ECOLOGY LETTERS

Ecology Letters, (2020) doi: 10.1111/ele.13667

LETTER

Forest responses to last-millennium hydroclimate variability are governed by spatial variations in ecosystem sensitivity

Christine R. Rollinson,[1]* Andria Dawson,[2] Ann M. Raiho,[3] John W. Williams,[4] Michael C. Dietze,[5] Thomas Hickler,[6,7] Stephen T. Jackson,[8,9] Jason McLachlan,[3] David JP Moore,[10] Benjamin Poulter,[11] Tristan Quaife,[12] Jörg Steinkamp[13,14] and Mathias Trachsel[15]

Abstract

Forecasts of future forest change are governed by ecosystem sensitivity to climate change, but ecosystem model projections are under-constrained by data at multidecadal and longer timescales. Here, we quantify ecosystem sensitivity to centennial-scale hydroclimate variability, by comparing dendroclimatic and pollen-inferred reconstructions of drought, forest composition and biomass for the last millennium with five ecosystem model simulations. In both observations and models, spatial patterns in ecosystem responses to hydroclimate variability are strongly governed by ecosystem sensitivity rather than climate exposure. Ecosystem sensitivity was higher in models than observations and highest in simpler models. Model-data comparisons suggest that interactions among biodiversity, demography and ecophysiology processes dampen the sensitivity of forest composition and biomass to climate variability and change. Integrating ecosystem models with observations from timescales extending beyond the instrumental record can better understand and forecast the mechanisms regulating forest sensitivity to climate variability in a complex and changing world.

Keywords

Climate change, drought, ecosystem modelling, palaeoecology, stability, vulnerability.

Ecology Letters (2020)

INTRODUCTION

Exposure to 21st-century climate change is expected to profoundly impact global forest composition, diversity and structure (Dawson *et al*., 2011; Keeley *et al*., 2019), but the sensitivity of ecosystems to climate variability at multi-decadal to centennial time scales is poorly constrained by instrumental observations. Multiple observational studies that employ subcontinental- to continental-scale data networks across a broad range of timescales have sought to empirically estimate the sensitivity of forest ecosystems to climate variability. The sensitivity of tree growth rates, biomass accumulation and ecophysiological processes to interannual climate variability is well-documented by dendroecological data, with compelling evidence that forest sensitivity to climate depends on forest age and is non-stationary across space and time (Charney *et al*., 2016; Klesse *et al*., 2018; Thom *et al*., 2019; Peltier and Ogle, 2020). On glacial-interglacial timescales, networks of fossil pollen records show that temperature variations are the primary driver of forest composition and species distributions (Shuman *et al*., 2004; Nolan *et al*., 2018), while over the last several thousand years, hydroclimate variability has strongly affected forest composition and structure in temperate forests of the northeastern and upper midwestern United States (Booth *et al*., 2012; Shuman *et al*., 2019).

Terrestrial ecosystem models used to forecast responses to climate change often have difficulty reproducing broad-scale and long-term responses to environmental variability, despite being well-grounded in empirical evidence and ecological theory (Friedlingstein *et al*., 2006, 2014; Matthes *et al*., 2016). These models mechanistically connect ecophysiological processes and climate variability to past and present changes in forest composition and structure but are subject to uncertainty in external forcings (e.g., drivers), process representation and parametrisation that complicate data-model comparisons (Figure 1) (LeBauer *et al*., 2013; Matthes *et al*.,

[1]*Center for Tree Science, The Morton Arboretum, 4100 Illinois Route 53, Lisle, IL 60532, USA*
[2]*Department of General Education, Mount Royal University, Calgary, Alberta T3E 6K6, Canada*
[3]*Department of Biological Sciences, University of Notre Dame, 100 Galvin Life Science Center, Notre Dame, IN 46556, USA*
[4]*Department of Geography and Center for Climatic Research, University of Wisconsin-Madison, Madison, WI 53704, USA*
[5]*Department of Earth and Environment, Boston University, 685 Commonwealth Ave, Boston, MA 02215, USA*
[6]*Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Senckenberganlage 25, Frankfurt/Main 60325, Germany*
[7]*Department of Physical Geography, Goethe University, Frankfurt/Main, Germany*
[8]*US Geological Survey, Southwest and South Central Climate Adaptation Centers, Denver, DE, USA*
[9]*Department of Geosciences, University of Arizona, Tucson, AZ 85721, USA*
[10]*School of Natural Resources, University of Arizona, Tucson, AZ 85721, USA*
[11]*NASA GSFC, Biospheric Sciences Lab, Greenbelt, MD 20771, USA*
[12]*Department of Meteorology, University of Reading, Reading RG6 6BB, UK*
[13]*Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Frankfurt/Main, Germany*
[14]*Johannes Gutenberg University, Mainz, Germany*
[15]*Department of Geography, University of Wisconsin-Madison, Madison, WI 53704, USA*
**Correspondence: E-mail: crollinson@mortonarb.org*

2016; Dietze, 2017; McLachlan & PalEON Project 2018). Each model includes hypotheses about the primary processes and ecosystem characteristics governing forest change, various simplifying assumptions and trade-offs between computational tractability and process complexity (De Kauwe *et al.*, 2013; Walker *et al.*, 2014; Medlyn *et al.*, 2015). Previous data-model comparisons have returned mixed evidence about whether models underestimate or overestimate the sensitivity of forest processes such as net primary productivity (NPP) and mortality to climate change (Schimel *et al.*, 2015; Walker *et al.*, 2015; Rollinson *et al.*, 2017). As a result, projections of forest compositional and structural responses to climate change have high uncertainty, which propagates to increased uncertainty in science-based adaptation planning (Friedlingstein *et al.*, 2014).

Several challenges have traditionally hindered the joint analysis and integration of terrestrial ecosystem models and palaeoecological data to better constrain modelled responses to climate variations at multi-decadal and longer timescales. First, the raw observations collected from fossil pollen records (counts of individual pollen taxa) have no direct counterparts in ecosystem models. Bayesian hierarchical models are providing new process-based approaches to infer emergent ecosystem properties from fossil pollen records, such as forest composition, diversity, percent cover and biomass (Blarquez and Aleman, 2016; Dawson *et al.*, 2016), but the number of state variables that can be estimated from palaeoecological data remains small relative to the number of latent (i.e. unobservable) variables simulated by ecosystem models (Fig. 1). Second, pre-instrumental model-data comparisons are complicated by reliance on driver datasets derived from general circulation models (GCMs). GCMs generally capture macroscale spatial patterns and low-frequency trends in climate but are unable to fully capture the complexity and stochasticity of local to regional-scale weather phenomena at the subdaily resolution needed to drive ecosystem models, resulting in systematic spatial and temporal biases in model simulations (Anav *et al.*, 2013; Matthes *et al.*, 2016; Dietze *et al.*, 2018). Third, the native temporal resolution varies between palaeodata and models and requires a temporal standardisation. Due to these challenges, the predicted sensitivity of ecosystem model state variables such as forest composition and biomass to climate change is largely unvalidated by observations at multidecadal and longer timescales, resulting in wide divergence among terrestrial ecosystem models in their 21st-century projections (Friedlingstein *et al.*, 2006, 2014). Fourth, terrestrial ecosystem models vary widely in represented processes, which can challenge intermodel comparisons but also provide insight into key governing ecological processes when data-model discrepancies emerge.

Here, we seek to establish the patterns of forest ecosystem and climate variability in the north-central and northeastern United States for the last millennium (850–1850 C.E.) and identify the mechanisms underpinning both forest ecosystem sensitivity and observed data-model discrepancies. In these analyses, we test hypotheses about the relative importance of hydroclimate exposure, defined as the magnitude of drought variability, and ecosystem sensitivity as determinants of the variability seen in forest ecosystems. We also hypothesise that ecosystem models will be overly sensitive to hydroclimate variability due to insufficient representation of ecophysiological and demographic processes that can dampen climate responses. To this end, we present a novel series of data-model and model-model comparisons that are designed to overcome traditional barriers to data-model intercomparison

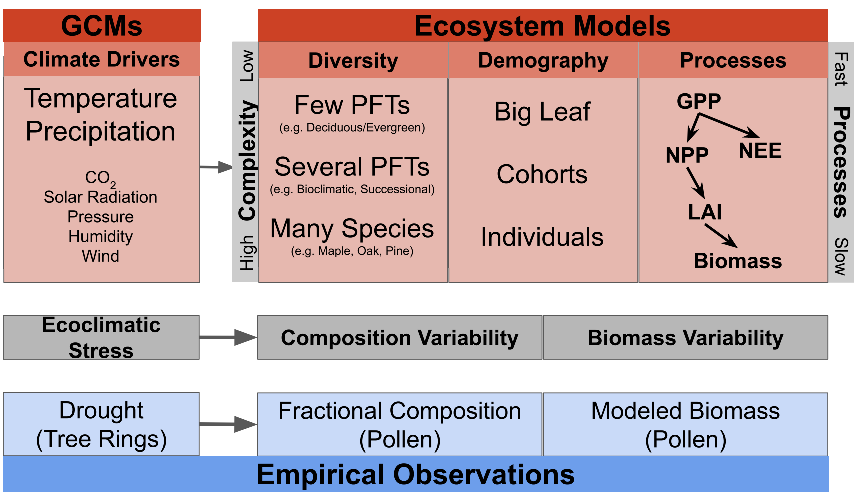

Figure 1 Overview of the unified conceptual framework (grey boxes) for parallel analysis of empirical data (blue boxes) and model output (red boxes). For ecosystem models, we describe the latent climatic and ecosystem processes that are unobservable in palaeoecological data and differences among models in complexity. Complexity here is organized into three categories: 1) diversity, ranging from a few plant functional types (PFTs) to many species; 2) demography, ranging from 'big leaf' models with no explicit treatment of forest demography to models with individual trees and 3) ecophysiological processes. Changes in forest biomass emerge from latent ecophysiological processes including gross primary productivity (GPP), net primary productivity (NPP), net ecosystem exchange (NEE) and leaf area index (LAI). Ecophysiological processes are controlled by model representation of higher-level vegetation processes (Table 1). Latent model drivers, processes and states (red boxes) result in estimates of forest composition and biomass that can be compared to palaeoecological data products (blue boxes). Models vary in complexity due to design philosophy and trade-offs between model complexity and computational speed.

for pre-instrumental times. Our analyses combine dendroclimatic indices of drought, recently published Bayesian spatiotemporal estimates of forest composition and biomass derived from pollen that provide independent checks on last-millennium simulations from five terrestrial ecosystem models for the northeastern and upper midwestern United States. The data-model comparisons discriminate among different representations of forest processes such as productivity and demography, whereas the model-model comparisons help diagnose causal relationships among ecological processes, changes in forest states and climate variability (Fig. 1). To test hypotheses while also overcoming known geographical biases in the model simulations of ecosystem state such as forest composition that source back to biases in the climate model drivers (Matthes *et al.*, 2016), we develop a new variability metric that we apply to the data and model-derived products that focuses on comparisons among variability of hydroclimate, composition and biomass (Fig. 1). Our results indicate that at centennial timescales, spatial patterns in the variability of forest composition and biomass are regulated by ecological factors such as ecotonal position and complexity rather than climate exposure as defined by the local magnitude of climate variability.

MATERIALS AND METHODS

Overview

We employ a combination of data-model and model-model comparisons (Fig. 1) in which we combine palaeoclimatic and palaeoecological data sets to draw inferences about past variations in hydroclimate and forest composition and biomass. The temporal domain of this study is 850–1850 AD and is bounded by the temporal extent of the climate drivers available for our model simulations (850 AD) and time of EuroAmerican settlement-era tree surveys (*c.* 1850 AD). In our study, 'data' refer to observation-based statistical models of past drought, forest composition and biomass, reconstructed from tree rings, historical tree surveys and networks of fossil pollen records. These data-based inferences are fully independent of the ecosystem model simulations. Model-based comparisons are from the PalEON Ecosystem Model Intercomparison Project (PEMIP) (Rollinson *et al.*, 2017), which used spatially and temporally downscaled past climate simulations from the Fifth Coupled Model Intercomparison Project (CMIP5) as drivers. Comparisons among ecosystem model simulations and empirical data rely on normalised values compared in environmental space, rather than geographic space, in order to reduce the effects of any bias in the climate drivers in our analyses and to focus on the sensitivity of ecosystems to climate variability (Supplemental Figure 1).

Observational data sets

The empirically inferred data sets leverage recent advances in pollen-vegetation modelling (Dawson *et al.*, 2016), a form of proxy system modelling (Evans *et al.*, 2013) in which ecosystem state variables such as composition and biomass are estimated along with associated observational uncertainties. Of the three inferred data sets used here, two were derived from networks of fossil pollen records provided by individual data contributors and the Neotoma Paleoecology Database and were calibrated against historical surveys of forest composition and structure from the early stages of EuroAmerican settlement (Liu *et al.*, 2011; Dawson *et al.*, 2016; Goring *et al.*, 2016; Kujawa *et al.*, 2016; Paciorek *et al.*, 2016). Pollen-based inferences are based on statistical pollen-vegetation models (PVMs) called STEPPS and ReFAB, and represent fractional vegetation composition and total woody biomass, respectively, for 12 tree genera that are common elements of upper Midwest forests. STEPPS is a Bayesian hierarchical spatio-temporal model that infers fractional forest composition from networks of fossil pollen records (Paciorek and McLachlan, 2009; Dawson *et al.*, 2016, 2019b; Trachsel *et al.*, 2020b). STEPPS employs a process-based representation of pollen dispersal and production, with taxon-specific parameterisations. STEPPS is calibrated using spatial data sets of pollen samples and forest composition data, here from the settlement era (Paciorek and McLachlan, 2009; Dawson *et al.*, 2016), then run for fossil pollen assemblages for other time intervals to produce posterior estimates of past forest composition. Using this framework, STEPPS: (1) explicitly characterises uncertainty in data and processes, with posterior distributions of process parameters and state variables such as forest composition, and (2) borrows information across space and time, allowing for spatially comprehensive estimates of composition. For both the upper Midwestern USA (UMW; Minnesota, Wisconsin, Michigan) (Dawson et al. 2019a) and the northeastern USA (NEUS) (Trachsel et al. 2020a,b), STEPPS has been used to estimate centennially resolved forest composition for the late Holocene (250 B.C. to 1750 A.D) at a 24 km grid; here we use the results from 850 to 1750 AD.

ReFAB also employs a similar approach to STEPPS but focuses specifically on estimating total aboveground woody biomass (Raiho et al., 2020). ReFAB is calibrated using the relationship between settlement-era multivariate pollen counts and biomass from PLS surveys (Paciorek *et al.*, 2019). Parameter estimates from calibration are then used to reconstruct centennially resolved biomass for 77 sites in the UMW for the last 10,000 years. ReFAB can characterise the uncertainty in sediment pollen age estimates, calibration parameters, the relationship between species composition and total aboveground woody biomass and species-level allometries.

The Living Blended Drought Atlas (LBDA) provides yearly estimates of summer (mean June, July, August) Palmer Severity Drought Index (PDSI) for North America, based on networks of tree-growth chronologies (Cook *et al.*, 2010; Woodhouse *et al.*, 2010). We used PDSI as our measure of hydroclimate variability because it is an important predictor of forest dynamics in this domain and can also be calculated directly from the meteorological forcings used for the ecosystem model simulations (Clifford and Booth, 2015; Cook *et al.*, 2015). LBDA PDSIs are provided at 0.5-degree spatial grid resolution. Due to varying temporal extent of tree-growth chronologies, the temporal extent of the LBDA varies. The earliest years in this spatial domain ranged from 0 to 1671 AD, while the latest year was 2005 (Supplemental Figure 1).

Modelling data sets

PEMIP model simulations here are composed of five ecosystem models with dynamic vegetation (ED2; LINKAGES; LPG-WSL; LPJ-GUESS; and JULES-TRIFFID) run at 254 locations across the eastern and midwestern United States at 0.5-degree spatial resolution (Rollinson *et al*., 2020). These models vary in how they characterise forest composition and carbon dynamics and range from species-based with little ecophysiological process representation (e.g. LINKAGES) to detailed ecophysiology and cohort representation, but reliance on plant functional types (PFTs; e.g. ED2, Table 1). LPJ-GUESS and LPJ-WSL both included stochastic fire disturbances in their simulations, whereas other models such as ED and LINKAGES include processes of tree mortality that assume landscape-scale equilibrium (Rollinson *et al*., 2017).

PEMIP climate drivers were temporally downscaled and bias-corrected from existing past climate simulations to meet the external forcing needs of the ecosystem model ensemble (Figure S1) (Kumar *et al*., 2012; Rollinson *et al*., 2017). CCSM4 output from the Paleoclimate Modeling Intercomparison Project, Phase III (PMIP3) past millennium simulations and the Coupled Model Intercomparison Project, Phase 5 (CMIP5) historical simulations were downscaled to 0.5-degree spatial resolution and 6-hourly temporal resolution using standard protocols (Kumar *et al*., 2012; Rollinson *et al*., 2017). After the 6-hourly PEMIP climate driver data sets were created, they were then temporally averaged to meet the specific driver requirements of individual ecosystem models, which vary in temporal resolution. ED2 and JULES-TRIFFID use the full suite of 6-hourly drivers for temperature, precipitation, shortwave radiation, longwave radiation, surface pressure, specific humidity, wind speed and carbon dioxide concentration. Meteorological drivers for the two LPJ variants include daily temperature, precipitation and shortwave radiation plus longwave radiation for LPJ-WSL. LINKAGES only requires monthly average temperature and precipitation. Soil texture used to parameterise locations in model simulations was extracted from the Harmonised World Soil Database (Wei *et al*., 2014). Monthly temperature and precipitation were combined with soil water holding capacity computed from model driver soil texture and depth to calculate PDSI, following (Cook *et al*., 2015), but using the Thornthwaite equation for evapotranspiration (Thornthwaite and Mather, 1957; Pelton *et al*., 1960). We used the Thornthwaite equation so that the calculation of PDSI was independent of internal model dynamics, including evapotranspiration, which can vary widely among ecosystem models, even when given the same temperature and precipitation drivers, due to differences in model structure and parameterisation. From the ecosystem models, we extracted fractional forest composition and total aboveground biomass, which can be directly compared to palaeoecological observations, and four variables that are latent, i.e. unobservable in the palaeoecological record (Fig. 1): gross primary productivity (GPP), net primary productivity (NPP), net ecosystem exchange (NEE) and leaf area index (LAI).

Analyses

Analyses focused on the comparison of empirical data and ecosystem model outputs of centennial-scale variability in forest composition and biomass driven by drought variability over the last 1,000 years. Our analytical approach involved three key stages to maximise commensurability between observations and model output: 1) temporal homogenisation of all

Table 1 Comparison of 1) ecosystem model complexity, based on representation of diversity, demographic and ecophysiological processes with 2) variability in forest composition (Comp) and biomass (Biom) and sensitivity to hydroclimate variability

Data source & model name	Tree diversity representation	Demographic representation	Vegetation processes	Comp. Var. (SD)	Comp. Sens. (SE)	Biom. Var. (SD)	Biom. Sens. (SE)
Pollen: STEPPS, ReFAB	Genera: 12 trees	relative abundance	[implicit]	−2.032 (0.617)	0.026 (0.019)	−7.798 (0.770)	−0.156 (0.119)
ED2	PFTs: 5 tree	cohort	photosynthesis, allocation, cross-PFT competition, cross-cohort competition	−7.156 (0.514)	0.118 (0.018) *†	−7.505 (0.446)	−0.079 (0.027)*
LINK-AGES	Species: 15 tree	individual	cross-PFT competition, cross-cohort competition	−6.598 (0.478)	0.074 (0.018) *	−6.741 (0.999)	0.230 (0.028) *†
LPJ-GUESS	PFTs: 6 tree, 1 grass	cohort	photosynthesis, allocation, cross-PFT competition, cross-cohort competition	−7.290 (0.452)	0.056 (0.018) *	−7.379 (0.597)	−0.069 (0.027)*
LPJ-WSL	PFTs: 5 tree, 1 grass	PFT	photosynthesis, allocation, cross-PFT competition, cross-PFT competition	−7.829 (0.943)	0.252 (0.018) *†	−7.106 (0.964)	−0.020 (0.027)
JULES-TRIFFID	PFTs: 2 Tree, 2 grass, 1 shrub	PFT	Photosynthesis, allocation, cross-PFT competition	−8.633 (1.075)	0.411 (0.022) *†	−8.639 (0.952)	0.203 (0.033) *†

Variability is a normalised metric of total change in the centennially resolved time series. Sensitivity is presented as the slope and standard error of linear regression between composition or biomass variability and hydroclimate variability. PFT, plant functional types. For sensitivity columns, *Indicates slopes significantly different from zero ($P < 0.05$); †Indicates model slope significantly different from pollen ($P < 0.05$).

variables to a common centennial resolution; 2) development of a common normalised variability metric for ecosystem and drought variability to facilitate comparison across different variables and 3) use of hydroclimate sensitivity as the basis for all model–data and model–model comparisons to minimise the potential effects of biases in the climate model drivers.

Temporal homogenisation

For annually resolved data sets in our study, including the LBDA and all model output and drivers, a generalised additive model (GAM) was used to generate time series with the similar centennial-scale smoothing as the pollen inferred observational data sets. In this process, the response variable for analysis (e.g. drought, biomass, GPP) was modelled as a function of time (year) using a thin-plate regression spline with one knot per 100 years (e.g. 10 knots for a 1,000 year window) using the *gam* function in the *mgcv* package in R (Wood, 2017; Simpson, 2018). To capture the temporal uncertainty similar to that generated in the PVMs, we generated a 1000-member posterior distribution of each predicted variable through time using the error and covariance of the intercept and spline parameters. We then extracted the predicted values at 100-year intervals corresponding to the windows captured by the STEPPS and ReFAB output.

Variability metric

To facilitate comparisons among variables with different units such as composition and biomass, we developed a base metric for all analyses, consisting of the normalised mean temporal variability of each data set (eq. 1).

$$variability_i = \ln\frac{\bar{d}_i}{\bar{x}} \quad (1)$$

$$_{d} = \left|x_{i,t} - x_{i,t-1}\right| \quad (2)$$

Mean temporal variability at each location ($\bar{d}_i$) for each variable (e.g. composition, biomass, PDSI) was calculated as the mean of the absolute first differences between adjacent time points $(t, t-1)$ extracted from centennially resolved time series for each location (i) (eqn 2). The use of first differences is a discretisation of the first derivative and describes the rate of change at each timestep. Each first-difference calculation was based on the mean of the posterior draws from the STEPPS or ReFAB PVM or to the GAMs fitted to the LBDA data and ecosystem model variables. We normalised variability by dividing the mean first differences for each location ($\bar{d}_i$) by the variable mean for that data set across the entire spatiotemporal domain ($\bar{x}$). For forest compositional data, the variability metric was calculated using the taxon or plant functional type (PFT) with the highest fractional composition at each location, with the choice of taxon or PFT allowed to vary among sites. For all analyses and presented results, normalised variability is log-transformed to meet standard statistical assumptions of Gaussian distributions and homoscedasticity (eqn 1).

Hydroclimate sensitivity

After the normalised temporal variability was calculated for PDSI and all ecosystem variables, sensitivity to hydroclimate variability was defined as the slope of a linear regression between variability as the independent variable and variability of the ecosystem response variable such as composition or biomass. These analyses always used the appropriate observational or modelled PDSI variability (i.e. LBDA for the pollen-inferred compositional variability; calculated PEMIP driver PDSI variability for the model-simulated compositional variability) to ensure internal consistency between climatic forcing and ecosystem response.

RESULTS

In the observational data, variability in forest composition or biomass in the northeastern United States (NEUS) and upper midwestern United States (UMW), did not correlate to drought variability (Table 1, Figs 2 and 3) in contrast with the hypothesis that high exposure to climate variability should lead to increased compositional variability. Neither the full spatiotemporal domain (Table 1) nor the UMW (Fig. 3, sensitivity slope = 0.010 SE 0.018) showed a significant relationship between reconstructed drought and composition variability, although the NEUS showed weak sensitivity (Fig. 3, sensitivity slope = 0.065 SE 0.027). Reconstructed biomass variability (Fig 2., biomass reconstructions not available for the NEUS, (Paciorek *et al.*, 2019)) also was uncorrelated to drought variability (Table 1) and instead showed the highest variability at the historic prairie-forest ecotone (Fig. 2) (Goring & Williams 2017). In pollen-based reconstructions, composition and biomass variability were weakly but positively related (Fig. 3c, $R^2 = 0.09$, slope = 0.479 SE 0.187) and locations with higher taxonomic richness tended to have higher variability (Fig. S1).

Modelled ecosystem sensitivity to drought variability was generally similar to or higher than observations, with less-complex models tending to have a too-high predicted sensitivity relative to the empirical reconstructions (Fig. 3). Composition variability was more sensitive to drought variability than in reconstructions for three of five ecosystem models (ED2, LPJ-WSL and TRIFFID), with the data-model discrepancy most pronounced in models with fewer plant types or taxa (Fig. 3a, Table 1). JULES-TRIFFID, which had only two tree PFTs (deciduous and evergreen), had the highest drought sensitivity (composition slope = −8.633 SE = 1.075, composition sensitivity slope 0.411 SE = 0.022). LPJ-WSL and ED2, with, respectively, six and five PFTs, had similar mean compositional variability (LPJ-WSL slope = −7.829 SE = 0.943, ED2 slope = −7.156 SE = 0.514), although LPJ-WSL was approximately twice as sensitive to hydroclimate variability as ED2 (Fig. 3a, Table 1, LPJ-WSL slope = 0.252 SE = 0.018, ED2 slope = 0.118 SE = 0.018). LINKAGES, which simulated 15 individual species, had among the lowest sensitivity to drought variability (Fig. 3a, Table 1, composition slope = −6.598 SE = 0.478, composition sensitivity slope 0.074 SE = 0.018).

Ecosystem models with simpler representation of vegetation ecophysiology (LINKAGES, JULES-TRIFFID) also had a too-high sensitivity of biomass to drought variability relative to empirical reconstructions (Table 1, Fig. 3b). Both LINKAGES and JULES-TRIFFID showed a tight positive coupling of biomass sensitivity to drought variability, which

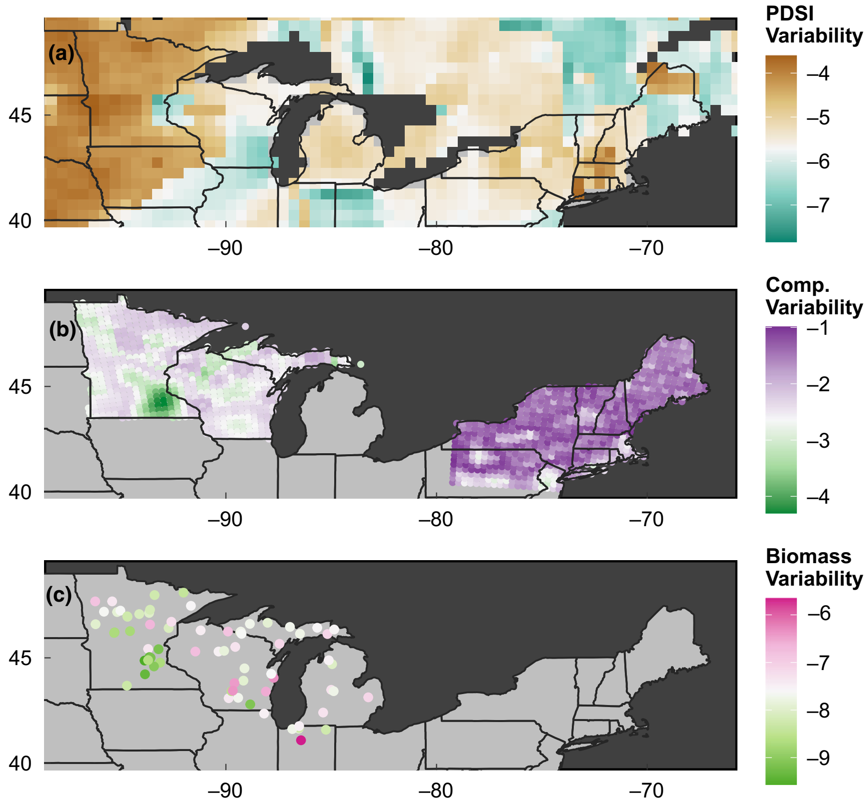

Figure 2 Spatial distribution of inferred temporal variability for 850–1850 AD for a) drought (PDSI) from the Living Blended Drought Atlas (44), b) forest composition from the STEPPS pollen-vegetation model (8, 24) and c) forest aboveground biomass from the ReFab pollen-biomass model (7). All variability estimates were divided by mean to facilitate inter-variable comparison (*Methods*). Spatial extents of compositional and biomass reconstructions are uneven across the study domain, as is the temporal extent of reconstructed drought variability (Figure S1). Empirical comparisons of composition or biomass variability with drought variability are restricted to the common temporal extents for each location. In the log scale, negative values indicate locations with low variability whereas more positive values indicate high variability.

corresponded to strong correlations between biomass and composition variability (Fig. 3c). LINKAGES showed a one-to-one relationship between composition and biomass variability, which is much stronger than reconstructions (Fig. 3c). Of all the models, only LPJ-WSL was consistent with the data in showing a weakly negative relationship between biomass and PDSI variability (Fig. 3b), whereas also showing a positive correlation between biomass and composition variability (Fig. 3c).

Further analysis of latent variables in the ecosystem models confirmed that variations in modelled ecosystem sensitivity to hydroclimate variability is linked to model complexity of ecosystem composition and processes (Fig. 4). There is a cascading series of linkages in physiological variables within and among taxa (Figs 1 and 4), in which gross primary productivity (GPP) is directly influenced by temperature and moisture availability, whereas other state variables such as net primary productivity (NPP), leaf area index (LAI) and aboveground biomass (AGB) are regulated by additional downstream processes that may decouple their variability from climate variability (Fig. 1). Hence, in most models, GPP variability is the most sensitive to drought variability (Fig. 4, Table S1). In all models, the sensitivity of forest composition to drought variability seems to be most closely linked to the sensitivity of NPP. NPP sensitivity tended to be higher in low-diversity models such as JULES-TRIFFID (Figure 4, Table S1). Higher diversity through more tree types or taxa was associated with higher compositional variability and reduced sensitivity to drought (Figure 3, Table 1, Figure S2).

Models with a more detailed representation of plant ecophysiology and either demography or disturbance (e.g. ED2, LPJ-GUESS, LPJ-WSL) also tended to have lower biomass sensitivity to hydroclimate variability (Fig. 4) and agree more closely with observations (Fig. 3). Biomass sensitivity to drought variability in our model ensemble was similar to NEE sensitivity in all models except LPJ-GUESS (Fig. 4, Table S1). LINKAGES and JULES-TRIFFID may be overly sensitive to hydroclimate variability for entirely different reasons. LINKAGES has a fairly simple representation of ecophysiological processes while being able to represent species-level demographic dynamics (Table 1). In contrast, JULES-TRIFFID contains a sophisticated representation of ecophysiology but for only two tree PFTs and five PFTs total (Table 1). The other models tend to be more intermediate cases, with intermediate to more sophisticated representations of both ecophysiology and vegetation dynamics.

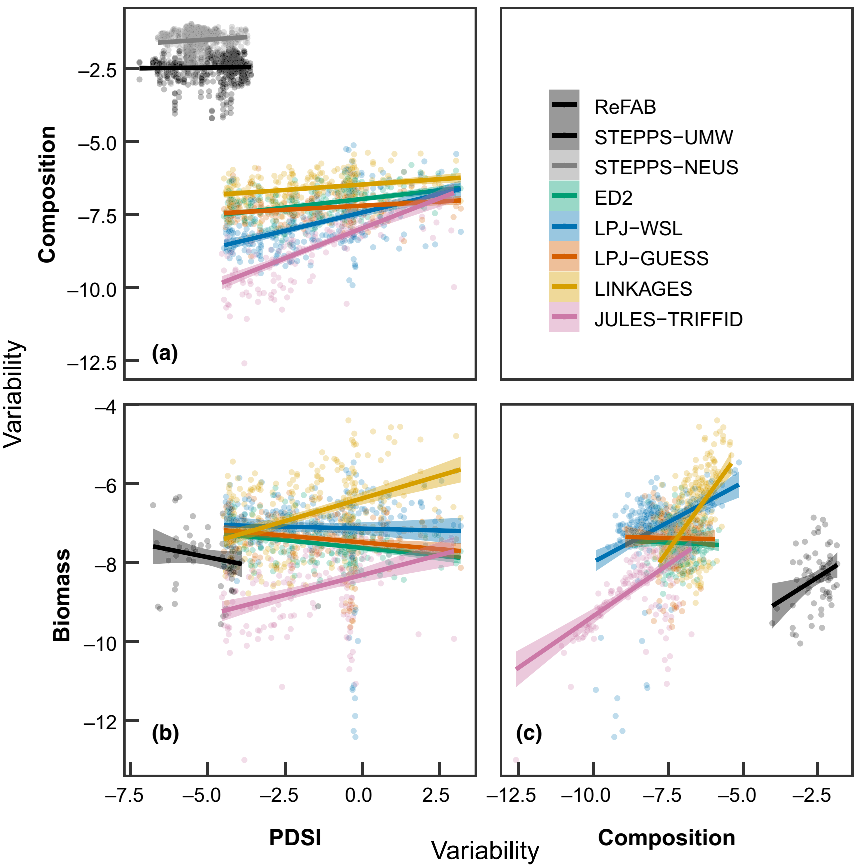

Figure 3 Inferred (black, grey) and simulated (colours) sensitivity of variability of forest composition and biomass to ecohydrological variability (PDSI) (a and b) and of biomass variability to compositional variability (c). Inferred variables suggest weak to no correlation (low sensitivity) between climate variability and ecosystem variability (composition and biomass). In contrast, ecosystem models generally simulate higher sensitivity of ecosystems to climate variability. Inferred compositional (STEPPS) and biomass (ReFAB) variability are positively correlated, whereas this relationship varied among models. In the log scale, negative values indicate locations with low variability, whereas more positive values indicate high variability.

DISCUSSION

Over the last millennium (850–1850 A.D.), both palaeodata networks and model simulations suggest that spatial patterns in forest composition and biomass variability in northeastern and upper midwestern United States are governed more by spatial variations in ecosystem sensitivity and less by spatial variations in exposure to climate variability. Ecotonal regions such as the prairie-forest border have higher variability in composition and structure than areas of high PDSI variability (Fig. 2). The intermodel comparisons suggest that added complexity allows slow-to-change variables such as composition and biomass to be insensitive to climate variability at centennial scales despite sensitivity of fast-changing ecophysiological processes such as gross and net primary productivity (Fig. 4). Incorporation of ecological processes and characteristics such as diversity and demography all tend to reduce simulated climate sensitivity and better align simulations with observations (Figs 3 and 4).

These analyses represent a milestone towards the goal of more comprehensive and rigorous data-model comparisons for timescales and time periods extending beyond the instrumental record. Common challenges for multi-centennial data-model comparisons include 1) a need for process-informed statistical models of inference for palaeoecological data, 2) generally lower temporal resolution in palaeoecological data than in model simulations and with more latent variables than for the instrumental period, 3) biases in palaeoclimatic simulations leading to biases in ecosystem model simulations and 4) differences among models in driver data sets and represented processes. The pollen-vegetation models used in our study include processes for pollen productivity and dispersal that translates relative pollen abundances into metrics of forest composition and biomass that can be directly compared to those produced by ecosystem models (Paciorek and McLachlan, 2009; Dawson *et al.*, 2016). We further increased the commensurability between centennially resolved pollen-based quantifications of forest change and higher frequency information from tree rings and ecosystem models using GAMs to achieve time series with similarly temporally smoothed properties (Simpson, 2018). By focusing on time series variability rather than directly comparing magnitude and timing of the change in specific geographical locations or taxonomic groupings we were able to overcome documented ecosystem model biases arising from driver, process and parameter limitations (Matthes *et al.*, 2016; Dietze, 2017). Finally, we leveraged

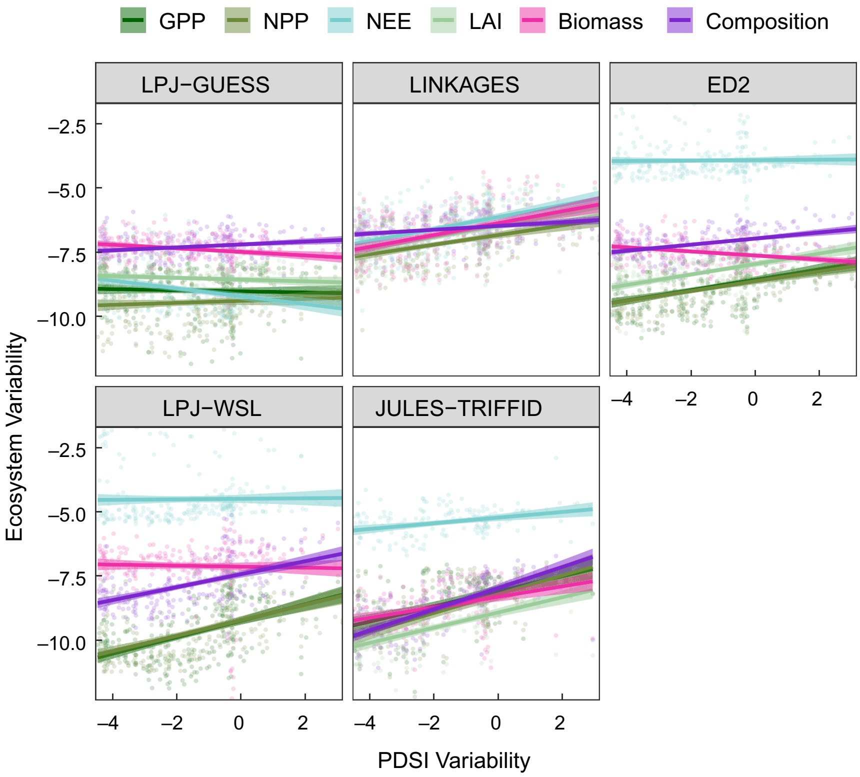

Figure 4 Diagnosing the observed and latent relationships among ecohydrological variability and variability in forest composition, structure and function in five terrestrial ecosystem models. Models in the top row (LPJ-GUESS, LINKAGES, ED2) have individual- or cohort-level representation of demography whereas those in the bottom row (LPJ-WSL, JULES-TRIFFID) do not simulate demography. All models showed positive correlations between composition and drought variability, but some models showed positive biomass sensitivities (LINKAGES, JULES-TRIFFID), whereas others were negative (ED2, LPJ-WSL, LPJ-GUESS). In all models, composition sensitivity to hydroclimate variability was most similar to NPP, whereas biomass sensitivity tended to mirror NEE. In the log scale, negative values indicate locations with low variability, whereas more positive values indicate high variability.

differences in process representation among models as a means of evaluating the importance of specific ecosystem processes for producing emergent patterns of climate sensitivity that are consistent with palaeoecological data (Medlyn *et al.*, 2015; McLachlan & PalEON Project 2018).

Prior studies have indicated that forest composition and growth is sensitive to climate variability at annual to centennial scales (Shuman *et al.*, 2004; Allen *et al.*, 2010; Thom *et al.*, 2019), yet there is also increasingly strong evidence that tree-climate relationships are non-stationary and subject to multiple interacting factors, leading to spatially complex forest responses to climate change (Girardin *et al.*, 2016) and variations in climatic sensitivity across space and time (Rollinson *et al.* 2021; Thom *et al.*, 2019; Peltier and Ogle, 2020; Wilmking *et al.*, 2020). Several possible explanations exist for the reporting here of generally low sensitivity of forest composition and biomass to hydroclimate in reconstructions (Fig. 2). First, this apparent insensitivity may be due to the temporal grain of this study. The centennially resolved temporal grain of our analyses limits detection of annual-scale growth variations, the effects of stochastic or short-lived extreme events such as sub-decadal to decadal drought (Breshears *et al.*, 2005; Allen *et al.*, 2010; Seidl *et al.*, 2011), or disturbance events such as fire and pest outbreaks, unless these are large enough to cause stand-replacing mortality events. Disturbance processes are often unrepresented in ecosystem models or treated as purely stochastic and with implicit assumptions of landscape-scale equilibria (Seidl *et al.*, 2011; Fisher *et al.*, 2018; McCabe and Dietze, 2019). Of the ecosystem models used here, LPJ-WSL and LPJ-GUESS included fire in their simulations as a semi-mechanistic process following GLOBFIRM (Thonicke *et al.*, 2001), which estimates burned area as a function of daily fire probabilities that are a function of fuel moisture and fuel load threshold. These models showed dampened biomass sensitivity to hydroclimate variability that was more closely aligned with observations (Fig. 4), but so did ED2, which lacked fire. Hence, process representation of fire or similar semi-stochastic disturbances is not a clear differentiator among modelled estimates of ecosystem climate sensitivity.

Second, apparent climate sensitivity might increase if the temporal extent was increased to include larger climate variations during the Holocene and last deglaciation. Although the last millennium includes climatic events such as the Medieval Climate Anomaly and Little Ice Age (PAGES 2k Consortium, 2013), these climate variations appear to have been muted relative to earlier hydroclimate and temperature variations (Fischer *et al.*, 2018). During the Holocene, hydroclimatic variability around the North Atlantic appears to have been an important driver of forest compositional changes and the collapses of individual tree species (Shuman *et al.*, 2019). Large vegetation changes associated with the abrupt temperature variations of the Younger Dryas and last deglaciation are well documented (Williams *et al.*, 2011), but the temporal extent of this study was constrained by the temporal extent of the last-millennium PMIP3/CMIP5 simulations used to drive ecosystem models (Braconnot *et al.*, 2011; Taylor *et al.*, 2012). As the next generation of transient Holocene simulations become available, the conclusions reached here about low apparent sensitivity can be revisited.

Third, this paper focuses on spatial patterns of climate and ecosystem variability, whereas most prior palaeoecological studies have tended to focus on temporal variations (Shuman *et al*., 2004; Booth *et al*., 2012). Our analyses of low sensitivity are consistent with recent dendroecological studies of climate-driven rates of tree growth, which are quickly shifting from assumptions of stationary tree-climate relationships to demonstrations of spatially complex forest responses (Girardin *et al*., 2016) and variations in climatic sensitivity varies across space and time (Rollinson et al. 2021; Thom *et al*., 2019; Peltier and Ogle, 2020; Wilmking *et al*., 2020). By focusing on spatial variations in ecosystem variability over the last millennium, our analyses suggest spatial variation in ecosystem properties are a more important regulator than spatial variations in climate exposure. Finally, uncertainties in the proxy-based reconstructions may lower correlations as detrending techniques used to remove non-climatic signals such as age effects may dampen estimates of centennial-scale variability (Allen *et al*., 2018; Esper *et al*., 2018). Despite lower PDSI variability in the LBDA than model drivers, we do not think that spatial variability in hydroclimate variability in the empirical data set is too low to detect effects on ecosystem variability. For example hydroclimate data syntheses for the last 2000 years suggest opposite patterns of hydroclimate variations between Minnesota/Wisconsin and New England, which explain 30% of variance in the hydroclimate records (Shuman *et al*., 2019).

Process-based ecosystem models are the main vehicle for forecasting climate-driven ecosystem dynamics across a range of timescales and in principle are better able to accommodate past and future no-analog climates (Williams and Jackson, 2007; Veloz *et al*., 2012). However, all ecosystem models face trade-offs in their ability to represent taxonomic or functional diversity versus detailed ecophysiological processes that drive ecosystem change (Fisher *et al*., 2018). Process-based ecosystem models will never be able to capture the full complexity of ecosystems nor perfectly reproduce the patterns of climatological or ecological variability observed in the past due to observational uncertainties and incomplete constraints of many processes and parameterisations (Dietze, 2017). This paper has shown how multiple palaeoecological data streams can be combined with harmonised palaeoclimatic simulations and multiple terrestrial ecosystem models to gain new insight into a) how diversity and biological processes can dampen ecosystem sensitivity to drought variability at broad spatial scales and b) the importance of complex representations of these aspects of ecosystems to achieve better agreement with the data. Nevertheless, these analyses followed a traditional approach in which past ecosystem reconstructions and simulations were run independently and compared at the final stage of analysis. The next major step forward is to move to a full data-assimilation framework, in which palaeoecological observations and simulations are combined to overcome systematic biases in model drivers, parameterisation and output to better evaluate palaeoecological change using mechanistic process-based frameworks (McLachlan & PalEON Project 2018). Through this iterative process that draws upon an ever-growing and diversifying suite of observational data streams (Farley *et al*., 2018), we can better understand the mechanisms regulating forest sensitivity to climate variability across a broad range of timescales and thereby better forecast future forest dynamics in a complex and rapidly changing world.

ACKNOWLEDGEMENTS

This work reflects the efforts of the Paleoecological Observatory Network (PalEON Project), funded by the National Science Foundation MacroSystems Biology under grants DEB-1241891, DEB-1241868, DEB-1241874 and DEB-1241851 and special thanks to Jody Peters, PalEON Project coordinator. PDSI calculations from ecosystem model drivers were derived from code graciously provided by Ben Cook. Fossil pollen data were obtained from the Neotoma Paleoecology Database (http://www.neotomadb.org) and its constituent database the North American Pollen Database. The work of the data contributors, data stewards and the Neotoma community is gratefully acknowledged. Any use of trade, firm or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

AUTHORSHIP

CRR, AD, AMR and JWW designed the study. AD, AMR, JWW, STJ, JM and MT created pollen reconstructions and aided in interpretation (STEPPS, ReFAB). AD and AMR wrote the pollen methods. CRR, AMR, MCD, JM, DJPM, BP, TQ and JS performed ecosystem model simulations and aided in interpretation. CRR and JWW wrote the manuscript with additional input from AD, AMR and all authors.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ele.13667.

DATA AVAILABILITY STATEMENT

Pollen data are already available or will be made available upon acceptance on the EDI data portal as an msb-paleon product. Composition Products are archived with the following DOIs: https://doi.org/10.6073/pasta/c941fc6bd9eb959384592766dc0e9dbb (UMW), https://doi.org/10.6073/pasta/06e149ccece9015a23fc332d8f7de47 (NEUS). The ReFAB biomass product has the following https://doi.org/10.6073/pasta/9e22f05a75e697e20c227f42c2c72d09. The Environmental Data Initiative is an NSF-funded program tailored towards environmental data and works closely with the US Long-Term Ecological Research (LTER) Network, NSF Macrosystems Biology program (which funded our work) and DataONE. Terrestrial ecosystem model drivers are being archived on the ORNL DAAC and will be available at the following https://doi.org/10.3334/ORNLDAAC/1779. The Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC) is managed by NASA's Earth Science Data and Information Systems program and is well suited to archive ecosystem model output, which is often large and has converged on netcdf as a standard file format. These repositories have been approved by *Ecology Letters* editorial staff. All code for analyses is publicly available on Github: https://github.com/Pa lEON-Project/EcosystemVariability. Readers should get in contact with the journal if they discover issues with data.

REFERENCES

Allen, C.D., Macalady, A.K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M. *et al.* (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *For. Ecol. Manag*., 259, 660–684.

Allen, K.J., Villalba, R., Lavergne, A., Palmer, J.G., Cook, E.C., Fenwick, P. *et al.* (2018). A comparison of some simple methods used to detect unstable temperature responses in tree-ring chronologies. *Dendrochronologia*, 48, 52–73.

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P. *et al.* (2013). Evaluating the land and ocean components of the global carbon cycle in the CMIP5 earth system models. *J. Clim*., 26, 6801–6843.

Blarquez, O. & Aleman, J.C. (2016). Tree biomass reconstruction shows no lag in postglacial afforestation of eastern Canada. *Can. J. For. Res*., 46, 485–498.

Booth, R.K., Jackson, S.T., Sousa, V.A., Sullivan, M.E., Minckley, T.A. & Clifford, M.J. (2012). Multi-decadal drought and amplified moisture variability drove rapid forest community change in a humid region. *Ecology*, 93, 219–226.

Braconnot, P., Otto-Bliesner, B., Jungclaus, J. & Peterschmitt, J.-Y. (2011). The paleoclimate modeling intercomparison project contribution to CMIP5. *CLIVAR Exch*., 16, 15–19.

Breshears, D.D., Cobb, N.S., Rich, P.M., Price, K.P., Allen, C.D., Balice, R.G. *et al.* (2005). Regional vegetation die-off in response to global-change-type drought. *Proc. Natl. Acad. Sci*., 102, 15144–15148.

Charney, N.D., Babst, F., Poulter, B., Record, S., Trouet, V.M., Frank, D. *et al.* (2016). Observed forest sensitivity to climate implies large changes in 21st century North American forest growth. *Ecol. Lett*., 19, 1119–1128.

Clifford, M.J. & Booth, R.K. (2015). Late-Holocene drought and fire drove a widespread change in forest community composition in eastern North America. *The Holocene*, 25, 1102–1110.

Cook, B.I., Ault, T.R. & Smerdon, J.E. (2015). Unprecedented 21st century drought risk in the American Southwest and Central Plains. *Sci. Adv*., 1, e1400082.

Cook, E.R., Seager, R., Heim, R.R., Vose, R.S., Herweijer, C. & Woodhouse, C. (2010). Megadroughts in North America: placing IPCC projections of hydroclimatic change in a long-term palaeoclimate context. *J. Quat. Sci*., 25, 48–61.

Dawson, A., Paciorek, C.J., McLachlan, J.S., Goring, S., Williams, J.W. & Jackson, S.T. (2016). Quantifying pollen-vegetation relationships to reconstruct ancient forests using 19th-century forest composition and pollen data. *Quat. Sci. Rev*., 137, 156–175.

Dawson, A., Goring, S., Paciorek, C.J., Jackson, S.T., McLachlan, J.S. & Williams, J.W. (2019a). STEPPS 2000 Year Forest Composition Estimates, Upper Midwest US, Level 2. Environmental Data Initiative. https://doi.org/10.6073/pasta/c941fc6bd9eb959384592766dc0e9dbb

Dawson, A., Paciorek, C.J., Goring, S.J., Jackson, S.T., McLachlan, J.S. & Williams, J.W. (2019b). Quantifying trends and uncertainty in prehistoric forest composition in the upper Midwestern United States. *Ecology*, 100, e02856. Available at: https://esajournals.onlinelibrary.wiley.com/doi/full/10.1002/ecy.2856. Last accessed 21 December 2020.

Dawson, T.P., Jackson, S.T., House, J.I., Prentice, I.C. & Mace, G.M. (2011). Beyond predictions: biodiversity conservation in a changing climate. *Science*, 332.

Dietze, M.C. (2017). Prediction in ecology: a first-principles framework. *Ecol. Appl*., 27, 2048–2060.

Dietze, M.C., Fox, A., Beck-Johnson, L.M., Betancourt, J.L., Hooten, M.B., Jarnevich, C.S. *et al.* (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proc. Natl. Acad. Sci*., 115, 1424–1432.

Esper, J., George, S.S., Anchukaitis, K., D'Arrigo, R., Ljungqvist, F.C., Luterbacher, J. *et al.* (2018). Large-scale, millennial-length temperature reconstructions from tree-rings. *Dendrochronologia*, 50, 81–90.

Evans, M.N., Tolwinski-Ward, S.E., Thompson, D.M. & Anchukaitis, K.J. (2013). Applications of proxy system modeling in high resolution paleoclimatology. *Quat. Sci. Rev*., 76, 16–28.

Farley, S.S., Dawson, A., Goring, S.J. & Williams, J.W. (2018). Situating ecology as a big-data science: current advances, challenges, and solutions. *Bioscience*, 68, 563–576.

Fischer, H., Meissner, K.J., Mix, A.C., Abram, N.J., Austermann, J., Brovkin, V. *et al.* (2018). Palaeoclimate constraints on the impact of 2 °C anthropogenic warming and beyond. *Nat. Geosci*., 11, 474–485.

Fisher, R.A., Koven, C.D., Anderegg, W.R.L., Christoffersen, B.O., Dietze, M.C., Farrior, C.E. *et al.* (2018). Vegetation demographics in Earth System Models: A review of progress and priorities. *Glob. Change Biol*., 24, 35–54.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V. *et al.* (2006). Climate-carbon cycle feedback analysis: Results from the C^4 MIP model intercomparison. *J. Clim*., 19, 3337–3353.

Friedlingstein, P., Meinshausen, M., Arora, V.K., Jones, C.D., Anav, A., Liddicoat, S.K. *et al.* (2014). Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *J. Clim*., 27, 511–526.

Girardin, M.P., Bouriaud, O., Hogg, E.H., Kurz, W., Zimmermann, N.E., Metsaranta, J.M. *et al.* (2016). No growth stimulation of Canada's boreal forest under half-century of combined warming and CO_2 fertilization. *Proc. Natl. Acad. Sci*., 113, E8406–E8414.

Goring, S.J., Mladenoff, D.J., Cogbill, C.V., Record, S., Paciorek, C.J., Jackson, S.T. *et al.* (2016). Novel and lost forests in the upper midwestern united states, from new estimates of settlement-era composition, stem density, and biomass. *PLoS One*, 11, e0151935.

De Kauwe, M.G., Medlyn, B.E., Zaehle, S., Walker, A.P., Dietze, M.C., Hickler, T. *et al.* (2013). Forest water use and water use efficiency at elevated CO 2: a model-data intercomparison at two contrasting temperate forest FACE sites. *Glob. Change Biol*., 19, 1759–1779.

Keeley, J.E., van Mantgem, P. & Falk, D.A. (2019). Fire, climate and changing forests. *Nat. Plants*, 5, 774–775.

Klesse, S., Babst, F., Lienert, S., Spahni, R., Joos, F., Bouriaud, O. *et al.* (2018). A combined tree ring and vegetation model assessment of european forest growth sensitivity to interannual climate variability. *Global Biogeochemical Cycles*, 32(8), 1226–1240.

Kujawa, E.R., Goring, S., Dawson, A., Calcote, R., Grimm, E.C., Hotchkiss, S.C. *et al.* (2016). The effects of anthropogenic land cover change on pollen-vegetation relationships in the American Midwest. *Anthropocene*, 15, 60–71.

Kumar, J., Brooks, B.-G.-J., Thornton, P.E. & Dietze, M.C. (2012). Sub-daily statistical downscaling of meteorological variables using neural networks. *Procedia Comput. Sci*., 9, 887–896.

LeBauer, D.S., Wang, D., Richter, K.T., Davidson, C.C. & Dietze, M.C. (2013). Facilitating feedbacks between field measurements and ecosystem models. *Ecol. Monogr*., 83, 133–154.

Liu, F., Mladenoff, D.J., Keuler, N.S. & Moore, L.S. (2011). Broadscale variability in tree data of the historical Public Land Survey and its consequences for ecological studies. *Ecol. Monogr*., 81, 259–275.

Matthes, J.H., Goring, S., Williams, J.W. & Dietze, M.C. (2016). Benchmarking historical CMIP5 plant functional types across the Upper Midwest and Northeastern United States. *J. Geophys. Res. Biogeosciences*, 121, 523–535.

McCabe, T.D. & Dietze, M.C. (2019). Scaling contagious disturbance: A spatially-implicit dynamic model. *Frontiers in Ecology and Evolution*, 7, 64. Available at: https://www.frontiersin.org/articles/10.3389/fevo.2019.00064/full. Last accessed 21 December 2020.

McLachlan, J.S. & PalEON Project (2018). Forecasting long-term ecological dynamics using open paleodata. *Past Glob. Change Mag*., 26, 76.

Medlyn, B.E., Zaehle, S., De Kauwe, M.G., Walker, A.P., Dietze, M.C., Hanson, P.J. *et al.* (2015). Using ecosystem experiments to improve vegetation models. *Nat. Clim. Change*, 5, 528–534.

Nolan, C., Overpeck, J.T., Allen, J.R.M., Anderson, P.M., Betancourt, J.L., Binney, H.A. *et al.* (2018). Past and future global transformation of terrestrial ecosystems under climate change. *Science*, 361, 920–923.

Paciorek, C.J., Goring, S.J., Thurman, A.L., Cogbill, C.V., Williams, J.W., Mladenoff, D.J. *et al.* (2016). Statistically-estimated tree composition for the Northeastern United States at Euro-American Settlement. *PLoS One*, 11, e0150087.

Paciorek, C.J., Cogbill, C.V., Peters, J.A., Goring, S.J., Williams, J.W., Mladenoff, D.J. *et al.* (2019). Statistically-estimated tree biomass, stem density, and basal area for the upper Midwestern United States at the time of Euro-American settlement (preprint). *bioRxiv*. https://doi.org/10.1101/856526

Paciorek, C.J. & McLachlan, J.S. (2009). Mapping ancient forests: Bayesian inference for spatio-temporal trends in forest composition using the fossil pollen proxy record. *J. Am. Stat. Assoc.*, 104, 608–622.

PAGES 2k Consortium (2013). Continental-scale temperature variability during the past two millennia. *Nat. Geosci.*, 6, 339–346.

Peltier, D.M.P. & Ogle, K. (2020). Tree growth sensitivity to climate is temporally variable. *Ecol. Lett.*, 23(11), 1–20.

Pelton, W.L., King, K.M. & Tanner, C.B. (1960). An evaluation of the Thornthwaite and mean temperature methods for determining potential evapotranspiration. *Agron. J.*, 52, 387–395.

Raiho, A., McLachlan, J. & Paciorek, C. (2020). ReFAB 10,000 year statistical estimate of aboveground wood biomass. Midwest US, Level 2 ver 0. Environmental Data Initiative. https://doi.org/10.6073/pasta/9e22f05a75e697e20c227f42c2c72d09

Rollinson, C.R., Finley, A.O., Alexander, M.R., Banerjee, S., Dixon Hamil, K.-A. & Koenig, L. et al. (2021). Working across space and time: nonstationarity in ecological research and application. *Front. Ecol. Environ.*

Rollinson, C.R., Liu, Y., Raiho, A., Moore, D.J.P., McLachlan, J., Bishop, D.A. *et al.* (2017). Emergent climate and CO_2 sensitivities of net primary productivity in ecosystem models do not agree with empirical data in temperate forests of eastern North America. *Glob. Change Biol.*, 23, 2755–2767.

Rollinson, C., Dietze, M., Mathes, J.H., Steinkamp, J., Hickler, T., Poulter, B. *et al.* (2020). PalEON: Terrestrial Ecosystem Model Ensemble for the Northeastern U.S., 850–2010. ORNL Distributed Active Archive Center.

Schimel, D., Stephens, B.B. & Fisher, J.B. (2015). Effect of increasing CO_2 on the terrestrial carbon cycle. *Proc. Natl. Acad. Sci.*, 112, 436–441.

Seidl, R., Fernandes, P.M., Fonseca, T.F., Gillet, F., Jönsson, A.M., Merganičová, K. *et al.* (2011). Modelling natural disturbances in forest ecosystems: a review. *Ecol. Model.*, 222, 903–924.

Shuman, B., Newby, P., Huang, Y. & Webb, T. (2004). Evidence for the close climatic control of New England vegetation history. *Ecology*, 85, 1297–1310.

Shuman, B.N., Marsicek, J., Oswald, W.W. & Foster, D.R. (2019). Predictable hydrological and ecological responses to Holocene North Atlantic variability. *Proc. Natl. Acad. Sci.*, 116, 5985–5990.

Simpson, G.L. (2018). Modelling palaeoecological time series using generalised additive models. *Frontiers in Ecology and Evolution*, 6, 149.

Taylor, K.E., Stouffer, R.J. & Meehl, G.A. (2012). An Overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.*, 93, 485–498.

Thom, D., Golivets, M., Edling, L., Meigs, G.W., Gourevitch, J.D., Sonter, L.J. *et al.* (2019). The climate sensitivity of carbon, timber, and species richness covaries with forest age in boreal–temperate North America. *Glob. Change Biol.*, 25, 1446–2458.

Thonicke, K., Venevsky, S., Sitch, S. & Cramer, W. (2001). The role of fire disturbance for global vegetation dynamics: coupling fire into a Dynamic Global Vegetation Model. *Glob. Ecol. Biogeogr.*, 10, 661–677.

Thornthwaite, C.W. & Mather, J.R. (1957). Instructions and tables for computing potential evapotranspiration and the water balance. *Publ. Climatol.*, 10, 185–311.

Trachsel, M., Dawson, A., Paciorek, C.J., Jackson, S.T., McLachlan, J.S., Williams, J.W., et al. (2020a). STEPPS 8000 Year Forest Composition Estimates, Northeastern US, Level 2. Environmental Data Initiative. https://doi.org/10.6073/pasta/06e149ccece9015a23fc332d8f7de47

Trachsel, M., Dawson, A., Paciorek, C.J., Williams, J.W., McLachlan, J.S., Cogbill, C.V. *et al.* (2020b). Comparison of settlement-era vegetation reconstructions for STEPPS and REVEALS pollen–vegetation models in the northeastern United States. *Quat. Res.*, 95, 23–42.

Veloz, S.D., Williams, J.W., Blois, J.L., He, F., Otto-Bliesner, B. & Liu, Z. (2012). No-analog climates and shifting realized niches during the late quaternary: implications for 21st-century predictions by species distribution models. *Glob. Change Biol.*, 18, 1698–1713.

Walker, A.P., Hanson, P.J., De Kauwe, M.G., Medlyn, B.E., Zaehle, S., Asao, S. *et al.* (2014). Comprehensive ecosystem model-data synthesis using multiple data sets at two temperate forest free-air CO_2 enrichment experiments: Model performance at ambient CO_2 concentration: FACE MODEL-DATA SYNTHESIS. *J. Geophys. Res. Biogeosciences*, 119, 937–964.

Walker, A.P., Zaehle, S., Medlyn, B.E., De Kauwe, M.G., Asao, S., Hickler, T. *et al.* (2015). Predicting long-term carbon sequestration in response to CO_2 enrichment: How and why do current ecosystem models differ? *Glob. Biogeochem. Cycles*, 29, 476–495.

Wei, Y., Liu, S., Huntzinger, D.N., Michalak, A.M., Viovy, N., Post, W.M. *et al.* (2014). The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project – Part 2: Environmental driver data. *Geosci. Model Dev.*, 7, 2875–2893.

Williams, J.W., Blois, J.L. & Shuman, B.N. (2011). Extrinsic and intrinsic forcing of abrupt ecological change: case studies from the late Quaternary: Extrinsic and intrinsic abrupt ecological change. *J. Ecol.*, 99, 664–677.

Williams, J.W. & Jackson, S.T. (2007). Novel climates, no-analog communities, and ecological surprises. *Front. Ecol. Environ.*, 5, 475–482.

Wilmking, M., Maaten-Theunissen, M., Maaten, E., Scharnweber, T., Buras, A., Biermann, C. *et al.* (2020). Global assessment of relationships between climate and tree growth. *Glob. Change Biol.*, 26, 3212–3220.

Wood, S.N. (2017). *Generalized Additive Models: An Introduction with R*, 2nd edn. Chapman and Hall/CRC, New York.

Woodhouse, C.A., Meko, D.M., MacDonald, G.M., Stahle, D.W. & Cook, E.R. (2010). A 1,200-year perspective of 21st century drought in southwestern North America. *Proc. Natl. Acad. Sci.*, 107, 21283–21288.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Editor, Xavier Morin
Manuscript received 3 April 2020
First decision made 18 November 2020
Manuscript accepted 23 November 2020