Observational constraints on the response of high-latitude northern forests to warming

Junjie Liu\textsuperscript{1,2}, Paul O. Wennberg\textsuperscript{2}, Nick Parazoo\textsuperscript{1}, Yi Yin\textsuperscript{2}, Christian Frankenberg\textsuperscript{2,1}

\textsuperscript{1} Jet Propulsion Laboratory, Caltech, United States
\textsuperscript{2} Caltech, United States
Observational constraints on the response of high-latitude northern forests to warming

Junjie Liu$^{1,2,*}$, Paul O. Wennberg$^{2,*}$, Nick Parazoo$^{1}$, Yi Yin$^{2}$, Christian Frankenberg$^{2,1}$

1. Jet Propulsion Laboratory, Caltech, US
2. Caltech, US

*Correspondence to: junjie.liu@jpl.nasa.gov; wennberg@caltech.edu

Key Points:

- 3/4 of the spatial variations in productivity of high-latitude northern forests can be explained by growing season mean temperature (GSMT).
- Substituting space for time, we estimate that the increase in GSMT between 1960–2010 enhanced productivity and net carbon uptake by ~20%.
- The calculated enhancement of net ecosystem exchange due to temperature alone contributes 56–72% of the trend in the CO$_2$ seasonal cycle amplitude at high latitudes.
Abstract
Since the 1960s, carbon cycling in the high-latitude northern forest (HLNF) has experienced dramatic changes: most of the forest is greening and net carbon uptake from the atmosphere has increased. During the same time period, the CO$_2$ seasonal cycle amplitude (SCA) has increased by ~50% or more. Disentangling complex processes that drive these changes has been challenging. In this study, we substitute spatial sensitivity to temperature for time to quantify the impact of temperature increase on Gross Primary Production (GPP), total ecosystem respiration (TER), the fraction of Photosynthetic Active Radiation (fPAR), and the resulted contribution of these changes in amplifying the CO$_2$ SCA over the HLN F since 1960s. We use the spatial heterogeneity of GPP inferred from solar-induced chlorophyll Fluorescence in combination with net ecosystem exchange (NEE) inferred from column CO$_2$ observations made between 2015 and 2017 from NASA’s Orbiting Carbon Observatory -2. We find that three quarters of the spatial variations in GPP can be explained by the spatial variation in the growing season mean temperature (GSMT). The long term hindcast captures both the magnitude and spatial variability of the trends in observed fPAR. We estimate that between 1960 and 2010, the increase in GSMT enhanced both GPP and the SCA of NEE by ~20%. The calculated enhancement of NEE due to increase in GSMT contributes 56–72% of the trend in the CO$_2$ SCA at high latitudes, much larger than simulations by most biogeochemical models.

Plain Language Summary
The carbon cycling in the high-latitude northern forest has experienced dramatic changes during the last 5-6 decades along with temperature increase almost double that of the global mean temperature. Using observations, we quantify the impact of temperature increase alone on the changes of gross primary production, greenness, and CO$_2$ seasonal cycle amplitude. Our results show a dominant temperature effect on the changes of carbon cycle over the region. Our results
imply that an additional 2-5 °C temperature increase by 2100 (e.g. RCP6.0, IPCC AR5) would dramatically alter the distribution of tree species over the region: evergreen trees will invade the shrublands, and much of the evergreen forests will transition to deciduous trees. GPP over the region has the potential to double.
1 Introduction

The High-Latitude Northern Forest (HLNF, 50°N to 75°N) is composed of needle leaf trees (both evergreen and deciduous), deciduous broad-leafed trees, and shrubland (Figure 1). It is responsible for ~20% of global Gross Primary Productivity (GPP) (Tramontana et al., 2017; Jung et al., 2017; Jung et al., 2020). Over the region northward of 50 degrees, precipitation substantially exceeds evaporation due to moisture convergence by the large-scale eddy circulation and low net radiation, which results in a generally well-watered, thermally-limited ecosystem (Nemani et al., 2003). The northern extent of the forest is limited by the Arctic ocean and/or where summertime temperatures are too low. Because of its seasonality (days with nearly 24 hours of sunlight followed by days with 24 hours of darkness), uptake and release of carbon from the atmosphere is responsible for a large fraction of the seasonal cycle amplitude (SCA) in atmospheric CO$_2$. Without the HLFN, the SCA of CO$_2$ in the northern hemisphere would be only about half as large (Graven et al., 2013).

Since the 1960s, carbon cycling in the HLFN has experienced dramatic changes: most of the forest is greening (Zhu et al., 2016; Keenan and Riley, 2018) and net carbon uptake from the atmosphere has increased (Ciais et al., 2019). Although the underlying processes driving these trends are uncertain (Randerson et al., 1997; Zeng et al., 2014; Gray et al., 2014; Forkel et al., 2016; Piao et al., 2018; Bastos et al., 2019), it is generally agreed that both the increase in temperature and CO$_2$ are responsible (Randerson et al., 1997; Forkel et al., 2016; Piao et al., 2018; Bastos et al., 2019). Global mean CO$_2$ concentration has increased from 320 ppm in 1958 to ~400 ppm in 2010s (https://www.esrl.noaa.gov/gmd/ccgg/trends/). Biogeochemical processes ranging from photosynthesis, net primary production (NPP), to the net carbon balance in general respond positively to the CO$_2$ increase, which is called CO$_2$ fertilization (Kimball et al., 1983;
Long et al., 2004). The optimal temperature for photosynthesis is currently higher than the mean annual temperature over high-latitude biomes (Huang et al., 2019), thus, we expect that increase in temperature accelerates gross primary production (GPP), unlike tropics where the increase in temperature can suppress growth (Sullivan et al., 2020). With the increase of GPP and temperature, total ecosystem respiration (TER) is also expected to increase (Davidson et al., 2006; Davidson and Janssens, 2006; Baldocchi et al., 2008). The temperature increase over the NH high latitudes is nearly double that of the mean temperature increase over land (IPCC AR5). Quantifying the individual role of temperature and CO₂ increases on photosynthesis, respiration, and net carbon uptake is crucial for projecting future climate change (Friedlingstein et al., 2014). However, because the temporal changes in temperature and CO₂ have been so correlated, disentangling their individual contribution with observations is difficult, especially over HLNF where conventional observations are sparse (Schimel et al., 2015). Previous studies primarily make use of factorial simulations of biogeochemical models to isolate the impact of change in climate from other factors (e.g., Bastos et al., 2019). Studies using observation constraint are scarce.

In this study, we provide observational constraints on the impact of temperature change on GPP, TER, and the resulting net ecosystem exchange (NEE) of the HLNF. We use the spatial heterogeneity of GPP and TER inferred from solar-induced chlorophyll Fluorescence (SIF) (Frankenberg et al., 2011) and a combination of NEE and GPP, respectively. The NEE is inferred from column CO₂ observations retrieved from spectra of NASA’s Orbiting Carbon Observatory -2 (OCO-2) (Crisp et al., 2017; Eldering et al., 2017) between 2015 and 2017.
Substituting the observed spatial-sensitivity of GPP and TER to temperature for time, we estimate the change in carbon cycling in the HLNF between 1960s and 2010s due to the warming this region has experienced. We diagnose what contribution these changes have made to the observed increase in the CO₂ seasonal cycle amplitude (SCA) between 1958-1963 (International Geophysical Year (IGY) aircraft campaign) (Keeling et al., 1968) and 2009-2011 (Hiaper Pole-to-Pole observations (HIPPO) aircraft campaign) (Wofsy et al., 2011), and over NH high latitude CO₂ flask sites with long-record (>25 years) (Alert (Canada), Cold Bay (Alaska, US), Barrow (Alaska, US), and Shemya Island (Alaska, US)). GPP is a function of fraction of photosynthetic active radiation (PAR) absorbed by plants (fPAR) (Farquhar et al., 1980), which in turn reflects forest structure (leaf area index, LAI) (Zhu et al., 2016). To evaluate whether space and time can be substituted, we compare the trend and interannual variability in fPAR and GPP calculated with this assumption to observations of the changes over space and time.

2 Materials and Methods
We calculate monthly mean GPP (2.1) and NEE (2.2) for 2015–2017 with OCO-2 SIF and column CO₂ retrievals, respectively. TER is calculated as the difference between NEE and GPP (2.2) (Figure S1). We substitute seasonal (i.e., spring, summer, and fall (2.3)) spatial GPP–T and TER–T relationship for time to predict the temporal changes of GPP and TER and the resulting changes in NEE from increase in temperature for the time period of 1980–2011 and between 1958-1963 and 2009-2011 (2.4) (Figure S1). To evaluate the space-for-time substitution, we compare predictions of the change in fPAR between 1982 and 2012 based on the modern spatial fPAR–T relationship with the changes in GIMMS3g fPAR (Zhu et al., 2013) (2.4). We further compare predictions of seasonal cycle and interannual variability in GPP with observations of SIF-constrained GPP and data from FLUXNET2015 dataset (2.4). Finally, we calculate the CO₂
seasonal cycle amplitude (SCA) changes due to changes in NEE between 1958–1963 and 2009–2011 and for the time period of 1980–2011 (2.5) (Figure S1).

2.1 Solar Induced chlorophyll Fluorescence (SIF) – constrained monthly mean GPP calculation

We use SIF retrieved from OCO-2 to calculate monthly mean GPP. SIF is fluorescence from plants that occurs during photosynthesis. It is sensitive to both plant structure and photosynthetic function, and is linearly correlated with GPP under high-light conditions at seasonal time scale (Frankenberg et al., 2011; Sun et al., 2018). OCO-2 SIF retrievals have footprint size ~1.2 x 2 km². Each SIF retrieval is scaled based on solar zenith angle, date, latitude, and local passing time, ~ 1:30pm, to produce a daily mean (Frankenberg et al., 2011). The monthly mean value at each 4° x 5° grid is a simple averaging of the daily mean SIF values. The 4° x 5° spatial resolution is chosen for the consistency with the optimized NEE from top-down atmospheric CO₂ flux inversion. We use OCO-2 level 2 SIF observations from land nadir mode only between Jan 2015 and Dec 2017. The time frame is constrained by the availability of the FLUXCOM GPP product, an upscaling GPP product based on GPP from flux towers, ancillary satellite, and meteorology (Tramontana et al., 2016; Jung et al., 2017, Jung et al., 2020).

The growing season mean SIF value increases exponentially with temperature over the grid points with tree cover larger than 40% (Figure S2). The SIF–temperature sensitivity is within the sensitivity range of the three FLUXCOM GPP products (Figure S2). There is no obvious discontinuity in the sensitivities of both SIF and FLUXCOM GPP products to temperature, so we calculate monthly mean GPP over 50°N-75° N at 4° x 5° spatial resolution as the product of OCO-2 SIF and a scaling factor derived using the growing season (GS) mean GPP from three
FLUXCOM products and mean SIF. The scaling factor, 22.9 ± 2.6 (gram carbon micron sr per day per watts (gC μm sr day⁻¹ W⁻¹)), is consistent with (Sun et al., 2018). In using the same conversion factor across the region, we have assumed that different tree types have the same relationship between SIF and GPP (see Section 4).

2.2 Top-down NEE estimation with CMS-Flux and Total Ecosystem Respiration (TER) calculation during 2015 and 2017

We optimize monthly mean NEE with CMS-Flux system (Liu et al., 2014; 2017; 2018; Bowman et al., 2017) at each 4° x 5° grid with OCO-2 B9 land nadir observations (Eldering et al., 2017) during Sep 2014–April 2018. In spite of increasing CO₂ observational coverage from OCO-2, the spatial resolution of NEE from top-down inversion is still limited by observations. For readers interested in more details on atmospheric CO₂ flux inversion, please refer to (Liu et al., 2014; 2015; 2017). The prior fluxes include climatology terrestrial biosphere carbon fluxes from CASA-GFED 3 (Randerson et al., 1997), biomass burning from GFED4, fossil fuel emissions from ODIAC (Oda et al., 2018), and ocean carbon fluxes from ECCO-Darwin (Menemenlis et al., 2008). Except the climatology terrestrial biosphere fluxes from CASA-GFED3, the other types of carbon fluxes are yearly dependent.

We calculate TER as the difference between the optimized NEE and GPP between 2015–2017. The right three panels in Figure S3 show the dependency of TER on temperature during the spring, summer, and fall. The estimated TER shows strong dependence on temperature, which is expected, since TER’s two components, autotrophic and heterotrophic respiration, depend on the amount of carbon fixed through photosynthesis (Baldocchi, 2008). The temperature dependence
of TER is not related to the amount of above ground biomass, which has a weak relationship
with TER on seasonal timescale (not shown), with $R^2$ less than 0.07. In Figure S3, we remove
points that have nonphysical negative TER values.

We validate the top-down flux inversion results by comparing the CO$_2$ SCA sampled at Mauna
Loa against observed CO$_2$ SCA, and comparing posterior CO$_2$ concentrations between April and
September against independent aircraft CO$_2$ observations north of 30°N. The SCA of posterior
CO$_2$ concentration sampled at Mauna Loa is within the uncertainty range of the observed CO$_2$
SCA (Figure S4). The random error and bias of regional mean monthly posterior CO$_2$
concentration is 0.4 ppm and 0.1 ppm, respectively (Figure S4) (Supplementary Text).

2.2 Definition of growing season

We define spring, summer, and fall with SIF-optimized GPP product (Parazoo et al., 2014),
which optimally combines OCO-2 SIF observations with climatology mean state from TRENDY
models (Sitch et al., 2015). This product is only used in defining seasons, since the product
maintains essentially the same seasonality as original SIF observations during growing season,
while filling the gaps of SIF observation during the late growing season and winter. At each grid
point, the growing season is defined as the time period when the GPP values are larger than 20%
of the maximum GPP at that grid point. Spring and fall are defined as the time period when GPP
is between 20% and 75% of maximum GPP, and summer is the time period between spring and
fall. We choose 20% as the threshold to ensure that the plants have started growing and the start
of the spring is well-defined. We find that 10% of grid points would have no spring if the
threshold is 30% instead of 20%, since GPP increases sharply during spring, especially in Eurasia.

2.4 Prediction of growing season monthly mean GPP, TER, and fPAR

We substitute the spatial sensitivity to temperature for time to predict the temporal changes in GPP, TER, and fPAR due to changes in temperature. Note that the spatial CO₂ gradient is typically less than 15 ppm, and we find that there is no co-varying relationship in space between GPP and CO₂ on seasonal time scale over the HLNF (not shown). Our approach of substitution of space-for-time to predict changes in carbon states due to climate has been used by others. For example, Sullivan et al. (2020) predicted changes in tropical biomass from changes in temperature using the spatial biomass–temperature sensitivity. Figure S5 shows that the fPAR – T spatial sensitivity is relatively constant, with variability less than 5%, implying that the spatial sensitivity to temperature represents a semi-equilibrium state. We use the spatial – temperature sensitivity shown in Figure S3 and S6 to predict temporal changes of GPP, TER, and fPAR.

In predicting historical GPP, we use GPP/PAR – T spatial sensitivity (Figure S3) to calculate the temporal changes of GPP/PAR with changes in temperature, and then multiply PAR to get final GPP prediction. Since PAR is only a function of latitude and time of year, we use the same PAR averaged over 2015-2017 to calculate GPP.

With the spatial – temperature sensitivity derived from the observation-constrained GPP and TER between 2015 and 2017 and CRU monthly mean air temperature during 2009-2011 and 1958-1963, we predict the monthly GPP and TER during growing season of these two time periods (Figure S7). In predicting the GPP for 2009-2011 and 1958-1963 time periods, we
assume that the temporal range of spring, summer, and fall in each grid cell is the same as in
2015-2017, which is a reasonable assumption since the temporal resolution of our study is month
and the growing season length only increases by ~0.25 month for every ~1 degree increase of
temperature (Figure S8). We generate 10-member ensemble GPP and TER predictions with the
95% confidence interval shown in Figure S3. We calculate the uncertainty from the ensemble
GPP and TER hindcasts shown in Figure S7.

We calculate NEE during growing season as the difference between ensemble TER and GPP,
and assume that non-growing season NEE is proportional to the growing season NEE (Kauppi
and Posch, 1985). The proportional assumption may be a conservative estimation of non-
growing season NEE changes (see Section 4).

To calculate NEE between 1980 to 2011, we predict growing season GPP and TER with four-
year frequency, where we use monthly mean temperature averaged over every four years. We
find that the results are not sensitive to the prediction frequency. In total, we have eight groups of
NEE hindcast between 1980 and 2011. We choose 2011 as the ending year to be consistent with
the time frame of HIPPO aircraft campaign.

With the same seasonally-dependent GPP/PAR – T spatial sensitivity and monthly mean
temperature, we predict monthly mean GPP between 1960 and 2014. Following the same method
described above, we predict monthly mean fPAR with the seasonally-dependent relationship
calculated for 2013-2016 (Figure S6). The sensitivity shown in Figure S6 is based on fPAR from
GIMMS3g. Since fPAR from GIMMS3g has more than 30-year record, the trend calculation is
more robust than that is based on MODIS. In the main text, we use fPAR from GIMSS3g in trend calculation, but use fPAR from MODIS in Figure 2, since MODIS has higher spatial resolution than AVHRR used in the earlier record of GIMMS3g (https://nsidc.org/support/21524371-What-are-the-MODIS-sensor-s-strongest-attributes-and-how-do-they-compare-to-other-sensors-).

2.5 CO₂ simulations with GEOS-Chem transport model

We use GEOS-Chem transport model to simulate 3-hourly 10-member ensemble of CO₂ concentrations during 1958-1963 (IGY), 2009-2011 (HIPPO) and from 1980 to 2011. The only surface boundary conditions are the predicted ensemble monthly NEE (i.e., TER-GPP). Therefore, the differences in simulated CO₂ are only due to the differences in the predicted monthly NEE between 50°N and 75°N. We assume the same diurnal biosphere fluxes during these two time periods, which are from CASA-GFED3.

We run GEOS-Chem transport model at 2° x 2.5° spatial resolution driven by MERRA2 reanalysis after regridding 4°x5° NEE to 2°x2.5° spatial resolution. The transport model has smaller transport errors over high latitudes at 2° x 2.5° resolution than at 4° x 5° resolution. For both IGY and HIPPO CO₂ simulations, we run ensemble transport model simulations continually from 2006 to 2011, and we then analyze CO₂ simulation results over the last three-year simulation. The difference between IGY and HIPPO CO₂ simulations is only from the boundary forcing, not from the meteorology forcing. For each of the eight-group NEE from 1980 to 2011, we run ensemble transport model simulations continually for six years with meteorology from 2006 to 2011, and then analyze CO₂ simulation results for each group over the last three-year
simulation. As a result, the changes of CO\(_2\) SCA are only due to changes of NEE, not due to changes of transport. We follow the same method described in Graven et al. (2013) (Supplementary Text) to calculate CO\(_2\) SCA for both surface sites and aircraft observations and the corresponding model simulations.

3 Results
3.1 The spatial GPP – temperature relationship
Over the HLFN, both the spatial distribution and seasonal evolution of GPP closely follow those of temperature (Figure 1). From April to July, increases in GPP propagate from the west to the east of each continent following the spatial increase of monthly mean temperature (Figure S9). From July to October, GPP decreases as cooler temperature propagates from the east to the west of each continent (Figure S9). The spatial changes illustrate how the onset and length of growing season (GS) are closely coupled to the surface temperature (Figure S8). The growing season length increases ~one week with each degree increase in surface temperature, consistent with Park et al. (2016).
Figure 1 Temperature controls on the spatial distribution of gross primary production (GPP) over high latitude northern forest (50°N – 75°N). (a). Tree percentage and dominant tree types at every 4° x 5° grid. Dots: shrubland; triangles: needleleaf forest; and squares: deciduous boreal forest. The tree type is based on condensed MODIS International Geosphere Biosphere Program (IGBP) map. (b). Mean temperature (unit: °C) between April and September over 2015-2017. (c). Mean GPP (unit: kgCm⁻²) integrated between April and September over 2015-2017. The temperature is 2-meter air temperature from ERA-interim reanalysis (Table S1). GPP is scaled SIF.

Quantitatively, the spatial variation in GSMT explains 70% of GPP/PAR (Photosynthetic Active Radiation) variability (Figure 1b and c and Figure 2a) (and 75% of the variability in GPP, Figure S10). The ratio between seasonally-averaged GPP and the seasonally-averaged PAR removes the influence of the somewhat larger solar flux at lower latitude. While PAR is a function of latitude over this domain, it contributes less than 10% to the GPP – temperature spatial relationship (Figure S10). The correlation between GPP/PAR and GS mean temperature is not simply due to variations in latitude: trees growing at the same latitude have large variability in productivity due to the longitudinal-variation in GSMT. For example, at 62° ± 2° N latitude (highlighted by
closed circles in Figure 2a), GPP/PAR varies from 0.10 to 0.45 gram carbon per mega Joules (gC MJ$^{-1}$), highly correlated with the variation in GSMT.

The distribution of vegetation cover is also highly correlated with GSMT (Figure 2a). The mean GS temperatures are 9.2°C, 11.5°C, and 13.6°C for shrubland, needleleaf forest, and deciduous boreal forest, respectively. Between the transitional zones, mixed forest types co-exist. The correlation illustrates the complex role that temperature likely plays on many aspects of the ecology of the HLNF. From this correlation, we might anticipate that warming would lead to succession of the dominant trees in a fashion consistent with the patterns observed spatially (e.g. evergreen forests will transition to deciduous forests as the GSMT reaches 13°C). Limited field measurements are consistent with this hypothesis (Kharuk et al., 2007; Bjorkman et al., 2018).

Why is GPP/PAR so sensitive to temperature in the HLNF? GPP/PAR is a product of fPAR and the light use efficiency (the efficiency of conversion of absorbed sunlight to assimilated carbon) (Farquhar et al., 1980). The fPAR is observed to increase with temperature in the HLNF (Figure 2b). The fraction of absorbed sunlight resulting in fixation of CO$_2$ (light use efficiency) depends on, among other things, plant functional types. Field experiments (Ahl et al., 2004) show that the light use efficiency of deciduous forest is ~30% higher than needle leaf trees, consistent with the behavior seen in Figure 2c. In the absence of water or nutrient limitations, light use efficiency is also expected to increase with temperature in cold forests (Figure 2c) (Horn and Schulz, 2011). Due to the high rates of precipitation, snow melt, and low net radiation, water limitations – especially earlier in the growing season – are not expected. As shown in Figure S11, GPP increases with the increase in vapor pressure deficit, opposite to the expected relationship for
water-limited ecosystems. Because the rate of nitrogen (N) fixation also increases rapidly with temperature (Schimel et al., 1997; Welter et al., 2015), N limitation may not change as plants increase their demand as temperature increases. Thus, the sensitivity of GPP/PAR to GSMT results from the sensitivity of both forest structure (fPAR) and light use efficiency, with similar contributions (0.090 vs. 0.100) from each (Figure 2 b and c).

Figure 2 The spatial relationship between temperature and Gross Primary Production (GPP) and its contributing factors between 50°N and 75°N. The x-axis is the growing season mean temperature (°C), GSMT. (a) GSMT vs. the ratio between GPP and photosynthetic active
radiation (PAR) (unit: gram carbon per mega Joules (gC MJ\(^{-1}\)); (b) GSMT vs. fraction of Photosynthetic Active Radiation absorbed by the forest (fPAR) (%); (c) temperature vs. light use efficiency (unit: gC MJ\(^{-1}\)). Each point is an average over growing season at individual grid cell. Green: needleleaf forest; brown: deciduous boreal forest; blue: shrubland; The size of circles is proportional to the percentage of tree coverage. Filled circles are the points with latitude equal to 62\(^\circ\)N. The minimum of tree coverage for each type is 40%. The uncertainty is 3\(\sigma\) of the fitting. The temperature 2-meter air temperature from ERA-interim reanalysis (Table S1)

### 3.2 Hindcast of the fPAR and GPP/PAR from the changes in GSMT since 1960

With more than 30-year (1982-2016) of fPAR record from GIMMS3g (Zhu et al., 2013), we can test how well the temporal change in fPAR is explained by the observed spatial temperature sensitivity. We find that the modern correlation of fPAR with GSMT accurately captures the variability and trend in fPAR over the past several decades. Remarkably, even the interannual variations in fPAR (Figure 3) are well predicted by the spatial fPAR – T correlations, with \(R^2\) between observation and predicted fPAR equal to 0.62 (Figure 3d). The calculated temporal trend of fPAR during the period of fPAR observation is 0.20 ± 0.07 %/year, same as the observed fPAR trend (0.20 ± 0.06 %/year). Spring has the largest increasing trend, and \(R^2\) between observation and predicted fPAR is the largest (\(R^2 = 0.84\)) (Figure S12). The comparisons between fPAR observations and prediction indicate that the current spatial correlation between fPAR and GSMT can explain the historical changes in HLNF mean fPAR without invoking a significant role for CO\(_2\) fertilization or other factors.

The hindcast fPAR and GPP show larger increasing trend during the later time period (after 1983). The fPAR trend is 0.05 ± 0.08 %/year from 1960-1983, while it is 0.20 ± 0.07 %/year during 1983-2014. The GPP increasing trend is 0.06 ± 0.19 %/year and 0.45 ± 0.17 %/year respectively during these two time periods. The larger increasing trend originates from the changing trend in the temperature. The trend in temperature is 0.01 ± 0.02 °C/year during 1960-
1983, while it is 0.04 ± 0.02 °C/year during 1983-2014. The acceleration of temperature increasing rate after 1980s agrees with Sánchez-Lugo et al. (2018), which shows more than two times greater of temperature increasing rate over the high latitudes after 1980s.

The hindcast of fPAR also captures the spatial inhomogeneity in the observed fPAR changes. Between 2007-2011 and 1983-1987, the northeast Eurasia and western Europe have experienced larger increases in GSMT than North America (NA) (Figure 3a). The changes in both the observed and hindcast GS fPAR show similar spatial pattern as the temperature differences between these two time periods (Figure 3b and c). This, again, is consistent with temperature playing the dominant role in controlling fPAR. For example, if CO$_2$ fertilization was responsible for enhanced fPAR, we would not expect the fractional change to be so well correlated with the spatial change in temperature given that the change in CO$_2$ was identical across both space and time. The overestimate in the calculated fPAR changes over NA may be caused by the land cover changes from disturbance (Wang et al., 2019) that the hindcast does not consider.

The trend in the hindcast GPP during GS is 0.36 ± 0.07%/year, about two times the hindcast fPAR trend during 1960-2014, consistent with Figure 2 showing that only half of the GPP/PAR – T spatial sensitivity is attributed to change in fPAR. The trend in GPP during spring is larger than summer or fall (0.43% ± 0.09% vs. 0.30% ± 0.08% and 0.29% ± 0.10%) (Figure S12), as a result of a larger increase in temperature in spring (Figure S12) and larger PAR value in spring than late summer/fall. Consequently, the increasing trend in spring temperature has a larger impact on GPP and fPAR (Zhang et al., 2020).
Our analysis hinges upon the validity of space-for-time assumption. The comparison between fPAR observations and the corresponding hindcast provides compelling support for this assumption. In the Supplementary Material (SM.2), we further test the space-for-time assumption with GPP constrained with SIF and GPP from 11 sites of FLUXNET2015 dataset (https://fluxnet.org/data/fluxnet2015-dataset/). We compare observations and predictions in terms of seasonality and interannual variability (Figures S13 - S15). Both tests support the space-for-time assumption.

We use fPAR – T spatial relationship and its temporal prediction to test whether the conclusion depends on spatial resolution. The fPAR at 1° × 1° resolution has almost the same spatial sensitivity to temperature as at 4° × 5° resolution at seasonal time scale (0.050 vs. 0.054, 0.074 vs. 0.079, 0.056 vs. 0.055) (S16). As a result, the predicted fPAR trend is essentially identical to that at 4° × 5° resolution (0.021 ± 0.007%/year vs. 0.020 ± 0.07%/year) (Figure S17). This test indicates our method and conclusion is robust at the spatial scale of hundred kilometers. An additional test with site-level data from FLUXNET dataset (Figure S14 and S15), which has spatial representativeness of 1-10 kilometers², further supports the robustness of the conclusion with respect to spatial resolution. Whether the method and conclusion are still valid at even smaller spatial scale needs further test with high-resolution data in the future.

We have assumed that the temporal range of spring, summer, and fall in each grid cell is the same between 1960-2014 as in 2015-2017. There is a reasonable possibility that a given month may change from “spring” in the 1960s to “summer” in later time period, since our temporal resolution is only monthly. To test the sensitivity of the prediction to the separation of seasons,
we move the ending month of spring at each grid cell one month later between 1960-1979. Figure S18 shows that the trend of predicted fPAR between 1960-2014 is somewhat higher (0.16 ± 0.03%/year), but within the uncertainty of the control prediction (0.14 ± 0.03%/year) where the same season definition is used throughout the time period. With the available of high spatiotemporal SIF observations from TROPOspheric Monitoring Instrument (TROPOMI, launched in October 2017) (Köehler, et al., 2018), future studies can increase the temporal resolution and take into account the changes of growing season length in the prediction.

Figure 3. Annual hindcast of GPP and fPAR between 1960 and 2014. (A) Growing season (GS) mean temperature changes between 2007-2011 and 1983-1987. (B) Observed growing season fPAR change (%) between 2007-2011 and 1983-1987. (C) Hindcast growing season
fPAR change (%) between 2007-2011 and 1983-1987 calculated from the fPAR-T spatial relationship (SI Appendix, Figure S13). (D) Annual fPAR and GPP anomaly and trend. Black: observed fPAR anomaly and trend (dashed line); blue: hindcast fPAR anomaly and trend (%); green: hindcast GPP anomaly (%) and trend (dashed line); red: GS temperature anomaly. The mean values for each variable are the average between 1983 and 2014. The uncertainty is 3σ of the linear fitting. The fPAR observations are from GIMMS3g. The temperature is from Climatic Research Unit (CRU) (Table S1).

3.3 The impact of increasing temperature on the NH CO₂ seasonal cycle amplitude

The increase of GPP will increase both net carbon uptake during growing season and net carbon release during non-growing season (due to increased carbon pool size and increases in the rate of ecosystem respiration with temperature) (Zhao et al., 2014). As a result, it is expected that the increase in temperature will enhance the SCA of NEE. In turn, the increase of NEE SCA may have contributed to the documented large increase of NH CO₂ SCA (Graven et al., 2013). Between 1960 and 2010, the GPP and TER are calculated to have increased 20 ± 12% and 18 ± 12%, respectively, between April and Sep (Figure S7). The NEE between April and Sep increased by 23 ± 12%. The NEE amplitude, which is defined as the difference between the net source and net sink, increases by 21 ± 8.0% over 50°N-75°N (Figure S7). The net carbon sink increased by 0.25 ± 0.17 GtC between 1960 and 2010. The estimate of net carbon sink has, however, large uncertainties that require further study to evaluate.

We sample the CO₂ concentrations along the IGY and HIPPO aircraft locations, respectively, in the mid-troposphere (~500 hPa) and those four high latitude sites. Figure S19 summarizes the CO₂ SCA changes from observations and predictions. The increase in GSMT is estimated to contribute 19-42% of the total observed CO₂ SCA increase north of 45°N between two aircraft campaigns, and 34 -76% of the CO₂ SCA increase due to changes in carbon cycling in the HLNF (Figure 4). The observed CO₂ SCA changes attributed to HLNF (dashed line in Figure 4) is
calculated as the multiplication between observed CO\textsubscript{2} SCA and the ratio between model simulated CO\textsubscript{2} SCA forced by NEE over 50°N-75°N and the observed CO\textsubscript{2} SCA (Figure S19).

The surface sites may more directly reflect the changes of NEE over the HLNF, and frequent measurements make it more precise compared to aircraft observations. At the four surface sites, increase in the GSMT is calculated to contribute 56 - 72% of the observed trends in the CO\textsubscript{2} SCA (Figure 4). The NEE over 50°N-75°N accounts for ~60% of the CO\textsubscript{2} SCA at these surface sites. Assuming that a similar or even higher proportion (60-70%) of the trend in the observed CO\textsubscript{2} SCA can be attributed to HLNF and the rest is due to changes over other regions, then the increase in GSMT can explain almost the total observed CO\textsubscript{2} SCA trend attributed to HLNF, which agrees with the dominant temperature impact on the fPAR trend discussed earlier. Since our focus is the temperature impact on the carbon exchange over the HLNF, the exact proportion of the CO\textsubscript{2} SCA trend attributed to HLNF is out of scope of this study.

Figure 4 Temperature driven change of NH CO\textsubscript{2} seasonal cycle amplitude. Solid black line: the observed CO\textsubscript{2} seasonal cycle amplitude (SCA) changes between HIPPO (2009-2011) and
IGY (1958-1963) aircraft campaigns at 500 hPa. Dashed black line: the approximate CO$_2$ SCA change between HIPPO and IGY at 500 hPa attributed to the HLNF, which is the multiplication between observed CO$_2$ SCA and the ratio between model simulated CO$_2$ SCA forced by forest NEE over 50°N-75°N and the observed CO$_2$ SCA (Figure S19). Red: the calculated CO$_2$ SCA changes due to the temperature increase of the HLNF. Blue: observed (solid dot) and predicted (diamonds) CO$_2$ SCA trend (unit: ppm/decade) at Shemya Island (Alaska, US) (52.71°N), Cold Bay (Alaska, US) (55.2°N), Barrow (Alaska, US) (71.3°N), and Alert (Canada) (82.4°N). The observation source is listed in Table S1.

4 Discussion

Though the predicted changes in carbon cycling in the HLNF is supported by the fPAR and GPP observations, further study is needed to test some assumptions used in the study. We have assumed a single scaling factor between the observed SIF and GPP. While this assumption is supported by (Sun et al., 2018; Li et al., 2018), it requires additional testing within these biomes. In particular, about 50% of the observed correlation between SIF and temperature is not explained by fPAR, but by the “light use efficiency”. The estimated change in light use efficiency is correlated with biome (evergreen vs. shrub vs. deciduous trees, Figure 2c) and further work is needed to test whether all the differences in SIF of these biomes is indeed related to the differing photosynthetic efficiency of these tree types, or may (partially) reflect differences in N content, or the optics of SIF within canopies differing in structure. Though the hypothesis that N fixation increases with warming has some support in the literature (e.g., Schimel et al., 1997), it is clear that more observations are needed to test such hypothesis.

The spatial covariation between TER and temperature may include some confounding factors, such as the available soil organic carbon and sensitivity of microbial activity to temperature (e.g., Davidson et al., 2006; Davidson and Janssens, 2006, Nottingham et al., 2019). Here, TER – temperature spatial sensitivity is a statistical relationship of the net impact of temperature on
TER. A detailed process understanding on factor controls of TER still has large uncertainties and requires more research.

Though studies show that the magnitude of non-growing season NEE is positively correlated with GS NEE (Yuan et al., 2011), the assumption that the change in non-growing season NEE is proportional to the changes in GS NEE may not always be valid. In using this assumption, we have ignored the impact of changes of snow coverage and freezing/thaw on ecosystem respiration (Natali et al., 2019) and nonhomogeneous changes in temperature throughout the year. There is evidence that the rate of respiration outside the GS may have increased faster than during GS because of larger temperature increase during fall and winter (Schuur et al., 2015).

Remarkably, the hindcast based on the spatial temperature sensitivity captures both the long-term trend and some fraction of short-term interannual variability (Figure 3 and Figure S12). We attribute this skill in part to the use of a single curve of the spatial data across multiple plant functional types and climate zones. The fit thus can capture possible species succession, which takes several decades, since the single curve assumes the same sensitivity to temperature among different forest types (Figure 2). At the same time, it also predicts with some skill the faster changes in GPP and other quantities associated with interannual temperature differences, because the single curve spans large range in temperature (Figure S3 and S6) and the temperature anomaly of any single year likely sits within the range. This is especially true of springtime when the initiation of photosynthesis is be highly sensitive to small changes in temperature.
The rapid increase in GPP with GSMT is certainly limited: the data shown in Figure 2 extends only to a GSMT of 16°C. It is expected that at some point, photosynthetic, water, and other nutrient limitations will limit further increase in GPP with temperature, especially during summer. Even today there is evidence that by mid/late summer such limitations come into play (e.g., Buermann et al., 2018), and browning trend has been observed over limited regions (Beck and Goetz, 2011). Radiation may also limit the response of GPP to temperature increase in the future, especially in the fall (Zhang et al., 2020). In addition, the increase in abiotic (e.g., fire) and biotic (e.g., herbivory) disturbance with climate will also impact productivity (Sulla-Menashe et al., 2018).

Despite the limitations noted above, the spatial variations in fPAR and GPP in the HLNF are remarkably well correlated with temperature (Figure 2 and Figure S2), consistent with a thermally-limited ecosystem. The strong temperature control on changes in fPAR and GPP and its large contributions to changes in CO₂ SCA are in contrast to results obtained from most (Piao et al., 2018; Bastos et al., 2019), but not all biogeochemical models (Forkel et al., 2016), which generally suggest a dominant CO₂ fertilization effect on the changes of NEE over HLNF (and thereby the CO₂ SCA). While Zhu et al., 2016 show dominant climate impact on the greening trend (i.e., leaf area index) over high latitudes, later studies (Piao et al., 2018; Bastos et al. 2019) suggest a dominant CO₂ fertilization effect on CO₂ SCA changes with a similar set of TRENDY models. The high degree of spatial correlation between fPAR, GPP, and GSMT provides a key observational measure that can be used to evaluate and improve upon corresponding model simulated properties. A factorial simulation with changes only in temperature drivers for biogeochemical models will be able to isolate temperature effect from other factors, and thereby test the model behavior against observations to identify possible model deficiencies.
That warming alone accounts for nearly all the observed fPAR trend is in agreement with prior observational studies that suggest CO$_2$ fertilization is not a major contributor to carbon exchange in mature forests (Jiang et al., 2020) and, more broadly, in the boreal (Girardin et al., 2011). Over mature forest, even under warmer climate (when CO$_2$ fertilization is expected to contribute much more than in cooler climates), GPP increases ~12% with 150 ppm CO$_2$ enrichment (Jiang et al., 2020). Extrapolating the results in Jiang et al. (2020), the ~80 ppm increase in CO$_2$ between 1960s and 2010s implies a ~6% increase of GPP over boreal forest, where a large proportion of forests is undisturbed, especially over Eastern and Northern Canada and Eurasia (https://daac.ornl.gov/NACP/guides/NA_Tree_Age.html). Since fPAR accounts for half of GPP increases, fPAR may only increases by ~3% from CO$_2$ fertilization between 1960s and 2010, well within the uncertainty estimates shown in Figure 3. The more important role of temperature implied by the current spatial dependence of fPAR on GSMT is consistent with warming explaining most of the temporal trend in fPAR.

The ensemble of CMIP5 models suggest that by 2100, absent of efforts to mitigate climate forcing, these forests will experience an additional 2-5 °C increase relative to 1986-2005 (RCP6.0, IPCC AR5) in GSMT. Figure 2 implies that such a change may dramatically alter the distribution of tree species over the region: evergreen trees will invade the shrublands all the way to the Arctic ocean, and much of the current evergreen forests will transition to deciduous trees. GPP over the region has the potential to double. Long-term consistent observation records are essential to monitor such possible changes over HLNF.

5. Conclusion
Three quarters of the spatial variations in gross primary production and fPAR over high-latitude northern forest can be explained by the spatial variation in growing season mean temperature. Substituting observed spatial sensitivity to temperature for time, we estimate the change in carbon cycling in the HLNF between 1960s and 2010s, and diagnose what contribution these changes have made to the observed increase in the CO₂ SCA. The hindcast captures both the magnitude and spatial variability of the trends in observed fPAR. We calculate that the observed high-latitude warming results in a 20 ± 12% and 18 ± 12% increase in GPP and TER, respectively, during growing season. The calculated changes in GPP and TER result in an increased seasonality of NEE by 21 ± 8% between the 1960s and 2010s. In turn, the increase of the NEE SCA contributes 56-72% of the observed increase in the CO₂ SCA at high latitude continental surface sites and 17-43% of the observed change in the SCA in the mid-troposphere.

Data Availability

The sources of all the data used in this paper are listed in Table S1.

Acknowledgments

We thank Editor Eric Davidson, reviewer Shilong Piao, and two other anonymous reviewers for their constructive comments. We also thank James Randerson, Dave Schimel, and TRENDY modelers (Stephen Sitch, Pierre Friedlingstein, Vivek Arora, Atul Jain, Markus Kautz, Danica Lobardozzi, Sebastian Lienert, Julia Nabel, Benjamin Poulter, Nicolas Vuichard, Andy Wiltshire, and Ning Zeng) for their comments and suggestions in preparing this paper. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) division at Ames Research Center. All authors declare no financial conflicts of interest. We acknowledge the funding support from NASA.
Carbon Cycle Science program and OCO-2/3 Science Team program. Part of the research was carried out at Jet Propulsion Laboratory, Caltech.

References


62. Zhu, Z., Bi, J., Pan, Y., et al. (2013). Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of Photosynthetically Active Radiation (FPAR)3g Derived from Global Inventory Modeling
(a) GPP/PAR = 0.030 e^{(0.19 \pm 0.02 \, \text{T})} 
\quad R^2 = 0.73

(b) fPAR = 16.911 \, e^{(0.09 \pm 0.01 \, \text{T})} 
\quad R^2 = 0.73

(c) GPP/(PAR \times fPAR) = 0.002 \, e^{(0.10 \pm 0.02 \, \text{T})} 
\quad R^2 = 0.56
(a) Growing season (GS) T difference between 2007–2011 and 1983–1987

(b) Observed GS fPAR change (%)

(c) Predicted GS fPAR change (%)

(d) Hindcast GS GPP and fPAR

\[
trend(f\text{PAR}(\text{obs})) = 0.20 \pm 0.06\%/\text{year}
\]

\[
R^2 = 0.64
\]

\[
trend(f\text{PAR}(\text{mod (1983-2014)})) = 0.20 \pm 0.07\%/\text{year}
\]

\[
trend(f\text{PAR}(\text{mod (1960-2014)})) = 0.14 \pm 0.03\%/\text{year}
\]

\[
trend(GPP(\text{mod (1960-2014)})) = 0.33 \pm 0.07\%/\text{year}
\]
Solid: observation
Dashed: approximate observed SCA change from the forest (50N-75N)
Predicted changes due to warming (50N-75N)
Flask sites SCA changes; dot: obs; diamond: prediction