Post-print of: Piao, S.L., et al. "Leaf onset in the northern hemisphere triggered by daytime temperature" in Nature communications (Nature publishing), april 2015. The final version is available at DOI: 10.1038/ncomms7911

Leaf onset in the northern hemisphere triggered by daytime temperature

Shilong Piao[1,2*], Jianguang Tan[2], Anping Chen[3], Yongshuo Fu[2,4], Philippe Ciais[5], Qiang Liu[2], Ivan A. Janssens[4], Sara Vicca[4], Zhenzhong Zeng[2], Su-Jong Jeong[6], Yue Li[2], Ranga B Myneni[7], Shushi Peng[2,5], Miaogen Shen[1], Josep Peñuelas[8,9]

[1] Key Laboratory of Alpine Ecology and Biodiversity, Institute of Tibetan Plateau Research, Center for Excellence in Tibetan Earth Science, Chinese Academy of Sciences, Beijing 100085, China

[2] Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing 100871, China

[3] Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544-1003, USA.

[4] Department of Biology, University of Antwerp, Universiteitsplein 1, 2610 Wilrijk, Belgium

[5] LSCE, UMR CEA-CNRS, Bat. 709, CE, L'Orme des Merisiers, F-91191 Gif-sur-Yvette, France

[6] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91011, USA.

[7] Department of Earth and Environment, Boston University, 675 Commonwealth Avenue, Boston, MA 02215, USA

[8] CREAF, Cerdanyola del Valles, Barcelona 08193, Catalonia, Spain.

[9] CSIC, Global Ecology Unit CREAF-CEAB-CSIC-UAB, Cerdanyola del Valles, Barcelona 08193, Catalonia, Spain.

Revised Manuscript for *Nature Communications*

February 26, 2015

* Author for correspondence: Shilong Piao (slpiao@pku.edu.cn)

Abstract

Recent warming significantly advanced leaf onset in the Northern Hemisphere. This signal cannot be accurately reproduced by current models parameterized by daily mean temperature (T_{mean}). Here, using in situ observations of leaf unfolding dates (LUD) in Europe and the United States, we show that the interannual anomalies of LUD during 1982-2011 are triggered by daytime (T_{max}) more than by nighttime temperature (T_{min}). Furthermore, an increase of 1^{o}C in T_{max} would advance LUD by 4.7 days in Europe and 4.3 days in the United States, more than the conventional temperature sensitivity estimated from T_{mean}. The triggering role of T_{max}, rather than T_{min} or T_{mean} variable, is also supported by analysis of the large-scale patterns of satellite-derived vegetation green-up in spring in the Northern Hemisphere (>30^{o}N). Our results suggest a new conceptual framework of leaf onset using daytime temperature to improve the performance of phenology module in current earth system models.

Introduction

Phenology, the timing of periodic events in the life cycle of living organisms, is sensitive to climate[1-4]. Phenological changes induced by climate change can alter species interactions[5,6] and ecosystem functioning, resulting into changes in the carbon, water and energy balances and, hence, climatic feedbacks[7]. Data from satellite greenness indices, field observations, and atmospheric CO_2 observations all show a trend toward an earlier spring green-up for northern vegetation over the last decades, super-imposed on high inter-annual variability[1,2,8,9]. Spring temperature correlates well with this trend and with the interannual variability of spring

green-up[2, 8]. Mean temperature is the principal variable used by dynamic global vegetation models (DGVMs) for calculating leaf onset in temperature limited biomes. These models, however, simulate onset dates that have large systematic errors compared with the in situ and satellite observations[3, 10], suggesting limitations in their equations describing phenology.

In cold and temperate regions, plants generally require accumulation of a certain amount of heat to trigger spring leaf onset. Several studies also outline the need for plant to endure cold conditions during their dormancy, which defines chilling requirements[11, 12]. Yet, evidence for a widespread chilling requirement is thin, and statistical models without chilling can predict the leaf onset date. Growing degree days (GDD), the sum of daily mean temperature (T_{mean}) above a fixed threshold value, is a common surrogate for the accumulation of heat needed to unfold leaves[13]. Current phenological models that use daily mean temperature ignore potentially different responses of plants to daytime and nighttime warming (e.g., ref. 14, 15). In other words, if daytime and nighttime temperatures impact distinctly the heat requirement of GDD, statistical and conceptual models of leaf onset must carefully distinguish which temperature should be used. In addition, global warming is increasing nighttime temperatures more than daytime temperatures, which makes the use of mean daily temperature likely impractical for modeling phenology[16].

Here using vegetation green-up date (VGD) diagnosed from satellite observations and in situ observations of leaf unfolding dates (LUD) in Europe and the United States, we demonstrate that the interannual anomalies of the timing of leaf onset are triggered by daytime more than by nighttime temperature across the Northern Hemisphere.

Results

Evidence from in situ observation

We first compared daytime versus nighttime temperature accumulation for predicting in situ observations of leaf unfolding dates (LUD) in Europe and the United States (US) over the past 30 years (1982-2011). Twenty-four plant species from 2400 phenology sites in Europe and lilac (*Syringa* L.) shrubs from 35 phenology sites in the US were selected from the European Pan European Phenological Database (hereafter EU) and the USA National Phenology Network (hereafter US) (Supplementary Figure 1; see Methods), respectively. Temperature data included monthly averaged daily maximum (T_{max}) and minimum temperature (T_{min}) with a spatial resolution of 0.5° obtained from the Tyndall Centre Climate Research Unit (CRU TS 3.20; see Methods). Both precipitation and cloudiness were included in the partial correlation analyses, while other variables such as soil moisture and soil temperature were not included, since temperature co-varies with these variables.

Many lines of evidence show that spring leaf unfolding date is strongly correlated with temperature of the preceding months (preseason)[8]. Here we define the length of the preseason for T_{max} by L_{max} (and by L_{min} for T_{min}). The value of L_{max} is calculated for each site as the period before LUD for which the partial-correlation coefficient between LUD and T_{max} is maximized in absolute value (controlling for the effects of T_{min}, precipitation and cloudiness ; note that the correlation is negative; see Methods) (Supplementary Figure 2). L_{max} ranges from 0 to 3 months across most of the species-site-year combinations for both phenology datasets

(68% for EU and 83% for US) (Fig. 1a and 1d). For the EU network, the partial interannual correlation between LUD and T_{max} averaged during L_{max} is negative and significant (P<0.05) at 33% of the species-site combinations. By contrast the significantly negative partial interannual correlation between LUD and L_{max}-averaged T_{min} occurs at less than 8% of the species-site combinations, as limited as the significantly positive counterpart (Fig. 1b and 1c). Similarly, the partial interannual correlations between LUD and L_{max}-averaged T_{max} were found significantly negative at 54% of the US Lilac sites (n = 35) compared to only 14% if L_{max}-averaged T_{min} is used (Fig. 1e and 1f). Similar results were also found with L_{min}-averaged variables (Supplementary Figure 3; see Methods). This observation suggests a pre-dominant role of T_{max} rather than T_{min} in controlling the interannual variations of LUD in both EU and the US phenological in-situ data.

To further test the robustness of the results shown in Fig. 1, we performed the same analyses with climate data of weekly and biweekly resolution (see Methods). All analyses produced similar results as shown in Fig. 1 (Supplementary Figure 4-6), confirming the stronger relationship of LUD with daytime temperature rather than with nighttime temperature, which was not affected by the temporal resolution of the climate datasets. Furthermore, we also extended the analyses to include winter temperature as a predictor to account for chilling effects (see Methods). Here winter temperature was defined as the average T_{mean} during the period from the onset of the preceding dormancy (the time at which daily mean temperature falls below 0 °C, or the default date of 1 November in the year preceding LUD) to the beginning of the T_{max}-preseason. Including winter temperature did not

alter the conclusion that T_{max} is a stronger predictor of LUD than T_{min} (Supplementary Figure 7).

Evidence from satellite observation

In situ phenology observations cover only a small fraction of world's vegetation types, geographic ranges and climate gradients. To evaluate the generality of the in situ observed asymmetric temperature effects on leaf onset in Europe and parts of the US, we further analyzed the effects of daytime and nighttime temperature changes on satellite-derived vegetation green-up date (VGD) in the terrestrial Northern Hemisphere (> 30°N) over the past 30 years (1982-2011) (see Methods). VGD at 0.5×0.5 degree resolution was estimated from time series of the NDVI3g data set (1982-2011) developed by the Global Inventory Modeling and Mapping Studies (GIMMS) group (see Methods). Note the reported VGD here is the average value from four different VGD algorithms: Spline-Midpoint, Hants-Maximum, Polyfit-Maximum, and Timesat-SG (see Methods). Similar to the in situ observation results, satellite derived L_{max} ranged between 0 and 3 months across 76% of the study area (Fig. 2a and 2b), also in agreement with earlier findings[17]. Statistically significant (P<0.05) negative partial correlations between VGD and L_{max}-averaged T_{max} were found in 42% of the study area (Fig. 2c). In contrast, over the same preseason only 11% of the study area showed significantly negative partial correlations between VGD and preseason averaged T_{min}, mostly in temperate dry regions (Fig. 2d). Even when using the T_{min}-preseason, the negative partial correlation between T_{min} and VGD remained less prevailing (13% of the study area exhibited significant correlation coefficients, Supplementary Figure 8b) than that between T_{max} and

VGD (32% of the study area; Supplementary Figure 8a).

In addition, we tested the robustness of the satellite-derived results using the four different satellite-derived VGD algorithms instead of the mean VGD from all algorithms (Supplementary Figure 9), using different climatic data sets that had different time resolutions (Supplementary Figure 10-12) and taking chilling effects into account (Supplementary Figure 13). All the tests returned similar results. In particular, we analyzed data from individual meteorological stations (Supplementary Figure 14) thereby avoiding any potential bias of spatial extrapolation like in the gridded climatic data sets. Examining the relationship between VGD and weekly, biweekly or monthly T_{max} and T_{min} at the locations of the 2510 meteorological stations with more than 15 years of climatic data available for 1982-2011 (see Methods), we confirmed the statistically significant and negative partial correlations between VGD and T_{max} in the preceding 0-3 months at about 43% of the stations, against only 14-16% of the stations between VGD and T_{min} for the same period (Supplementary Figure 14). Additionally, to determine whether the temporal binning of the GIMMS NDVI3g could bias the results, we performed the same analysis using VGD estimated from MODIS NDVI (2000-2010) (Supplementary Figure 15). For comparison, results from GIMMS NDVI3g during the same period (2000-2010) were also presented in Supplementary Figure 15. The significantly negative correlation between VGD and preseason T_{max} was still unambiguously more prevailing (22% of area for GIMMS and 24% for MODIS) than that between VGD and T_{min} for the same period (8% of area for GIMMS and 12% for MODIS) although the results were not as apparent as that in Fig. 2 due to the shorter period in the MODIS NDVI time

series (Supplementary Figure 15).

Temperature sensitivity of spring phenology

The stronger relationship between LUD (VGD) and T_{max} compared to T_{min} suggests that T_{max} is a better indicator of spring phenology. We therefore calculated the sensitivity (linear regression slope) of both in situ observed LUD and satellite-derived VGD to preseason T_{max} (SVT_{max}) using multiple linear regressions in which LUD (VGD) is regressed against T_{max}, T_{min}, precipitation and cloudiness (see Methods). On average, an increase of $1^{o}C$ in T_{max} would advance LUD by 4.7 days in Europe and 4.3 days in the US (Fig. 3d and 3e) during 1982-2011. As for VGD, an increase in T_{max} of 1 °C was associated with a 3-day earlier VGD in the Northern Hemisphere during 1982-2011 (95% CI: -12.0 days $^{o}C^{-1}$ ~ 9.1days $^{o}C^{-1}$). The highest T_{max} sensitivities of VGD were observed in Europe, northern Siberia and northwestern Canada, where they exceeded -10 days $°C^{-1}$ in some regions (Fig. 3a). Estimates of the T_{max} sensitivity of VGD were robust across the different satellite-derived VGD algorithms used (Supplementary Figure 16). We note that the satellite-derived T_{max} sensitivity is not fully consistent with that derived from in situ observations at the same geographical locations (satellite pixel containing the site) in Europe (-8 ±7 vs. -5±9 days $°C^{-1}$) and in the US (-3±3 vs. -4±6 days $°C^{-1}$). This small discrepancy between satellite and in situ observations sensitivities may be due to their different temporal and spatial footprint. Compared with discrete, in situ data, GIMMS NDVI3g satellite observations provide more homogeneous phenological records over 8 by 8 km areas (including different species and sometimes different land cover types) [4].

Discussion

Our results show that both in situ-observed and satellite-derived spring leaf onset is more closely associated with the inter-annual variation of T_{max} than with T_{min}. Three potential mechanisms may account for these results. First, plants in temperate and boreal regions need a critical level of forcing temperature (e.g., growing degree days; GDD) to trigger spring phenology[13]. Only temperatures above this specific threshold (commonly set at 0 or 5 °C) count in GDD formation[11,18]. Before the onset of green-up, T_{min} is more likely to be below the threshold temperature than T_{max} and thus contribute less to fulfill the GDD requirement for green-up. This hypothesis is supported by the multi-year averaged values of T_{max} and T_{min} during L_{max} for VGD (Supplementary Figure 17, a and b). Averaged T_{max} was generally above 5 °C, while averaged T_{min} remained below 0 °C in many areas of the Northern Hemisphere. Hence, daytime rather than nighttime warming in spring fulfills more efficiently the GDD requirement that triggers leaf onset. Second, photoperiod may also co-regulate spring phenology[19]. For example, the synchronicity of the daily cycles of light and temperature in spring makes daytime temperature an important determinant of *Arabidopsis* phenology[20]. The combined effects of photoperiod and daytime temperature in early spring could, hence, contribute to the stronger relationship with T_{max}. Thirdly, since most plant photosynthesis occurs during the daytime but is suspended during the nighttime, daytime temperature rather than nighttime temperature would be more responsible for plant carbon fixation and energy capture and thus produces a stronger effect on the onset of green-up.

In temperate dry regions, by contrast, a weak negative or even a positive interannual correlation between T_{max} and VGD is observed, which may be related to spring phenology being delayed by T_{max} regulated water stress. It has been suggested that spring phenology of temperate grasslands is co-determined by soil water availability and temperature[21]. Daytime warming is observed to reduce soil water content by enhancing evaporation[14], which may partly or totally offset its advancing effect on VGD. On the other hand, significant negative correlations between T_{min} and VGD are observed in temperate dry regions (Fig. 2d), which could be partly attributed to decreased frost risk at higher nighttime temperature. It is also noted that in those areas the preseason average T_{min} is at 0°C or above (Supplementary Figure 17c) and thus T_{min} could contribute to fulfill the heat requirement for spring green-up. In addition, changes in plant community structure and composition in response to rising T_{min}[22] may also help explain this positive response of satellite derived VGD to T_{min} variations, which need to be further tested.

Daily mean temperature (T_{mean}) is currently used as the driver of spring phenology in models. Considering the unequal contribution of T_{max} versus T_{min} to spring leaf onset, models based on T_{mean} which includes an ineffective or less effective component of T_{min} may give questionable performance in analyzing the responses of spring phenology to temperature changes, considering the recent faster warming rate at night than at daytime. For example, in Europe and in the US, we found that the correlation between LUD and T_{mean} was weaker than that between LUD and T_{max} at more than 55% of the species-site combinations, although both correlations were significant for comparable percentages of species-site combinations

(Supplementary Figure 18). Similarly, VGD in northern hemisphere also shows a weaker correlation with T_{mean} than with T_{max} in 65% of the study area, suggesting that T_{max} outperforms T_{mean} as a predictor of spring leaf onset variation. Furthermore, the absolute value of the LUD sensitivity to T_{mean}, (e.g. ref. 3, 23), was smaller than its sensitivity to T_{max} in both Europe (-3.2 vs -4.7 days $^{o}C^{-1}$) and the US (-3.8 vs -4.3 days $^{o}C^{-1}$) (Fig. 3d and 3e). The higher LUD sensitivity to T_{max} (SVT_{max}) than to T_{mean} (SVT_{mean}) obtained from in situ observations is also corroborated by satellite observations in 60% of the study area, with spatial variation in the magnitude of the positive differences between SVT_{max} and SVT_{mean} (Fig. 3c). A larger SVT_{max} than SVT_{mean} is found for the area north of 50°N compared to south of 50°N (Fig. 3c). The largest positive differences between SVT_{max} and SVT_{mean} were observed in Eastern Europe and north-central Siberia, where SVT_{max} was one to two times larger than SVT_{mean}. On the other hand, regions where SVT_{max} is similar or even lower than SVT_{mean} are temperate dry ecosystems, where T_{min} rather than T_{max} is controlling the interannual variation of VGD as shown in Fig.2.

The findings suggest that spring phenology GDD models parameterized by daily mean temperature can be problematic in current vegetation models. To directly translate our findings into spring phenology predictions, we estimated and compared the changes of VGD under future climate and CO_2 scenarios using T_{mean}-based (as used in current vegetation models) and T_{max}-based (proposed by this study) GDD models (see methods). The scenarios include 24 climate models and 3 radiative forcing trajectories, RCP2.6, RCP4.5, and RCP8.5 (IPCC 2013). We found significant difference between T_{max}- and T_{mean}-based phenology

predictions. The advance of VGD caused by warming was larger in the T_{mean}-based prediction than in the T_{max}-based one for 85% of the NH across all climate scenarios (Fig. 4 and Supplementary Figure 19). This is because in all the climate models analyzed, the projected increase of T_{min}, and hence of T_{mean}, is faster than that of T_{max} (IPCC 2013). This result suggests that T_{mean} based GDD models may overestimate changes in leaf onset, highlighting the need to incorporate the asymmetric phenology effects of daytime and nighttime temperature changes in earth system models.

In summary, our results provide information that can be used to improve the performance of current phenological module in DGVMs. The statistical analyses presented in this study, however, require more information for the accumulation of triggering energy and acclimation mechanisms. In this study, most of the preseason nighttime temperature does not contribute to the accumulation of a critical heat amount required for triggering spring phenology in the mid- and high-latitudes of the Northern Hemisphere. The heat requirement for spring green-up calculated based on daily mean temperature becomes problematic when one wants to predict future phenology changes based on past heat requirement, given the asymmetric warming rates between daytime and nighttime. While here we used maximum daytime and minimum nighttime temperature, which are not the same as the average daytime and nighttime temperature, our work suggests that temperature accumulation for spring green-up calculated at finer temporal resolutions, such as hourly or every three hours, may be more appropriate. In addition, the impact of temperature on spring phenology has been found to be non-linear[24], which further adds to the difficulty in using a statistical relationship established

between current temperature and phenology to predict phenology under future climate scenarios. The non-linear impacts have been noticed and incorporated in early model development, such as the Spring Indices phenological models[25]. Finally, it should be noted that while daily weather can be a very random event, some synoptic-scale unusually warm daily events may also be critical in determining the timing of spring phenology[25,26]. The underlying mechanism through daytime and nighttime temperature affects spring phenology remains poorly understood in temperate and boreal ecosystems. Well-designed manipulation experiments therefore are needed to improve our understanding of the interaction between spring leaf unfolding phenology and daytime temperature, and ultimately result in more accurate simulations of spring phenology and better understanding of global carbon balance and ecosystem feedbacks to the ongoing climate change.

Methods

In situ observation data set

We used in situ observations of leaf unfolding date (LUD) from two independent phenology datasets. One is the Pan European Phenological Database (PEP725; http://www.pep725.eu/), which is an open-access database with long-term plant phenological observations from 19608 sites and 78 species across 25 European countries. This data set has been widely used for studying the relationships between spring phenology and climatic changes, especially global warming[8, 27]. To exclude potential biases caused by outliers and inadequate degrees of freedom, we removed species-site compositions with the dates of LUD later than June (180 DOY: day of the year) and focused on the sites with more than 15 years

records over the period 1982-2011. In total, 2400 phenological sites and 24 plant species from PEP725 were used in this study. We also used in situ phenology observations from the USA National Phenology Network (USA-NPN) (https://www.usanpn.org/results/data) [28], and only shrubs of the *Lilac* genus had sufficient station records for LUD. Similarly, after excluding the data with the dates of LUD later than June (180 DOY) or with less than 15 years of records for 1982-2011, we analyzed *Lilac* LUD data from 35 phenological sites in the US. The distribution of selected phenological stations is shown in Fig. S1. It should be noted that the first leafing date from USA-NPN was regarded as an equivalent of LUD here since USA-NPN does not include the exact phenological event of LUD as those defined in PEP725.

Satellite-derived date of onset of green-up

The temporal cycle of NDVI is an indicator of the seasonal growth of vegetation and can be used for investigating vegetation phenology over large regions[1,29,30]. The GIMMS NDVI3g data set (1982-2011) with a spatial resolution of 1/12° and a 15-day interval has been used to monitor the phenological cycle of ecosystems[2, 30]. Areas with sparse vegetation, i.e. multi-year NDVIs less than 0.1, were excluded from the analyses. Using NDVI3g data sets, we applied four methods (Spline-Midpoint, HANTS-Maximum, Polyfit-Maximum and Timesat-SG) to estimate the VGD. Detailed information about the four VGD deriving algorithms and the uncertainty in VGD estimation from the 15-day interval NDVI data set have been documented by White et al. (2009) and Fu et al. (2014). Follow the previous study[31], we applied a Bayesian constraint in each method to rule out the influence of snow cover and limit the VGD within the thermal growing season (5 day average temperature larger

than 0^{o}C). The average VGDs of the four algorithms were used in this study (Supplementary Figure 20), unless otherwise noted.

Climatic data

Monthly data for T_{max}, T_{min}, T_{mean}, precipitation and cloudiness were obtained from the Tyndall Centre Climate Research Unit (CRU TS 3.20) and are available for a regular 0.5° latitude/longitude grid for 1982-2011[32]. Due to the lack of solar-radiation data in the CRU data set during the study period, we used cloudiness data. To independently validate the results based on the CRU data set, we also used 0.5×0.5° latitude/longitude gridded 3-hourly climatic data applying WATCH Forcing Data Methodology to ERA-Interim data (WFDEI, 1982-2011)[33], a 3-hourly global meteorological forcing data set (1982-2008)[34] and the station-level global-surface summary of day product (GSOD) by the National Climatic Data Center (NCDC). Daily weather can be a very random event, and the satellite data are of biweekly resolution, so the climatic data from WFDEI, Sheffield and GSOD were then rescaled to weekly, biweekly and monthly resolutions. For the GSOD data, we only considered the 2510 stations with more than 15 years of available data for 1982-2011 and with NDVIs larger than 0.1 for the 0.5° latitude/longitude grids containing the stations. The data for short-wave radiation was obtained from WFDEI.

Analyses

We used partial-correlation analyses to explore the effects of T_{max} and T_{min} on observed LUDs. With this approach, we could exclude the confounding effects of other climatic

variables (precipitation and solar radiation) and of covariate effects between T_{max} and T_{min}[14]. Temperature during the pre-season dormancy period is arguably the most dominant factor for spring phenology[12], and current phenology models in most DGVMs are solely based on temperature. It is therefore the aim of this study to identify the most appropriate temperature variables for use in phenology models. However, other environmental drivers in addition to temperature, such as precipitation, can also be involved in controlling the complex vegetation seasonality. Hence, in exploring the temperature effect on spring phenology, we have excluded the confounding effects of precipitation and cloudiness (radiation) in the partial correlation analyses; while other variables such as soil moisture and soil temperature were not explicitly excluded, since temperature also indirectly influence them.

Spring phenological changes are highly associated with the temperatures in the preceding months[8]. To determine the length of the preseason whose average T_{max} had the largest influence on LUD, we calculated partial-correlation coefficients between LUD and mean T_{max} during the 0, 1, 2, 3 … k months preceding LUD, controlling for corresponding average T_{min}, accumulated precipitation and cloudiness (all variables non-detrended). The maximum k corresponded to the length of the period from the month of mean LUD (1982-2011) to the onset of preceding dormancy, defined as the month when the multi-year averaged mean temperature dropped to 0 °C, or November as a default value. The preceding months with the highest absolute partial-correlation coefficients were then considered as the T_{max}-derived "preseason", in which T_{max} had the largest influence on the timing of green-up. Similarly, by replacing T_{max} with T_{min}, we also obtained the T_{min}-derived preseason.

To assess the robustness of our results, we also used grided climatic data sets at different temporal resolutions instead of the CRU monthly climatic data sets. In addition, to determine if winter chilling affected the responses of LUD to T_{max} and T_{min}, we performed the same partial-correlation analysis with T_{max}, T_{min}, precipitation, cloudiness and winter temperature as independent variables. Winter temperature was defined as the average T_{mean} during the period from the onset of the preceding dormancy (the time at which daily mean temperature falls below 0 °C, or the default date of 1 November in the year preceding LUD) to the beginning of the T_{max}-preseason.

Our results indicated that the interannual variation in LUD was more strongly associated with changes in T_{max} than that in T_{min}, so we then only estimated the sensitivity of LUD to T_{max} based on multiple linear regressions with LUD as the dependent variable and T_{max}, T_{min}, precipitation and cloudiness as independent variables (all variables non-detrended). We used the monthly averaged values of each independent variable during the T_{max}-derived preseason. We estimated the sensitivity of LUD to T_{mean} based on the same multiple linear regressions but replacing T_{max} and T_{min} with T_{mean}. Accordingly, the monthly averaged values of each independent variable during the T_{mean}-derived preseason were used in this analysis.

The same partial-correlation and sensitivity analyses were applied to satellite derived observations, with preseason defined separately. To spatially match satellite data (1/12° spatial resolution) with climatic data (0.5° spatial resolution), we used averaged VGDs within each

grid of the climatic data set. Besides the same robustness tests as those in the species-site level analysis (see above), we also performed additional robustness tests by using VGDs derived from individual algorithms instead of the multi-method averaged VGD, using station-level climate dataset at different time resolution instead of CRU monthly climatic data sets, as well as using VGDs derived from MODIS NDVI instead of AVHRR NDVI estimated VGDs. The partial-correlation coefficients and temperature sensitivities derived from satellite and in situ observations were further compared for all pixels covered by both data sources.

Future prospects

In order to directly translate our findings into spring phenology predictions, we performed phenology prediction tests with T_{mean} (used in current DGVMs) and T_{max} (proposed by this study) approaches, respectively. First, we calculated the mean growing degree-days (GDD) requirement for each pixel over the period of 1991-2010 using both daily T_{max} and T_{mean} from WFDEI climate datasets (Supplementary Figure 21). The GDD requirement here is defined as an integration of temperature above $0^{o}C$ from 1 January to the satellite derived VGD of each year. Second, we applied these two mean GDD values (GDD_{Tmax}, GDD_{Tmean}) separately as the threshold to predict the VGD of each year over two periods, i.e., 1991-2010 and 2081-2100, using 24 climate models and 3 climate scenarios (RCP2.6, RCP4.5 and RCP8.5). For each climate model and RCP scenario, the difference between the mean VGD of the two periods ($Mean\_VGD_{2081-2100}$ minus $Mean\_VGD_{1991-2010}$) was then calculated for both T_{max}- and T_{mean}-based GDD models. Finally, under each RCP scenario, the mean values of those differences across all climate models were calculated for each vegetated pixel. For

comparison, we calculated the ratio of VGD changes predicted by GDD_{Tmax} to that predicted by GDD_{Tmean}. The spatial distribution of the ratios is shown in Supplementary Figure 19.

References

1. Schwartz, M. D., Ahas, R. & Aasa, A. Onset of spring starting earlier across the Northern Hemisphere. *Global Change Biol.* **12**, 343–51 (2006).

2. Barichivich, J. *et al.* Large-scale variations in the vegetation growing season and annual cycle of atmospheric CO_2 at high northern latitudes from 1950 to 2011. *Global Change Biol.* **19**, 3167-3183 (2013).

3. Richardson, A. D. *et al*. Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis. *Global Change Biol.* **18**, 566-584 (2012).

4. Cleland, E. E. *et al.* Shifting plant phenology in response to global change. *Trends Ecol. Evol.* **22**, 357-365 (2007).

5. CaraDonna, P. J., Iler, A. M. & Inouye D. W. Shifts in flowering phenology reshape a subalpine plant community. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4916–4921 (2014).

6. Tylianakis, J. M., Didham, R. K., Bascompte, J. & Wardle, D. A. Global change and species interactions in terrestrial ecosystems. *Ecol. Lett.* **11**, 1351–1363 (2008).

7. Richardson, A. D. *et al.* Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agricult. Forest Metero.* **169**, 156-173 (2013).

8. Menzel, A. *et al*. European phenological response to climate change matches the warming pattern. *Global Change Biol.* **12**, 1969-1976 (2006).

9. Keeling, R. F., Piper, S. C., & Heimann, M. Global and hemispheric CO_2 sinks deduced from changes in atmospheric O_2 concentration. *Nature* **381**, 218-221 (1996).

10. Kucharik, C. J. *et al.* A multiyear evaluation of a Dynamic Global Vegetation Model at three AmeriFlux forest sites: Vegetation structure, phenology, soil temperature, and CO_2 and H_2O vapor exchange. *Ecol. Model.* **196**, 1-31 (2006).

11. Harrington, C. A., Gould, P. J. & St. Clair, J. B. Modeling the effects of winter environment on dormancy release of Douglas-fir. *For. Ecol. Manage.* **259**, 798–808 (2010).

12. Hänninen, H. & Kramer, K. A framework for modelling the annual cycle of trees in boreal and temperate regions. *Silva Fenn.* **41**, 167-205 (2007).

13. Chuine, I. A unified model for budburst of trees. *J. Theor. Biol.* **207**, 337-347 (2000).

14. Peng, S. S. *et al.* Asymmetric effects of daytime and night-time warming on Northern Hemisphere vegetation. *Nature* **501**, 88-92 (2013).

15. Fu, Y. H. *et al*. Unexpected role of winter precipitation in determining heat requirement for spring vegetation green-up at northern middle and high latitudes. *Global Change Biol.* (2014), DOI:10.1111/gcb.12610.

16. IPCC, Climate Change 2013: The Physical Science Basis. Contribution of Working Group

I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (Cambridge University Press, Cambridge, 2013).

17. Jeong, S. J., Ho, C. H., Gimi, H. J. & Brown, M. E. Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008. *Global Change Biol.* **17,** 2385-2399 (2011).

18. Hänninen, H. Modelling bud dormancy release in trees from cool and temperate regions. *Acta For. Fenn.* **213**, 1- 47 (1990).

19. Körner, C. & Basler, D. Phenology Under Global Warming. *Science* **327**, 1461-1462 (2010).

20. Chew, Y. H. *et al*. An augmented Arabidopsis phenology model reveals seasonal temperature control of flowering time. *New Phytol.* **194**, 654-665 (2012).

21. Yu, F., Price, K. P., Ellis, J. & Shi, P. Response of seasonal vegetation development to climatic variations in eastern central Asia. *Remote Sens. Environ.* **87**, 42–54 (2003).

22. Alward, R. D., Detling, J. K. & Milchunas, D. G. Grassland Vegetation Changes and Nocturnal Global Warming. *Science* **283**, 229-231 (1999).

23. Wolkovic, E. M. *et al*. Warming experiments underpredict plant phenological responses to climate change. *Nature* **485,** 494-497 (2012).

24. Pope, K. S. *et al*. Detecting nonlinear response of spring phenology to climate change by Bayesian analysis. *Global Change Biol.* **19**, 1518-1525 (2013).

25. Schwartz, M. D., Ahas, R. & Aasa, A. Onset of Spring Starting Earlier Across the Northern Hemisphere. *Global Change Biol.* **12**, 343-351 (2006).

26. Schwartz, M. D. & Marotz, G. A. Synoptic Events and Spring Phenology. *Phys. Geogr.* **9**, 151-161 (1988).

27. Cook, B. I. *et al*. Sensitivity of spring phenology to warming across temporal and spatial climate gradients in two independent databases. *Ecosystems* **15**, 1283-1294 (2012).

28. Schwartz, M. D., Betancourt, J. L. & Weltzin, J. F. From Caprio's lilacs to the USA National Phenology Network. *Front. Ecol. Environ.* **10**, 324-327 (2012).

29. Piao, S. L. *et al*. Evidence for a weaking relationship between interannual temperature variability and northern vegetataion activity. *Nat. Comm.* doi:10.1038/ncomms6018 (2014).

30. White, M. A. *et al.* Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982-2006. *Global Change Biol.* **15**, 2335-2359 (2009).

31. Fu, Y.S. *et al.* Unexpected role of winter precipitation in determing heat requirement for spring vegetation green-up at northern middle and high latitudes. *Global Change Biol.* **20**, 3743-3755 (2014)

32. Mitchell, T. D. & Hulme, M. New, Climate data for political areas. *Area* **34**, 109-112 (2002).

33. Weedon, G. P. *et al*. The WFDEI meteorological forcing dataset: WATCH forcing data methodology applied to ERA-Interim reanalysis data. *Water Resour. Res.* (2014), DOI:10.1002/2014WR015638.

34. Sheffield, J. G. & Wood, E. F. Development of a 50-yr high-resolution global dataset of meteorological forcings for land surface modeling. *J. Climate* **19**, 3088-3111 (2006).

Acknowledgements

This study was supported by a Strategic Priority Research Program (B) of the Chinese Academy of Sciences (Grant No.XDB03030404), the National Basic Research Program of China (Grant No. 2013CB956303), National Natural Science Foundation of China (41125004), and the 111 Project (B14001).

Author contributions

S.L.P. designed research; J.T. and Q.L. performed analysis; S.L.P., J.T. and A.P.C. wrote the draft; and all authors contributed to the interpretation of the results and the writing of the paper. S.L.P. and J.T contributed equally to this work.

Author Information

Reprints and permissions information is available at www.nature.com/reprints. The authors have no competing financial interests. Correspondence and requests for materials should be addressed to S.L.P. (slpiao@pku.edu.cn)

1

2

Figure Legends

Figure 1. Responses of in situ observed leaf unfolding dates to T_{max} and T_{min} in Europe and the US during 1982-2011. The frequency distributions of the length (in months) of T_{max}_preseason in (**a**) Europe and (**d**) the US are shown. The T_{max}_preseason is defined as the period with the highest negative partial correlation between LUD and averaged T_{max} for the months preceding LUD. Frequency distributions of the highest partial-correlation coefficients between leaf unfolding dates (LUD) and preseason T_{max} in (**b**) Europe and (**c**) the US after controlling for corresponding T_{min}, cloudiness and precipitation. Frequency distributions of partial-correlation coefficients between LUD and T_{min} in (**e**) Europe and (**f**) the US during the same preseason as in a after controlling for corresponding T_{max}, cloudiness and precipitation. Note that leaf unfolding dates of multiple species in Europe and only lilacs (*Syringa* L.) in the US were analyzed. The mean values of significant (P<0.05) partial-correlation coefficients across all phenological stations, the percentages of significantly negative partial correlations and the percentages of significantly positive partial correlations (in parentheses) are provided in (**b**), (**c**), (**e**) and (**f**).

Figure 2. The relationship of the satellite derived onset dates of vegetation green-up with T_{max} and T_{min} in the Northern Hemisphere during 1982-2011. (**a**) The spatial pattern of the length (in months) of the preseason defined as the period with the highest negative partial correlation between VGD and averaged T_{max} for the months preceding VGD. An example of the length of the "preseason" is also shown in the left part. (**b**) The frequency distribution of the length of the preseason shown in (**a**). (**c**) Partial-correlation coefficients (R) between

preseason-T_{max} and VGD after controlling for corresponding T_{min}, cloudiness and precipitation. (**d**) Partial-correlation coefficients (*R*) between T_{min} and VGD during the T_{max}-derived preseason after controlling for corresponding T_{max}, cloudiness and precipitation. The 1% and 5% significance levels of the partial correlations correspond to ±0.49 and ±0.38, respectively.

Figure 3. The sensitivity of LUD and VGD to T_{max} and T_{mean} during 1982-2011. The frequency distributions of the temperature sensitivity of LUD to (**a**) T_{max} (SVT_{max}) and (**b**) T_{mean} (SVT_{mean}) and (**c**) the ratio between SVT_{max} and SVT_{mean} (SVT_{max}/SVT_{mean}) in Europe and the US. In the right panel of figure, the spatial distributions of the sensitivity of VGD are shown (**d**) T_{max} (SVT_{max}) and (**e**) T_{mean} (SVT_{mean}) and (**f**) the ratio between SVT_{max} and SVT_{mean} (SVT_{max}/SVT_{mean}) in the Northern Hemisphere during 1982-2011.

Figure 4. The ratios of future VGD changes predicted by a T_{max}-based GDD concept model to that predicted by a T_{mean}-based GDD concept model. Both T_{mean}-based GDD approaches and T_{max}-based GDD models were applied to predict the VGD changes ($\triangle$VGD) between 1991-2010 and 2081-2100, using 24 climate models and different climate change scenarios (RCP2.6, RCP4.5 and RCP8.5). For each RCP, the T_{max}-based predictions and the T_{mean}-based predictions were averaged across all models and the distributions of their ratio (T_{max}-based predictions / T_{mean}-based predictions) are shown in (**a**), (**b**) and (**c**). The ratio smaller than 1 (blue bar) represents that the future VGD changes predicted by T_{mean}-based approaches are larger than those predicted by T_{max}-based approaches, and vice versa (red bar).

The percentage of ratios smaller than 1 and the percentage of ratios larger than 1 are both provided in (**a**), (**b**) and (**c**).

Figure 1

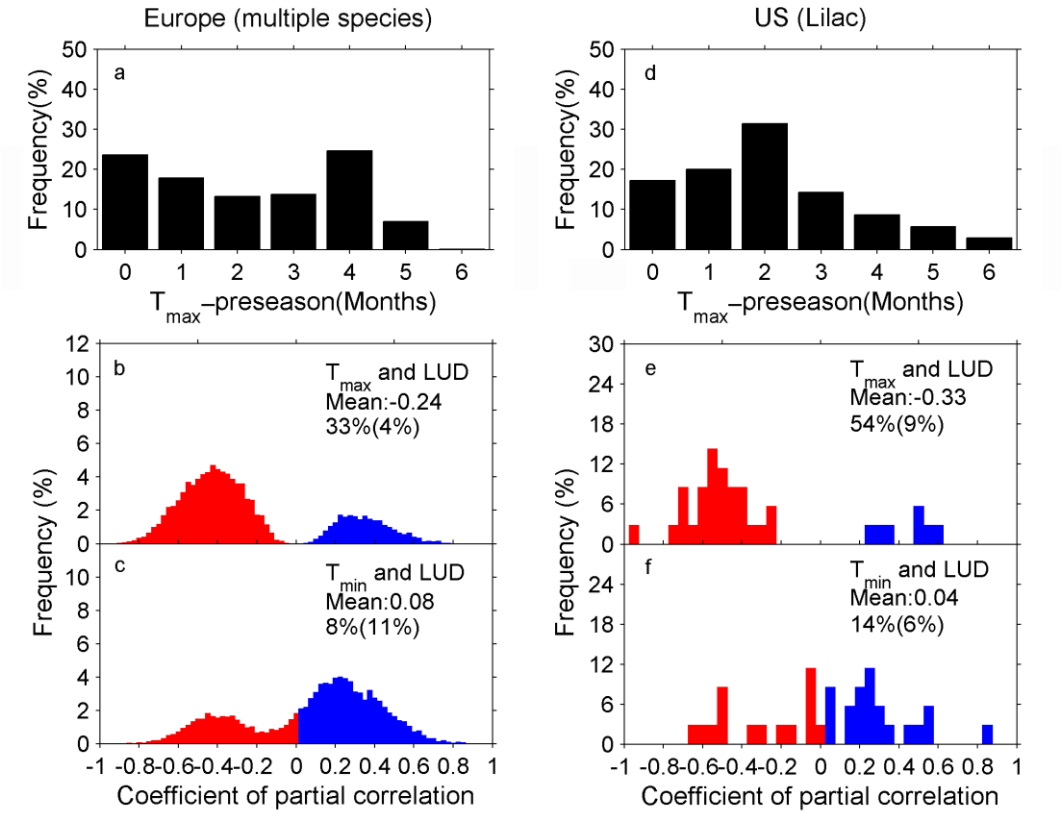

Europe (multiple species)
US (Lilac)
a
d
Frequency(%)
T_{max}–preseason(Months)
b
e
T_{max} and LUD
Mean:-0.24
33%(4%)
Mean:-0.33
54%(9%)
c
f
T_{min} and LUD
Mean:0.08
8%(11%)
Mean:0.04
14%(6%)
Frequency (%)
Coefficient of partial correlation

Figure 2

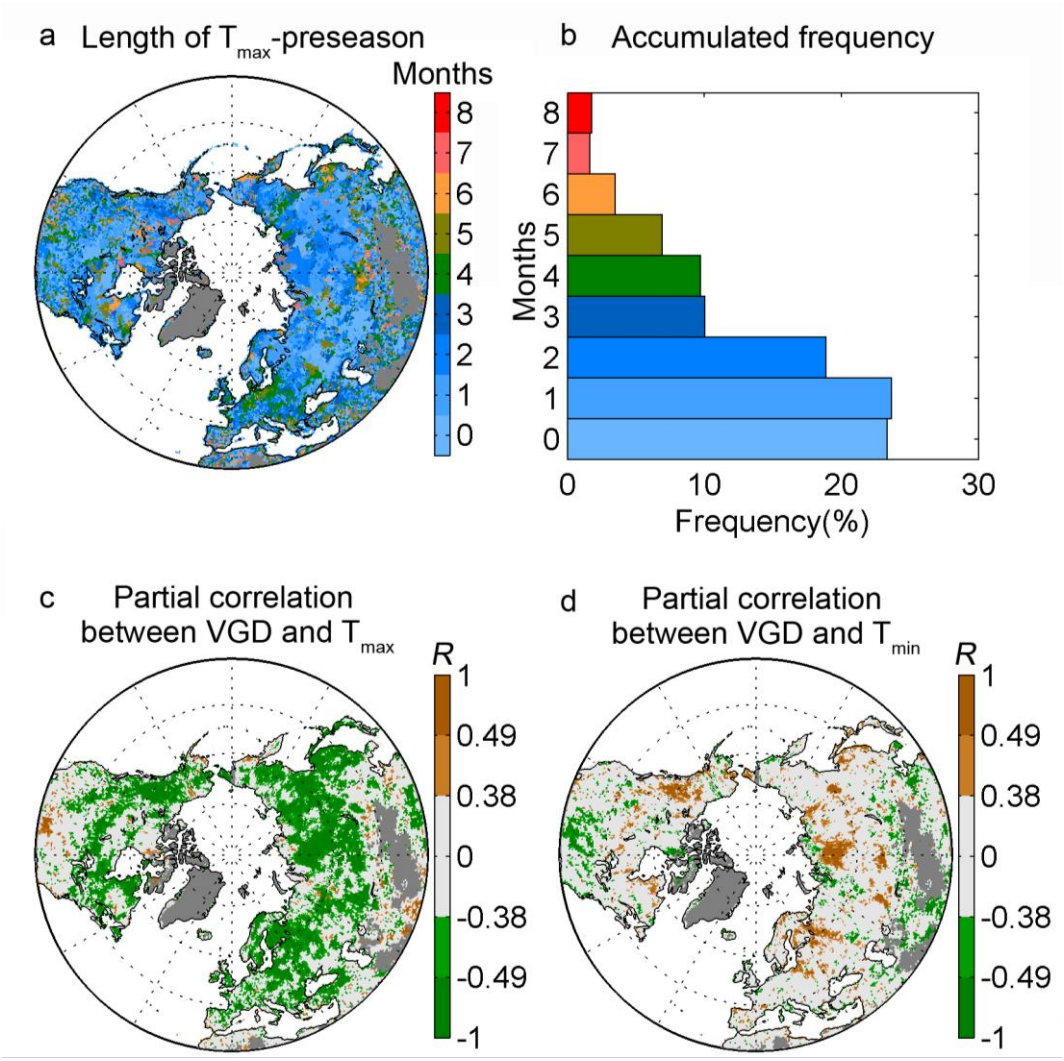

a Length of T$_{max}$-preseason
Months
8
7
6
5
4
3
2
1
0
b Accumulated frequency
Months
0
10
20
30
Frequency(%)
c Partial correlation between VGD and T$_{max}$
d Partial correlation between VGD and T$_{min}$
R
1
0.49
0.38
0
-0.38
-0.49
-1

Figure 3

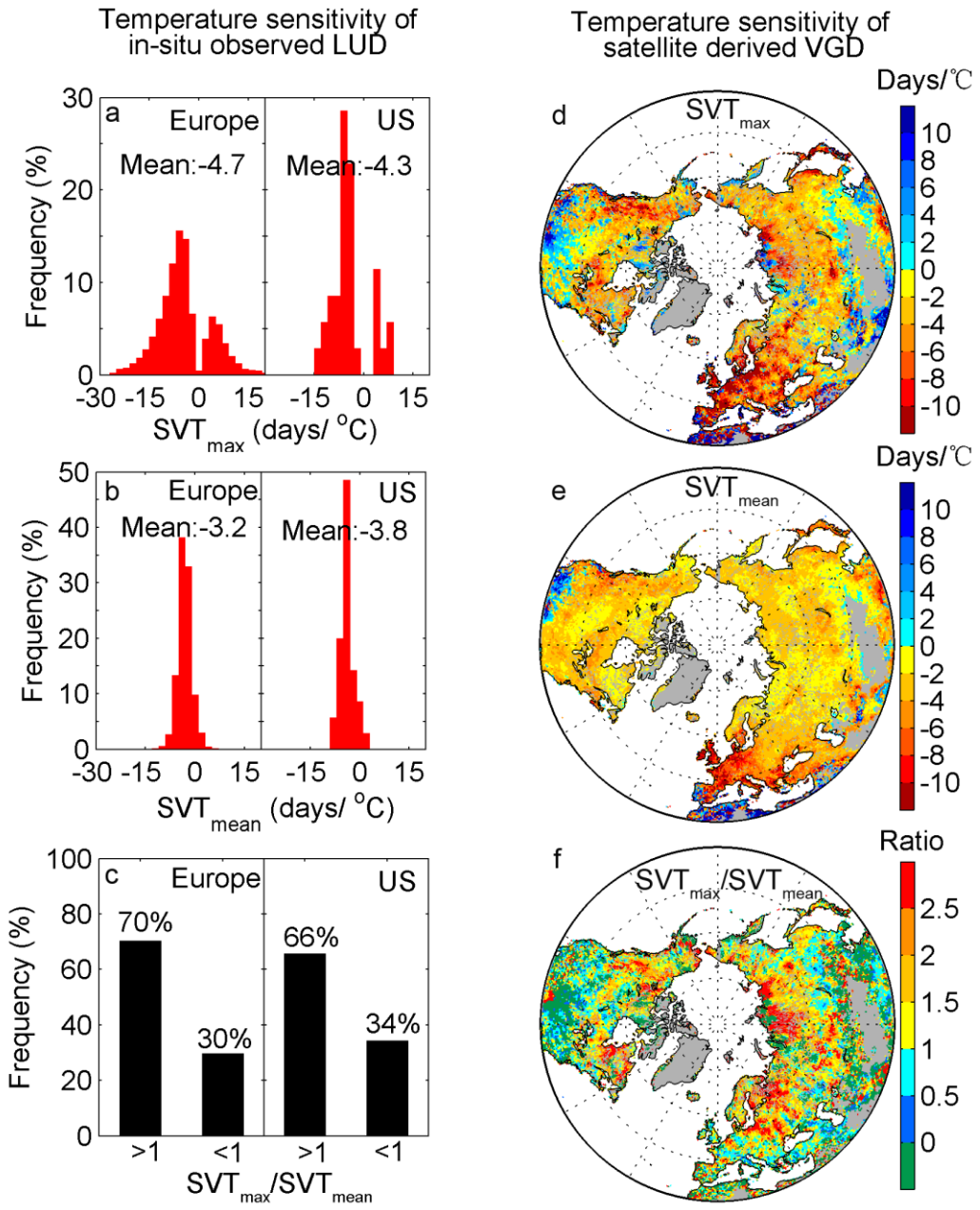
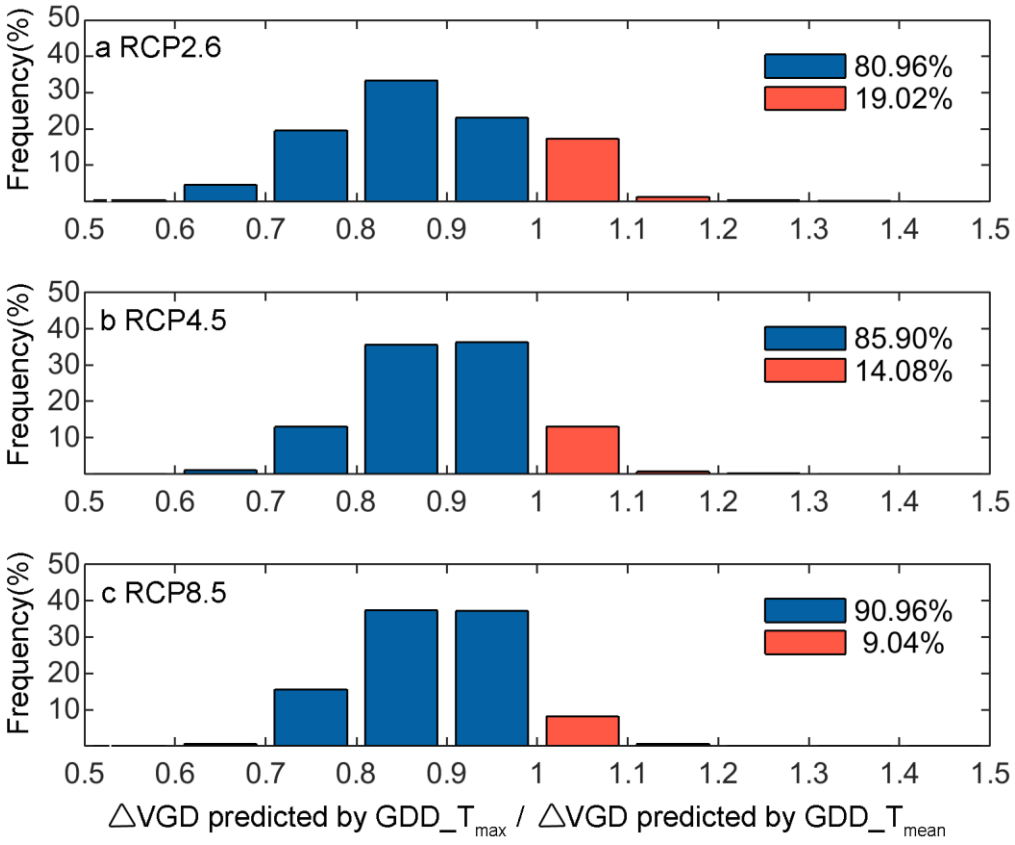

Temperature sensitivity of in-situ observed LUD
Temperature sensitivity of satellite derived VGD
a
Europe
US
Mean:-4.7
Mean:-4.3
Frequency (%)
SVT max (days/ °C)
b
Mean:-3.2
Mean:-3.8
SVT mean (days/ °C)
c
70%
30%
66%
34%
>1
<1
SVT max/SVT mean
d
SVT max
Days/℃
e
SVT mean
f
SVT max/SVT mean
Ratio

Figure 4


a RCP2.6
80.96%
19.02%
b RCP4.5
85.90%
14.08%
c RCP8.5
90.96%
9.04%
Frequency(%)
△VGD predicted by GDD_T$_{max}$ / △VGD predicted by GDD_T$_{mean}$