nature

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41586-018-0555-7

In the format provided by the authors and unedited.

Widespread seasonal compensation effects of spring warming on northern plant productivity

Wolfgang Buermann[1,2*], Matthias Forkel[3], Michael O'Sullivan[1], Stephen 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] & Andrew D. Richardson[19,20]

[1]Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds, UK. [2]Institute of the Environment and Sustainability, University of California, Los Angeles, Los Angeles, CA, USA. [3]Climate and Environmental Remote Sensing Group, Department for Geodesy and Geoinformation, TU Wien, Vienna, Austria. [4]College of Life and Environmental Sciences, University of Exeter, Exeter, UK. [5]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK. [6]CSIRO Oceans and Atmosphere, Canberra, Australian Capital Territory, Australia. [7]Department of Atmospheric Sciences, University of Illinois, Urbana, IL, USA. [8]Institute of Applied Energy, Tokyo, Japan. [9]Forest Research Institute Baden-Württemberg, 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, USA. [13]Max Planck Institute for Meteorology, Hamburg, Germany. [14]International Center for Climate and Global Change Research, School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL, USA. [15]Research Center for Eco-Environmental Sciences, State Key Laboratory of Urban and Regional Ecology, Chinese Academy of Sciences, Beijing, China. [16]Met Office Hadley Centre, Exeter, UK. [17]Laboratoire des Sciences du Climat et de l'Environnement, LSCE CEA-CNRS-UVSQ, Gif sur Yvette, France. [18]School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, USA. [19]School of Informatics, Computing and Cyber Systems, Northern Arizona University, Flagstaff, AZ, USA. [20]Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ, USA. *e-mail: wolfgang.buermann@geo.uni-augsburg.de

SUPPLEMENTARY INFORMATION (SI)

Section 1: Robustness of analysis framework

A) Confounding climate effects

For estimation of lagged plant productivity responses, our analysis framework relies on a correlative approach linking annual spring temperatures (as an independent phenological indicator) to subsequent annual summer and fall satellite greenness (or simulated GPP) as described in details in the Methods section in the main text. However, associations between spring temperature and subsequent plant productivity later in the season may also be a result of covarying climate over seasonal time scales. For example, warmer springs may also be associated with warmer summers, whereby the latter could then drive changes in summer greenness.

To check for such confounding effects in our analysis framework, we first estimated correlations between annual spring temperature and subsequent summer as well as autumn temperature and similarly between annual spring temperature and summer as well as autumn precipitation. Results show that in the case for spring temperature vs summer and autumn temperature robust correlations do exist at regional scales (Fig. S1A and C). For example, robust positive correlations are seen over large parts of northern North America and in central temperate Asia. The North American pattern resembles in part an ENSO pattern, which is well known to cause long-ranging effects on climate through atmospheric teleconnection pattern. In the case of summer and autumn precipitation, corresponding correlations with spring temperatures are generally much weaker and also tend to occur more at local scales (Fig. S1B and D).

In order to isolate these confounding effects of concurrent climate in the lagged correlations between spring temperature and satellite greenness (or GPP) during subsequent seasons, we opted to use partial correlation analysis in our framework (see Methods in main text). The overall influence of such effects are however relatively small: For example comparisons of patterns between spring temperature and summer greenness based on correlations and partial correlations (controlling for summer temperature and precipitation) are largely consistent (Fig. S2A and B).

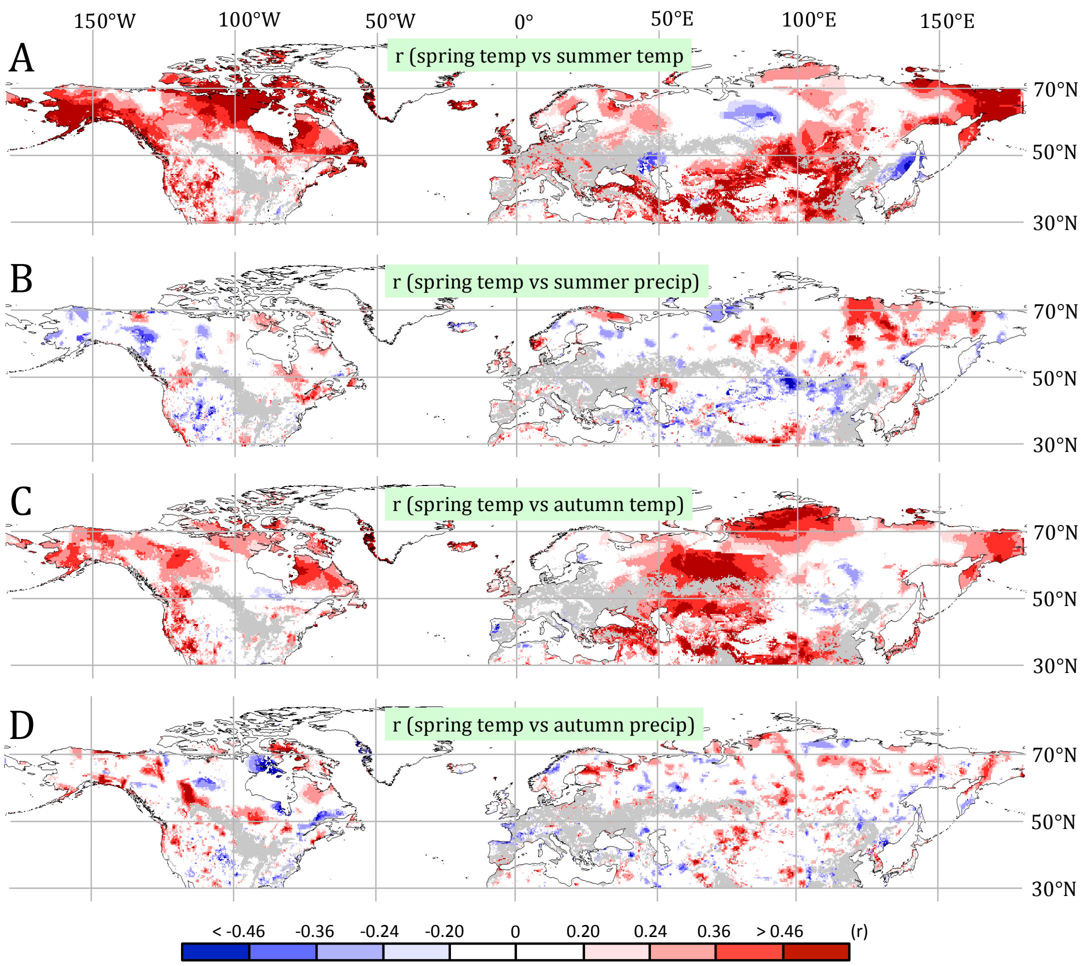

Figure S1: Spatial pattern of correlations between annual spring temperatures and subsequent annual summer and fall temperatures and precipitation. Pattern show grid-cell correlations between spring temperature and (A) summer temperature as well as (B) summer precipitations, and between spring temperature and (C) autumn temperature as well as (D) autumn precipitation. Correlations are based on the study period 1982-2011, and 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. Areas cultivated or managed (light grey) are not included in the analysis.

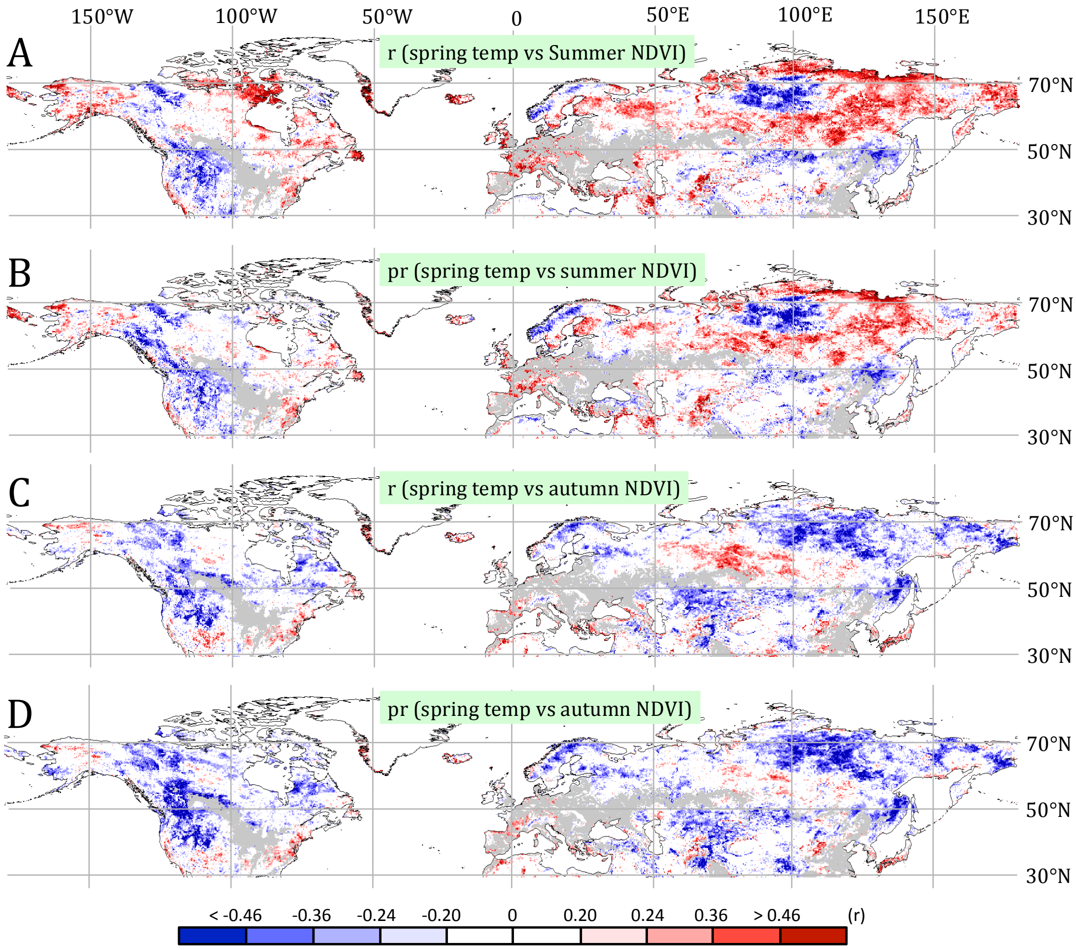

Figure S2: Spatial pattern of lagged productivity responses to spring warming based on satellite greenness observations. Panels (A) and (C) show correlations (r) between annual spring temperature and subsequent summer as well as autumn NDVI over the study period 1982-2011. Panels (B) and (D) show partial correlations (pr) between annual spring temperature and summer as well as autumn NDVI, whereby covarying effects of temperature and precipitation during summer ` (Panel B) and temperature and precipitation during autumn (Panel D) have been controlled for. Cultivated or managed areas (light grey) are not included in the analysis.

B) Analysis on detrended data

Satellite greenness long-term trends:

Satellite greenness data are known to exhibit long-term trends in part related to a response to a changing climate (e.g. ref. 3 in main text). However, such trends may also be a result of non-vegetation artifacts in the raw satellite measurements (see Methods in main text). Here we assess the degree by which such trends influence our correlation-based analysis through first detrending (removal of a linear trend) the original underlying data prior correlations analysis.

We find that the correlation pattern based on detrended data of spring temperature and seasonal NDVI (Fig. S3) are very similar to those based on original data (Fig. 1 in main text) which suggests a dominant influence of interannual to quasi-decadal variability on the correlation pattern between spring temperature and satellite greenness during subsequent seasons.

Trendyv6 GPP long-term trends:

Simulated GPP fluxes show generally considerable positive long-term trends over the satellite era largely because of the effect of CO_2 fertilization (ref. 21 in main text). To account for corresponding influences on our estimates of lagged effects, we again repeated the analysis with detrended productivity and climate time series. Results indicate an overall decrease (increase) in areal coverage of positive (negative) lags which suggests some influence of long-term trends in GPP on the underlying correlation pattern that form the basis for estimating the areal extent of lagged effects, but the diverging pattern between satellite- and model-based analyses in the areal extent of positive and negative lags still remains (Fig. S4 and Fig. 2 in main text).

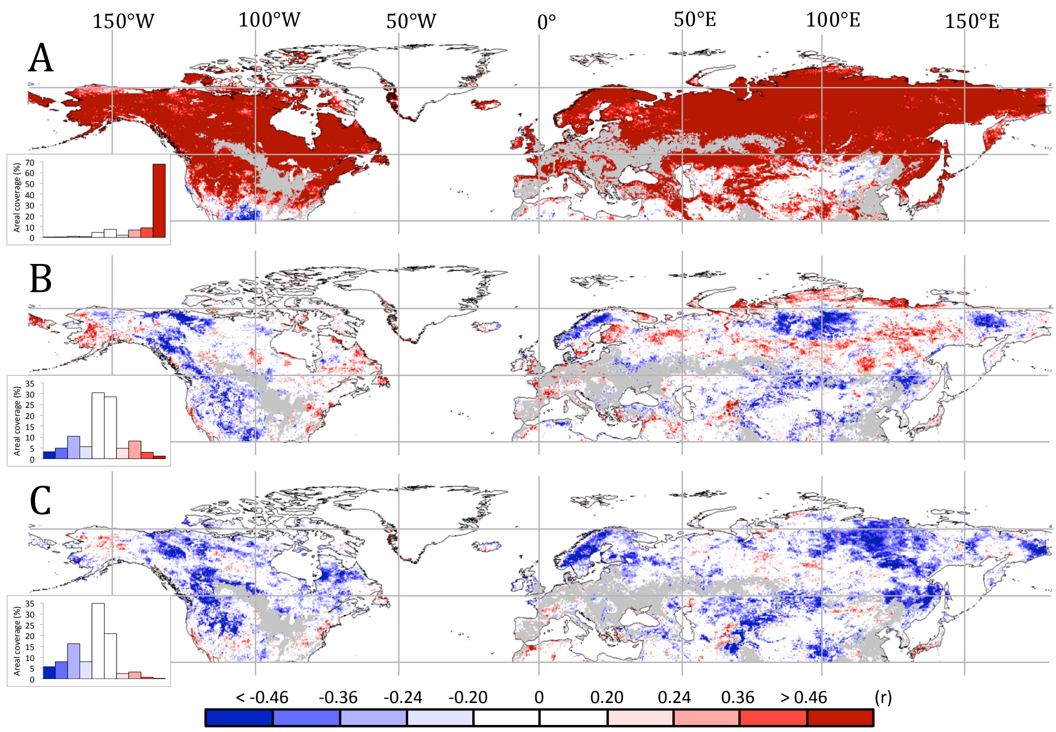

Figure S3: Spatial pattern of lagged productivity responses based on detrended satellite greenness observations. Panel (A) shows detrended (linear trends removed prior correlations) grid-cell correlations between yearly spring temperatures and spring satellite vegetation greenness (expressed through the NDVI, normalized difference vegetation index), respectively, for the study period 1982-2011. Panels (B) and (C) illustrate detrended partial correlations between annual spring temperatures and summer NDVI as well autumn NDVI, respectively. Here, the covarying influence of detrended summer temperatures and precipitation (Panel B) and autumn temperature and precipitation (Panel C) on the detrended spring temperature vs summer/autumn NDVI correlations has been removed. Seasons are defined through a local adaptive algorithm (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 are also provided (see inset panels). Areas cultivated or managed (light grey) are not included in the analysis.

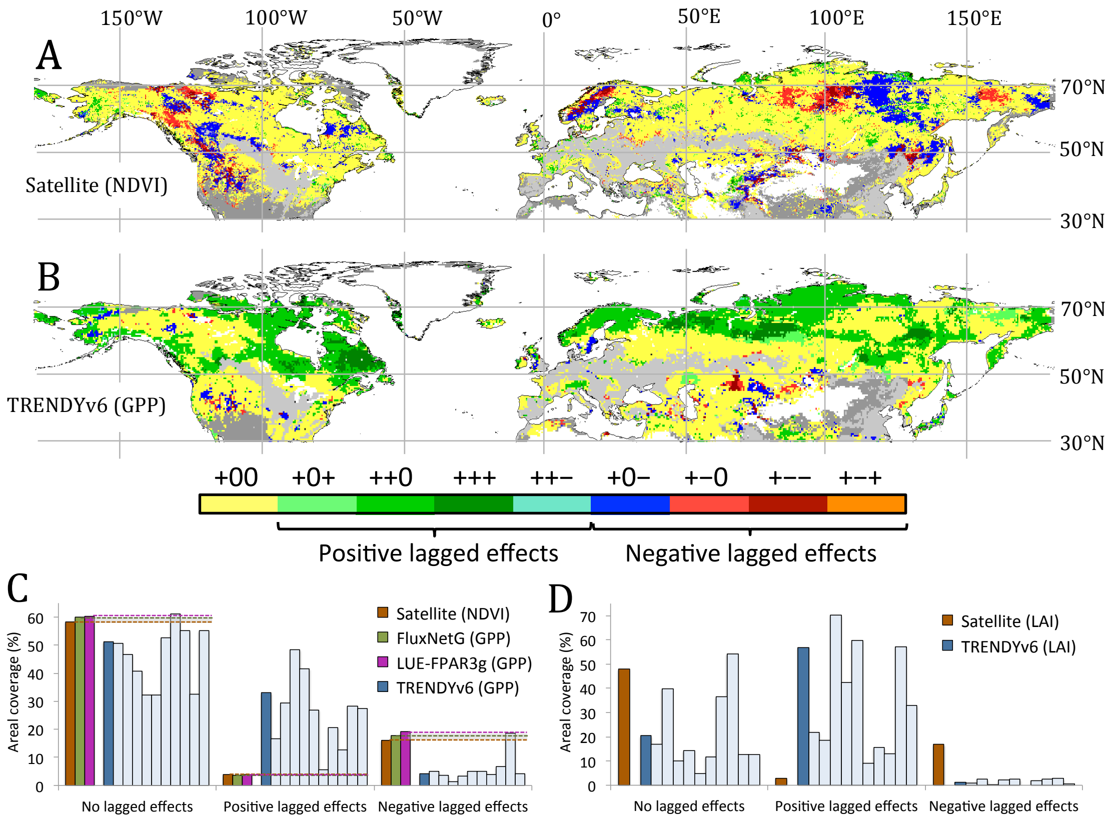

Figure S4: Spatial pattern of lagged productivity response scenarios based on detrended satellite observations and modelling approaches. In (A) pattern are based on satellite GIMMS NDVI3g data spanning 1982-2011. In (B) pattern are based on GPP from the TRENDYv6 ensemble spanning 1982-2011. The maps summarize direction of statistically significant ($P < 0.05$) correlation between detrended (linear trend removed) annual spring temperature and spring, summer and autumn NDVI as well as GPP, respectively. For details on utilized datasets and models, scenario classifications and contour labels see Figure 2 in main text.

C) Comparison of lagged and concurrent climate influences on greenness

In a complementary analysis, we also compared the importance of spring temperatures as a driver of changes in subsequent summer and autumn greenness to those related to concurrent climatic influences. Across northern land, robust ($P < 0.05$) partial correlations (including positive and negative correlations) between summer greenness and summer temperatures are spatially most widespread (26% areal proportions), followed by concurrent influences of summer precipitation (14%) and effects of preceding spring temperatures (12%) (Table S1 and Fig. S5). However, at regional scales we show here that the influence of preceding spring temperatures on summer and autumn greenness can be equally important or even dominant (Table S1 and Fig. S5).

Table S1: Comparison of preceding and direct climatic influences on variability in northern summer and autumn vegetation greenness. Table shows areal extent (%) of regions showing statistically significant ($P < 0.05$; sum of positive and negative) partial correlations between selected climate drivers and summer as well as autumn NDVI, respectively, for the study period 1982–2011. Estimates are provided for northern vegetated land (encompassing all vegetated non-agricultural land north of 30°N) as well as for two focal regions: Siberia (60–70°N; 80–125°E) and Western US (40–50°N; 120–105°W). The partial correlation maps on which these estimates of areal extents are based are shown in Fig. 1b and c in the main text as well in Fig. S5.

Climate driver	Northern vegetated land		Western US		Siberia	
	$NDVI_{Summer}$	$NDVI_{Autumn}$	$NDVI_{Summer}$	$NDVI_{Autumn}$	$NDVI_{Summer}$	$NDVI_{Autumn}$
$Temperature_{Spring}$	11.7	12.3	25.6	40.5	26.8	32.2
$Temperature_{Summmer}$	26.2	—	11.3	—	31.1	—
$Precipitation_{Summer}$	13.9	—	16.3	—	9.2	—
$Temperature_{Autumn}$	—	21.5	—	28.6	—	7.8
$Precipitation_{Autumn}$	—	19.7	—	20.4	—	28.2

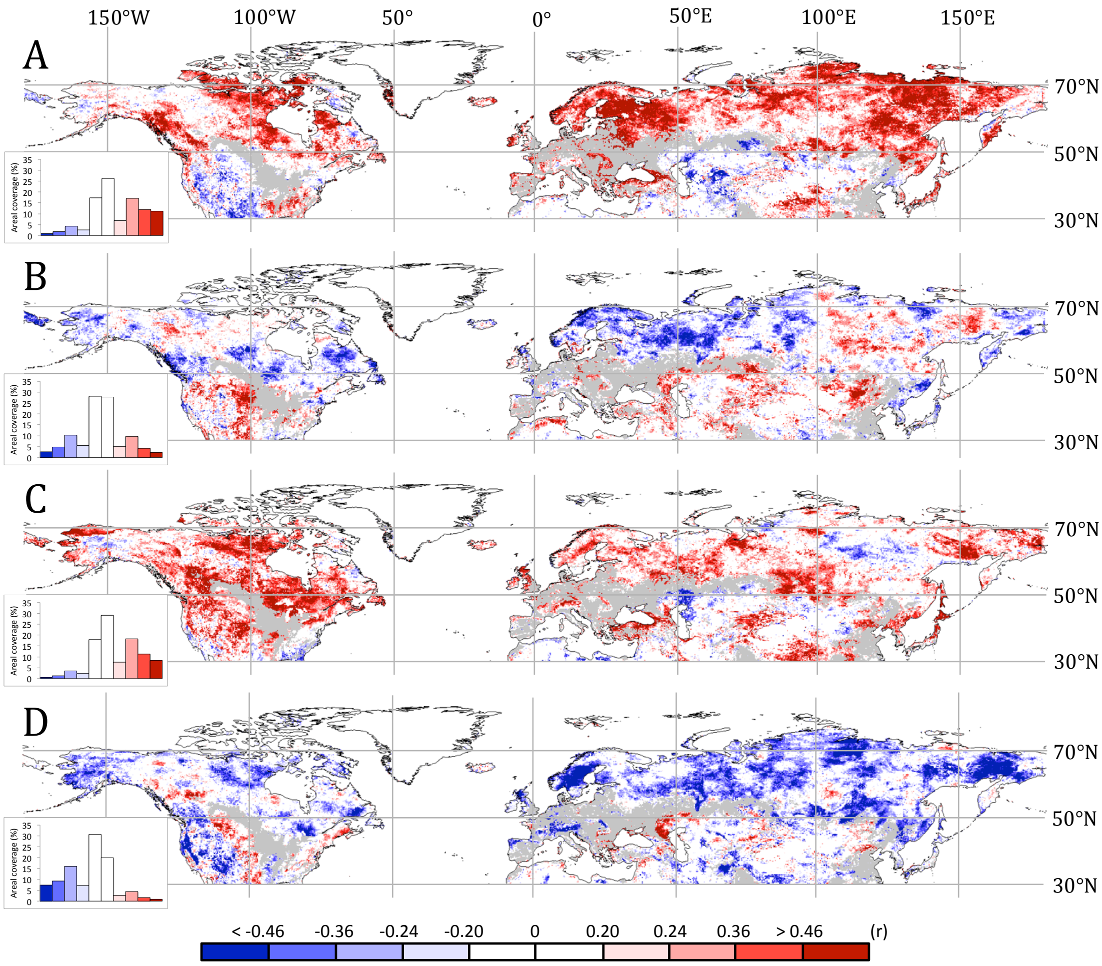

Figure S5: Partial correlations between changes in concurrent climate and satellite greenness during northern summer and autumn seasons. Panels show partial grid-cell correlations between annual (A) summer temperatures and summer NDVI as well as (B) summer precipitation and summer NDVI, respectively, for the study period 1982-2011. Also shown are the partial correlations between annual (C) autumn temperature and autumn NDVI as well as (D) autumn precipitation and autumn NDVI. Variables whose covarying effects on the correlation pattern are controlled for include spring temperatures (Panels A-D), as well as summer precipitation (panel A), summer temperature (Panel B), autumn precipitation (Panel C) and autumn temperatures (Panel D). Areas cultivated or managed (light grey) are not included in the analysis.

Negative correlations between changes in concurrent precipitation and NDVI may be interpreted as the effect of increased cloudiness and reduced radiation levels (which correlates with increasing precipitation) diminishing vegetation activity.

Section 2: Random forest analysis

We used a random forest machine learning approach to possibly explain the spatial patterns of the lagged correlations between and spring temperature and plant productivity later in the growing season. Here, we specifically aim to attribute underlying drivers of the partial correlation patterns between spring temperatures and summer NDVI (See Fig. 1b in main text), hereinafter called “correlation” or the response variable. Therefore, we computed a non-linear regression based on the random forest (RF) machine learning algorithm (Breiman 2001) to predict the spatial patterns of the correlation from a set of explanatory variables.

Explanatory variables are:

1. Elevation (SRTM);
2. The fractional coverage of the following land cover classes within 0.25° grid cells from the GLC2000 land cover map: sparse vegetation, evergreen shrubs, deciduous shrubs, mosaic vegetation, grasses, needle-leaf evergreen forests, needle-leaf deciduous forests, mixed forests, and broad-leaved deciduous forests;
3. And the following bioclimatic variables from the WorldClim 2.0 dataset (Fick & Hijmans 2017): temperature seasonality (TSEAS), annual temperature range (TRANGE), mean temperature in warm quarter (TMEAN.warmqua), mean temperature in cold quarter (TMEAN.coldqua), precipitation in warm quarter (P.warmqua), precipitation in cold quarter (P.coldqua), precipitation seasonality (PSEAS), annual precipitation (P), mean diurnal temperature range (MDTR), and mean annual temperature (MAT).

The RF was trained at 18275 randomly sampled 0.25° grid cells from northern ecosystems (excluding managed land cover). The RF model showed a very good performance in reproducing the correlation coefficient for these training grid cells ($R2 = 0.96$, $RMSE = 0.06$; Fig. S6). We further evaluated the performance of the trained RF. Therefore, we applied the trained RF model to all other grid cells in northern non-managed ecosystems ($n = 73103$), which resulted in good performances (R2 = 0.68, RMSE = 0.13; Fig. S6). These results demonstrate that bioclimatic indices, land cover, and elevation can explain the spatial patterns of the partial correlations between spring temperature and summer satellite greenness (NDVI).

We further assessed the importance of explanatory variables within the RF. Therefore we computed the increment in mean squared error (IncMSE) of the RF predictions after permuting each explanatory variable. The five most important variables were precipitation during the warmest quarter, precipitation seasonality, elevation, annual total precipitation and mean annual temperature (see Extended Data Fig. 2 in main text). Of medium importance were the other bioclimatic variables and the coverage of grasses. The coverage of all other land cover classes was of medium to low importance.

To identify the sensitivity of the correlation to the explanatory variables, we computed individual conditional expectations (ICE) from the random forest model (Goldstein et al. 2014). ICE curves are an approach to visualize the sensitivity of RF predictions to unique explanatory variables. To compute ICE curves, we sampled 0.25° grid cells from the observations whereby each grid cell has a specific combination of all explanatory variables. Each explanatory variable was then varied over its range (and all other variables were held constant) and the correlation was computed for each individual case from the fitted RF model. This results in a response function (ICE curve). ICE curves were computed for a random sample of all grid cells representing the hemispheric sensitivity (all northern non-managed vegetated ecosystems) and for sampled grid cells in the focus regions Siberia and Western US (see Fig. 2 in main text), respectively. We used the R software (version 3.4.2) with the RandomForest package (version 4.6-12) (Liaw & Wiener 2002) to compute RF models and the ICEbox package (version 1.1.2) (Goldstein *et al.* 2014) to compute ICE curves.

Extend Data Fig. 2 in the main text shows ICE curves for the five most important explanatory variables. At hemispheric scales (all northern non-agricultural land), correlations in respect to elevation are on average positive across all grid cells but become smaller towards higher elevation. Siberia shows in average increasing correlations from negative to close-to-zero values, i.e. the sensitivity of the correlation to elevation has an opposite direction than the hemispheric response or the response in western US. Correlations are on average higher (r = 0.1) at mean annual temperature below -12°C than above this temperature threshold. Sensitivities with precipitation differ in Siberia and Western US. Correlations decrease to negative values with increasing warm-quarter precipitation in Siberia but increase with increasing warm-quarter precipitation in Western US. A similar pattern but in opposite directions are observed for precipitation seasonality. These diverging pattern at regional scales in terms of directional dependencies of key drivers suggest region-specific peculiarities (e.g. annual rainfall and how it is distributed throughout the season) are important in determining the lagged productivity responses to spring warming.

Interpretation of ICE curves

As described, trained random forest (RF) models can be used to get insight into the shape of relationship between a predictor and the target variable and for illustrations, partial dependence (PDP) (Friedman 2001) and ICE (Goldstein *et al.* 2014) plots can be obtained from the RF model. PDPs show the partial relationship between the predicted target variable and one predictor variable when other predictor variables are set to their mean value (mean over all grid cells). ICE curves show the relationship between the predicted target variable and one predictor variable for individual cases of the predictor dataset (Goldstein *et al.* 2014); in our application, an individual case is a specific combination of bioclimatic, land cover, and elevation data for a given grid cell. The average of all ICE curves (from all grid cells) equals the PDP. The ICE plots shown in Extended Data Fig. 2B-F show the PDP (thick lines) and the quantile range (q0.05 - q0.95) of ICE curves (shaded areas). This means that the shaded areas are not uncertainties of the RF but reflect distributions of individual (grid cell-specific or local) response curves.

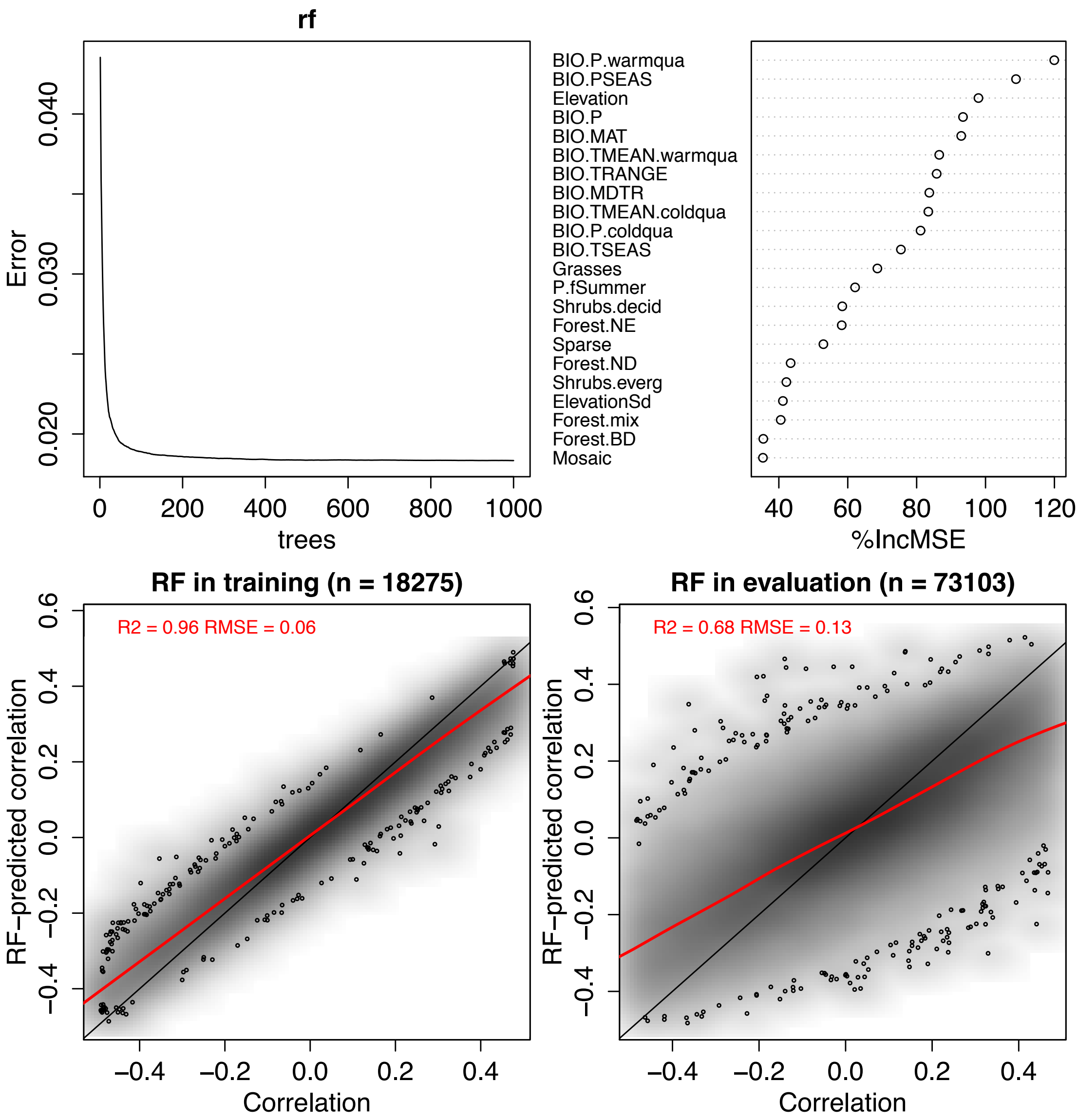

Figure S6: Performance of a random forest model to explain the partial correlation pattern between spring temperature and summer NDVI. Shown are the observed partial correlations of annual spring temperature and subsequent summer NDVI (x-axis) plotted against the predicted partial correlation (y-axis) as determined by the RF model. This validation scatterplot is based on 73171 grid cells (0.25° resolution) that were not used to train the RF model. High R^2 and low RMSE (see inset) suggest that the RF model is able to reproduce the spatial patterns of the partial correlations between annual spring temperature and summer NDVI.

Section 3: Regional analysis

To better understand why current-generation terrestrial carbon cycle models underestimate the spatial extent of negative lagged effects (and also overestimate the positive lagged effects) in respect to spring warming so profoundly, we performed a regional analysis for the Western USA and Siberia where the observation-based and simulated lagged productivity responses show more converging and diverging pattern (see Fig. 2 in main text). Results show that for Western US seasonal trajectories in aggregated satellite-based and modelled leaf area index (LAI; a key variable in the model's phenology schemes) 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 in main text). 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, whereas for ET seasonal trajectories the observation-based and modeled estimates for years with anomalous spring temperatures are more in agreement, although there is evidence that the models tend to underestimate water stress in summer in warm spring years (Fig. 3c and d in main text).

An analysis for Siberia at sub-regional scales provides more clear evidence for negative lagged effects linked to water stress: For warm spring years, both LAI and ET show negative anomalies mimicking the situation for the Western US. In addition, over eastern Siberia (which is covered predominantly by needleleaf deciduous forests) negative lagged effects may be potentially linked to fixed leaf life spans (Figure S7). This is inferred from the dominant pattern of positive spring temperature-spring greenness correlations coinciding with negative spring temperature-autumn greenness partial correlations (labelled as '+0–' lagged productivity response scenario in Fig. 2 in main text) and also supported by the absence of water stress during the peak growing season in warm spring years (Figure S7).

The Western USA and Siberia both are climatologically relatively dry and are vulnerable to summer drought (and reduced plant productivity) in response to warmer springs (see ref. 7 and refs. 27-29 in main text). The ability of the TRENDYv6 models to realistically simulate seasonal hydrological (ET) and vegetation growth (LAI) responses to spring warming for the Western USA, suggests that the model's hydrology and phenology schemes are generally fit for purpose. The strong divergence between observation-based and modelled vegetation growth cycles over Siberia may suggest that the models substantially underestimate the effects of water stress on seasonal canopy development in these colder environments. In addition, the often relatively poor representation of the phenology (lack of fixed leaf life span in deciduous phenology schemes (ref. 18 in main text) and physiology of needleleaf deciduous forests in the current models (see Extended Data Table 1 in main text) may also be an important factor for explaining differences in the lagged productivity responses between models and observations at regional levels.

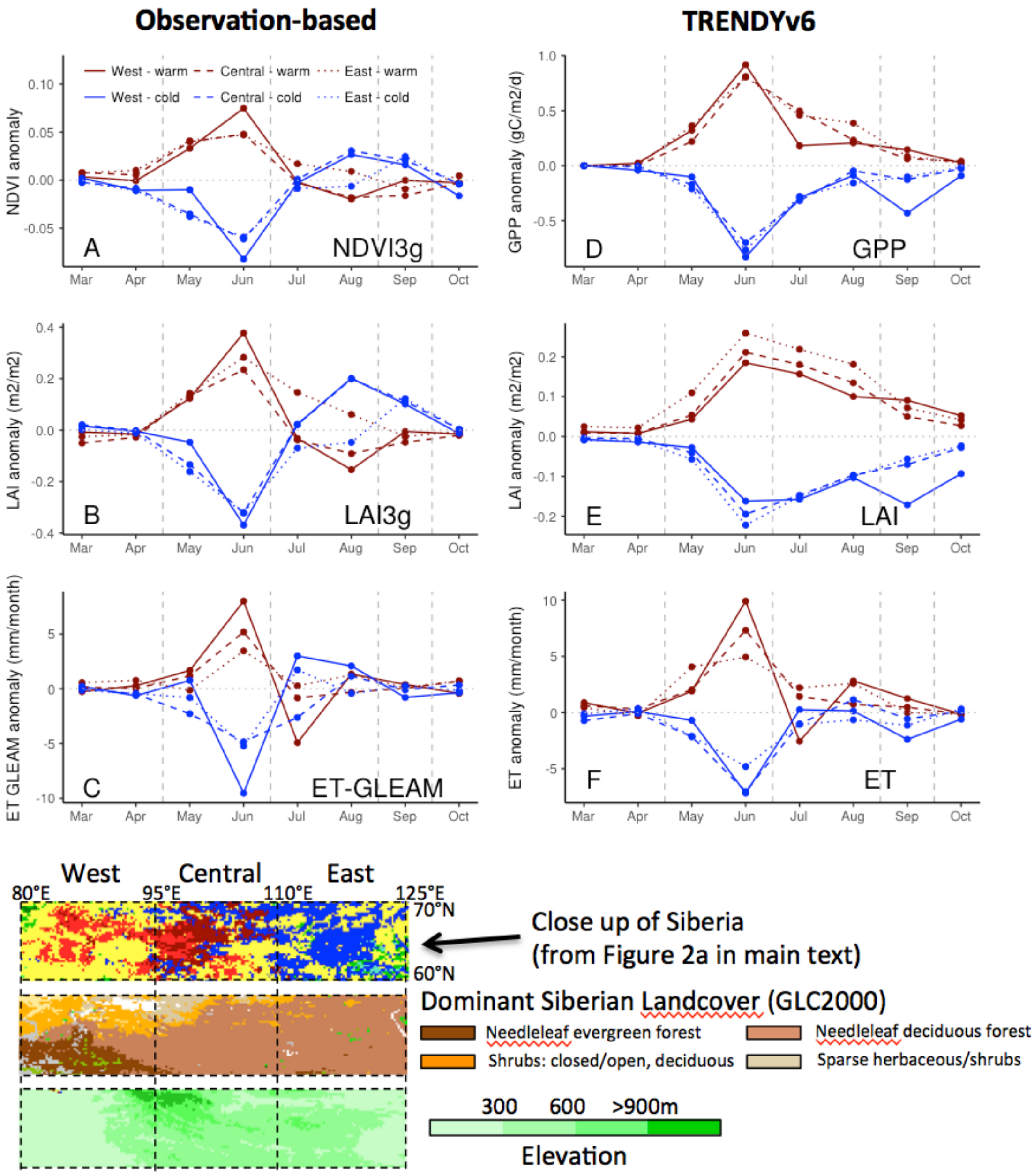

Figure S7: Regionally averaged anomalies in plant carbon fluxes and biophysical variables based on satellite observations and modelling approaches for years with warm and cold spring temperatures. The panels depict anomalies in composited variables for West, Central and East Siberia (see bottom panel) relative to the study period 1982-2011 based on (A-C) observation-constrained approaches and the (D-F) TRENDYv6 multi-model mean. Maximum composites represent monthly means of the seven warmest (red lines) and coldest (blue lines) spring years within the study period.

References:

Breiman, L. Random forests. *Mach. Learning* **45**, 5–32 (2001).

Fick, S. E. & Hijmans, R. J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol*. **37**, 4302–4315 (2017).

Friedman, J. Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics* **29**, 1189–1232 (2001).

Goldstein, A. *et al*. Peeking Inside the Black Box: Visualizing Statistical Learning with Plots of Individual Conditional Expectation*. J. Comput. Graph. Stats*. DOI: 10.1080/10618600.2014.907095 (2014).

Liaw, A. & Wiener, M. Classification and Regression by randomForest. *R News* **2**, 18–22 (2002).