

LETTER • OPEN ACCESS

## The signature of internal variability in the terrestrial carbon cycle

To cite this article: Gordon B Bonan *et al* 2021 *Environ. Res. Lett.* **16** 034022

View the [article online](#) for updates and enhancements.

You may also like

- [Uncertainty in carbon budget estimates due to internal climate variability](#)  
Katarzyna B Tokarska, Vivek K Arora, Nathan P Gillett et al.
- [The effect of spatial averaging and glacier melt on detecting a forced signal in regional sea level](#)  
Kristin Richter, Ben Marzeion and Riccardo Riva
- [Pattern formation in spatially forced thermal convection](#)  
S Weiss, G Seiden and E Bodenschatz

ENVIRONMENTAL RESEARCH  
LETTERS

## LETTER

## The signature of internal variability in the terrestrial carbon cycle

## OPEN ACCESS

RECEIVED  
25 September 2020REVISED  
16 December 2020ACCEPTED FOR PUBLICATION  
24 December 2020PUBLISHED  
17 February 2021Gordon B Bonan<sup>1</sup> , Danica L Lombardozzi<sup>1</sup> and William R Wieder<sup>1,2</sup> <sup>1</sup> Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, United States of America<sup>2</sup> Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, United States of AmericaE-mail: [bonan@ucar.edu](mailto:bonan@ucar.edu)**Keywords:** internal variability, carbon cycle, large ensemble, irreducible uncertaintySupplementary material for this article is available [online](#)Original content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](#).Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the title  
of the work, journal  
citation and DOI.**Abstract**

Uncertainty in model initial states produces uncertainty in climate simulations because of unforced variability internal to the climate system. Climate scientists use initial-condition ensembles to separate the forced signal of climate change from the unforced internal variability. Our analysis of an 11-member initial-condition ensemble from the Community Earth System Model Version 2 that spans the period 1850–2014 shows that a similar ensemble approach is needed to robustly assess trends in the terrestrial carbon cycle. Uncertainty in model initialization gives rise to internal variability that masks trends in carbon fluxes, and also creates spurious unforced trends, during the period 1960–2014 across North America, meaning that a single model realization can diverge from the observational record or from other models simply because of random behavior. The forced response is, however, evident in the ensemble mean and emerges from the noise of unforced variability at decadal timescales. Our results suggest that trends in the observational record must be interpreted with caution because of multiple possible histories that would have been observed if the sequence of internal variability had unfolded differently. Furthermore, internal variability produces irreducible uncertainty in the carbon cycle, leading to ambiguity in the magnitude and sign of carbon cycle trends, especially at small spatial scales and short timescales. The small spread in initial land carbon pools at 1850 suggests that internal climate variability arising from atmospheric and oceanic initialization, not the biogeochemical initialization, is the predominant cause of carbon cycle variability among ensemble members. Initial-condition ensembles with other Earth system models are needed to develop a multi-model understanding of internal variability in the terrestrial carbon cycle.

**1. Introduction**

Human-caused climate change is the outcome of forced change from greenhouse gas emissions, aerosols, and other anthropogenic forcings superimposed upon unforced random variations internal to the climate system (termed internal or natural variability) (Collins *et al* 2013). Internal variability arises from the chaotic behavior of the atmosphere at timescales of several days to a few weeks (Lorenz 1963) and from longer term variations in the coupled atmosphere-ocean system such as the El Niño/Southern Oscillation, Pacific decadal variability, and Atlantic multi-decadal variability (Deser *et al* 2012a, 2014). It manifests as climate variability in the observational record at subseasonal-to-seasonal timescales, and also as interannual variability and

decadal or longer fluctuations. Internal variability is seen in climate models in that imprecision in initial atmospheric and oceanic states produce different climate trajectories over the historical record and in future projections, each with its own random sequence of unforced variability and each an equally plausible realization of climate change. Research in recent years utilizing single-model large ensembles, with 20–100 ensemble members that differ only in initial conditions, shows that internal variability generates considerable ensemble spread in climate, particularly at regional spatial scales, and has produced new insights for interpreting the observational record, for model evaluation in comparison with observations and other models, and to separate the forced signal of climate change from unforced variability (Deser *et al* 2012b, 2020, Kay *et al* 2015,

Maher *et al* 2019). In particular, internal variability can mask anthropogenic influences on climate, and a single model realization may diverge from the observational record, or from other models, simply because of random behavior. Furthermore, internal variability represents irreducible uncertainty in climate simulations and limits the accuracy of climate predictions, especially at regional spatial scales and timescales up to a few decades (Hawkins *et al* 2016).

Use of initial-condition ensembles to investigate internal variability has gained considerable traction among climate scientists, seen, for example, in analysis of ensemble spread in temperature and precipitation trends (Deser *et al* 2012a, 2012b, 2014, 2016, 2020, Kay *et al* 2015, Martel *et al* 2018, Swain *et al* 2018, Bengtsson and Hodges 2019, Dai and Bloecker 2019, Maher *et al* 2019). Internal variability is a large source of uncertainty in projections of future climate at regional and decadal scales (Hawkins and Sutton 2009, 2011, Lehner *et al* 2020). Moreover, the forced anthropogenic signal may not emerge from the noise within the climate system for several decades, as seen in the concept of time of emergence (Hawkins and Sutton 2012, Lehner *et al* 2017). Internal variability affects not just physical climate; for example, initial-condition ensembles reveal variability in Arctic sea ice (Swart *et al* 2015, Kirchmeier-Young *et al* 2017, Duvivier *et al* 2020, Landrum and Holland 2020) and snowmelt runoff (Mankin *et al* 2015). Initial-condition ensembles have also been used to study unforced variability in ocean biogeochemistry and the time of emergence of the forced signal (Frölicher *et al* 2009, 2016, Rodgers *et al* 2015, Long *et al* 2016, Lovenduski *et al* 2016, McKinley *et al* 2016, Schlunegger *et al* 2019, 2020).

The internal variability of the terrestrial carbon cycle in Earth system models and the associated ensemble spread are less well studied. The divergent climate trajectories arising from uncertainty in atmospheric and oceanic initialization should manifest in unforced variability in carbon cycle trends. Uncertainty in the initialization of land biogeochemical pools may also contribute to spread among ensemble members. An analysis of a six-member initial-condition ensemble of the Community Climate System Model found differences among ensemble members in annual land carbon uptake, especially at small spatial scales, though the spread among members decreased when averaging over decadal or continental scales (Lombardozi *et al* 2014). The study found that internal variability precludes detection of forced changes in the carbon cycle at timescales shorter than several decades in many regions of the world. Earth system models show considerable differences in the terrestrial carbon cycle (Arora *et al* 2020), but the extent to which the differences reflect model structural uncertainty or internal variability is unknown. Initial studies of uncertainty partitioning in terrestrial carbon cycle projections did not

utilize large ensembles to estimate the contribution of internal variability (Hewitt *et al* 2016, Lovenduski and Bonan 2017). Here, we analyze an initial-condition ensemble for the Community Earth System Model (CESM) over the historical era (Danabasoglu *et al* 2020). We focus on North America to illustrate uncertainty in the terrestrial carbon cycle arising from model initialization so as to complement previous analyses of internal variability in North American temperature and precipitation trends (Deser *et al* 2012b, 2014, 2016, 2020). This work contributes to growing literature that climate uncertainty is a key component of carbon cycle uncertainty (Wu *et al* 2017, Bonan *et al* 2019).

## 2. Methods

We analyzed an 11-member ensemble of simulations with the CESM Version 2 (CESM2) that spans the historical period 1850–2014 (Danabasoglu *et al* 2020). Simulations have a nominal 1° horizontal resolution with active atmosphere, ocean, sea ice, and land component models. The atmospheric model for these simulations is the Community Atmosphere Model Version 6, and the simulations are referred to as CESM2(CAM6) in Danabasoglu *et al* (2020). All simulations use the same historical forcings, and atmospheric CO<sub>2</sub> concentration is specified. Danabasoglu *et al* (2020) describe the simulations and the initialization of the component models. The 11 simulations differ only in initial conditions for 1850, which were obtained from different years of a 1200 year preindustrial control simulation (years 501, 601, 631, 661, 691, 721, 751, 781, 811, 841, and 871). Each year represents equally plausible model states for initialization. Danabasoglu *et al* (2020) provide further details on the spinup of the preindustrial control. Uncertainty in atmosphere, ocean, sea ice, and land initial states propagates throughout the system as a result of physical, chemical, and biological processes in the component models and interactions among component models.

The Community Land Model Version 5 (CLM5), the CESM2 component land model, simulates surface fluxes of energy, mass, and momentum for coupling with the atmosphere; the associated hydrologic cycle and runoff of freshwater to the ocean; and terrestrial biogeochemistry so that leaf area index, canopy height, and vegetation and soil carbon pools are prognostic (Lawrence *et al* 2019). In addition to soil temperature, soil moisture, and snow cover, land biogeochemical states must be initialized for 1850. For much of North America, initial vegetation and soil carbon states in 1850 differ by less than ±5% of the 11-member mean (figures S1 and S2 (available online at [stacks.iop.org/ERL/16/034022/mmedia](https://stacks.iop.org/ERL/16/034022/mmedia))). The largest percentage spread in vegetation carbon occurs in regions with herbaceous vegetation and small initial carbon pools. Analyses of individual grid cells

and ecoregions of the US further show that vegetation carbon is in quasi-steady state and the initial conditions differ on the order of  $100\text{--}200\text{ g C m}^{-2}$  (figure S3). Soil carbon has a small trend (relative to pool size) over the preindustrial control, but initial conditions for the ensemble members differ by only  $100\text{--}200\text{ g C m}^{-2}$  (figure S4). Land surface variables related to the energy, water, and carbon cycles are improved in CLM5 compared with previous versions of the model (Lawrence *et al* 2019, Danabasoglu *et al* 2020). We analyzed gross primary production (GPP) and net ecosystem production (NEP). NEP is the net carbon gain after subtracting ecosystem respiration from GPP.

We analyze the period 1960–2014, which overlaps with the Global Carbon Project analysis of carbon trends (Friedlingstein *et al* 2019). We refer to an individual simulation as an ensemble member and the mean of the 11 members as the ensemble mean. We analyzed the time series of each member and the ensemble mean to obtain the trend over the 55 year period from 1960 to 2014. The trend was assessed using a linear fit to the 1960–2014 time series, and statistical significance was determined by regression slopes with  $p < 0.05$ .

The National Ecological Observatory Network (NEON) characterizes the US in terms of 20 ecoclimatic regions across the continental US, Hawaii, and Puerto Rico. Each region has a core terrestrial site (table S1) and an associated geographic domain (figure S5). We mapped these ecoclimatic domains onto the  $1^\circ$  CLM5 surface grid and analyzed trends for 18 ecoregions (excluding Atlantic Neotropical and Pacific Tropical) at the core terrestrial site (defined as the model grid cell with the corresponding latitude and longitude of the site) and for the domain average.

A standard definition for time of emergence is the year when the forced signal  $S(t)$  (i.e. the change from a reference period) exceeds the noise  $N$  by a specified amount ( $S(t)/N > 2$ ), but methods differ in the temporal filtering of data to estimate the forced signal (e.g. running mean, polynomial fit, or linear trend) and how to calculate the noise (e.g. standard deviation for a reference period, standard deviation of residuals after extracting the forced signal, or ensemble standard deviation of trend estimates) (Hameau *et al* 2019). Similar to other studies (Keller *et al* 2014, Rodgers *et al* 2015, McKinley *et al* 2016, Schlunegger *et al* 2019, 2020), we defined the forced signal by the linear fit to the 1960–2014 time series of GPP and NEP for the ensemble mean, using the slope of the linear regression as the signal  $S$ . Because of the small ensemble size, we estimated the noise not as the variability in slope estimates among ensemble members (Rodgers *et al* 2015, McKinley *et al* 2016, Schlunegger *et al* 2019, 2020), but rather as the standard deviation for a reference period representative of the unperturbed state. Long *et al* (2016), for

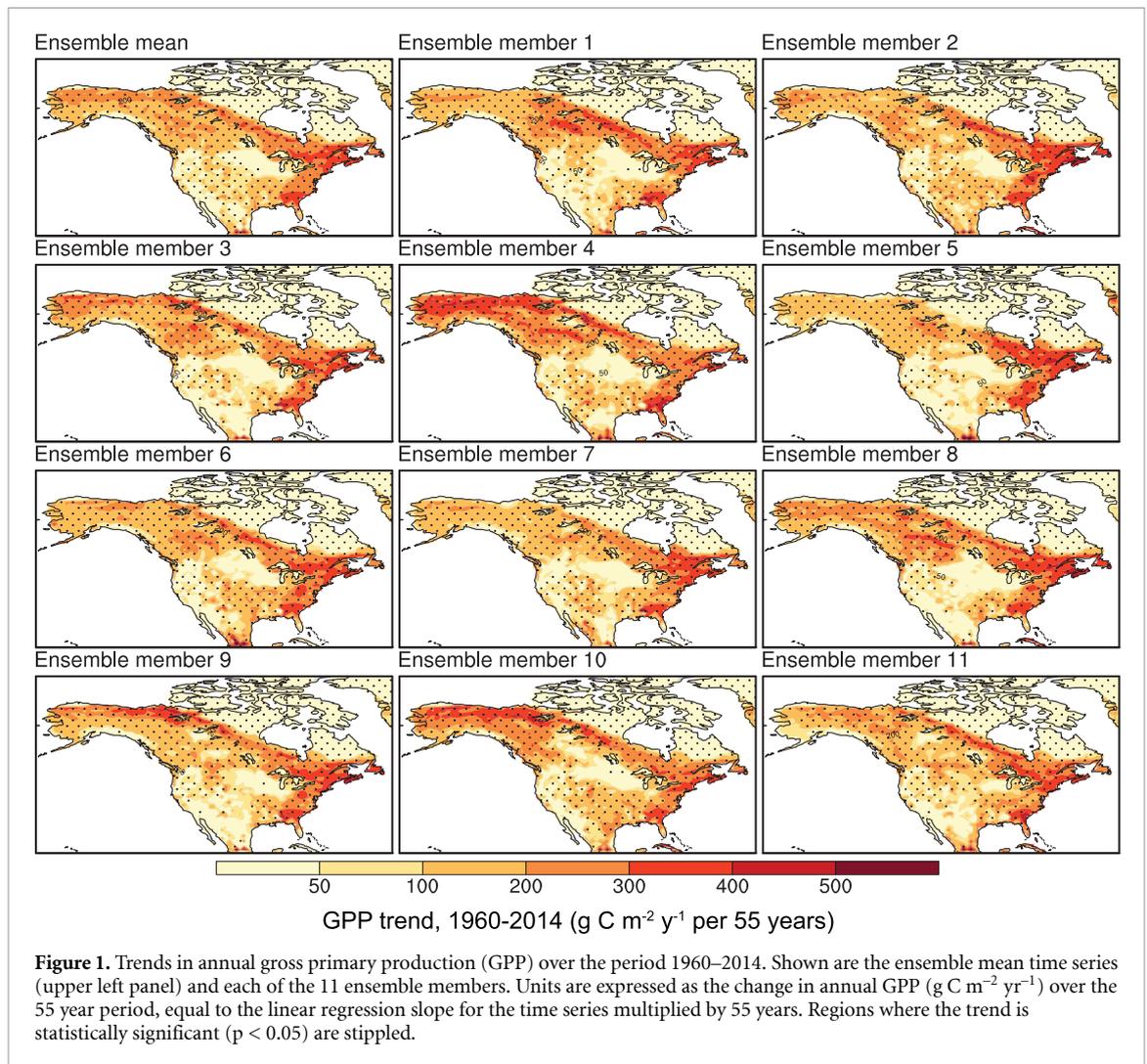
example, used the period 1920–1950 to estimate the noise in oceanic dissolved oxygen in their 24-member ensemble. In our analyses, the noise  $N$  is the standard deviation of annual GPP and NEP over the 30 year period 1930–1959 for the 11 ensemble members (a total of 330 simulated years). The time of emergence is  $2N/S$  and is the number of years since 1960 for the forced signal to emerge within the 1960–2014 time frame. Values greater than 55 years indicate no emergence of the forced signal within the time period.

### 3. Results

The ensemble mean time series has a statistically significant increase in annual GPP throughout much of North America between 1960 and 2014 (figure 1). Annual GPP increases by  $200\text{--}300\text{ g C m}^{-2}\text{ yr}^{-1}$  or more over the 55 year period throughout portions of Alaska, Canada, and the eastern US. Portions of the central and western US have a smaller increase in GPP. The magnitude of the trend varies among ensemble members, most prominently in Alaska and western Canada with some ensemble members having a small GPP increase (e.g. members 5 and 7) and others having a large increase in the same region (e.g. members 4 and 10).

Analysis of GPP trends for ecoclimatic regions across the US, defined by the NEON core terrestrial sites (table S1) and geographic domains (figure S5), shows regional differences in ensemble variability. The model grid cell corresponding to the Northeast (D01) site has a statistically significant increase in annual GPP over the period 1960–2014 in all 11 ensemble members, ranging from  $266$  to  $363\text{ g C m}^{-2}\text{ yr}^{-1}$  over the 55 years (figure 2(a)). The Mid-Atlantic (D02) site has larger spread in GPP trends: the trend is not statistically significant in one ensemble member, and GPP increases by  $136\text{--}451\text{ g C m}^{-2}\text{ yr}^{-1}$  over the 55 year period in the other ten members (figure 2(b)). The Ozarks (D08) site has large variability among ensemble members: GPP is statistically unchanged in three members, decreases in one member ( $-111\text{ g C m}^{-2}\text{ yr}^{-1}$  over the 55 years), and increases by  $88\text{--}591\text{ g C m}^{-2}\text{ yr}^{-1}$  over the 55 years in the other seven ensemble members (figure 2(c)). GPP increases in all ensemble members at the Taiga (D19) site, with a range of  $159\text{--}303\text{ g C m}^{-2}\text{ yr}^{-1}$  per 55 years (figure 2(d)).

The Mid-Atlantic (D02) and Ozarks (D08) sites show that although some ensemble members may not have a statistically significant GPP trend, significance is evident in the ensemble mean time series. The Prairie Peninsula (D06), Central Plains (D10), and Southern Rockies & Colorado Plateau (D13) sites have a similar result (table S2). Conversely, the Northern Plains (D09) and Desert Southwest (D14) sites have a small number of ensemble members with a statistically significant increase in GPP, but not in the ensemble mean.

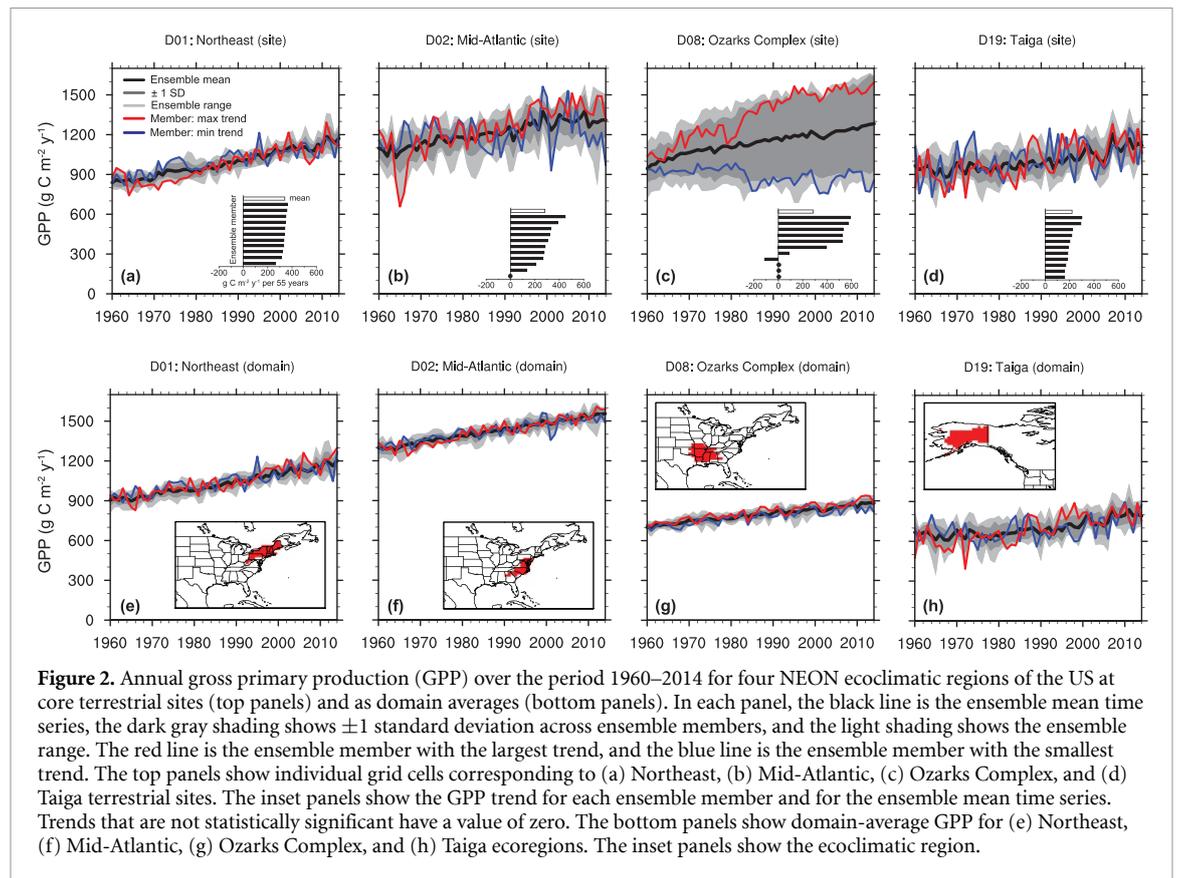


Ensemble variability decreases when GPP is spatially averaged over the ecoclimatic domains, most noticeably for the Mid-Atlantic (D02) and Ozark (D08) domains (figures 2(e)–(h)). Domain-average trends are generally less variable than site trends (table S2). However, while all domains have a statistically significant increase in the ensemble mean GPP, some domains such as the Prairie Peninsula (D06) and Northern Plains (D09) have a large number of members in which GPP does not increase. Two domains where the ensemble variability is essentially unchanged with spatial scale are Tundra (D18) and Taiga (D19).

NEP has a statistically significant increase in the ensemble mean time series across Alaska, Canada, the eastern US, and portions of the Central Plains and Pacific Northwest, but similar statistical significance is mostly lacking in any ensemble member (figure 3). Only the Canadian Arctic, eastern Canada, and northeastern US have a statistically significant NEP increase in the individual ensemble members. Analysis of the ecoclimatic regions further highlights the difference between the ensemble mean and individual ensemble members (figure S6, table S3). NEP

increases significantly in all 11 ensemble members for the Northeast (D01) site and in seven members for the Ozarks (D08) site, but in only one member for the Mid-Atlantic (D02) site and none at the Taiga (D19) site. At each of these sites, the ensemble mean NEP does significantly increase. The Northeast (D01), Great Lakes (D05), and Pacific Northwest (D16) sites are one end of the uncertainty spectrum, with all members showing positive NEP trends. The Central Plains (D10) and Taiga (D19) sites are the opposite extreme: NEP has no statistically significant trend in any ensemble member, but a positive NEP trend in the ensemble mean. The spread among ensemble members generally decreases when NEP is spatially averaged over the domain, but, consistent with figure 3, there are many ecoregions where the domain-average NEP increases in the ensemble mean whereas the ensemble members have instances of non-significant trends. Decadal averaging generally decreases the variability among ensemble members (figure S7).

The time at which the forced signal in GPP emerges, based on trends over the 1960–2014 period, varies among ecoclimatic regions and with spatial



**Figure 2.** Annual gross primary production (GPP) over the period 1960–2014 for four NEON ecoclimatic regions of the US at core terrestrial sites (top panels) and as domain averages (bottom panels). In each panel, the black line is the ensemble mean time series, the dark gray shading shows  $\pm 1$  standard deviation across ensemble members, and the light shading shows the ensemble range. The red line is the ensemble member with the largest trend, and the blue line is the ensemble member with the smallest trend. The top panels show individual grid cells corresponding to (a) Northeast, (b) Mid-Atlantic, (c) Ozarks Complex, and (d) Taiga terrestrial sites. The inset panels show the GPP trend for each ensemble member and for the ensemble mean time series. Trends that are not statistically significant have a value of zero. The bottom panels show domain-average GPP for (e) Northeast, (f) Mid-Atlantic, (g) Ozarks Complex, and (h) Taiga ecoregions. The inset panels show the ecoclimatic region.

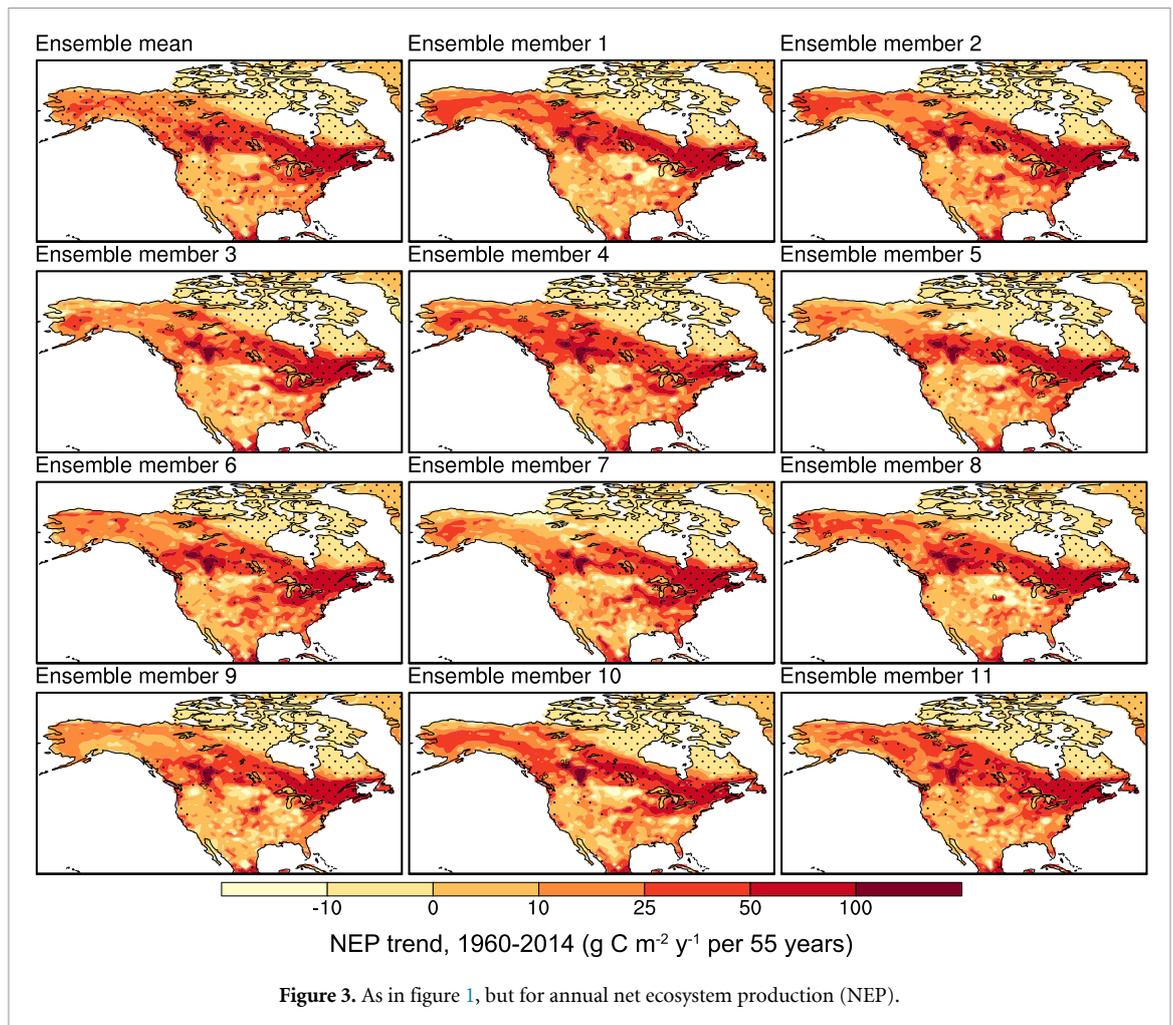
scale. At the site scale, the increase in GPP emerges throughout the eastern US and in the Pacific Northwest, where the trend in GPP is large and/or variability is small (figure 4(a), table S4). The Northeast (D01) site has the shortest time of emergence (1978; 18 years). The longest time of emergence (2008; 48 years) is at the Mid-Atlantic (D02) site, where the forced signal is large but variability is also large. Although the GPP trend at the Great Basin (D15) site is small, the signal emerges because of low variability. The signal emerges over a larger area of the US when GPP is spatially averaged over the domains, and in most domains more rapidly than at the site scale (figure 4(b), table S4). In central and southwestern regions of the US, the increase in GPP does not emerge during the 1960–2014 period at either the site or domain scales. The increase in NEP seen at many sites and most domains (table S3) does not emerge above the internal variability in any of the ecoregions at the site scale and in only the Northeast (D01) and Pacific Northwest (D16) at the domain scale (figure S8).

Concomitant with carbon cycle variability in the 11-member ensemble is variability in climate trends, such as 2 m surface air temperature (figure S9). Although the ensemble mean time series has a statistically significant warming between 1960 and 2014 over all of North America, differences among ensemble members are evident. Over Alaska and northwest Canada, for example, some ensemble members have small to non-significant warming

while other members have more pronounced warming. Temperature trends for NEON ecoclimatic regions further highlight variability among ensemble members (figure S10, table S5). At the sites in the Northeast (D01), Mid-Atlantic (D02), Ozarks (D08), and Taiga (D19), the 55 year temperature trend varies by  $1^\circ\text{C}$  or more among ensemble members. Ensemble variability does not necessarily decrease when spatially averaged for the entire domain. Notably, the Tundra (D18) and Taiga (D19) have the largest ensemble spread at both the site and domain scales. Ensemble variability in annual GPP correlates with variability in temperature and terrestrial water storage (figure S11). In northern and eastern regions of North America, years that are anomalously warm also have anomalously high GPP, whereas warm years correlate with low GPP in central and southwestern regions. Central and southwestern regions have high GPP during wet years.

#### 4. Discussion

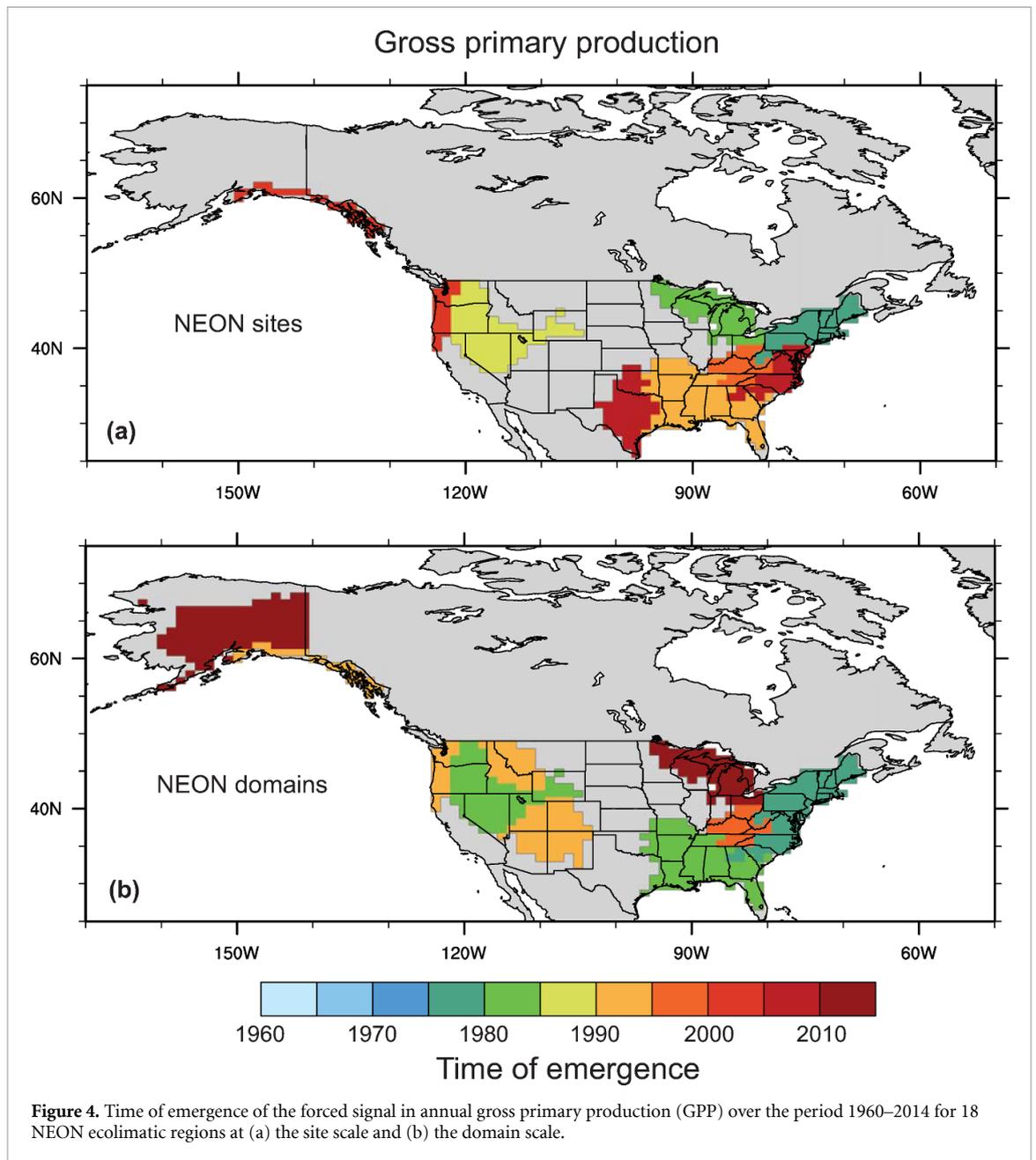
Large initial-condition ensembles are necessary to separate the forced signal of climate change from unforced variability internal to the climate system (Deser *et al* 2012b, 2020, Kay *et al* 2015, Maher *et al* 2019). Our analyses show that multimember initial-condition ensembles are also necessary when considering temporal changes in the terrestrial carbon cycle. Throughout North America, the time series of annual GPP and NEP over the period



1960–2014 for any one ensemble member can deviate from the ensemble mean. Ensemble variability is particularly evident at the site scale and is generally reduced for spatial averages across ecoclimatic domains. Decadal averages reduce ensemble variability compared to annual fluxes. Ensemble variability raises important considerations for the detection and attribution of forced changes in the carbon cycle; for comparisons of models to observations and to other models; and for our understanding of how to reduce uncertainty in Earth system model simulations of the carbon cycle.

Many modeling studies (e.g. Sitch *et al* 2015, Fernández-Martínez *et al* 2019, Friedlingstein *et al* 2019, Haverd *et al* 2020) report that plant productivity and net carbon uptake are changing, possibly as a result of elevated  $\text{CO}_2$  concentration, climate change, increased nitrogen deposition, disturbance legacy, or some combination of these factors. Observational analyses using eddy covariance flux towers with a small footprint or biogeochemical studies of small stands also suggest a changing terrestrial carbon cycle (e.g. Dragoni *et al* 2011, Pilegaard *et al* 2011, Froelich *et al* 2015, Baldocchi *et al* 2018, 2020, Finzi *et al* 2020). Our results show that modeled and observational trends in the carbon cycle must

be interpreted with caution. Use of one ensemble member to discern trends in GPP at the site scale could, by chance, result in masked trends (i.e. a non-significant trend even though the ensemble mean trend is significant) (figure 2, table S2). The ensemble also reveals the possibility of spurious trends in which some members have a significant trend even though the ensemble mean trend is not significant. Even at the domain scale, there is still a possibility of masked GPP trends in some ecoclimatic regions if only one ensemble member is analyzed. The need to consider the ensemble mean rather than any one ensemble member is particularly apparent for NEP (figure 3). In many ecoclimatic regions, a statistically significant increase in NEP is found in the ensemble mean time series at both the site and domain scales but not in most, or sometimes in any, of the individual ensemble members (table S3). Averaging across ensemble members decreases the random internal variability such that the forced response is apparent. Analysis of observed carbon cycle trends, or of carbon cycle models forced with meteorological observations, is subject to similar uncertainty in that the observational record is but one realization of many possible alternatives in which the sequence of internal variability could have unfolded differently.



The forced signal in GPP emerges after a few decades in eastern regions of the US, is not evident in central regions, and emerges in western regions mostly at the domain scale (figure 4). The implication is that only decadal-scale changes in GPP can be attributed to anthropogenic forcings. Our findings pertain to the period 1960–2014, but a previous analysis also found multi-decadal time of emergence for forced changes in the terrestrial carbon cycle over the 21st century (Lombardozzi *et al* 2014). Trends in NEP do not emerge over the 55 year period in any ecoclimatic region at the site scale and in only two regions at the domain scale (figure S8). Ensemble variability decreases when analyzing decadal averages (figure S7), suggesting that NEP trends might be detectable when averaged over long timescales. Annual assessments of the carbon cycle such as the Global Carbon Project (Friedlingstein *et al* 2019) are more likely to

express internal variability than forced change, but decadal assessments are likely more robust.

Another implication of our findings relates to ecological studies of carbon cycle trends. Ecologists inherently focus on site research to discern changes in the carbon cycle (e.g. Dragoni *et al* 2011, Pilegaard *et al* 2011, Froelich *et al* 2015, Finzi *et al* 2020). Because carbon cycle trends are more readily detected at the domain scale than the site scale, bridging the gap between site-scale, *in situ* measurements and the regional to continental scale is essential. Eddy covariance measurements provide insights to the environmental drivers of interannual variability in carbon fluxes (Baldocchi *et al* 2018, 2020). Further research is needed to bridge the understanding of carbon cycle variability gained from eddy covariance studies with the notion of internal variability gained from Earth system models.

Atmospheric internal variability means that when comparing Earth system models to temperature and precipitation records, a single model realization may not replicate the observations simply because of unforced variability. Likewise, the observational record is but one realization of many possible outcomes. Climate scientists have begun to consider internal variability as it pertains to observations, using statistical models to generate an ensemble of unforced climate histories so as to estimate the uncertainty in temperature and precipitation trends due to internal variability (McKinnon *et al* 2017, McKinnon and Deser 2018). A similar consideration pertains to the terrestrial carbon cycle. Our results suggest that the observational record has multiple histories that would have been realized if the sequence of internal variability had unfolded differently.

A further implication of our findings relates to uncertainty reduction. The ensemble spread produced by internal variability represents irreducible uncertainty in climate projections (Deser *et al* 2012b, 2020, Hawkins *et al* 2016). The uncertainty is an inherent property of the model and differs from model structural uncertainty, which can be reduced. Internal variability is a particularly large contribution to climate uncertainty at regional and smaller spatial scales and at decadal and shorter timescales (Hawkins and Sutton 2009, 2011, Lehner *et al* 2020). Ocean biogeochemistry has similarly large irreducible uncertainty at regional scales (Frölicher *et al* 2016, Lovenduski *et al* 2016). Our results show that the notion of irreducible uncertainty extends to the terrestrial carbon cycle and creates ambiguity in carbon cycle projections. This uncertainty is an inherent property of the model and will not be reduced with better process representation of the carbon cycle. The uncertainty decreases when averaging over larger spatial regions (figures 2 and S6) or decadal timescales (figure S7). Better understanding of the irreducible uncertainty of the terrestrial carbon cycle, and better communication of the uncertainty, is critical to informing carbon cycle research.

It is well understood that small differences in initial atmosphere or ocean states produce different simulated climate trajectories. Indeed, uncertainty in initial conditions is the rationale for using initial-condition ensembles to assess uncertainty in climate simulations (Deser *et al* 2012a, 2012b, 2014, 2020, Kay *et al* 2015, Maher *et al* 2019). Our results show that the initial model states in 1850, and the resulting climate variability among ensemble members, affect the magnitude and sign of carbon cycle trends over the historical era. Ensemble variability in annual temperature and water storage on land strongly correlates with ensemble variability in annual GPP (figure S11). A previous study with a different Earth system model correlated variability in net primary production with anomalies in temperature and precipitation and found a similar qualitative pattern over

North America (Lombardozi *et al* 2014). The CLM5 used in the CESM2 simulations differs from previous versions of the model (CLM4, CLM4.5) in the sensitivity of the carbon cycle to meteorological forcings (Bonan *et al* 2019, Lawrence *et al* 2019), and the model has low interannual variability in carbon fluxes compared with previous versions (Wozniak *et al* 2020). Our results are likely model dependent, meaning that other modeling centers need to consider initial-condition large ensembles of the terrestrial carbon cycle so as to develop a multi-model understanding of internal variability. Analyses of climate uncertainty also point to the need to consider internal variability in a multi-model context (Lehner *et al* 2020).

Uncertainty in the land biogeochemical states used to initialize the ensemble members in 1850 could also contribute to ensemble variability in carbon cycle trends. However, the initial vegetation and soil carbon pools used for each ensemble member differ by less than  $\pm 5\%$  of the ensemble mean across much of North America (figures S1 and S2). If the  $\pm 5\%$  spread in biogeochemical initial conditions contributes greatly to the ensemble variability, this is a level of uncertainty that would be difficult to reduce. It is possible, too, that different time-evolving biogeochemical states contribute to ensemble variability over the period 1960–2014, but vegetation and soil carbon pools still differ by only about  $\pm 5\%$  across ensemble members in 1960 (figures S12 and S13). It is more likely that ensemble variability in climate contributes to the variability in carbon cycle trends and that biogeochemical states are at most a secondary source of uncertainty. Simulations that vary only atmosphere/ocean initial states or land biogeochemical initial states could help separate the two influences on carbon cycle variability, but care must be taken to prevent introduction of spurious carbon cycle trends.

Finally, further research is needed to determine how many ensemble members are needed to accurately sample internal variability. The 11-member ensemble is too small to provide reliable probability distributions, such as can be obtained with larger ensembles of up to 100 members (Deser *et al* 2020). Nonetheless, we find more instances of masked trends at the site and domain scale for GPP (table S2) and NEP (table S3) than for temperature warming (table S5), suggesting that more ensemble members are needed to capture carbon cycle variability than temperature variability.

### Data availability statement

The data that support the findings of this study are openly available at the following URLs: <https://esgf-node.llnl.gov/search/cmip6/> and <https://www.cesm.ucar.edu>.

## Acknowledgments

This material is based upon work supported by the National Center for Atmospheric Research (NCAR), which is a major facility sponsored by the National Science Foundation under Cooperative Agreement 1852977. Computing and data storage resources, including the Cheyenne supercomputer (doi:10.5065/D6RX99HX), were provided by the Computational and Information Systems Laboratory (CISL) at NCAR. The authors acknowledge funding from the National Institute of Food and Agriculture/US Department of Agriculture (2015-67003-23485). WRW was also supported by a grant from the US National Science Foundation (DEB-1926413). GBB, DLL, and WRW designed the analyses. DLL provided the time of emergence analysis. WRW prepared the NEON domain mask and the correlations with climate variables. GBB performed the remaining model analyses. GBB, DLL, and WRW wrote the manuscript. The authors declare no conflict of interest.

## ORCID iDs

Gordon B Bonan  <https://orcid.org/0000-0003-0830-6437>

Danica L Lombardozzi  <https://orcid.org/0000-0003-3557-7929>

William R Wieder  <https://orcid.org/0000-0001-7116-1985>

## References

- Arora V K *et al* 2020 Carbon-concentration and carbon-climate feedbacks in CMIP6 models and their comparison to CMIP5 models *Biogeosciences* **17** 4173–222
- Baldocchi D D 2020 How eddy covariance flux measurements have contributed to our understanding of Global Change Biology *Glob. Change Biol.* **26** 242–60
- Baldocchi D, Chu H and Reichstein M 2018 Inter-annual variability of net and gross ecosystem carbon fluxes: a review *Agric. For. Meteorol.* **249** 520–33
- Bengtsson L and Hodges K I 2019 Can an ensemble climate simulation be used to separate climate change signals from internal unforced variability? *Clim. Dyn.* **52** 3553–73
- Bonan G B, Lombardozzi D L, Wieder W R, Oleson K W, Lawrence D M, Hoffman F M and Collier N 2019 Model structure and climate data uncertainty in historical simulations of the terrestrial carbon cycle (1850–2014) *Glob. Biogeochem. Cycles* **33** 1310–26
- Collins M *et al* 2013 Long-term climate change: projections, commitments and irreversibility *Climate change 2013: the physical science basis. contribution of working group I to the fifth assessment report of the Intergovernmental Panel on climate change* ed T F Stocker *et al* (Cambridge: Cambridge University Press) pp 1029–136
- Dai A and Bloecker C E 2019 Impacts of internal variability on temperature and precipitation trends in large ensemble simulations by two climate models *Clim. Dyn.* **52** 289–306
- Danabasoglu G *et al* 2020 The community Earth system model version 2 (CESM2) *J. Adv. Model. Earth Syst.* **12** e2019MS001916
- Deser C *et al* 2020 Insights from Earth system model initial-condition large ensembles and future prospects *Nat. Clim. Change* **10** 277–86
- Deser C, Knutti R, Solomon S and Phillips A S 2012b Communication of the role of natural variability in future North American climate *Nat. Clim. Change* **2** 775–9
- Deser C, Phillips A S, Alexander M A and Smoliak B V 2014 Projecting North American climate over the next 50 years: uncertainty due to internal variability *J. Clim.* **27** 2271–96
- Deser C, Phillips A, Bourdette V and Teng H 2012a Uncertainty in climate change projections: the role of internal variability *Clim. Dyn.* **38** 527–46
- Deser C, Terray L and Phillips A S 2016 Forced and internal components of winter air temperature trends over North America during the past 50 years: mechanisms and implications *J. Clim.* **29** 2237–58
- Dragoni D, Schmid H P, Wayson C A, Potter H, Grimmond C S B and Randolph J C 2011 Evidence of increased net ecosystem productivity associated with a longer vegetated season in a deciduous forest in south-central Indiana, USA *Glob. Change Biol.* **17** 886–97
- Duvivier A K, Derepentigny P, Holland M M, Webster M, Kay J E and Perovich D 2020 Going with the floe: tracking CESM large ensemble sea ice in the Arctic provides context for ship-based observations *Cryosphere* **14** 1259–71
- Fernández-Martínez M *et al* 2019 Global trends in carbon sinks and their relationships with CO<sub>2</sub> and temperature *Nat. Clim. Change* **9** 73–79
- Finzi A C *et al* 2020 Carbon budget of the Harvard Forest long-term ecological research site: pattern, process, and response to global change *Ecol. Monogr.* **90** e01423
- Friedlingstein P *et al* 2019 Global carbon budget 2019 *Earth Syst. Sci. Data* **11** 1783–838
- Froelich N, Croft H, Chen J M, Gonsamo A and Staebler R M 2015 Trends of carbon fluxes and climate over a mixed temperate–boreal transition forest in Southern Ontario, Canada *Agric. For. Meteorol.* **211–212** 72–84
- Frölicher T L, Joos F, Plattner G-K, Steinacher M and Doney S C 2009 Natural variability and anthropogenic trends in oceanic oxygen in a coupled carbon cycle–climate model ensemble *Glob. Biogeochem. Cycles* **23** GB1003
- Frölicher T L, Rodgers K B, Stock C A and Cheung W W L 2016 Sources of uncertainties in 21st century projections of potential ocean ecosystem stressors *Glob. Biogeochem. Cycles* **30** 1224–43
- Hameau A, Mignot J and Joos F 2019 Assessment of time of emergence of anthropogenic deoxygenation and warming: insights from a CESM simulation from 850 to 2100 CE *Biogeosciences* **16** 1755–80
- Haverd V, Smith B, Canadell J G, Cuntz M, Mikaloff-Fletcher S, Farquhar G, Woodgate W, Briggs P R and Trudinger C M 2020 Higher than expected CO<sub>2</sub> fertilization inferred from leaf to global observations *Glob. Change Biol.* **26** 2390–402
- Hawkins E, Smith R S, Gregory J M and Stainforth D A 2016 Irreducible uncertainty in near-term climate projections *Clim. Dyn.* **46** 3807–19
- Hawkins E and Sutton R 2009 The potential to narrow uncertainty in regional climate predictions *Bull. Am. Meteorol. Soc.* **90** 1095–107
- Hawkins E and Sutton R 2011 The potential to narrow uncertainty in projections of regional precipitation change *Clim. Dyn.* **37** 407–18
- Hawkins E and Sutton R 2012 Time of emergence of climate signals *Geophys. Res. Lett.* **39** L01702
- Hewitt A J, Booth B B B, Jones C D, Robertson E S, Wiltshire A J, Sansom P G, Stephenson D B and Yip S 2016 Sources of uncertainty in future projections of the carbon cycle *J. Clim.* **29** 7203–13
- Kay J E *et al* 2015 The community Earth system model (CESM) large ensemble project: a community resource for studying

- climate change in the presence of internal climate variability *Bull. Am. Meteorol. Soc.* **96** 1333–49
- Keller K M, Joos F and Raible C C 2014 Time of emergence of trends in ocean biogeochemistry *Biogeosciences* **11** 3647–59
- Kirchmeier-Young M C, Zwiers F W and Gillett N P 2017 Attribution of extreme events in Arctic sea ice extent *J. Clim.* **30** 553–71
- Landrum L and Holland M M 2020 Extremes become routine in an emerging new Arctic *Nat. Clim. Change* **10** 1108–15
- 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** 4245–428
- Lehner F, Deser C, Maher N, Marotzke J, Fischer E M, Brunner L, Knutti R and Hawkins E 2020 Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6 *Earth Syst. Dyn.* **11** 491–508
- Lehner F, Deser C and Terray L 2017 Toward a new estimate of ‘time of emergence’ of anthropogenic warming: insights from dynamical adjustment and a large initial-condition model ensemble *J. Clim.* **30** 7739–56
- Lombardozzi D, Bonan G B and Nychka D W 2014 The emerging anthropogenic signal in land atmosphere carbon-cycle coupling *Nat. Clim. Change* **4** 796–800
- Long M C, Deutsch C and Ito T 2016 Finding forced trends in oceanic oxygen *Glob. Biogeochem. Cycles* **30** 381–97
- Lorenz E N 1963 Deterministic nonperiodic flow *J. Atmos. Sci.* **20** 130–41
- Lovenduski N S and Bonan G B 2017 Reducing uncertainty in projections of terrestrial carbon uptake *Environ. Res. Lett.* **12** 044020
- Lovenduski N S, McKinley G A, Fay A R, Lindsay K and Long M C 2016 Partitioning uncertainty in ocean carbon uptake projections: internal variability, emission scenario, and model structure *Glob. Biogeochem. Cycles* **30** 1276–87
- Maher N et al 2019 The Max Planck institute grand ensemble: enabling the exploration of climate system variability *J. Adv. Model. Earth Syst.* **11** 2050–69
- Mankin J S, Viviroli D, Singh D, Hoekstra A Y and Diffenbaugh N S 2015 The potential for snow to supply human water demand in the present and future *Environ. Res. Lett.* **10** 114016
- Martel J-L, Mailhot A, Brissette F and Caya D 2018 Role of natural climate variability in the detection of anthropogenic climate change signal for mean and extreme precipitation at local and regional scales *J. Clim.* **31** 4241–63
- McKinley G A, Pilcher D J, Fay A R, Lindsay K, Long M C and Lovenduski N S 2016 Timescales for detection of trends in the ocean carbon sink *Nature* **530** 469–72
- McKinnon K A and Deser C 2018 Internal variability and regional climate trends in an observational large ensemble *J. Clim.* **31** 6783–802
- McKinnon K A, Poppick A, Dunn-Sigouin E and Deser C 2017 An ‘observational large ensemble’ to compare observed and modeled temperature trend uncertainty due to internal variability *J. Clim.* **30** 7585–98
- Pilegaard K, Ibrom A, Courtney M S, Hummelshøj P and Jensen N O 2011 Increasing net CO<sub>2</sub> uptake by a Danish beech forest during the period from 1996 to 2009 *Agric. For. Meteorol.* **151** 934–46
- Rodgers K B, Lin J and Frölicher T L 2015 Emergence of multiple ocean ecosystem drivers in a large ensemble suite with an Earth system model *Biogeosciences* **12** 3301–20
- Schlunegger S et al 2020 Time of emergence and large ensemble intercomparison for ocean biogeochemical trends *Glob. Biogeochem. Cycles* **34** e2019GB006453
- Schlunegger S, Rodgers K B, Sarmiento J L, Frölicher T L, Dunne J P, Ishii M and Slater R 2019 Emergence of anthropogenic signals in the ocean carbon cycle *Nat. Clim. Change* **9** 719–25
- Sitch S et al 2015 Recent trends and drivers of regional sources and sinks of carbon dioxide *Biogeosciences* **12** 653–79
- Swain D L, Langenbrunner B, Neelin J D and Hall A 2018 Increasing precipitation volatility in twenty-first-century California *Nat. Clim. Change* **8** 427–33
- Swart N C, Fyfe J C, Hawkins E, Kay J E and Jahn A 2015 Influence of internal variability on Arctic sea-ice trends *Nat. Clim. Change* **5** 86–89
- Wozniak M C, Bonan G B, Keppel-Aleks G and Steiner A L 2020 Influence of vertical heterogeneities in the canopy microenvironment on inter-annual variability of carbon uptake in temperate deciduous forests *J. Geophys. Res. Biogeosci.* **125** e2020JG005658
- Wu Z, Ahlström A, Smith B, Ardö J, Eklundh L, Fensholt R and Lehsten V 2017 Climate data induced uncertainty in model-based estimations of terrestrial primary productivity *Environ. Res. Lett.* **12** 064013