Widespread seasonal compensation effects of spring warming on northern plant productivity

Authors

Wolfgang Buermann[1,2]*, Matthias Forkel[3], Michael O'Sullivan[1], Stephen S. Sitch[4], Pierre Friedlingstein[5], Vanessa Haverd[6], Atul K. Jain[7], Etsushi Kato[8], Markus Kautz[9], Sebastian Lienert[10,11], Danica Lombardozzi[12], Julia E.M.S. Nabel[13], Hanqin Tian[14,15], Andrew J. Wiltshire[16], Dan Zhu[17], William K. Smith[18] and Andrew D. Richardson[19,20]

[1]Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK
[2]Institute of the Environment and Sustainability, University of California, Los Angeles, 90095 Los Angeles, USA
[3]Climate and Environmental Remote Sensing Group, Department for Geodesy and Geoinformation, TU Wien, 1040 Vienna, Austria.
[4]College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QF, UK
[5]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QE, UK
[7]Department of Atmospheric Sciences, University of Illinois, Urbana, IL 61821, USA
[8]Institute of Applied Energy, Minato, Tokyo 105-0003, Japan
[9]Forest Research Institute Baden-Württemberg, 79100 Freiburg, Germany
[10]Climate and Environmental Physics, Physics Institute, University of Bern, Bern, Switzerland
[11]Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland
[12] National Center for Atmospheric Research, Climate and Global Dynamics, Terrestrial Sciences Section, Boulder, CO 80305, USA
[13]Max Planck Institute for Meteorology, Hamburg, Germany
[14]School of Forestry and Wildlife Sciences, Auburn University, 602 Ducan Drive, Auburn, AL 36849, USA
[15]Research Center for Eco-Environmental Sciences, State Key Laboratory of Urban and Regional Ecology, Chinese Academy of Sciences, Beijing, 100085, Chin
[16]Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK
[17]Laboratoire des Sciences du Climat et de l'Environnement, LSCE CEA-CNRS-UVSQ, 91191 Gif sur Yvette, France
[18] School of Natural Resources and the Environment, University of Arizona, Tucson, AZ 85721, USA
[19]School of Informatics, Computing and Cyber Systems, Northern Arizona University, Flagstaff AZ 86011 USA
[20]Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff AZ 86011 USA

*Corresponding author

Climate change is causing plant phenological cycles to shift[1], thereby altering the functioning of ecosystems, which in turn induces feedbacks to the climate system[2]. In northern ecosystems, warmer springs lead generally to an earlier onset of the growing season[3,4] and increased ecosystem productivity early in the season[5]. But in-situ[6] and regional studies[7-9] also provide evidence for corresponding lagged effects of spring warmth on plant productivity during subsequent summer and autumn seasons. Yet, our present understanding of such lagged effects – including their direction and geographic distribution – is still very limited. Here we analyse long-term satellite data products, flux tower and model–based data and show that there are widespread and contrasting lagged productivity responses to spring warmth across northern ecosystems. We find that 13-16% and 4-6% of a total area of about 41 million km² show adverse and beneficial lagged effects, respectively. In contrast, current-generation terrestrial carbon cycle models show significantly lower areal fractions of adverse lagged effects (1–14%) and much higher areal expanses of beneficial lagged effects (9–54%). Furthermore, we find that elevation and seasonal precipitation patterns largely dictate the geographic pattern and direction of the lagged effects. Inadequate consideration of the effects of seasonal build-up of water stress on seasonal vegetation growth may be able to explain why current models do not adequately represent lagged effects associated with spring warming. Overall, our results suggest that for many northern ecosystems, the benefits of warmer springs on growing season ecosystem productivity are substantially reduced by the accumulation of

seasonal water deficits despite the fact that northern ecosystems are thought to be largely temperature and radiation limited[10].

Northern land regions have experienced substantial warming since the early 1970s and this has changed how ecosystems function[11]. One prominent example of such emerging ecosystem responses is shifts in plant phenological cycles: earlier spring onset and delayed autumn senescence are together lengthening the northern growing season[6,12]. These phenological shifts have altered ecosystem productivity[5,6,8,13,14] and the seasonality of important ecosystem feedbacks to the atmosphere and climate system[6,15].

Warmer and earlier springs may also influence ecosystem function later in the growing season through indirect or lagged effects[16,17]. For example, in-situ studies provide evidence for significant positive lagged effects on ecosystem productivity, whereby the influence of warmer springs may be conveyed to subsequent seasons through development of larger leaf area and/or increased foliar nitrogen[6]. In contrast, warmer/earlier springs may also cause earlier autumn senescence because of fixed leaf life spans[18] or adversely affect plant productivity later in the season through the building up of water deficits[7-9,19,20]. However, at present a more comprehensive understanding of such lagged productivity responses is still lacking.

Here, we exploit long-term satellite-based measures of vegetation greenness (as a proxy of potential photosynthesis)[21], flux tower- and model-based estimates of CO_2 uptake through photosynthesis (gross primary productivity (GPP))[22,23] and high-resolution climate data[24] to estimate the strength and geographic distribution of lagged effects that capture the influence of spring phenological transitions on plant productivity during subsequent summer and autumn seasons. Our analysis framework relies on identifying correlations between spring temperatures (which serve as an independent phenological indicator), and satellite greenness or simulated GPP during spring and subsequent seasons to estimate concurrent phenological responsiveness and linked lagged effects (see Methods).

Across northern lands, correlations between annual spring temperatures and spring greenness show significantly positive and spatially extensive pattern consistent with the notion of a tight control of spring temperature on concurrent plant productivity: 80% of northern (>30°N) vegetated non-agricultural land (total study area ~41 million km^2) exhibit statistically significant ($P < 0.05$ at grid cell level) positive correlations (Fig. 1a). To assess lagged effects on plant productivity associated with anomalous spring temperatures, we computed partial correlations between spring temperature and subsequent summer and autumn greenness, whereby covarying effects of concurrent climate on these correlations are controlled for (see Supplementary Information section 1). Partial correlations between annual spring temperature and subsequent summer greenness show widespread positive (6%, $P < 0.05$) as well as negative (6%, $P < 0.05$) pattern (Fig. 1b). Areas of positive partial correlations are predominantly situated in Eurasia covering vast regions north of 50°N, whereas areas displaying negative correlations are more localized in western North America, Siberia and temperate eastern Asia. The partial correlation pattern between spring temperature and autumn greenness indicate an extension of the 'summer' pattern of negative correlation (11%, $P < 0.05$;

positive correlations cover only 2%, $P < 0.05$) with additional coverage seen mainly in northeastern Eurasia and temperate central Asia (Fig. 1b and c). While long-term trends in temperature and greenness may potentially influence these correlations, a corresponding analysis on detrended data shows that the patterns are similar (Supplementary Information section 1) which suggests a dominant influence of interannual to quasi-decadal variability on the correlation pattern between spring temperature and satellite greenness during subsequent seasons. A comparison of the strength of these lagged relationships with concurrent climatic influences on greenness pattern shows that at regional scales the influence of preceding spring temperatures on summer and autumn greenness can be equally important or even dominant (Supplementary Information section 1).

To further assess the robustness of the identified satellite-based lagged productivity responses we also compared them to those inferred from tower-based measurements of land-atmosphere CO_2 flux (FLUXNET). Results show that across n=16 tower sites, the strength and direction of relationships between spring temperature and spring as well as summer greenness correspond well to those based on FLUXNET estimates of spring temperature and spring as well as summer GPP (Extended Data Fig. 1). The agreement between satellite- and tower-based relationships between spring temperature and autumn greenness and GPP, respectively, is however not as strong (Extended Data Fig. 1). This validation has several caveats including few available tower sites and differences in spatial scales for satellite (coarse) and tower (fine-scale) data, but the overall consistency in the estimated lagged productivity responses suggests that the satellite-based estimates are plausible.

The geographic distribution of the relationships between changes in spring temperature and subsequent summer greenness (see Fig. 1B) suggests that some combination of climate, elevation, and/or landcover may explain these patterns. To investigate this, we conducted a random forest (RF) analysis using a set of predictors that encapsulate such factors (see Supplementary Information section 2). Results show that the spring temperature-summer greenness partial correlation pattern can be explained with elevation and selected climate variables (e.g. summer precipitation and precipitation seasonality) acting as the most important variables (Extended Data Fig. 2 and Supplementary Information section 2). Across northern ecosystems, we find that these partial correlations tend to become more negative with higher elevation, but such well-defined directional relationships are not as apparent for important precipitation metrics (Extended Data Fig. 2).

Grouping the lagged productivity responses based on the direction of robust correlations between spring temperature and spring, summer and autumn greenness, respectively, shows large clusters of regions with negative lagged effects and more scattered areas with positive lagged effects (Fig. 2a). Hereby, negative 'lagged productivity response scenarios' associated with spring warming and greening coupled with summer and/or autumn greenness declines stretch over vast areas in western North America, Siberia and to some extent eastern temperate Asia, whereas positive lagged effects are more common in eastern Eurasia above 50°N (except Siberia).

Carbon cycle models must be able to simulate vegetation phenology responses and corresponding impacts on ecosystem productivity and net carbon uptake realistically to credibly estimate climate-carbon feedbacks [25]. We therefore assessed the ability of ten current-generation models contributing to TRENDYv6[22,23] to replicate the observed lagged productivity responses in respect to spring warming. Results show a substantially higher areal coverage of positive lagged effects on plant productivity for the multi-model mean (and significantly lower coverage of negative lagged effects) than for the satellite-based estimates (Fig. 2a and b). While there are marked differences among the individual models (Extended Data Fig. 3), a striking pattern in the ensemble is the near absence of any negative lagged effects across Siberia and the overall abundance of positive lagged effects that extend over summer and autumn seasons (Fig. 2a and b). Satellite greenness has been used extensively as a proxy for vegetation productivity[3,26] but direct comparisons between greenness- and GPP-based pattern may be limited (see Methods). However, a similar analysis with two satellite-constrained GPP datasets (based on upscaled FLUXNET data and a light use efficiency model; see Methods) shows nearly identical lagged productivity pattern to those based on satellite greenness (Extended Data Fig. 4).

Grouping these 'lagged productivity response scenarios' more broadly in positive and negative lagged effects yields an areal extent of regions showing positive lags of 36% for the TRENDYv6 ensemble (spanning 9–54% for the ten individual TRENDYv6 models) and 4–6% for the satellite-based estimates (Fig. 2c). The areal coverage of negative lagged effects based on the TRENDYv6 ensemble is only 2% (spanning 1–14% for the ten models) compared to the satellite-based estimates of 13-16%.

Why can present terrestrial carbon cycle models not adequately replicate the spatial pattern of observed lagged productivity responses in respect to warmer springs? One key factor may be how seasonal vegetation growth is represented in the models. To assess this, we performed a similar seasonal correlation analyses with satellite-based and modelled leaf area index (LAI) data (see Methods). These results show an even larger discrepancy in the areal proportions of positive and negative lagged LAI responses to spring warming between observation-based and modelling approaches compared to the results based on productivity metrics (Fig. 2c and d, Extended Data Fig. 4). The substantial overestimation of growing season LAI in the models in response to spring warmth may cause too much new carbon to be allocated in plant tissue, which then serves to enhance GPP.

Water availability may cause adverse lagged effects in response to spring warmth and could help to reconcile the differences in observations and models. To further investigate this we performed a regional analysis for the Western US and Siberia where observation-based and simulated lagged productivity responses show more converging and diverging pattern (see Fig. 2). For the Western US, we find that seasonal trajectories in aggregated satellite-based and modelled LAI and evapotranspiration (ET) both display positive anomalies during spring in years with warmer springs and corresponding negative anomalies later in the growing season (suggestive of negative lagged effects associated with a buildup of water stress) (Fig. 3a and b). For Siberia, however, the seasonal trajectories in observation-based and modelled LAI for warm spring years start to diverge substantially during summer and autumn with the observations displaying more negative anomalies during summer and autumn (again suggestive of water stress) and the opposite pattern for the models (Fig. 3c). Seasonal trajectories of observation-based and modeled ET for years with anomalous spring temperatures are more in agreement, although there is some indication that the models tend to underestimate water stress in summer in warm spring years (Fig. 3d). The consistent response in observed and modelled LAI and ET in respect to spring warmth over Western US, a region which is known for its vulnerability to drought in respect to spring warmth[27-29], suggests that the model's hydrology and phenology schemes are generally fit for purpose. The strong divergence between observation-based and modelled seasonal vegetation growth responses to spring warmth over Siberia (which is dominated by needleleaf deciduous forests) may be due to underestimating the effects of water stress on seasonal canopy development and omission of fixed leaf life spans in the models (Extended Data Table 1 and Supporting Information section 3). We estimate that due this observation-model mismatch across Siberia annual GPP for a warm spring year (relative to mean conditions) may be up to 4 times higher in the TRENDYv6 ensemble (1.7 PgC/yr) compared to an observation-constrained estimate based on upscaled FLUXNET data (0.4 PgC/yr) (Extended Data Fig. 5).

Our analysis based on satellite vegetation records over multiple decades provides first evidence for widespread positive and negative lagged plant productivity responses across northern ecosystems in association with warmer springs. The spatially extensive pattern of negative lagged effects identified here also implies substantially reduced benefits for ecosystem productivity and carbon sequestration from longer northern growing seasons under climate change. We also show that current terrestrial carbon cycle models substantially underestimate (overestimate) negative (positive) lagged effects associated with spring warming. This is possibly because of inadequately capturing the effects of seasonal buildup of water stress on seasonal vegetation growth. Continued monitoring of emerging ecosystem responses and improved modelling capabilities will thus be crucial to improve understanding of the complex interactions of a changing climate, shifts in phenological cycles and impacts on energy, water and carbon cycles.

References

1. Fu, Y.H. *et al*. Declining global warming effects on the phenology of spring leaf unfolding. *Nature* **526**, 104–107 (2015).
2. Peñuelas, J., Rutishauser, T. & Filella, I. Phenology feedbacks on climate change. *Science* **324**, 887–888 (2009).
3. Myneni, R., Keeling, C., Tucker, C., Asrar, G. & Nemani, R. R. Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* **386**, 698–702 (1997).
4. Menzel, A. *et al.* European phenological response to climate change matches the warming pattern. *Glob. Change Biol*. **12**, 1969–1976 (2006).
5. Keenan, T.F. *et al.* Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Clim Change* **4**, 598–604 (2014).

6. Richardson, A. D. *et al.* Climate change, phenology, and phenological control of vegetation feedbacks to the climate system *Agric. For. Meteor*. **169** 156–173 (2013).

7. Grippa, M. *et al.* The impact of snow depth and snowmelt on the vegetation variability over central Siberia, *Geophys. Res. Lett*. **32,** L21412 (2005).

8. Buermann, W., Bikash, P. R., Jung, M., Burn, D. H. & Reichstein, M. Earlier springs decrease peak summer productivity in North American boreal forests. *Environ. Res. Lett*. **8**, 024027 (2013).

9. Sippel, S. *et al.* Contrasting and interacting changes in simulated spring and summer carbon cycle extremes in European ecosystems. *Environ. Res. Lett.* **12**, 075006 (2017).

10. Nemani, R. R. *et al.* Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science* **300**, 1560–1563 (2003).

11. Settele, J. *et al.* Terrestrial and inland water systems. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the IPCC [Field, C.B. *et al.* (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 271–359 (2014).

12. Forkel, M. *et al.* Codominant water control on global interannual variability and trends in land surface phenology and greenness. *Glob. Change Biol.* **21**, 3414–3435 (2015).

13. Piao, S. L. *et al.* Net carbon dioxide losses of northern ecosystems in response to autumn warming. *Nature* **451**, 49–53 (2008).

14. Forkel, M. *et al.* Enhanced seasonal CO_2 exchange caused by amplified plant productivity in northern ecosystems. *Science* **351**, 696–699 (2016).

15. Forzieri, G. *et al.* Satellites reveal contrasting responses of regional climate to the widespread greening of Earth, *Science* **356**, 1180–1184 (2017).

16. Richardson, A. D. *et al*. Influence of spring and autumn phenological transitions on forest ecosystem productivity. *Philosoph. Trans. R. Soc. B* **365**, 3227–3246 (2010).

17. Wolf, S. *et al.* Warm spring reduced carbon cycle impact of the 2012 US summer drought. *Proc. Natl. Acad. Sci.* **113**, 5880–5885 (2016).

18. Keenan, T. & Richardson, A. D. The timing of autumn senescence is affected by the timing of spring phenology: implications for predictive models. *Global Change Biol.* **21**, 2634–41 (2015).

19. Barnett, T. P., Adam, J. C. & Lettenmaier, D. P. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature* **438**, 303–309 (2005).

20. Zhang, K., Kimball, J. S., Kim, Y. & McDonald, K. C. Changing freeze–thaw seasons in northern high latitudes and associated influences on evapotranspiration. *Hydrol. Proc.* **25**, 4142–4151 (2011).

21. Pinzon, J. E. & Tucker, C. J. A Non-Stationary 1981–2012 AVHRR NDVI3g Time Series. *Remote Sens.* **6**, 6929–6960 (2014).

22. Sitch, S. *et al.* Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences* **12**, 653–679 (2015).

23. LeQuéré, C. *et al.* Global Carbon Budget 2017. *Earth System Sci. Data Disc*.

https://doi.org/10.5194/essd-2017-123 (2017).

24. Harris, I., Jones, P. D., Osborna, T. J., & Lister, D. H. Updated high-resolution grids of monthly climate observations – the CRU TS3.10 dataset, *Int. J. Climatol. 34*, 623–642, (2014).
25. Cadule, P. *et al.* Benchmarking coupled climate-carbon models against long-term atmospheric CO2 measurements. *Glob. Biogeochem. Cycles* **24**, Gb2016 (2010).
26. Huang, M. *et al.* . Velocity of change in vegetation productivity over northern high latitudes. *Nature Eco. Evo.* **1**, 1649–1654 (2017).
27. Westerling, A. L., Hidalgo, H. G, Cayan, D. R. & Swetnam, T. W. Warming and earlier spring increase western U.S. forest wildfire activity. *Science* **313**, 940–943 (2006).
28. Sacks, W. J., Schimel, D. S. & Monson, R. K. Coupling between carbon cycling and climate in a high-elevation subalpine forest: a model-data fusion analysis. *Oecologia* **151**, 54–68 (2007).
29. Parida, B. R. & Buermann, W. Increasing summer drying in North American ecosystems in response to longer nonfrozen periods. *Geophys. Res. Lett*. **41**, 5476–5483 (2014).

Acknowledgements

M.O. is funded through an EU Marie Curie Integration grant to W.B.. M.F. is funded through the TU Wien Wissenschaftspreis 2015, a personal science award to Wouter Dorigo. V.H.'s contribution is supported through funding from the Earth Systems and Climate Change Hub of the Australian Government's National Environmental

Science Program. H.T. is supported by National Key R & D Program of China (no. 2017YFA0604702) and US National Science Foundation (1210360, 1243232).

A.D.R. is funded through the Macrosystems Biology Program of the National Science Foundation (award EF-1702697). This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia, and USCCC. The ERA-Interim reanalysis data are provided by ECMWF and processed by LSCE. The FLUXNET eddy covariance data processing and harmonization was carried out by the European Fluxes Database Cluster, AmeriFlux Management Project, and Fluxdata project of FLUXNET, with the support of CDIAC and ICOS Ecosystem Thematic Center, and the OzFlux, ChinaFlux and AsiaFlux offices. We also thank Martin Jung for providing upscaled FLUXNET GPP data.

Author contributions

W.B., M.F. and A.D.R. designed the research. W.B. and M.F. carried out the analysis and W.B. wrote the manuscript with contributions from all authors. S.S., P.F, V.H., A.K.J., E.K., M.K., S.L., D.L., J.E.M.S.N., H.T., A.J.W., and D.Z. contributed to TRENDY results. W.K.S. contributed to LUE-FPAR3g results.

Competing financial interests

The authors declare no competing financial interests.

Figure Captions

Figure 1 | Spatial pattern of concurrent and lagged productivity responses to spring warming based on satellite greenness observations. Panel (a) shows grid cell correlations between yearly spring temperatures and spring satellite vegetation greenness (expressed through the NDVI, normalized difference vegetation index), for our study period 1982-2011. Also shown are partial correlations between annual spring temperatures and subsequent (b) summer NDVI as well (c) autumn NDVI over this period. In these partial correlations, the covarying influences of summer temperatures and precipitation (panel b) and autumn temperature and precipitation (panel c) on the lagged spring temperature – summer/autumn NDVI correlations have been removed. Seasons are defined through a local adaptive procedure (see Methods). Absolute r-value categories correspond to significance levels $P = 0.3$ ($r = 0.20$), $P = 0.2$ ($r = 0.24$), $P = 0.05$ ($r = 0.36$) and $P = 0.01$ ($r = 0.46$), respectively. For each map frequency histograms showing areal coverage of corresponding positive and negative correlations, estimated as fraction of total study area, are also provided (see insets). Areas cultivated or managed[32] (light grey) are not included in the analysis.

Figure 2 | Spatial pattern of lagged productivity response scenarios based on satellite greenness observations and modelling approaches. The two maps summarize direction of robust ($P < 0.05$) grid cell correlations between annual spring temperature and spring, summer and autumn (a) satellite NDVI as well as (b) simulated GPP from the TRENDYv6 multi-model mean. For example, the 'lagged productivity response scenario' denoted as '+++' captures positive correlations between spring temperature and spring, summer and autumn NDVI (GPP), respectively. Here, the relationships between spring temperature and subsequent summer and autumn NDVI (GPP) are estimated through partial correlations whereby effects of covarying concurrent climate influences have been controlled for (see Fig. 1 and Methods). Corresponding pattern for individual models are shown in Extended Data Fig. 3. Areas with no robust link between spring temperature and spring NDVI or GPP (dark grey) and areas cultivated or managed (light grey) are outlined. The two focal regions (Western US and Siberia) in this study are also delineated (black-dashed rectangles). Panel (c) shows extent of areas with either no, positive or negative lagged effects (see definition in panel a) within the study region for satellite NDVI and GPP based on TRENDYv6. In addition, corresponding results from a similar analysis for two satellite-constrained GPP datasets, based on upscaled FLUXNET data (FluxNetG) and a light use efficiency model (LUE-FPAR3g; see Methods), are also shown (see also Extended Data Fig. 4). Panel (d) shows the results from a complementary analysis for satellite-based and modelled LAI (see Methods). Heavy shaded columns represent satellite-constrained estimates and those based on the TRENDYv6 multi-model mean and light shaded columns represent estimates for the individual TRENDYv6 models. Results from the same analysis for detrended data show that the differences in satellite- and model-based estimates of areal proportions of positive and negative lagged effects are similar (Supplementary Information section 1).

Figure 3 | Seasonal trajectories of regionally averaged LAI and ET anomalies based on observation-constrained and modeling approaches for warm and cold spring years. The panels depict anomalies in spatially averaged and composited LAI and ET based on satellite-constrained estimates (LAI3g and ET-GLEAM) and model simulations (TRENDYv6 multi-model mean) for (a, b) Western US and (c, d) Siberia. Western US encompasses the non-agricultural regions from 120°W to 105°W and 40°N to 50°N, whereas Siberia is defined from 80°E to 125°E and 60°N to 70°N (see also Fig. 2). Anomalies are relative to the study period 1982-2011. The monthly maximum composites are based on the mean LAI (ET) of the seven warmest and coldest spring years within the study period. Start and end of climatological spring, summer and autumn season are also outlined (vertical grey dashed lines). Uncertainty bounds (shaded area) reflect the spread in the monthly LAI (ET) anomalies within the compositing period (± 1 s.d., n=7).

Methods

Data sources. For satellite vegetation data, we used the GIMMS-NDVI version 3g (NDVI3g)[21] and the LAI3g[30] products both available at 8-km spatial and 15 days temporal resolution covering our study period 1982–2011. The NDVI3g data stem from optical surface reflectance measurements from a series of NOAA-AVHRR satellites, and in its generation effects of orbital drifts, inter-sensor calibration and stratospheric aerosols from volcanic eruption have been corrected for making it presently the most consistent long-term satellite vegetation data[21]. The LAI3g fields are derived from the NDVI3g data using an artificial neural network model[30]. Gridded monthly climate data were obtained from the Climatic Research Unit (CRU TS3.23) at 0.5° spatial resolution[24] for our study period (1982-2011). As an estimate for observation-constrained ET, we included the Global Land Evaporation Amsterdam Model (GLEAM) data set, which has a 0.25° spatial resolution at daily time steps[31]. While the GLEAM approach is based on an empirical model, it is heavily constrained by observations through assimilating satellite microwave vegetation optical depth data as a proxy for water stress[31]. In addition, land cover data used in this study are based on the GLC2000 land cover classification[32]. For complementary analyses, we also used two observation-constrained monthly GPP data sets: 1) We used GPP data (0.5° spatial resolution and available for 1982-2008) derived from upscaled carbon observations based on the global FLUXNET tower network (termed FluxNetG in this study)[33]. Note, FluxNetG is different from the previously published version (FluxNet-MTE)[33] since it has been produced with inputs from only a single satellite vegetation data set (NDVIg; a predecessor of NDVI3g) to reduce artefacts from usage of multiple satellite data (the FLuxNetG data set was also used in ref. 8). 2) We used GPP data (0.5° spatial resolution and available for 1982-2011) derived using the light use efficiency (LUE) MODIS GPP algorithm driven by bi-monthly GIMMS FPAR3g (termed LUE-FPAR3g in this study)[34]. Additional meteorological driver data required as input into the MODIS GPP algorithm were derived from NCEP-DOE Reanalysis II (http://www.esrl.noaa.gov). For more information on the GIMMS3g GPP dataset, see Smith et al. (2016)[34].

TRENDYv6 models. We also analyzed monthly GPP, LAI and ET simulation outputs for 1982 to 2011 from ten terrestrial carbon cycle models that were part of a recent model intercomparison project: TRENDYv6[22,23]. The models included in the analysis here are the LPX-Bern, LPJ-GUESS, ISAM, CABLE, VISIT, CLM4.5, DLEM, JSBACH, ORCHIDEE-MICT and JULES. In TRENDYv6, the models were forced with CRUNCEPv6 climate data, which is based on a merged product of the monthly CRU climate data and to be consistent with the TRENDYv6 ensemble we also used this climate dataset in this study. In addition, a set of factorial simulations[22] were performed and we analyzed outputs from a simulation in which only atmospheric CO_2 and climate were varied (land use change held fixed; experiment 'S2') since our study focus was on non-agricultural ecosystems. For an overview of the processes included in the models with relevance for this study see Extended Data Table 1. For a more general overview of the models see Table 4a and Table 5 in Le Quéré et al.[23].

Analysis framework. The satellite-based bi-monthly GIMMS NDVI3g and LAI3g vegetation data were averaged to a monthly temporal resolution (to be consistent with the TRENDYv6 model outputs). Then the fine-scale satellite vegetation and coarse-scale CRU temperature fields were (dis)aggregated to a common 0.25° spatial grid on which all correlation analyses were performed. The motivation for this spatial aggregation step is two-fold: (i) it retains a certain level of spatial information inherent in the satellite products and (ii) aligns more closely with the coarser spatial resolutions of the TRENDY carbon cycle models. Model outputs from TRENDYv6 were either analyzed at their native model resolutions spanning grid cell dimensions from 0.5° to 1.9°[22] or resampled to a common 0.5° grid through nearest neighbors (e.g. for estimation of multi-model means of GPP, LAI and ET at grid-cell levels).

To estimate lagged vegetation growth and productivity responses we first divided the mean seasonal cycle of NDVI or simulated GPP (based on the 30 year study period) into spring, summer and autumn periods for each grid cell. Hereby, the start of spring and end of autumn periods are defined by the month in which corresponding temperatures are closest to 0°C, whereas the start and end of the summer periods are defined by the month in which the NDVI (GPP) is closest to 95% (85%) of the annual maximum NDVI (GPP), respectively. Alternative approaches for characterizing phenological cycles involving start- and end-dates of the growing season are more ambiguous if it is based solely on optical vegetation indices[35,36] or when the underlying data have relatively low temporal resolutions as in this study[12].

In a next step, we (building on the conceptual model of Richardson *et al*.[16]) classified 'lagged productivity response scenarios' for each grid cell as follows: First, as a minimum requirement for phenological responsiveness to spring warming, we require the springtime temperature and the response variable of interest (NDVI, LAI or GPP) to be significantly ($P < 0.05$) and positively correlated. Second, we then define a lagged productivity (NDVI, GPP) or phenology (LAI) response scenario on the basis of the direction of robust ($P < 0.05$) partial correlations between annual spring temperatures (as independent phenological indicator) and subsequent summer as well as autumn seasonal means of the response variable of interest; for example if at a given locality annual spring temperature is positively correlated with spring NDVI but negatively correlated with subsequent summer NDVI and not robustly correlated with autumn NDVI the assigned scenario label would be '+–0' where the type and sequence of symbols denotes the direction of correlations between spring-spring, spring-summer and spring-autumn relationships, respectively (see Figure 2). Partial correlations are used to control for covarying effects of climate over seasonal time scales, which can confound the correlations between annual spring temperatures and subsequent summer and autumn response variables (see Supplementary Information section 1).

As indicated, the long-term satellite vegetation data (NDVI3g, LAI3g) exploited here stem from a series of satellites and while this record has been carefully assembled and also validated to some extent[30] remaining non-vegetation artefacts in the data cannot be ruled out[37]. Further, satellite greenness (or NDVI) captures the amount of light absorbed by chlorophyll in green leaves[38] and has been exploited extensively as a proxy for spatially-resolved vegetation productivity at continental and multi-decadal scales[3,26]. However, to overcome the limited comparability of directly observed NDVI-based and simulated GPP-based pattern we also analysed observation-constrained GPP data and corresponding results show good agreement in lagged productivity pattern at both site level (using GPP flux tower data) and across northern ecosystems (using gridded GPP data from upscaled FLUXNET and a LUE model) providing further support for the robustness of our results (see Extended Data Fig. 4). Finally, we also use satellite-based and modelled LAI data to probe the mismatch between lagged greenness and modelled (TRENDYv6) GPP responses to spring warmth.

Data availability. The satellite NDVI3g data that support the findings of this study were downloaded from http://ecocast.arc.nasa.gov/data/pub/gimms/3g.v0/. The satellite LAI3g data that support the findings from this study are available from Ranga B. Myneni (rmyneni@bu.edu) upon request. The 'LUE-FPAR3g' GPP data that support the findings of this study can be requested from W.K.S. (wksmith@email.arizona.edu), whereas the 'FluxNetG' GPP data can be obtained from Martin Jung (mjung@bgc-jena.mpg.de). The TRENDYv6 data that support the findings of this study are available from S.S.S. (s.a.sitch@exeter.ac.uk) upon reasonable request.

30. Zhu, Z. *et al.* Global data sets of vegetation Leaf Area Index (LAI) 3g and Fraction Of Photosynthetically Active Radiation (FPAR) 3g derived from Global Inventory

Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the period 1981 to 2011. *Rem. Sens.* **5**, 927–948, (2013).

31. Martens, B. *et al.* GLEAM v3: satellite-based land evaporation and root-zone soil moisture*. Geosci. Model Dev.* **10**, 1903–1925 (2017).
32. Bartholome, E., & Belward, A. S. GLC2000: A new approach to global land cover mapping from Earth observation data. *Int. J. Rem. Sens*. **26**, 1959–1977 (2005).
33. Jung, M. *et al*. Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *J. Geophys. Res. Biogeosci*. **116**, G00J07 (2011).
34. Smith,W. K. *et al*. Large divergence of satellite and Earth system model estimates of global terrestrial CO2 fertilization. Nat. Clim. Change 6, 306–310 (2016).
35. Walter, S. *et al*. Satellite chlorophyll fluorescence measurements reveal large-scale decoupling of photosynthesis and greenness dynamics in boreal evergreen forests. *Glob. Change Biol*. **22**, 2979–2996 (2016).
36. Wu, C. *et al*. Land surface phenology derived from normalized difference vegetation index (NDVI) at global FLUXNET sites. Agr. For. Met. 233, 171-182 (2017).
37. Jiang, C. *et al*. Inconsistencies of interannual variability and trends in long-term satellite leaf area index products**.** *Glob. Change Biol*. **23**, 4133–4146 (2017).
38. Myneni, R. B. *et al*. The interpretation of spectral vegetation indexes. *IEEE Trans. Geosci. Rem. Sens*. **33**, 481–486 (1995).

Extended Data

Extended Data Table Captions

Extended Data Table 1 | Comparison of specific process representation in the TRENDYv6 carbon cycle models with relevance to this study

Extended Data Figure Captions

Extended Data Figure 1 | Comparison of lagged productivity responses based on satellite greenness observations and in-situ estimates of carbon fluxes across selected FLUXNET sites. Panels (a-c), show site-specific correlations between spring temperature (T) and spring, summer and autumn satellite NDVI (x-axis) plotted over the corresponding site-specific correlations between spring T and spring, summer and autumn in-situ based GPP (y-axis). In panels (b) and (c), relationships shown are based on partial correlations (pr) between spring T and subsequent summer as well as autumn NDVI/GPP, whereby covarying effects of summer T and precipitation (panel b) and autumn T and precipitation (panel c) have been removed. For this comparison, satellite NDVI time series at 8 km (native) spatial resolution have been extracted for the 16 included FluxNet tower sites with at least 10 year data records (Panel d). (Partial) correlations are shown for two estimates of GPP: GPP-N (based on nighttime partitioning) and GPP-D (daytime partitioning). In the maps (panel e), approximate location and name of FLUXNET tower sites are shown, along with forest type (ENF: Evergeen Needleleaf Forest, DBF: Deciduous Broadleaf Forest, MF: Mixed Forest) and record length (in brackets). FLUXNET data are from the FLUXNET2015 Dataset (Tier 1).

Extended Data Figure 2 | Random forest analysis to explain the partial correlation pattern between annual spring temperature and summer satellite greenness for hemispheric and regional scales. Panel (a) shows ranked importance of a set of explanatory variables in a random forest (RF) model for the whole northern ecosystem study region encompassing all northern vegetated non-agricultural land north of 30°N (see Supplementary Information section 2 for details on explanatory variables used). Ranking is based on the highest increment in mean squared error (IncMSE) between the observed and RF-predicted correlation after permuting this explanatory variable. Panels (b-f) show individual conditional expectation (ICE) lines of the RF-predicted partial correlation between spring T and summer NDVI. They encapsulate response curves for the five most important explanatory variables based on the RF analysis. Lines and shaded bands reflect the mean (i.e. regional average response) and the percentile range (5% to 95%, i.e. grid cell level responses to environmental predictors) of ICE curves for the entire northern hemisphere study region (red), and for the focus regions Siberia (blue) and Western US (green), respectively (see Supplementary Information section 2).

Extended Data Figure 3 | Spatial pattern of lagged productivity response scenarios based on the individual carbon cycle models included in TRENDYv6. All pattern are based on monthly GPP over the period 1982-2011 using outputs from the ten TRENDYv6 models included in the analysis. The maps summarize direction of statistically significant ($P < 0.05$) correlation between annual spring temperature and spring, summer and autumn GPP, respectively. For details on classification scenarios and contour labels see Figure 2 in main text. Areas with no robust link between spring temperature and spring GPP (dark grey) and areas cultivated or managed (light grey) are also outlined.

Extended Data Figure 4 | Spatial pattern of lagged productivity and vegetation growth response scenarios based on satellite-constrained and modelling approaches. The six maps summarize direction of robust ($P < 0.05$) correlations between annual spring temperature and spring, summer and autumn (a) satellite NDVI, (b) satellite LAI, (c) satellite-constrained upscaled GPP (FluxNetG), (d) satellite-driven LUE-modelled GPP (LUE-FPAR3g) as well as multi-model mean (e) GPP and (f) LAI based on the ten TRENDYv6 models. For details on scenario classifications and contour labels see Figure 2 in main text.

Extended Data Figure 5 | Changes in regional climate, satellite greenness and plant carbon fluxes from observation-constrained and modelling approaches for years with warm and cold spring temperatures. The panels depict anomalies in regionally-averaged composited climate, NDVI and GPP for the focus regions (a-c) Western US and (d-f) Siberia relative to the study period 1982-2011. The anomalies are based on maximum composites of monthly means of the seven warmest and coldest spring years within the study period. The observation-constrained GPP anomalies shown here (panels c and f) stem from an upscaled FLUXNET product (FluxNetG), which combined GPP estimates from flux towers with climate and satellite greenness in a machine learning framework (see Methods). Shown are also the respective (a, d) climate and (b, e) NDVI anomalies for warm and cold spring years. Start/end of climatological spring, summer and autumn seasons are indicated (vertical grey dashed lines). Uncertainty bounds (shaded area) reflect the spread in the respective monthly anomalies within the compositing period (± 1 s.d., n=7). On the basis of these anomalies, we estimate for a warm spring year (relative to mean conditions) in Siberia (2.5 Mill km^2) annual GPP increases of 0.4 PgC and 1.7 PgC for FluxNetG and the TRENDYv6 ensemble (see Panel F), respectively, which suggests a roughly 4 times higher plant carbon uptake in the TRENDYv6 ensemble. This is to a large part (~64%) because of the overestimation of positive lagged effects in the TRENDYv6 models , but another significant factor (36%) is also the higher sensitivity of concurrent carbon uptake to spring warming in the TRENDYv6 models (compared to FluxNetG).

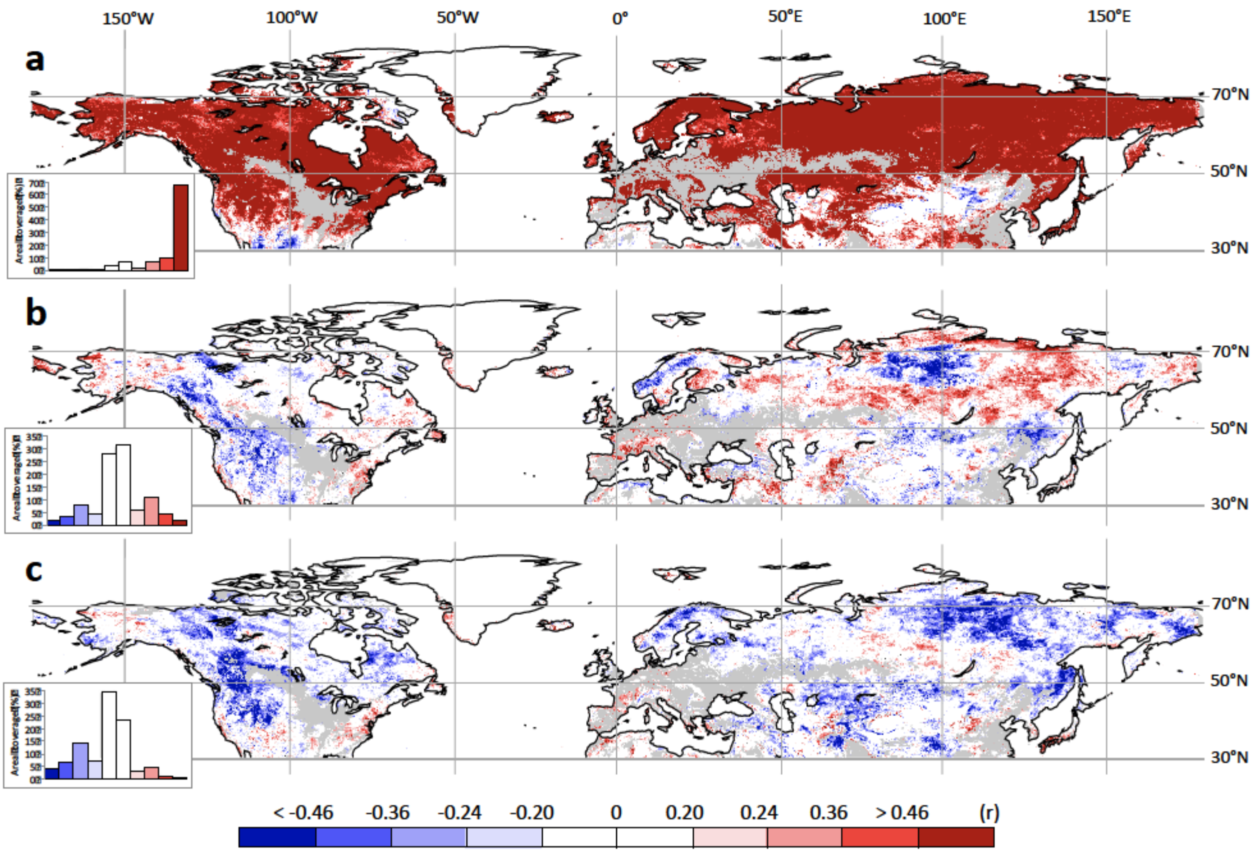
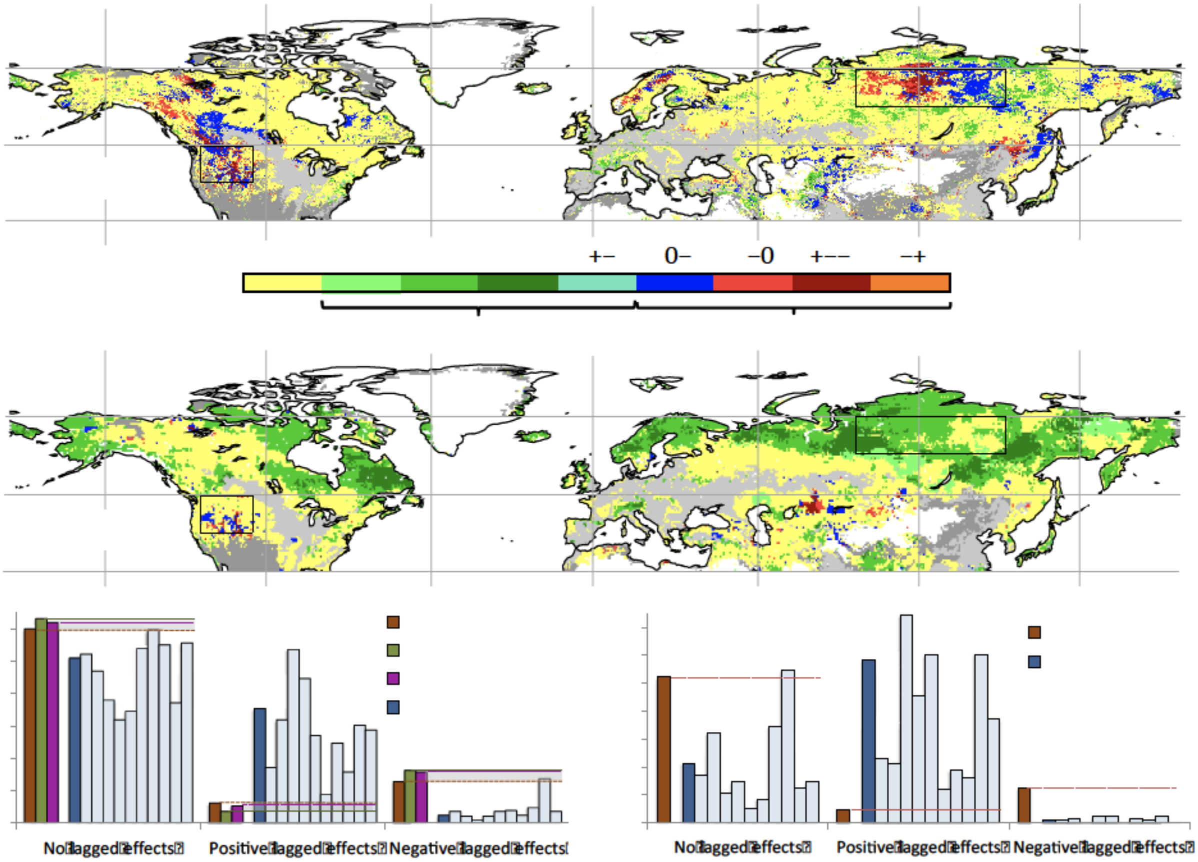

150°W
100°W
50°W
0°
50°E
100°E
150°E
a
b
c
70°N
50°N
30°N
Areal coverage (%)
< -0.46
-0.36
-0.24
-0.20
0
0.20
0.24
0.36
> 0.46
(r)

Figure 1



+−
0−
−0
+−−
−+
No lagged effects
Positive lagged effects
Negative lagged effects
No lagged effects
Positive lagged effects
Negative lagged effects