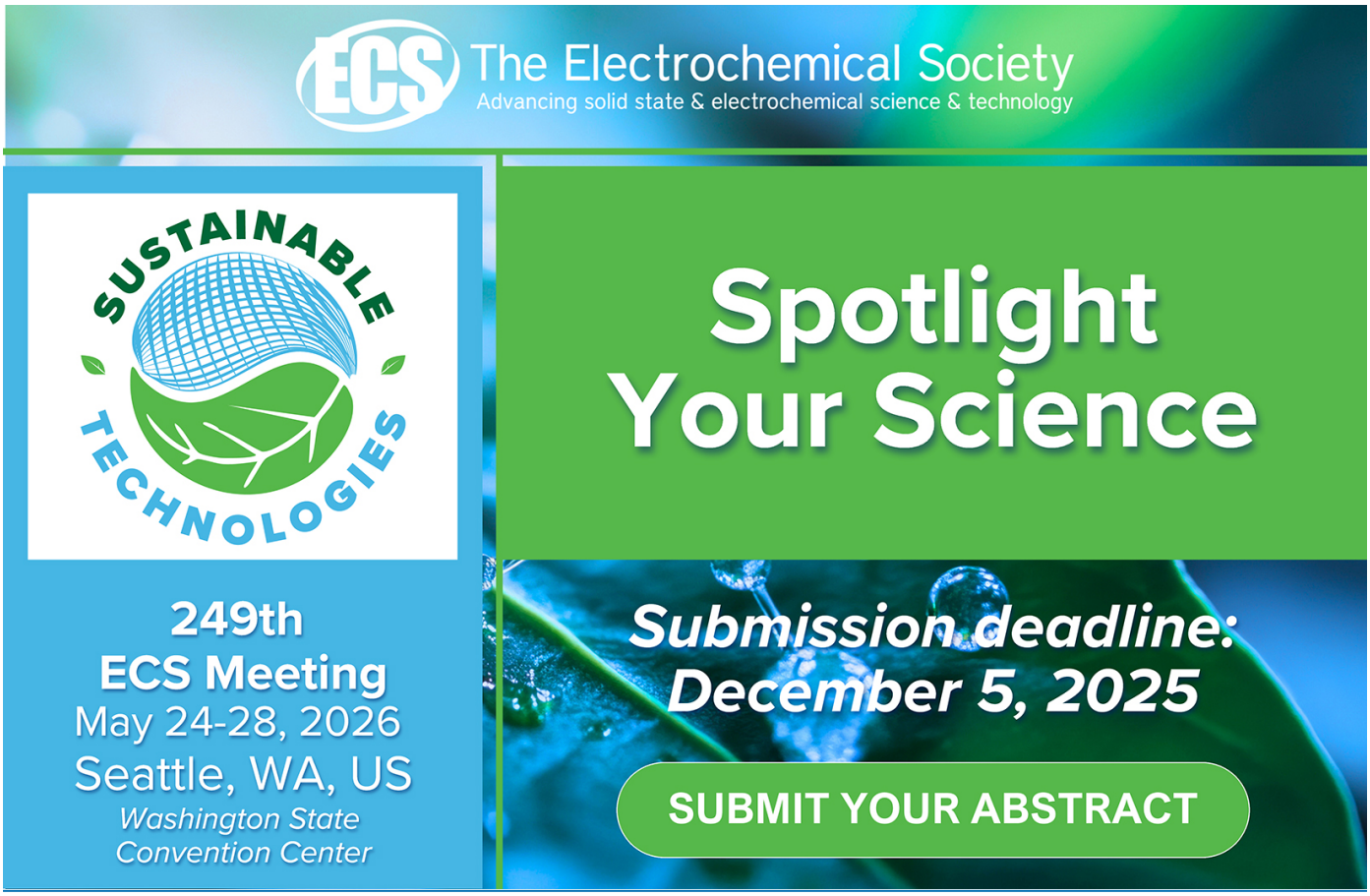
ENVIRONMENTAL RESEARCH
LETTERS

PURPOSE-LED PUBLISHING™ Part of the ENVIRONMENTAL RESEARCH SERIES

LETTER • OPEN ACCESS

Improving the quantification of global free-living and symbiotic nitrogen fixation in natural terrestrial ecosystems: present-day estimates and 21st century projections

To cite this article: Ye Yuan *et al* 2025 *Environ. Res. Lett.* **20** 124005



ENVIRONMENTAL RESEARCH
LETTERS

OPEN ACCESS

RECEIVED
20 March 2025

REVISED
24 October 2025

ACCEPTED FOR PUBLICATION
30 October 2025

PUBLISHED
21 November 2025

LETTER

Improving the quantification of global free-living and symbiotic nitrogen fixation in natural terrestrial ecosystems: present-day estimates and 21st century projections

Ye Yuan[1], **Qianlai Zhuang**[1,2,*], **Bailu Zhao**[3] and **Licheng Liu**[4]

[1] Department of Earth, Atmospheric, and Planetary Science, Purdue University, West Lafayette, IN 47906, United States of America
[2] Department of Agronomy, Purdue University, West Lafayette, IN 47907, United States of America
[3] High Meadows Environmental Institute, Princeton University, Princeton, NJ 08544, United States of America
[4] Department of Bioproducts and Biosystems Engineering, University of Minnesota, St. Paul, MN 55108, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: qzhuang@purdue.edu

Keywords: biological nitrogen fixation, free-living nitrogen fixation, modeling, process

Supplementary material for this article is available online

Abstract

Biological nitrogen fixation (BNF) is a critical natural nitrogen input that sustains terrestrial carbon cycling, yet it remains poorly represented in terrestrial ecosystem models (TEMs). Here, we refine the nitrogen cycle representation in a TEM by incorporating free-living and symbiotic nitrogen fixation (SNF) processes along with atmospheric nitrogen deposition effects. Our updated model provides a new assessment of present-day and future global natural ecosystem BNF rates and their spatial distribution. We estimate that present-day (1981–2020) free-living nitrogen fixation (FLNF) contributes to 36 Tg N yr^{-1}, ranging from 33 to 38.5 Tg N yr^{-1} and SNF is 88 Tg N yr^{-1} with a range of 80–95 Tg N yr^{-1}, resulting in a total BNF 124 Tg N yr^{-1}, ranging from 112 to 134 Tg N yr^{-1}. Under the SSP5-8.5 warming scenario, our projections indicate that total BNF could increase to 178 Tg N yr^{-1} by the end of the 21st century. Our model results highlight that the FLNF plays a dominant role in low-temperature environments, and plant functional type emerges as the primary factor for both symbiotic and FLNF pathways, thus precise global vegetation classification is highly necessary for N_2 fixation simulation. Our refined model advances global nitrogen cycle modeling and will improve future quantification of both global carbon and nitrogen cycles.

1. Introduction

Nitrogen availability is a key factor limiting plant growth across ecosystems worldwide, constraining plant productivity and terrestrial biosphere carbon storage (Dynarski and Houlton 2018, Cleveland *et al* 2013, Peng *et al* 2018, Meyerholt *et al* 2016). Biological nitrogen fixation (BNF), the reduction of atmospheric dinitrogen (N_2) to biologically available nitrogen (NH_3), is the primary natural input of N into terrestrial ecosystems (Galloway *et al* 2004).

BNF occurs across multiple terrestrial niches through two main pathways. Symbiotic nitrogen fixation (SNF) involves N-fixing bacteria living in root nodules of higher plants, while free-living nitrogen fixation (FLNF) is carried out by diverse organisms such as cyanobacteria, soil and litter bacteria, bryophytes, and lichens (Reed *et al* 2011, Dynarski and Houlton 2018, Davies-Barnard and Friedlingstein 2020, Bellenger *et al* 2020). SNF is primarily influenced by plant traits and climate factors including precipitation, temperature, soil properties, nutrients, and biomass (Yao *et al* 2024, Liu *et al* 2011, Wu and McGachan 1999, Thornley 2001), whereas FLNF is mainly controlled by temperature, moisture, nutrient availability (N, P, Mo), and oxygen availability (Reed *et al* 2011, Dynarski and Houlton 2018, Cleveland *et al* 2022, Rousk & Michelsen 2017, Perakis *et al* 2017, Smercina *et al* 2019).

Most C–N coupled terrestrial biosphere models represent BNF empirically, linking it to net primary productivity or evapotranspiration (Cleveland *et al*

1999, Xu and Prentice 2017, Davies-Barnard and Friedlingstein 2020, Hupperts *et al* 2021). These approaches, however, are limited in projecting responses to elevated CO_2 and other environmental changes (Davies-Barnard *et al* 2022). More recent models—such as CABLE (Wang *et al* 2007, Peng *et al* 2020), CLM5 (Lawrence *et al* 2019), ELMv1-ECA (Zhu *et al* 2019), and GFDL LM3-SNAP (Kou-Giesbrecht *et al* 2021)—incorporate mechanistic representations of symbiotic BNF. Yet FLNF, contributing at least one-third of global BNF (Davies-Barnard and Friedlingstein 2020, Kou-Giesbrecht and Arora 2022), remains underrepresented despite its importance in N-limited ecosystems such as temperate and boreal forests, where symbiotic fixers are scarce but ground mosses with N-fixing cyanobacteria are abundant (Dynarski and Houlton 2018, Hupperts *et al* 2021).

Recent FLNF representations include temperature and soil organic carbon effects (Kou-Giesbrecht and Arora 2022), but their linear formulation likely oversimplifies complex, non-linear interactions. Free-living fixers respond differently to temperature, while soil moisture and nitrogen stocks also strongly influence fixation (Reed *et al* 2011, Smercina *et al* 2019). FLNF further shows strong spatial and temporal variability, underscoring major knowledge gaps in process understanding (Cleveland *et al* 2022). Advancing understanding of these regulatory factors is critical for improving ecosystem models and predicting terrestrial productivity and biogeochemical cycles under environmental change.

Here we improve the process-based terrestrial ecosystem model (TEM; Yuan *et al* 2025) by incorporating both FLNF and SNF processes. The refined model is applied globally to quantify BNF in terrestrial ecosystems for the present-day period (1981–2020) and throughout the 21st century with temporal and spatial specificity. By explicitly representing FLNF—an underrepresented process in previous biosphere models—and projecting future fixation rates, this study provides new insights into how this critical input to the nitrogen cycle may respond to global change.

2. Method

2.1. Model description

TEM is a global-scale biogeochemical model designed to quantify the cycling of water, carbon (C) and nitrogen (N) in terrestrial ecosystems (Melillo *et al* 1993, McGuire *et al* 1997, Zhuang *et al* 2003, Yuan *et al* 2025). The major processes of nitrogen (N) dynamic module were inherited from McGuire *et al* (1997) and Yuan *et al* (2025), including nitrogen input from plant litters and the atmosphere, nitrogen uptake by vegetation, net soil mineralization, nitrogen lost from ecosystem, as well as the effect of physical conditions on both nitrification and denitrification, and the principles of the stoichiometry of carbon and nitrogen dynamics in soils. The model simulates soil C and N pools (e.g. soil mineral N, soil organic C) from the surface to a depth of 1 m. Detailed model description can be found in Melillo *et al* (1993), McGuire *et al* (1997) and Yuan *et al* (2025). Here we revised the N cycling algorithms in TEM by incorporating the input of nitrogen through both SNF and FLNF pathways and additional inputs of mineral atmospheric N deposition (NH_4^+ and NO_3^-) (Tang 2015). (figure 1). The detailed algorithms and parameterization schemes are provided in SI 1: Model Description.

2.2. Model calibration and validation

We compiled observational data through a broad literature search of peer-reviewed publications using the keywords *'nitrogen fixation', 'N fixation', 'N_2 fixation',* and *'dinitrogen fixation'*. A total of 102 observations from 92 natural ecosystem sites worldwide were used for model calibration and evaluation (figure S1; table S1). These sites encompass diverse ecosystems, including dry tundra ($n = 6$), wet tundra ($n = 10$), boreal forests ($n = 15$), temperate coniferous forests ($n = 11$), temperate deciduous forests ($n = 9$), temperate evergreen forests ($n = 9$), tropical forests ($n = 20$), grasslands ($n = 11$), xeric shrubland/woodlands ($n = 8$), and Mediterranean shrublands ($n = 3$). Site locations are shown in figure S1, with detailed site information provided in table S1. Measurements were obtained using either the acetylene reduction assay or the 15^N relative abundance method. To facilitate meaningful comparisons with observations, all data was aggregated to an annual scale, as most results are reported annually. The model output for the corresponding exact time periods was extracted to make the comparison.

In this study, 14 parameters defined in equations (SI 1 model description; tables S2 and S3) were calibrated with 102 observations. Calibration and validation were conducted separately for each plant functional type (PFT) and fixation type, using an 80/20 split of the observational dataset for training and testing, respectively. Parameter estimation was performed with the PEST software package (Parameter ESTimation; Welter *et al* 2015), applying the Gauss–Marquardt–Levenberg optimization algorithm to calibrate BNF parameters for major natural ecosystem types. Initial parameter values were derived from previous studies (table S2) and iteratively adjusted to minimize the sum of squared weighted residuals between simulated and observed values for each PFT and fixation type. When only a single observed parameter value was available, the allowable range was conservatively set to 0.5–2.0 times the initial value.

Figure S2 presents a comparison across all sites, while table S1 summarizes the mean and median values of simulations and observations for each plant functional and fixation type of calibration and

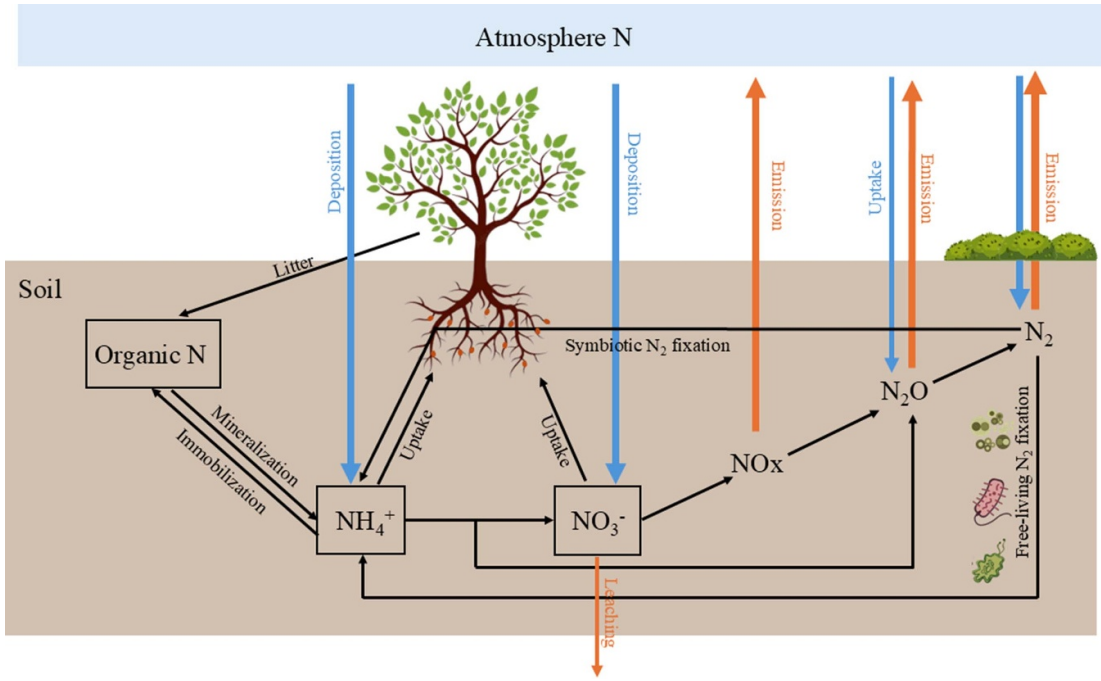


Figure 1. Nitrogen cycle and N_2 fixation pathways in the revised terrestrial ecosystem model. Blue arrows represent nitrogen (N) inputs, while orange arrows represent nitrogen outputs. Mineralization: organic N mineralized to inorganic N; immobilization: inorganic N to organic N; litter: organic N from plant litters; Uptake (black): inorganic N uptake by plants; Deposition: atmospheric deposition of N; Emission: N_2O emissions from soils; Uptake (blue): Atmospheric N_2O uptake in soils; Symbiotic N_2 fixation: N_2 fixation through root nodules; Free-living nitrogen fixation: N_2 fixation through soil microorganisms and moss, lichens and bryophytes.

validation. The calibrated parameter values (table S3) are used in subsequent global simulations. Other nitrogen cycle parameters in the model were defined and calibrated in previous studies (Melillo *et al* 1993, McGuire *et al* 1997, Zhuang *et al* 2003, Yuan *et al* 2025).

The average observed grassland SNF is 13.08 kg N ha^{-1} yr^{-1} (median: 12.3; table SI). Among the five grassland sites, only one lies below 40° latitude, whereas 65% of simulated grid cells fall within 40° S–40° N, where warmer temperatures likely promote higher fixation rates. This helps explain the higher simulated FLNF and SNF in tundra ecosystems (table 1) compared to Reed *et al* (2011), who reported FLNF of 0.4–3 and SNF of 1–4.9 kg N ha^{-1} yr^{-1}. Observations show FLNF ranging from 0.23 to 64 (mean: 6.03; median: 3.03) and SNF from 0.5 to 5.04 (mean: 1.89; median: 0.74). Of 17 tundra observation sites, only one is below 65° latitude, while 46% of tundra grid cells in our vegetation map fall below 65° N or 65° S, likely contributing to differences between observed and simulated fixation.

2.3. Regional extrapolation

To generate spatially and temporally explicit global BNF estimates, we used land cover, soil, and climate data at 0.5° × 0.5° resolution to drive TEM. For the historical period (1981–2020), monthly climate forcing was from CRU TS v4.05, while future projections (2021–2100) used CMIP6 GFDL-ESM4 under SSP1-2.6, SSP2-4.5, and SSP5-8.5. To ensure consistency, GFDL-ESM4 data (2016–2020) were bias-adjusted against CRU. CO_2 concentrations for 1981–2020 were obtained from NOAA's Global Monitoring Laboratory, and future values were digitized from Meinshausen *et al* (2011). Nitrogen deposition was taken from ACCMIP (Lamarque *et al* 2013), including NH_x and NO_y for 1980, 2000, 2030, and 2100 across scenarios; intermediate years were linearly interpolated and reformatted for TEM with model modifications to incorporate deposition. Land cover, soil bulk density, and pH data came from Melillo *et al* (1993), the Global Soil Bulk Density Map (Global Soil Data Task 2000), and the Global Database of Soil Properties (Carter & Scholes 2000). The model was spun up for at least 50 years using first-year climate data to reach steady state for all state and flux variables in each grid cell.

2.4. Statistical analysis

Annual results from 1981 to 2020 were used for statistical analysis and comparison with published studies. A random forest model (R package *ranger*) was applied to identify key environmental drivers of TEM-simulated FLNF and SNF. Differences in fixation rates among PFTs were tested using one-way ANOVA with Tukey's HSD for post-hoc comparisons, implemented in R 4.3.3 (*stats* package).

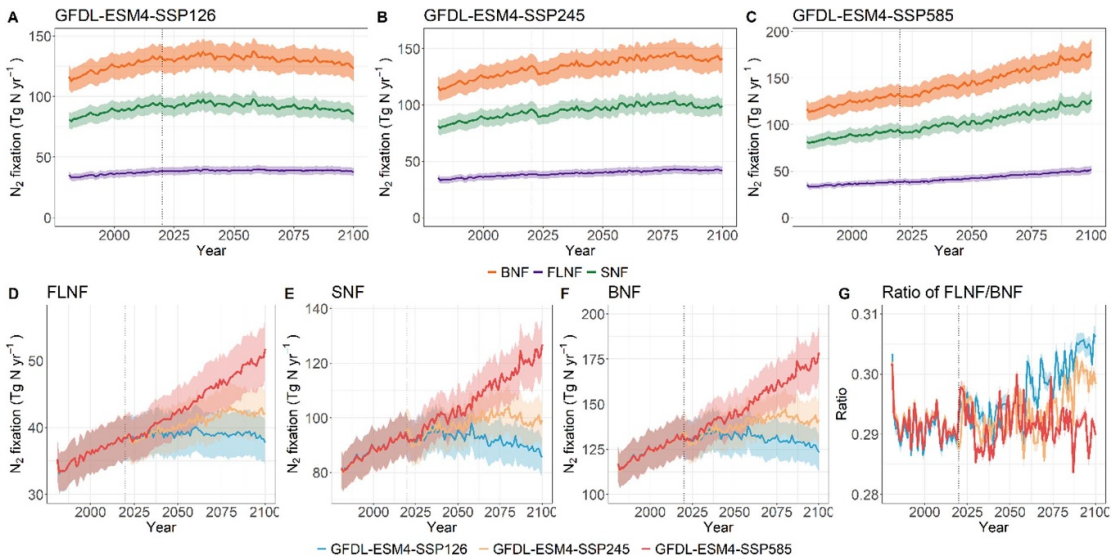


Figure 2. Annual global nitrogen fixation trends under different warming scenarios. Panels show FLNF, SNF and total BNF under SSP1-2.6 (A), SSP2-4.5 (B), and SSP5-8.5 (C). Panels D, E, and F present comparisons of FLNF, SNF, and total BNF, respectively, across the three warming scenarios, along with the ratio of FLNF to BNF. The vertical grey line marks the year 2020.

3. Results

3.1. Global present-day BNF

Our simulations estimated that annual FLNF, SNF, and total BNF in natural terrestrial ecosystems increased from 1980 to 2020. Specifically, FLNF rose from 33 to 38.5 Tg N yr^{-1}, averaging 36 Tg N yr^{-1}. SNF increased from 80 to 95 Tg N yr^{-1}, with an average of 88 Tg N yr^{-1}. As a result, the total BNF increased from 112 to 134 Tg N yr^{-1} with average of 124 Tg N yr^{-1}.

Spatially, global FLNF during present years (1981–1990 and 2011–2020) ranged from 0 to over 20 kg N ha^{-1} yr^{-1} (mean 2.61, median 2.32 kg N ha^{-1} yr^{-1}), SNF from 0 to over 50 kg N ha^{-1} yr^{-1} (mean 5.72, median 3.88 kg N ha^{-1} yr^{-1}), and total BNF from 0 to 60 kg N ha^{-1} yr^{-1} (mean 8.33, median 6.59 kg N ha^{-1} yr^{-1}) (figures 3(A) and (B)). From period 1981–1990 to 2011–2020, FLNF, SNF, and total BNF exhibit a widespread, gradual increase, except in South America, where SNF and total BNF show a slight decline (figure S3(D1)).

3.2. Major factors contributing to BNFs

Considering the nonlinear relationships, we used feature importance scores from the random forest model to assess the key environmental factors influencing TEM-simulated FLNF and SNF. These scores helped identify the most influential variables in predicting FLNF and SNF. The annual TEM input environmental factors listed in figure 4 from 1981 to 2020 are explicitly considered in the RF model, with TEM-simulated annual FLNF and SNF serving as the dependent variables. The models with 500 decision trees achieved R^2 of 0.92 for SNF and 0.9 for FLNF. PFT was the most influential factor in both cases, with an importance score greater than 0.4. For FLNF, the top-ranked factors after PFT were average air temperature, vapor pressure, NH_x deposition, and annual precipitation. Similarly, for SNF, the most important factors were annual precipitation, average temperature, vapor pressure, and NH_x deposition (figure 4). Soil characteristics (soil bulk density; soil sand, salt and clay content) are the least important for both FLNF and SNF.

3.3. BNF by PFT

Present-day simulations (1981–2020) of FLNF and SNF were used to calculate the PFT-specific fixation rates (table 1). Excluding Wet tundra and Mediterranean shrubland, FLNF and SNF differed significantly among PFTs. Grassland, tundra (dry and wet), and temperate deciduous forest had the highest FLNF (>3 kg N ha^{-1} yr^{-1}), while Grassland, Xeric shrubland/woodland, Mediterranean shrubland, and Wet tundra showed the highest SNF (>5 kg N ha^{-1} yr^{-1}). Mediterranean shrubland, Xeric shrubland/woodland, and temperate coniferous forest had the lowest FLNF (<2 kg N ha^{-1} yr^{-1}), whereas temperate deciduous forest, temperate coniferous forest, and dry/alpine tundra had the lowest SNF (<2 kg N ha^{-1} yr^{-1}).

3.4. Global future BNF

Under the SSP1-2.6 scenario, FLNF increased from 2020, peaking at 40.04 Tg N yr^{-1} in 2060, then declined to 33.05 Tg N yr^{-1} by 2100. SNF peaked earlier around 2037 at over 97 Tg N yr^{-1}, dropping to 87.5 Tg N yr^{-1} by 2096. Total BNF mirrored SNF, rising from 128.9 Tg N yr^{-1} in 2021 to 137 Tg N yr^{-1}

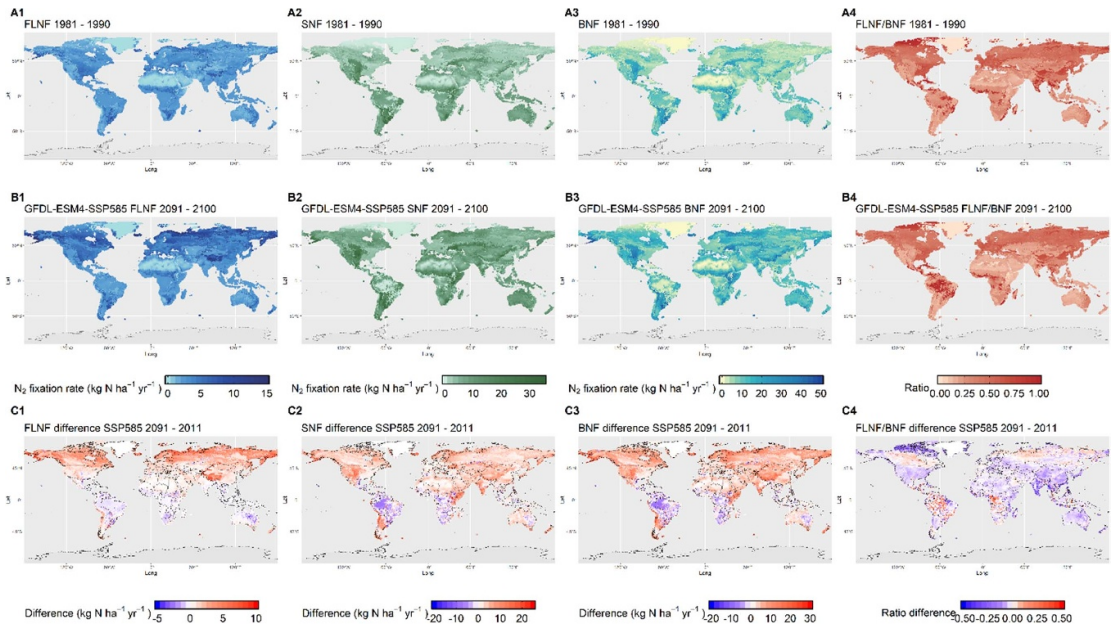


Figure 3. Spatial distribution of 10 year averaged nitrogen fixation, including FLNF (blue), SNF (green), total BNF (yellow-blue), and the FLNF/BNF ratio (red), for the periods 1981–1990 (A), SSP585 2091–2100 (B). Panels represent the differences between B and A.

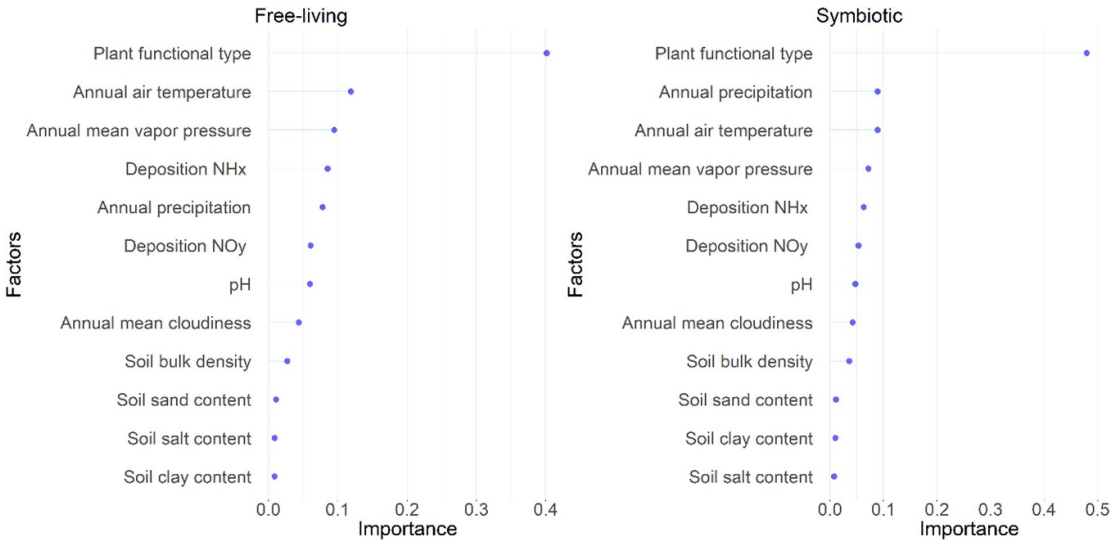


Figure 4. Feature importance scores of variables in the random forest model for free-living and symbiotic nitrogen fixation.

in 2037 before declining (figure 3). Under SSP2-4.5, all BNF components rose from 2020 to 2080: FLNF from 37.7 to 43.1 Tg N yr^{-1}, SNF from 89.3 to 104.1 Tg N yr^{-1}, and total BNF from 127.3 to 147.18 Tg N yr^{-1} (figure 2(B)). Under SSP5-8.5, FLNF, SNF, and total BNF steadily increased from 37.9 to 51.78 Tg N yr^{-1}, 91.1 to 126.67 Tg N yr^{-1}, and 129.8–178.65 Tg N yr^{-1}, respectively (figure 2(C)). Overall, BNF was highest under SSP5-8.5, followed by SSP2-4.5 and SSP1-2.6, whereas the FLNF/BNF ratio was greatest under SSP1-2.6, suggesting FLNF is more adapted to cooler temperatures (figures 2(D)–(G)).

In future projections (2091–2100), FLNF is expected to range from 0 to over 25 kg N ha^{-1} yr^{-1} (mean 3.36, median 2.63 kg N ha^{-1} yr^{-1}), SNF from 0 to over 60 kg N ha^{-1} yr^{-1} (mean 7.02, median 5.44 kg N ha^{-1} yr^{-1}), and total BNF from 0 to 80 kg N ha^{-1} yr^{-1} (mean 10.38, median 8.61 kg N ha^{-1} yr^{-1}, figure 2(B)). Changes between 2011–2020 and 2091–2100 are smallest under SSP1-2.6 and largest under SSP5-8.5, with increases mainly in the Northern Hemisphere and declines in the Southern Hemisphere (figure S3(D)). The largest SNF declines occur under SSP5-8.5 in the Amazon,

Table 1. Mean and median of free-living nitrogen fixation and symbiotic nitrogen fixation of each plant functional type. Q1 refers to the first quartile (25th percentile), and Q3 refers to the third quartile (75th percentile).

	Free-living (kg N ha^{-1} yr^{-1})		Symbiotic (kg N ha^{-1} yr^{-1})	
PFT	Mean ± SD	Median [Q1, Q3]	Mean ± SD	Median [Q1, Q3]
Dry/Alpine tundra	4.95 ± 3.5	4.40 [2.28,6.37]	2.11 ± 2.24	1.61 [0.45, 2.72]
Wet tundra	4.04 ± 2.51	3.46 [2.17, 5.65]	5.50 ± 3.86	4.62 [2.72, 7.65]
Boreal forest	1.92 ± 0.84	1.85 [1.22, 2.38]	2.57 ± 1.52	2.35 [1.55, 3.19]
Temperate coniferous forest	1.85 ± 0.7	1.74 [1.41, 2.19]	1.96 ± 0.74	1.86 [1.43, 2.39]
Temperate deciduous forest	3.26 ± 1.42	3.24 [2.5, 4.09]	1.47 ± 0.83	1.44 [0.92, 2.02]
Grassland	5.17 ± 2.6	4.66 [3.61, 6.64]	17.92 ± 9.82	16.23 [11.65, 22.91]
Xeric shrubland/woodland	1.45 ± 1.04	1.38 [0.56, 2.12]	7.49 ± 6.45	6.1 [2.68, 10.43]
Tropical evergreen forest	2.15 ± 0.43	2.31 [2.08, 2.40]	4.15 ± 3.76	3.79 [0.21, 7.30]
Temperate evergreen forest	2.53 ± 0.74	2.51 [2.1, 2.95]	4.6 ± 1.63	4.75 [3.56, 5.69]
Mediterranean shrubland	1.17 ± 0.31	1.24 [0.95, 1.28]	5.57 ± 2.19	5.41 [4.09, 7.05]

driven by reduced soil moisture and higher soil N, increasing the FLNF/BNF ratio locally.

Notable FLNF increases occur in northern high latitudes and the Tibetan Plateau, where the FLNF/BNF ratio exceeds 0.5 and continues to rise, indicating FLNF dominance. Amazon also shows a rising FLNF/BNF ratio due to SNF declines, while most other regions exhibit a decreasing trend.

3.5. Model sensitivity

We performed Morris sensitivity screening to rank parameter influence. The analysis shows that SNF rates are most sensitive to the distribution of N-fixers or N fixation rates, both of which are linearly related to SNF. FLNF is most sensitive to fl_N, parameter related to free-living BNF and soil N. We tested four levels of perturbation (±10% and ±20%) for these two factors separately. The sensitivity analysis reveals that changes in fl_N have a profound and asymmetric impact on future nitrogen fixation projections under climate change scenarios (table 2). A critical finding is that decreases in fl_N suppress total nitrogen fixation much more strongly than equivalent increases in fl_N stimulate it. This is evident across all SSP scenarios, where a 20% reduction in fl_N leads to a large decrease in total fixation (ranging from −14.79% to −21.16%), while a 20% increase (fl_N +20%) results in a comparatively modest positive change (ranging from +6.65% to +14.64%). This inhibitory effect is particularly pronounced under the SSP5-8.5 scenario. SNF demonstrates a more symmetric and positive sensitivity to the change of N fixer distribution of 10%. However, under a ±20% change, the model exhibits scenario-dependent responses. In SSP2-4.5, a 20% decrease and a 20% increase in parameter values result in changes of similar magnitude. In contrast, under SSP1-2.6, the 20% decrease leads to a more pronounced reduction in SNF compared to the effect of a 20% increase. Conversely, in SSP5-8.5, the response to a 20% decrease is less pronounced than that observed for a 20% increase.

4. Discussion

4.1. Comparison with other studies

4.1.1. Global present-day fixation

Our model estimated global present-day (1981–2020) FLNF, natural SNF, and total BNF consistent with observation-based estimates and other land surface models. TEM-estimated FLNF (36 Tg N yr^{-1}, range: 33–38.5 Tg N yr^{-1}) is higher than 23 Tg N yr^{-1} reported by Kou-Giesbrecht and Arora (2022) and 31 Tg N yr^{-1} (range: 21–66 Tg N yr^{-1}) by Davies-Barnard and Friedlingstein (2020), but lower than the 49 Tg N yr^{-1} attributed to cryptogamic covers by Elbert *et al* (2012). The Kou-Giesbrecht and Arora (2022) map shows no fixation across much of the northern high latitudes and Tibetan Plateau—regions generally considered FLNF-dominant—which likely explains their lower global estimate.

TEM-estimated total BNF (124 Tg N yr^{-1}, range: 112–134 Tg N yr^{-1}) falls well within the wide range of prior estimates (40–340 Tg N yr^{-1}; table 3). This spread reflects methodological and structural differences among studies. Empirical upscaling approaches (e.g. Cleveland *et al* 1999, Davies-Barnard and Friedlingstein 2020, Reis Ely *et al* 2025) extrapolate site-level measurements using vegetation maps, with uncertainties dominated by sparse and biased sampling. In contrast, process-based models simulate BNF mechanistically, making their outputs sensitive to assumptions about fixation drivers. Models linking BNF primarily to plant N demand (e.g. DyN-LPJ; Xu and Prentice 2017) produce the highest estimates, while those constraining fixation by carbon costs (e.g. CASACNP, CLM5; Wang and Houlton 2009, Lawrence *et al* 2019) yield lower values. Our TEM estimate (124 Tg N yr^{-1}) lies at the high end, consistent with the view that fixation is costly. Representation of environmental constraints also matters: models omitting explicit N limitation (e.g. CLM5, CABLE) generally simulate lower fixation in N-limited ecosystems.

Table 2. The sensitivity of global free-living, symbiotic and total nitrogen fixation rate to changing distribution of nitrogen fixers between 1981 and 2100.

	FLNF (%)				SNF (%)			
	fl_N −20%	fl_N −10%	fl_N +10%	fl_N +20%	Distribution −20%	Distribution −10%	Distribution +10%	Distribution +20%
SSP1−2.6	−21.16 ± 0.25	−13.9 ± 0.45	1.56 ± 0.29	7.68 ± 1.01	−20.41 ± 0.23	−8.35 ± 0.22	7.94 ± 0.2	12.87 ± 0.67
SSP2−4.5	−14.79 ± 0.59	−7.37 ± 0.3	7.26 ± 0.29	14.64 ± 0.61	−17.64 ± 0.19	−8.30 ± 0.17	8.22 ± 0.19	17.02 ± 0.38
SSP5−8.5	−19.71 ± 0.81	−13.19 ± 0.32	2.88 ± 0.64	6.65 ± 1.16	−13.56 ± 0.42	−8.41 ± 0.15	8.17 ± 0.21	23.37 ± 0.48

Table 3. Comparison of global biological nitrogen fixation (BNF) estimates from existing studies and this study. The table highlights differences in methodology, model structure, and resulting global BNF estimates. Key drivers represented in the models include T (temperature), M (soil moisture), C (carbon), and N (nitrogen).

BNF (Tg N yr^{-1})	Method	Model	SNF (Tg N yr^{-1})	Key SNF driver	FLNF (Tg N yr^{-1})	Key FLNF driver	Factors explicitly represented	References
58 (40–100)	Top down							Vitousek *et al* (2013)
195 (101–290)	Empirical upscaling							Cleveland *et al* (1999)
88 (52–130)	Empirical upscaling		57 (31–66)	Averaged biome value	31 (21–66)	Averaged biome value		Davies-Barnard and Friedlingstein (2020)
65 (52–77)	Empirical upscaling		42%–47%	N fixer abundance	53%–58%	N fixer abundance		Reis Ely *et al* (2025)
340	Process-based model	DyN-LPJ		N Demand				Xu and Prentice (2017)
125	Process-based model	CASACNP		C Cost of N uptake			T, C, N, P	Wang *et al* (2007); Wang and Houlton (2009)
69	Process-based model	CLM5		C Cost of N uptake		Evapotranspiration		Lawrence *et al* (2019); Davies-Barnard and Friedlingstein (2020)
73–122	Process-based model	CABLE		C Cost of N uptake			T, C	Peng *et al* (2020)
70	Process-based model	CLASSIC	47	Temperature and N stress of vegetation	23	Temperature and soil C	T, C, N	Kou-Giesbrecht and Arora (2022)
124 (112–134)	Process-Based Model	TEM	88 (80–95)	T, M, C, N	36 (33–38.5)	T, M, C, N	T, M, C, N	This study

Data biases further compound uncertainty. Reis Ely *et al* (2025) show that field measurements overrepresent ecosystems where N fixers are ∼17 times more abundant than the global mean, while Cleveland *et al* (1999) assumed fixer coverage of 5%–25% based on sparse data—later revised to only 2%–3% for symbiotic niches and 4%–9% for mosses and biocrusts. Because many models, including ours, reference these early distributions, such assumptions remain a persistent source of error.

In summary, the spread in global BNF estimates is systematic, driven by fixation drivers (N demand vs C cost), treatment of secondary limitations (e.g. N, P availability), and biased field constraints. Progress will require refining representations of physiological trade-offs, particularly the carbon cost of fixation and N–P interactions, alongside improved observations of N fixer distributions and rates in under-sampled ecosystems.

4.1.2. The influence of PFT on fixation

Across PFTs, TEM estimated there is higher SNF in grasslands (mean: 17.92, median: 16.23 kg N ha^{-1} yr^{-1}) compared to estimates (mean: 6.71, median: 8.1 kg N ha^{-1} yr^{-1}) by Davies-Barnard and Friedlingstein (2020) and the reported range for temperate grasslands (1–10 kg N ha^{-1} yr^{-1}) from Reed *et al* (2011). This discrepancy partially explains the higher total SNF values in our estimates. However, differences in land classification and definitions across studies complicate direct comparisons. For instance, our model distinguishes wet and dry tundra in northern high latitudes, whereas other studies categorize these regions as grasslands.

The classification of tropical forests and temperate evergreen forests in our simulation generally aligns with the evergreen broadleaf forest category in Davies-Barnard and Friedlingstein (2020), while temperate coniferous forests correspond to evergreen needleleaf forests. The estimated SNF and FLNF in our simulations are consistent with their reported values. Additionally, the FLNF estimates for temperate coniferous forests, temperate deciduous forests, and boreal forests align with the values reported in Reed *et al* (2011), and the SNF falls well within the range provided by their study.

4.1.3. Controlling factors of BNF

In our analysis, PFT is the primary factor contributing to both TEM-simulated FLNF and SNF, significantly outweighing other climate factors and soil properties in explaining the variations in BNF. We performed statistical tests with available observational data. The results indicate that both SNF and FLNF differ significantly ($p < 0.05$) among PFTs. This finding is consistent with the results of Yao *et al* (2024), who reported that host plant traits exerted a much stronger influence SNF rates than soil properties or climatic factors, accounting for 79% of the total relative importance. Similarly, in Davies-Barnard and Friedlingstein (2020), linear regression results also indicated that climate or variables are generally not the best predictors of BNF, with soil variables also lacking strong predictive power. Therefore, the classification and distribution of PFTs become key factors in accurately reporting global BNF, as they depend on the extent and distribution of land cover worldwide.

4.2. Study limitations and future directions

Key nutrients such as nitrogen (N), phosphorus (P), and molybdenum (Mo) regulate nitrogen fixation (Reed *et al* 2011, Dynarski and Houlton 2018). We excluded P and Mo due to unclear relationships with nitrogen fixation and limited spatial data, which may introduce uncertainties in BNF simulations, particularly in P-limited tropical regions (Elser *et al* 2007). Time lags between detecting nutrient limitation and BNF regulation may also affect fixation rates (Menge *et al* 2009).

Our new FLNF representation advanced the mechanistic depiction of free-living BNF in biogeochemical models. However, the underlying mechanisms driving FLNF remain poorly understood, and observational studies on factors influencing FLNF are limited. This knowledge gap poses challenges for model parameterization, emphasizing the need for additional observational and experimental studies to further refine BNF modeling. The substantial variability is spread within the same plant and fixation type highlights the importance of comprehensive measurements of fixation rates by non-symbiotic organisms in terrestrial ecosystems and a deeper exploration of the factors governing these rates.

The spatial variability in the abundance of nitrogen-fixing trees poses a major challenge for accurately quantifying BNF (Menge *et al* 2019, Staccone *et al* 2020). We adopted values from Cleveland *et al* (1999) and performed sensitivity analyses based on the distribution of nitrogen fixers. However, realistically representing this variability in land surface models remains difficult, particularly for estimating regional symbiotic BNF (SNF; Kou-Giesbrecht *et al* 2021). Recent evidence further highlights these uncertainties. Reis Ely *et al* (2025) demonstrated that field-based BNF measurements are disproportionately sampled in ecosystems where nitrogen fixers are ∼17 times more abundant than their global mean distribution, introducing strong sampling biases. Incorporating these updated distributions of nitrogen fixers into global models would help improve the robustness of BNF quantification.

PFT emerged as the main control of both FLNF and SNF, partly reflecting reliance on annual data for TEM calibration and random forest inputs. The

use of annual aggregates constrains the ability to capture interannual variability and limits the representation of transient or episodic dynamics. In addition, the observational dataset used for model evaluation is restricted to a single year and therefore reflects climatic conditions specific to that period. As such, the validation presented here should be regarded as a temporal 'snapshot', and the extent to which the model applies to years with substantially different climate regimes remains uncertain.

Although we aimed for global coverage, the available data remains sparse and are predominantly concentrated in North America and Europe. A limitation of this study is the relatively small number of observations (102 data points) available to constrain the 14 parameters, which may contribute to parameter uncertainty and affect model predictions. These observations are insufficient to fully constrain the parameters across diverse PFTs and fixation types, highlighting the need for more comprehensive and spatially representative datasets to improve parameterization and reduce uncertainty in future modeling efforts. Expanding observations worldwide and establishing standardized measurement protocols, particularly for PFTs with limited data, are crucial for improving model accuracy. Moreover, most existing observations are restricted to a single year, limiting our ability to capture long-term trends. Multi-year observations of both FLNF and SNF, as well as high-frequency (daily, weekly, or monthly) continuous measurements, would greatly enhance our capacity to represent the seasonal and temporal dynamics of nitrogen fixation in terrestrial ecosystems.

5. Conclusion

Our study provides a mechanistic representation of FLNF and SNF and new assessment of BNF. We estimate present-day FLNF at 36 Tg N yr^{-1} (range: 33–38.5 Tg N yr^{-1}) and SNF at 88 Tg N yr^{-1} (range: 80–95 Tg N yr^{-1}), resulting in a total BNF of 124 Tg N yr^{-1} (range: 112–134 Tg N yr^{-1}) in natural terrestrial ecosystems. Our findings indicate that FLNF contributes approximately 30% of total BNF and is particularly dominant under cooler temperature conditions. These estimates align with previous studies, supporting the reliability of our approach. By the end of the 21st century, under the SSP8.5 warming scenario, our model projects FLNF, SNF, and total BNF to increase to 51.78 Tg N yr^{-1}, 126.67 Tg N yr^{-1}, and 178.65 Tg N yr^{-1}, respectively. Our analysis highlights PFT as the primary driver of variability in both fixation pathways, emphasizing the need for precise global vegetation classification in BNF quantification. Additionally, expanding field measurements across diverse PFTs is essential, particularly for improving our understanding of FLNF. By refining BNF representation in global models, our study enhances estimates of nitrogen inputs, ultimately improving projections of future nitrogen cycle dynamics and their interactions with terrestrial ecosystems. This advancement will help reduce uncertainties in terrestrial carbon cycle predictions and facilitate more accurate integration of coupled carbon-nitrogen cycle processes in Earth system models.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://purr.purdue.edu/publications/4780/1.

Acknowledgment

This research has been supported by the National Science Foundation (Grant No. 1802832).

Open Research (availability statement)

The TEM codes, outputs, and samples of running directory can be accessed via the Purdue University Research Repository (Yuan & Zhuang 2005) https://purr.purdue.edu/publications/4780/1 and on GitHub at https://github.com/yeyuanc/TEMN2Fixation.

ORCID iDs

Ye Yuan 0000-0003-1330-4049
Licheng Liu 0000-0002-9649-1056

References

Bellenger J P, Darnajoux R, Zhang X and Kraepiel A M L 2020 Biological nitrogen fixation by alternative nitrogenases in terrestrial ecosystems: a review *Biogeochemistry* **149** 53–73

Carter A J and Scholes R J 2000 SoilData v2.0: generating a global database of soil Properties *CSIR* environmentek

Cleveland C C *et al* 1999 Global patterns of terrestrial biological nitrogen (N_2) fixation in natural ecosystems *Glob. Biogeochem. Cycles* **13** 623–45

Cleveland C C *et al* 2022 Exploring the role of cryptic nitrogen fixers in terrestrial ecosystems: a frontier in nitrogen cycling research *Ecosystems* **25** 1653–69

Cleveland C C, Houlton B Z, Smith W K, Marklein A R, Reed S C, Parton W, Del Grosso S J and Running S W 2013 Patterns of new versus recycled primary production in the terrestrial biosphere *Proc. Natl Acad. Sci.* **110** 12733–7

Davies-Barnard T, Zaehle S and Friedlingstein P 2022 Assessment of the impacts of biological nitrogen fixation structural uncertainty in CMIP6 earth system models *Biogeosciences* **19** 3491–503

Davies-Barnard T and Friedlingstein P 2020 The global distribution of biological nitrogen fixation in terrestrial natural ecosystems *Glob. Biogeochem. Cycles* **34** e2019GB006387

Dynarski K A and Houlton B Z 2018 Nutrient limitation of terrestrial free-living nitrogen fixation *New Phytol.* **217** 1050–61

Elbert W, Weber B, Burrows S, Steinkamp J, Büdel B, Andreae M O and Pöschl U 2012 Contribution of cryptogamic covers to the global cycles of carbon and nitrogen *Nat. Geosci.* **5** 459–62

Elser J J, Bracken M E S, Cleland E E, Gruner D S, Harpole W S, Hillebrand H, Ngai J T, Seabloom E W, Shurin J B and Smith J E 2007 Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems *Ecol. Lett.* **10** 1135–42

Galloway J N *et al* 2004 Nitrogen cycles: past, present, and future *Biogeochemistry* **70** 153–226

Global Soil Data Task 2000 Global gridded surfaces of Selected Soil Characteristics (IGBP-DIS) *ORNL Distributed Active Archive Center* (https://doi.org/10.3334/ORNLDAAC/569)

Hupperts S F, Gerber S, Nilsson M and Gundale M J 2021 Empirical and Earth system model estimates of boreal nitrogen fixation often differ: a pathway toward reconciliation *Glob. Change Biol.* **27** 5711–25

Kou-Giesbrecht S, Malyshev S, Martínez Cano I, Pacala S W, Shevliakova E, Bytnerowicz T A and Menge D N L 2021 A novel representation of biological nitrogen fixation and competitive dynamics between nitrogen-fixing and non-fixing plants in a land model (GFDL LM4.1-BNF) *Biogeosciences* **18** 4143–83

Kou-Giesbrecht S and Arora V K 2022 Representing the dynamic response of vegetation to nitrogen limitation via biological nitrogen fixation in the CLASSIC land model *Glob. Biogeochem. Cycles* **36** e2022GB007341

Lamarque J-F *et al* 2013 Multi-model mean nitrogen and sulfur deposition from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): evaluation of historical and projected future changes *Atmos. Chem. Phys.* **13** 7997–8018

Lawrence D M *et al* 2019 The Community Land Model Version 5: description of new features, benchmarking, and impact of forcing uncertainty *J. Adv. Model. Earth Syst.* **11** 4245–87

Liu Y, Wu L, Baddeley J A and Watson C A 2011 Models of biological nitrogen fixation of legumes. A review *Agron. Sustain. Dev.* **31** 155–72

McGuire A D, Melillo J M, Kicklighter D W, Pan Y, Xiao X, Helfrich J, Moore B, Vorosmarty C J and Schloss A L 1997 Equilibrium responses of global net primary production and carbon storage to doubled atmospheric carbon dioxide: sensitivity to changes in vegetation nitrogen concentration *Glob. Biogeochem. Cycles* **11** 173–89

Meinshausen M *et al* 2011 The SSP greenhouse gas concentrations and their extensions from 1765 to 2300 *Clim. Change* **109** 213–41

Melillo J M, McGuire A D, Kicklighter D W, Moore B, Vorosmarty C J and Schloss A L 1993 Global climate change and terrestrial net primary production *Nature* **363** 234–40

Menge D N L, Levin S A and Hedin L O 2009 Facultative versus obligate nitrogen fixation strategies and their ecosystem consequences *Am. Nat.* **174** 465–77

Menge D N *et al* 2019 Patterns of nitrogen-fixing tree abundance in forests across Asia and America *J. Ecol.* **107** 2598–610

Meyerholt J, Zaehle S and Smith M J 2016 Variability of projected terrestrial biosphere responses to elevated levels of atmospheric CO_2 due to uncertainty in biological nitrogen fixation *Biogeosciences* **13** 1491–518

NOAA Global Monitoring Laboratory n.d. *Trends in atmospheric carbon dioxide* (National Oceanic and Atmospheric Administration) (available at: https://gml.noaa.gov/ccgg/trends/) (Accessed 24 May 2024)

Peng J, Dan L, Wang Y-P, Tang X, Yang X, Yang F, Lu X and Pak B 2018 How biological nitrogen fixation and climate change contribute to future terrestrial carbon sequestration from the global to biome scale *J. Clean. Prod.* **202** 1158–66

Peng J, Wang Y, Houlton B Z, Dan L, Pak B and Tang X 2020 Global carbon sequestration is highly sensitive to model-based formulations of nitrogen fixation *Glob. Biogeochem. Cycles* **34** e2019GB006296

Perakis S S, Pett-Ridge J C and Catricala C E 2017 Nutrient feedbacks to soil heterotrophic nitrogen fixation in forests *Biogeochemistry* **134** 41–55

R Core Team 2024 R: a language and environment for statistical computing. Version 4.3.3 (R Foundation for Statistical Computing) (available at: www.R-project.org/)

Reed S C, Cleveland C C and Townsend A R 2011 Functional ecology of free-living nitrogen fixation: a contemporary perspective *Annu. Rev. Ecol. Evol. Syst.* **42** 489–512

Reis Ely C R *et al* 2025 Global terrestrial nitrogen fixation and its modification by agriculture *Nature* **643** 705–11

Rousk K and Michelsen A 2017 Ecosystem nitrogen fixation throughout the snow-free period in subarctic tundra: effects of willow and birch litter addition and warming *Glob. Change Biol.* **23** 1552–63

Smercina D N, Evans S E, Friesen M L and Tiemann L K 2019 To fix or not to fix: controls on free-living nitrogen fixation in the rhizosphere *Appl. Environ. Microbiol.* **85** e02546–18

Staccone A, Liao W, Perakis S, Compton J, Clark C M and Menge D 2020 A spatially explicit, empirical estimate of tree-based biological nitrogen fixation in forests of the United States *Glob. Biogeochem. Cycles* **34**

Tang J Y 2015 On the relationships between the Michaelis–Menten kinetics, reverse Michaelis–Menten kinetics, equilibrium chemistry approximation kinetics, and quadratic kinetics *Geosci. Model Dev.* **8** 3823–35

Thornley J 2001 Simulating grass-legume dynamics: a phenomenological submodel *Ann. Bot.* **88** 905–13

Vitousek P M, Menge D N L, Reed S C and Cleveland C C 2013 Biological nitrogen fixation: rates, patterns and ecological controls in terrestrial ecosystems *Phil. Trans. R. Soc. B* **368** 20130119

Wang Y-P and Houlton B Z 2009 Nitrogen constraints on terrestrial carbon uptake: implications for the global carbon-climate feedback *Geophys. Res. Lett.* **36** L24403

Wang Y-P, Houlton B Z and Field C B 2007 A model of biogeochemical cycles of carbon, nitrogen, and phosphorus including symbiotic nitrogen fixation and phosphatase production *Glob. Biogeochem. Cycles* **21** 2006GB002797

Welter D E, White J T, Hunt R J and Doherty J E 2015 Approaches in highly parameterized inversion—PEST++ version 3, a parameter ESTimation and uncertainty analysis software suite optimized for large environmental models (7-C12) (available at: https://pubs.usgs.gov/publication/)

Wu L and McGechan M B 1999 Simulation of nitrogen uptake, fixation and leaching in a grass/white clover mixture *Grass Forage Sci.* **54** 30–41

Xu R and Prentice I C 2017 Modelling the demand for new nitrogen fixation by terrestrial ecosystems *Biogeosciences* **14** 2003–17

Yao Y, Han B, Dong X, Zhong Y, Niu S, Chen X and Li Z 2024 Disentangling the variability of symbiotic nitrogen fixation rate and the controlling factors *Glob. Change Biol.* **30** e17206

Yuan Y and Zhuang Q 2005 Quantifying global biological nitrogen fixation with in-situ data and a process-based biogeochemistry model *Purdue University Research Repository* (https://doi.org/10.4231/PYV7-W139)

Yuan Y, Zhuang Q, Zhao B and Shurpali N 2025 Impacts of permafrost degradation on N_2O emissions from natural terrestrial ecosystems in northern high latitudes: a process-based biogeochemistry model analysis *Glob. Biogeochem. Cycles* **39** e2024GB008439

Zhu Q, Riley W J, Tang J, Collier N, Hoffman F M, Yang X and Bisht G 2019 Representing nitrogen, phosphorus, and carbon interactions in the E3SM land model: development and global benchmarking *J. Adv. Model. Earth Syst.* **11** 2238–58

Zhuang Q *et al* 2003 Carbon cycling in extratropical terrestrial ecosystems of the Northern Hemisphere during the 20th century: a modeling analysis of the influences of soil thermal dynamics *Tellus B* **55** 751–76