kretchun, A. M., Van Tuyl, S., Clark, K. L., Lucash, M. S., & Hom, J. (2012). Divergent carbon dynamics under climate change in forests with diverse soils, tree species, and land use histories. *Ecosphere*, *3*(11), 1. https://doi.org/10.1890/ES12-00241.1

Scheller, R. M., & Mladenoff, D. J. (2004). A forest growth and biomass module for a landscape simulation model, LANDIS: Design, validation, and application. *Ecological Modelling*, *180*(1), 211–229.

Scheller, R. M., Sturtevant, B. R., Gustafson, E. J., Ward, B. C., & Mladenoff, D. J. (2010). Increasing the reliability of ecological models using modern software engineering techniques. *Frontiers in Ecology and the Environment*, *8*(5), 253–260. https://doi.org/10.1890/080141

Shifley, S. R., He, H. S., Lischke, H., Wang, W. J., Jin, W., Gustafson, E. J., Thompson, J. R., Thompson, F. R., Dijak, W. D., & Yang, J. (2017). The past and future of modeling forest dynamics: From growth and yield curves to forest landscape models. *Landscape Ecology*, *32*(7), 1307–1325. https://doi.org/10.1007/s10980-017-0540-9

Smith, B. (2001). LPJ-GUESS-an ecosystem modelling framework. *Department of Physical Geography and Ecosystems Analysis*, *INES*, *Sölvegatan*, *12*(22), 362.

Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Seneviratne, S. I., & Peñuelas, J. (2019). Drought impacts on terrestrial primary production underestimated by satellite monitoring. *Nature Geoscience*, *12*(4), 264–270. https://doi.org/10.1038/s41561-019-0318-6

Teets, A., Bailey, A. S., Hufkens, K., Ollinger, S., Schädel, C., Seyednasrollah, B., & Richardson, A. D. (2023). Early spring onset increases carbon uptake more than late fall senescence: Modeling future phenological change in a US northern deciduous forest. *Oecologia*, *201*(1), 241–257. https://doi.org/10.1007/s00442-022-05296-4

Thompson, J. R., Foster, D. R., Scheller, R., & Kittredge, D. (2011). The influence of land use and climate change on forest biomass and composition in Massachusetts, USA. *Ecological Applications*, *21*(7), 2425–2444. https://doi.org/10.1890/10-2383.1

Zhou, Z., Ollinger, S. V., & Lepine, L. (2018). Landscape variation in canopy nitrogen and carbon assimilation in a temperate mixed forest. *Oecologia*, *188*(2), 595–606. https://doi.org/10.1007/s00442-018-4223-2

How to cite this article: Gao, X., Zhou, Z., Ollinger, S. V., Matthes, J. H., Jiao, W., & Thompson, J. R. (2025). *pnetr*: An R package for the PnET family of forest ecosystem models. *Methods in Ecology and Evolution*, *16*, 1378–1388. https://doi.org/10.1111/2041-210X.70076