Open access • Journal Article • DOI:10.1111/GCB.13954

Simulating the onset of spring vegetation growth across the Northern Hemisphere — Source link

Qiang Liu, Yongshuo H. Fu, Yongshuo H. Fu, Yongwen Liu ...+2 more authors

Institutions: Peking University, University of Antwerp

Published on: 01 Mar 2018 - Global Change Biology (John Wiley & Sons, Ltd)

Topics: Phenology and Temperate rainforest

Simulating the onset of spring vegetation growth across the Northern Hemisphere

Reference:

Liu Qiang, Fu Yongshuo, Janssens Ivan, Piao Shilong, Liu Yongw en.- Simulating the onset of spring vegetation grow th across the Northern Hemisphere
Global change biology - ISSN 1354-1013 - 24:3(2018), p. 1342-1356
Full text (Publisher's DOI): https://doi.org/10.1111/GCB.13954
To cite this reference: https://hdl.handle.net/10067/1496880151162165141

Title: Simulating the onset of spring vegetation growth across the Northern Hemisphere

Running head: simulating spring vegetation growth onset

Qiang Liu[1], Yongshuo H. Fu[1,2], Yongwen Liu[1], Ivan A. Janssens[2], Shilong Piao[1*]

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

[2] Centre of Excellence PLECO (Plant and Vegetation Ecology), Department of Biology, University of Antwerp, Universiteitsplein 1, B-2610 Wilrijk, Belgium

Keywords: Chilling, climate change, model evaluation, photoperiod, remote sensing, spring phenology model

Revised Manuscript for *Global Change Biology*

***Corresponding author:** Shilong Piao, Tel: +86 10 6275 3298, Email: slpiao@pku.edu.cn

Paper type: Primary Research Articles

Abstract

Changes in the spring onset of vegetation growth in response to climate change can profoundly impact climate–biosphere interactions. Thus, robust simulation of spring onset is essential to accurately predict ecosystem responses and feedback to ongoing climate change. To date, the ability of vegetation phenology models to reproduce spatiotemporal patterns of spring onset at larger scales has not been thoroughly investigated. In this study, we took advantage of phenology observations via remote sensing to calibrate and evaluated six models, including both one-phase (considering only forcing temperatures) and two-phase (involving forcing, chilling, and photoperiod) models across the Northern Hemisphere between 1982 and 2012. Overall, we found that the model that integrated the photoperiod effect performed best at capturing spatiotemporal patterns of spring phenology in boreal and temperate forests. By contrast, all of the models performed poorly in simulating the onset of growth in grasslands. These results suggest that the photoperiod plays a role in controlling the onset of growth in most Northern Hemisphere forests, whereas other environmental factors (e.g., precipitation) should be considered when simulating the onset of growth in grasslands. We also found that the one-phase model performed as well as the two-phase models in boreal forests, which implies that the chilling requirement is probably fulfilled across most of the boreal zone. Conversely, two-phase models performed better in temperate forests than the one-phase model, suggesting that photoperiod and chilling play important roles in these temperate forests. Our results highlight the significance of including chilling and photoperiod effects in models of the spring onset of forest growth at large scales, and indicate that the consideration of additional drivers may be required for grasslands.

Introduction

The timing of spring vegetation growth onset has significantly advanced due to global warming over the past decades, and this has substantially impacted interactions between the atmosphere and terrestrial ecosystems (Walther *et al*., 2002; Menzel *et al*., 2006; Piao *et al*., 2007; Peñuelas & Filella, 2009; Friedl *et al*., 2014; Piao et al., 2017). Numerous phenology models, typically parameterized with ground-based phenological observations under ambient climatic conditions (Cannell & Smith, 1983; Hänninen, 1990; Kramer, 1994a; Chuine, 2000), have been developed and embedded into state-of-the-art land surface models (LSMs) to simulate the response and feedback of vegetation to climate change (Botta *et al*., 2000; Sitch *et al*., 2003; Krinner *et al*., 2005). However, a recent study reported that current vegetation phenology models generate considerable uncertainties in these LSMs, consequently leading to large biases in the estimation of ecosystem carbon balances (Richardson *et al*., 2012). Therefore, improved modeling of vegetation spring phenology at large scales is essential to reliably simulate the impact of climate change on vegetation growth and ecosystem carbon, water, and nutrient cycles.

Previous studies have reported the complex role of environmental cues in regulating the onset date of vegetation growth in spring. Temperature is generally considered as the main driver of spring onset (Cannell & Smith, 1983; Hänninen, 1990; Menzel *et al*., 2006; Piao *et al*., 2015), whereas photoperiod (Körner & Basler, 2010; Laube *et al*., 2014a; Chuine *et al*., 2016; Zohner *et al*., 2016), air humidity (Laube *et al*., 2014b), and precipitation (Forkel *et al*., 2015; Fu *et al*., 2014) co-determine vegetation growth onset dates for specific plant functional types and/or in specific regions. Based on these findings, various process-based models have been proposed to simulate the spring onset of growth. These models can be categorized into two types. One-phase models (e.g., the Spring Warming model) only involve the ecodormancy phase, and consider the day when accumulated forcing temperatures reach a required threshold (i.e., the heat requirement) to be the date of leaf unfolding (Sarvas, 1974). Two-phase models, such as the Sequential, Parallel, Unified, UniChill, and DORMPHOT models (Hänninen, 1990; Kramer, 1994b; Chuine, 2000; Caffarra *et al.*, 2011a), involve both the endodormancy and ecodormancy phases. These models assume that a certain amount of chilling is required to break endodormancy, and that a certain amount of heat must be accumulated to break ecodormancy and induce leaf unfolding (Cannell & Smith, 1983; Murray *et al.*, 1989; Myking & Heide, 1995). The DORMPHOT model also integrates the photoperiod effect and assumes that increased photoperiod (longer days) promotes the rate of forcing accumulation during the ecodormancy phase (Caffarra *et al.*, 2011a). Overall, one-phase and two-phase models differ in their assumptions regarding the control over spring phenology processes, and consequently predict widely diverging spring growth onset dates under future, warmer conditions.

Phenology models have generally been calibrated using ground-based records for specific plant species, and have provided reasonable simulations of local conditions and similar climatic regions (Hunter & Lechowicz, 1992; Chuine, 2000; Migliavacca *et al.*, 2011; Jeong *et al.*, 2013; Melaas *et al.*, 2013a). However, these locally parameterized models cannot simply be extrapolated to larger scales (Cleland *et al.*, 2007; Rayner, 2010). In contrast to ground-based phenology observations, remotely sensed satellite imagery provides spring phenology estimates at the landscape scale (White *et al.*, 2009; Jeong *et al.*, 2011; Wang *et al.*, 2015), bridging plant physiology and ecosystem processes as they relate to global change (Díaz & Cabido, 1997) at the same resolution as state-of-the-art LSMs. Therefore, satellite-based phenological observations are ideal for optimizing phenology models (Peñuelas & Filella, 2009) and for scaling up site-level phenological models to regional and even global scales. Several studies have investigated vegetation phenology models at the regional scale using satellite-derived estimates of the starting date of the growing season (SOS) (Yang *et al.*, 2012; Melaas *et al.*, 2015; Xin *et al.*, 2015; Chen *et al.*, 2016); however, these studies only considered forcing and/or chilling, ignoring the effects of photoperiod, whose role is expected to become increasingly important under future climate warming conditions (Fu *et al.*, 2015).

In the current study, we calibrated and evaluated six widely used spring phenology models, including one-phase and two-phase chilling and photoperiod models, using remote-sensing based SOS inferred from a newly released GIMMS (Global Inventory Modelling and Mapping Studies) NDVI (Normalized Difference Vegetation Index) dataset from 1982 to 2012. The primary goals of this study were (1) to calibrate different vegetation spring phenology models across the Northern Hemisphere; (2) to evaluate and identify an optimal model for specific plant functional types (PFTs) in reproducing spatiotemporal SOS patterns; and (3) to explore the chilling, forcing, and photoperiod effects in modeling vegetation growth onset for each PFT.

Materials and Methods

Datasets

Climate data and vegetation map. Daily temperature data with a spatial resolution of 0.5° × 0.5° was retrieved from the CRU-NCEP v5 climate dataset (available at ftp://nacp.ornl.gov/synthesis/2009/frescati/model_driver/cru_ncep/analysis/readme.htm), spanning the study period from 1982 to 2012. This dataset was generated from NCEP reanalysis and CRU-TS climatology data (New *et al*., 2000; Mitchell & Jones, 2005). Vegetation type data used in this study was adapted from a land cover dataset based on PFT (Verant *et al*., 2004; MacBean *et al*., 2015). We focused on natural vegetation across the Northern Hemisphere (latitudes exceeding 30° N) including three temperate forests, two boreal forests, and two grasslands. We did not take regions dominated by cropland into consideration, because cropland did not exhibit clear seasonality and was more susceptible to interruption by human activities. To eliminate the effects of any spatial mismatch between the two datasets, we remapped the PFT map into the same spatial resolution as the temperature data and calculated the fraction of PFT located within each 0.5° × 0.5° pixel. Finally, seven PFTs were chosen, including temperate needleaf evergreen (TeNE), temperate broadleaved evergreen (TeBE), temperate broadleaved deciduous (TeBD), boreal needleaf evergreen (BoNE), boreal broadleaved deciduous (BoND), natural C3 grass (NC3), and natural C4 grass (NC4) (Fig. S1).

Satellite-derived phenology. The seasonal cycle of NDVI serves as a proxy for vegetation greenness and photosynthetic activity (Myneni & Hall, 1995; Myneni *et al*., 1997), and is thus widely applied in the extraction of vegetation phenology at the regional and global scales (Buitenwerf *et al*., 2015; Garonna *et al*., 2015; Piao *et al*., 2015). In this study, we estimated the satellite-derived SOS from the latest release of the GIMMS NDVI datasets ($NDVI_{3g.v1}$) at a spatial resolution of 1/12° × 1/12° during the period from 1982 to 2012. Apart from the update of satellite sensors, atmospheric interference, and non-vegetation dynamics addressed in $NDVI_{3g}$ (Pinzon & Tucker, 2014), artifacts associated with snow coverage and changes in calibration that took place after 2006 were processed (Pinzon & Tucker, 2016). Following the methods of previous studies (Jakubauskas *et al.*, 2001; Zhang *et al.*, 2003; Piao *et al.*, 2006; Julien & Sobrino, 2009), we estimated SOS using four commonly used phenology extraction methods (i.e., Hants-Mr, Polyfit-Mr, double logistic, and piecewise logistic), which depend on either predefined thresholds or changing characteristics of the NDVI curve (Liu *et al.*, 2016). To reduce the uncertainty resulting from different interpretations of SOS from NDVI data, we applied an average SOS value derived from the four methods in the following analysis. As with the resample of the PFT map, we categorized all of the SOS data located in 0.5°× 0.5° pixels into specific PFTs, and then calculated the average for each PFT in each pixel.

Spring phenology models

Six spring phenology models, including the one-phase Spring Warming model and two-phase Sequential, Parallel, Unified, UniChill, and DORMPHOT models, were employed in this study. These models were calibrated on a daily basis, using gridded SOS and temperature data. The modeled SOS was defined as the day when the state of forcing (S_f) reached its critical value (F_{crit}). A detailed description of these models is provided below.

Spring Warming model. The one-phase spring warming model only considers the influence of forcing and calculates the accumulated daily rate of forcing (R_f) according to a logistic function (Eq.

1) (Sarvas, 1974) starting from t_0 (set as January 1 of the current year). Therefore, the SOS date is defined as the date when S_f exceeds F_{crit}:

$$S_f = \sum_{t_0}^{t} R_f = \sum_{t_0}^{t} \frac{A_f}{1 + e^{alpha(T - beta)}} \qquad (1)$$

Sequential model. The Sequential model assumes that the accumulation of forcing will not start until a critical threshold (C_{crit}) of chilling state (S_c, a daily sum of chilling rates) is met (Kramer, 1994b). In this model, the rate of chilling (R_c) is described using a triangular function (Hänninen, 1990) (Eq. 2) and S_c begins to accumulate after September 1 of the previous year, according to the methods of previous studies (Eq. 3) using a 1-day time step (Chuine, 2000). This model uses an R_f similar to that of the Spring Warming model, but with a competence function (K) (Eqs. 4 and 5). The forcing begins to accumulate when the daily temperature (T) is greater than T_d and C_{crit} is fulfilled:

$$R_c = \begin{cases} 0, T \le T_a \\ \dfrac{T - T_a}{T_b - T_a}, T_a < T \le T_b \\ \dfrac{T - T_c}{T_b - T_c}, T_b < T < T_c \\ 0, T \ge T_c \end{cases} \qquad (2)$$

$$S_c = \sum_{t_c}^{t} R_c \qquad (3)$$

$$K = \begin{cases} 1, S_c \ge C_{crit} \\ 0, S_c < C_{crit} \end{cases} \qquad (4)$$

$$R_f = \begin{cases} 0, T \le T_d \\ K \dfrac{Af}{1 + e^{alpha(T + beta)}}, T > T_d \end{cases} \qquad (5)$$

$$S_f = \sum_{t_c}^{t} R_f \qquad (6).$$

Parallel model. Unlike the Sequential model, the Parallel model assumes that forcing functions even when C_{crit} has not been attained (Landsberg, 1974); however, the parameters R_c, S_c, and R_f are exactly the same as in the Sequential model. The competence function (K, Eq. 7), however, introduces another parameter (K_{min}), determining the minimum potential of an unchilled bud to respond to the forcing temperature (Hänninen, 1990; Kramer, 1994b). Therefore, the state of chilling and forcing could increase simultaneously over time:

$$K = \begin{cases} K_{\min} + \dfrac{1-K_{\min}}{C_{crit}} S_c, S_c < C_{crit} \\ 1, S_c \geq C_{crit} \end{cases} \quad (7).$$

Unified model. The Unified model (Chuine, 2000) applies paired sigmoid functions to describe R_c and R_f. The accumulation of chilling (S_c, Eq. 8) starts at t_c (fixed on September 1 of the previous year) and continues even after C_{crit} is attained. Forcing begins after the chilling critical state, and the forcing requirement for leaf-out (i.e., the critical state of forcing, F_{crit}), is an experimental function of the total chilling accumulation (C_{tot}) (Eq. 11):

$$S_c = \sum_{t_c}^{t} R_c = \sum_{t_c}^{t} \frac{1}{1+e^{T_a(T-T_c)^2+T_b(T-T_c)}} \quad (8)$$

$$C_{tot} = \sum_{t_c}^{t_c+deltaP} R_c = \sum_{t_c}^{t_c+deltaP} \frac{1}{1+e^{T_a(T-T_c)^2+T_b(T-T_c)}} \quad (9)$$

$$S_f = \sum_{t_0}^{t} R_f = \sum_{t_0}^{t} \frac{1}{1+e^{alpha(T-T_d)}} \quad (10)$$

$$F_{crit} = we^{zC_{tot}} \quad (11).$$

UniChill model. The Unichill model is simplified from the Unified model (Chuine, 2000), and is similar to the Sequential model. The difference between the UniChill and Sequential models is the rate function incorporating chilling and forcing accumulation. The UniChill model uses the same R_c and R_f as the Unified model. The accumulation of chilling begins September 1 of the previous year (Eq. 12), and continues until C_{crit} is reached. Forcing (S_f) begins once C_{crit} is attained (Eq. 13) and leaf-out begins when F_{crit} is reached (i.e., SOS):

$$S_c = \sum_{t_c}^{t} R_c = \sum_{t_c}^{t} \frac{1}{1+e^{T_a(T-T_c)^2+T_b(T-T_c)}} \qquad (12)$$

$$S_f = \sum_{t_1}^{t} R_f = \sum_{t_1}^{t} \frac{1}{1+e^{alpha(T-beta)}} \qquad (13).$$

DORMPHOT model. The DORMPHOT model integrates the photoperiod effect to improve the simulation of budburst (Caffarra *et al.*, 2011a), which involves dormancy induction, dormancy release, and growth resumption. In this model, dormancy induction is triggered by both short photoperiods (DR_P) and low temperatures (DR_T), and is complete when the state of dormancy induction (DS, Eq. 14) reaches a critical value (D_{crit}):

$$DS = \sum_{t_0}^{t} DR_T \times DR_P = \sum_{t_0}^{t} \frac{1}{1+e^{aD(T-bD)}} \times \frac{1}{1+e^{10(DL-DL_{crit})}} \,(14),$$

where t_0 is the start date of dormancy induction (fixed to September 1 of the year preceding budburst), T is the daily mean temperature, and DL is day length on day t. Dormancy release and growth resumption then begin. As in the Unified model, S_c is defined as the daily sum of R_c after t_d (the day when dormancy induction is complete) as follows:

$$S_c = \sum_{t_d}^{t} R_c = \sum_{t_d}^{t} \frac{1}{1+e^{aC(T-cC)^2+(T-cC)}} \qquad (15).$$

By contrast, R_f is determined by both photoperiod and S_c. Long days above a certain threshold increases R_f (i.e., photoperiod sensitivity), which is in turn promoted by a longer previous chilling exposure or accumulation by decreasing this threshold (Caffarra *et al.*, 2011b). This mechanism is characterized by Equations 16–18. The mid-response temperature (T_{50}) decreases as DL increases, and results in an increase in R_f (Eq. 16). DL_{50} corresponds to the critical day length at which T_{50} is 30°C and decreases with higher S_c (Eqs. 17–18). This negative relationship is set by parameters gT ($\geq$ 0) and hDL ($\geq$ 0), respectively.

$$S_f = \sum_{t_d}^{t} R_f = \sum_{t_d}^{t} \frac{1}{1+e^{dF(T-T_{50})}} \quad (16)$$

$$T_{50} = \frac{60}{1+e^{gT(DL-DL_{50})}} \quad (17)$$

$$DL_{50} = \frac{24}{1+e^{hDL(S_c-C_{crit})}} \quad (18).$$

Model calibrations using satellite-derived SOS

The minimum root mean square error (RMSE, Eq. 19) was used as the criterion to calibrate the parameters of the models given in Table 1. Bayesian optimization techniques (Mockus, 2012; Martinez-Cantin, 2014) were applied to determine the minimum RMSE and the corresponding parameter set for each PFT. The model parameters were constrained within a range based on the literature, and a Gaussian process model of the objective function (i.e., RMSE between model simulations and satellite observations) was maintained to train the model. Then the next point to evaluate was determined by an acquisition function, which could balance sampling at points with low-modeled objective functions and help to find an optimal boundary value. To evaluate the robustness of these models, PFT-specific samples for model internal (90%) and external (10%) validation were randomly sampled from satellite-derived SOS. In addition to RMSE, other metrics were introduced to assess model performance for specific PFT. For example, the Akaike Information Criterion (AIC, see Eq. 20) trades off the goodness of fit and the complexity of the models, with lower AIC values indicating better model simulation (Akaike, 1998). Following Chen (2017), the Nash–Sutcliffe Efficiency index (NSE, Eq. 21) (Nash & Sutcliffe, 1970) was applied to quantitatively describe the reliability of the calibrated models. Positive NSE (varies from 0 to 1, higher = better) indicates that the model explained more variance than the null model (i.e., mean SOS):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(obs_i - pre_i\right)^2}{n}} \quad (19)$$

$$AIC = n \times \ln\left(\frac{\sum_{i=1}^{n}\left(obs_i - pre_i\right)^2}{n}\right) + 2\left(k+1\right) \quad (20)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(obs_i - pre_i\right)^2}{\sum_{i=1}^{n}\left(obs_i - \overline{obs}\right)^2} \quad (21),$$

where obs_i and pre_i are pixel–year observed and predicted SOS, respectively; $\overline{obs}$ is the mean value of observed SOS; n is the number of pixel–year SOS; k is the number of parameters in each model.

Results

Best model for each PFT

Table 1 summarizes the optimal parameters for each of the six models for each PFT, as well as the internal and external model evaluations. Overall, the models performed better than the null model (mean SOS) across all of the PFTs (i.e., NSE > 0). The percentage of variance explained by the models ranged from 24% (NC4, Parallel model) to 82% (NC3, DORMPHOT model), and varied greatly among the PFTs and models. Models generally performed better for the boreal forest than the temperate forest (Table 1, Figs. 1 and S2). For the boreal forest, although the one-phase spring warming model performed similarly to the two-phase models, the UniChill and DORMPHOT models performed slightly better in BoND (RMSE: approximately 9 days) and BoNE (RMSE: approximately 12 days) than the other models, according to the lowest AIC values (Table 1). In temperate forests, the two-phase models generally outperformed the one-phase model. The DORMPHOT model performed best for TeNE, TeBE, and TeBD (RMSE: approximately 14 days) (Table 1, Figs. 1 and S2). In the two grassland ecosystems, these models could not successfully reproduce the SOS dates, particularly in NC4, because the RMSE was greater than 21 days (Table 1, Figs. 1 and S2).

Evaluation of spring phenology models at the regional scale

The spatial patterns of SOS predicted by the six calibrated models during the period 1982–2012 were broadly similar to those of the satellite-derived SOS, with spatial correlation coefficients ranging from 0.91 to 0.94 (Fig. 2). Nonetheless, model predictions differed greatly from satellite observations in western America and in arid/semi-arid regions (e.g., central Eurasia, Fig. 2). Moreover, SOS generally occurred earlier in low latitudes but later in the middle and higher latitudes (Fig. 2a). We found that such latitude gradients in satellite-derived SOS were captured precisely by model simulations except by the Parallel and Unichill models, which showed slight disagreement in latitudes lower than 35°N (Fig. 2e and i). Within boreal forests, all six models were in quite close agreement, with less than 10 days of difference in long-term mean SOS in more than 68% of BoNE and 85% of BoND (Fig. S3). In temperate forests, the spring warming model displayed the largest differences, whereas the DORMPHOT model agreed best with the satellite observations (Fig. S3). The largest discrepancies between model predictions and observations were found in the C4 grasslands (Fig. S3). Across the Northern Hemisphere, all of the six models predicted SOS with a median RMSE of 7–15 days for forests and 10–18 days for grasslands (Fig. 3). Within forests, boreal forest (e.g., BoNE and BoND) generally exhibited the highest overall accuracy (with a median RMSE of < 10 days), whereas within grasslands, the spring phenology of C4 grasses was very poorly captured by the models (Fig. 3). In temperate forests, the one-phase Spring Warming model displayed poorer accuracy than the two-phase models involving chilling and photoperiod effects (mainly for TeNE and TeBE, Fig. 3a). The spatial distribution of RMSE indicated better agreement with satellite observations at latitudes higher than 45°N (e.g., RMSE < 10 days, except for the one-phase spring warming model), whereas higher RMSE values (e.g., > 15 days) were obtained in lower latitudes, such as western America and central Eurasia (Figs. 4 and S4). Across the Northern Hemisphere (Fig. 4g), the two-phase DORMPHOT, Unified, and Sequential models showed the best accuracy (with median RMSE = 9.3, 9.5, 9.8 days, respectively), and the poorest performance was found in the one-phase Spring Warming model (median RMSE = 11.9 days).

Evaluation of predicted SOS trends in the past three decades

During the period of 1982–2012, the modeled advancing trends in SOS in the Northern Hemisphere were obtained by the six models, which were in good agreement with the observations (Fig. 5). The model predictions captured SOS trends across approximately 74% (Parallel) to 78% (DORMPHOT) of the study area, although their magnitudes were weaker than the observations (Fig. 5). The differences in trends between satellite-derived SOS and model predictions were mainly distributed in northeastern Europe and southwestern North America (Fig. 5), which are primarily covered by TeBE and grasslands (Fig. S5). At the PFT scale, satellite observations documented a significant trend towards earlier SOS across all of the PFTs (except for NC4, $P = 0.67$) during the period of 1982–2012; however, the one-phase Spring Warming model generated non-significant changes in SOS, particularly in temperate forests, and the two-phase models showed good agreement with satellite observations in boreal, but not temperate, forest. Overall, the DORMPHOT model provided a better performance in trend simulations across most PFTs (Figs. 5, S5, and 6).

Discussion

Our study, covering a large spatial and temporal scale, revealed an overall median RMSE of 7–15 days for forests and 10–18 days for grasslands between the model predictions and satellite-derived SOS (Fig. 3). The overall accuracy of the simulations was comparable with previous efforts limited to either fewer PFTs or much shorter study periods. For example, the models fitted for TeBD showed similar accuracy to those of Melaas *et al*. (2015) for the eastern United States (median RMSE: 5–10 days), although in our study, model predictions for TeBD in southern Europe showed larger differences with satellite-derived SOS (Figs. 4 and S4). The models reproduced SOS less accurately for grasslands than for forests in this study (e.g., NC4 in central America, NC3 in western America) and in previous research (Xin *et al*., 2015), suggesting that other environmental factors besides spring forcing temperatures, chilling, and photoperiod should be considered in the phenology models. These factors may include soil moisture or precipitation, particularly in arid and semi-arid regions (Schwartz *et al*., 2006; Forkel *et al*., 2015). In terms of model-predicted SOS trends during the period of 1982–2012, advanced SOS trends were well captured among most PFTs (except for TeBE), with consistent signs but weaker magnitudes. These results may be related to the fact that current phenology models do not predict phenology dates well under extreme climate conditions, and that a few years' poor model predictions during the study period (e.g. 1983-1984, 2005-2006 for TeBE and 1982-1983 for TeBD for most models) can result in a large bias in the trend estimation (Figure S6). Moreover, we found that this issue was less likely related to the model calibration, because the ensemble means of four phenology-extraction methods did not reduce the temporal variation as shown in the SOS at PFT scale using individual method (Figure S7).

Previous studies parameterized spring phenology models with ground phenological records and then extended the ground-based models to regional and hemispherical scale (Schwartz *et al.,* 2006; Jeong e*t al.,* 2013; Yang *et al.,* 2012). However, large unclear in understanding the mechanisms underlying leaf-out process, especially scaling them from species level to ecosystem level, is still existed and needs to be explored. In this study, we found that the Spring Warming model, which only involves thermal forcing, performed as well as two-phase models in the boreal forests (particularly for BoND), which is consistent with previous phenological modeling studies (Hunter & Lechowicz, 1992; Vitasse *et al*., 2011). This may be because the boreal forests experience long and cold winters, with chilling requirements largely satisfied even under the ongoing climate warming conditions (Chuine, 2000; Vitasse *et al*., 2011; Xu & Chen, 2013). For temperate forests, however, a lower accuracy in predicting SOS was found in the Spring Warming model than in the two-phase models, implying that chilling is generally insufficient, and thus that chilling is an important factor in modeling spring phenology in temperate forests. This finding is also consistent with those of previous studies in temperate regions (Chuine *et al*., 2010; Fu *et al*., 2015). Importantly, we found that the DORMPHOT model, which incorporated the photoperiod effect, performed better in temperate deciduous and evergreen broadleaf forests, but yielded similar results in boreal or temperate evergreen needle-leaf forests. This finding is also supported by recent experimental studies that documented the fundamental but species-specific role of photoperiod in temperate tree species (Körner & Basler, 2010; Caffarra *et al*., 2011b; Zohner *et al*., 2016). Therefore, we conclude that the spring phenology of temperate forests might rely more on chilling and photoperiodic cues than boreal forest, which mainly depends on spring forcing temperatures under current climate conditions. With continued climate warming, however, chilling requirements might not be completely fulfilled in boreal forests, such that chilling and/or photoperiod might become more important in the spring onset process of these forests, although this speculation remains to be tested experimentally.

Forecasts of spring vegetation growth onset and its variation under climate change scenarios depend largely on model structure, thereby requiring a better understanding of the interactions with environmental cues across different PFTs (Richardson *et al.*, 2012). In this study, models integrating chilling and photoperiod generally produced better agreement in spatiotemporal patterns with satellite-derived SOS, although their RMSEs remained high. This result illuminates the potential to further improve current temperature-driven models by considering more environmental and physiological factors. Previous studies based on either *in situ* or satellite observations reported statistically significant correlations between spring growth onset and other environmental cues, such as soil water content or precipitation (Peñuelas *et al.*, 2004; Bernal *et al.*, 2011; Fu *et al.*, 2014), snowfall (Chen *et al.*, 2015), daytime (rather than daily mean) temperature (Piao *et al.*, 2015; Fu *et al.*, 2016) and successional status (Jeong *et al.*, 2013). Undoubtedly, these cues co-determine vegetation spring phenology; however, the mechanisms via which they influence the leaf-out process remain unclear. Further investigation through manipulative experiments at species and ecosystem level is necessary to make the models more biologically realistic, thus enhancing their predictive ability. Another possible source of uncertainty between satellite-derived SOS and model predictions may stem from the coarse spatial resolution of the observations applied in this analysis. The inclusion of different plant species within a pixel or specific PFT could misrepresent the response of spring onset to climate change (Chuine *et al.*, 2000; Fu *et al.*, 2014; MacBean *et al.*, 2015), and recent studies have found substantial differences in chilling requirements and photoperiod sensitivities of different tree species (Körner & Basler, 2010; Laube *et al.*, 2014; Zohner *et al.*, 2016). Regional climate datasets and satellite-based spring phenology dates with finer spatial resolution, as phenology data inferred from Landsat imagery (Melaas *et al.*, 2013b) are becoming increasingly available, and thus would be profitable for modeling spring phenology. Besides, the parameterization of spring phenology models might be the third source of uncertainties in simulating SOS. We applied the optimal parameters in models, while large spatial variation of model parameters may exist and the Markov Chain Monte Carlo (MCMC) approach may be a good reference when applying these models within the PFTs (Jeong *et al.*, 2012).

In conclusion, we calibrated six frequently applied spring phenology models using satellite-derived SOS, and compared their simulations across PFTs in the Northern Hemisphere during the period of 1982–2012. Model performance relied strongly on their interpretations of the effects of forcing, chilling, and photoperiod. Across the Northern Hemisphere, forcing temperatures might play a prominent role in modeling boreal forest SOS (particularly for BoND), with the one-phase Spring Warming model (which considers only forcing temperature) showing similar accuracy to the two-phase models. By contrast, for temperate PFTs (e.g., TeNE and TeBE), models involving chilling and particularly photoperiod provided better simulations. Specifically, the DORMPHOT model (which represented the photoperiod effect) best captured the spatiotemporal pattern of SOS in most PFTs (except in TeBD). Our results suggest that chilling and photoperiod effects should be embedded in large-scale vegetation phenology simulations, albeit in a PFT-specific manner. Several issues, however, still need to be addressed: 1) incorporating soil moisture or precipitation in arid/semi-arid regions and snowmelt in higher altitudes to improve simulations in grasslands; and 2) applying climatic and phenology observational datasets with finer spatial and temporal resolutions. Moreover, manipulative experiments aimed toward elucidating environmental and biological mechanisms underlying spring vegetation growth onset are required to better predict the date of

SOS, and thus ecosystem responses to ongoing climate change.

Acknowledgements

This study was supported by the National Key R&D Program of China (2017YFA0604702), National Natural Science Foundation of China (41530528 and 31770516), the BELSPO STEREO project ECOPROPHET (SR00334) and National Youth Top-notch Talent Support Program in China. Y.S.H acknowledge support by the National Key Research and Development Program of China (2017YFA06036001).

References

Akaike, H. (1998) Information theory and an extension of the maximum likelihood principle. In: *Selected Papers of Hirotugu Akaike.* pp 199-213. Springer.

Bernal, M., Estiarte, M., & Peñuelas, J. (2011) Drought advances spring growth phenology of the Mediterranean shrub Erica multiflora. Plant Biology, 13, 252-257.

Botta, A., Viovy, N., Ciais, P., Friedlingstein, P., & Monfray, P. (2000) A global prognostic scheme of leaf onset using satellite data. Global Change Biology, 6, 709-725.

Buitenwerf, R., Rose, L., & Higgins, S.I. (2015) Three decades of multi-dimensional change in global leaf phenology. Nature Climate Change, 5, 364-368.

Caffarra, A., Donnelly, A., & Chuine, I. (2011a) Modelling the timing of Betula pubescens budburst. II. Integrating complex effects of photoperiod into process-based models. Climate Research, 46, 159-170.

Caffarra, A., Donnelly, A., Chuine, I., & Jones, M.B. (2011b) Modelling the timing of Betula pubescens budburst. I. Temperature and photoperiod: a conceptual model. Climate Research, 46, 147-157.

Cannell, M., & Smith, R. (1983) Thermal time, chill days and prediction of budburst in Picea sitchensis. Journal of Applied Ecology, 20, 951-963.

Chen, M., Melaas, E.K., Gray, J.M., Friedl, M.A., & Richardson, AD (2016) A new seasonal-deciduous spring phenology submodel in the Community Land Model 4.5: impacts on carbon and water cycling under future climate scenarios. Global Change Biology, 22, 3675-3688.

Chen, X., An, S., Inouye, D.W., & Schwartz, M.D. (2015) Temperature and snowfall trigger alpine vegetation green-up on the world's roof. Global Change Biology, 21, 3635-3646.

Chen, X. (2017) *Spatiotemporal Processes of Plant Phenology: Simulation and Prediction,* Berlin Heidelberg, Springer-Verlag Berlin Heidelberg.

Chen, X., Wang, L., & Inouye, D.W. (2017) Delayed response of spring phenology to global warming in subtropics and tropics. Agricultural and Forest Meteorology, 234, 222-235.

Chiang, J-M., & Brown, K.J. (2007) Improving the budburst phenology subroutine in the forest carbon model PnET. Ecological Modelling, 205, 515-526.

Chuine, I. (2000) A unified model for budburst of trees. Journal of Theoretical Biology, 207, 337-347.

Chuine, I., Cambon, G., & Comtois, P. (2000) Scaling phenology from the local to the regional level: advances from species - specific phenological models. Global Change Biology, 6, 943-952.

Chuine, I., Morin, X., & Bugmann, H. (2010) Warming, Photoperiods, and Tree Phenology. Science, 329, 277-278.

Chuine, I., Bonhomme, M., Legave, J.M., García De Cortázar - Atauri I, Charrier, G., … Améglio, T. (2016) Can phenological models predict tree phenology accurately in the future? The unrevealed hurdle of endodormancy break. Global Change Biology, 22, 3444-3460.

Cleland, E.E., Chuine, I., Menzel, A., Mooney, H.A., & Schwartz, M.D. (2007) Shifting plant phenology in response to global change. Trends in ecology & evolution, 22, 357-365.

Díaz, S., & Cabido, M. (1997) Plant functional types and ecosystem function in relation to global change. Journal of Vegetation Science, 463-474.

Donnelly, A., Salamin, N., & Jones, M.B. (2006) Changes in tree phenology: an indicator of spring warming in Ireland? In: *Biology and Environment: Proceedings of the Royal Irish Academy.* pp 49-56, JSTOR.

Forkel, M., Migliavacca, M., Thonicke, K., Reichstein, M., Schaphoff, S., … Carvalhais, N. (2015)

Co-dominant water control on global inter-annual variability and trends in land surface phenology and greenness. Global Change Biology, 21, 3414-3435.

Friedl, M.A., Gray, J.M., Melaas, E.K., Richardson, AD., Hufkens, K., Keenan, T.F., . . . O'keefe, J. (2014) A tale of two springs: using recent climate anomalies to characterize the sensitivity of temperate forest phenology to climate change. Environmental Research Letters, 9, 054006.

Fu, Y.H., Piao, S., Op De Beeck, M., Cong, N., Zhao, H., Zhang, Y., . . . Janssens, I.A. (2014) Recent spring phenology shifts in western Central Europe based on multiscale observations. Global Ecology and Biogeography, 23, 1255-1263.

Fu, Y.H., Piao, S., Zhao, H., Jeong, S.J., Wang, X., Vitasse, Y., . . . Janssens, I.A. (2014) Unexpected role of winter precipitation in determining heat requirement for spring vegetation green-up at northern-middle and high latitudes. Global Change Biology, 20, 3743-3755.

Fu, Y.H., Piao, S., Vitasse, Y., Zhao, H., De Boeck, HJ., Liu, Q., . . . Janssens, I.A. (2015) Increased heat requirement for leaf flushing in temperate woody species over 1980-2012: effects of chilling, precipitation and insolation. Global Change Biology, 21, 2687–2697.

Fu, Y.H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G., . . . Janssens, I.A. (2015) Declining global warming effects on the phenology of spring leaf unfolding. Nature, 526, 104-107.

Fu, Y.H., Liu, Y., De Boeck, HJ., Menzel, A., Nijs, I., Peaucelle, M., . . . Janssens, I.A. (2016) Three times greater weight of daytime than of night - time temperature on leaf unfolding phenology in temperate trees. New Phytologist, 212, 590-597.

Garonna, I., De, Jong. R., & Schaepman, M. (2015) Variability and evolution of global land surface phenology over the past three decades (1982 - 2012). Global Change Biology, 22, 1456-1468.

Hänninen, H. (1990) Modelling bud dormancy release in trees from cool and temperate regions. Acta

Forestalia Fennica, 213, 1-47.

Hunter, A.F., & Lechowicz, M.J. (1992) Predicting the Timing of Budburst in Temperate Trees. Journal of Applied Ecology, 29, 597-604.

Jakubauskas, M.E., Legates, D.R., & Kastens, J.H. (2001) Harmonic analysis of time-series AVHRR NDVI data. Photogrammetric Engineering and Remote Sensing, 67, 461-470.

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

Jeong, S.J., Medvigy D, Shevliakova E, Malyshev S (2012) Uncertainties in terrestrial carbon budgets related to spring phenology. Journal of Geophysical Research: Biogeosciences, 117, G01030, doi:10.1029/2011JG001868.

Jeong, S.J., Medvigy, D., Shevliakova, E., & Malyshev, S. (2013) Predicting changes in temperate forest budburst using continental‐scale observations and models. Geophysical Research Letters, 40, 359-364.

Julien, Y., & Sobrino, J. (2009) Global land surface phenology trends from GIMMS database. International Journal of Remote Sensing, 30, 3495-3513.

Körner, C., & Basler, D. (2010) Phenology under global warming. Science, 327, 1461-1462.

Kramer, K. (1994a) A modelling analysis of the effects of climatic warming on the probability of spring frost damage to tree species in The Netherlands and Germany. Plant, Cell & Environment, 17, 367-377.

Kramer, K. (1994b) Selecting a Model to Predict the Onset of Growth of Fagus sylvatica. Journal of Applied Ecology, 31, 172-181.

Krinner, G., Viovy, N., De Noblet‐Ducoudré N, Ogée, J., Polcher, J., Friedlingstein, P., . . . Prentice, I.C. (2005) A dynamic global vegetation model for studies of the coupled atmosphere - biosphere system. Global Biogeochemical Cycles, 19, doi: 10.1029/2003GB002199.

Landsberg, J. (1974) Apple fruit bud development and growth; analysis and an empirical model. Annals of Botany, 38, 1013-1023.

Laube, J., Sparks, T.H., Estrella, N., Hofler, J., Ankerst, D.P., & Menzel, A. (2014a) Chilling outweighs photoperiod in preventing precocious spring development. Global Change Biology, 20, 170-182.

Laube, J., Sparks, T.H., Estrella, N., & Menzel, A. (2014b) Does humidity trigger tree phenology? Proposal for an air humidity based framework for bud development in spring. New Phytologist, 202, 350-355.

Liu, Q., Fu, Y.H., Zhu, Z., Liu, Y., Liu, Z., Huang, M., . . . Piao, S. (2016) Delayed autumn phenology in the Northern Hemisphere is related to change in both climate and spring phenology. Global Change Biology, 22, 3702-3711.

Macbean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F.M., & Ciais, P. (2015) Using satellite data to improve the leaf phenology of a global terrestrial biosphere model. Biogeosciences, 12, 7185-7208.

Martinez-Cantin, R. (2014) BayesOpt: a Bayesian optimization library for nonlinear optimization, experimental design and bandits. Journal of Machine Learning Research, 15, 3735-3739.

Melaas, E.K., Richardson, AD., Friedl, M.A., Dragoni, D., Gough, C.M., Herbst, M., . . . Moors, E. (2013a) Using FLUXNET data to improve models of springtime vegetation activity onset in forest ecosystems. Agricultural and Forest Meteorology, 171-172, 46-56.

Melaas, E.K., Friedl, M.A., & Zhu, Z. (2013b) Detecting interannual variation in deciduous broadleaf forest phenology using Landsat TM/ETM+ data. Remote Sensing of Environment, 132, 176-185.

Melaas, E.K., Friedl, M.A., & Richardson, AD. (2015) Multi-scale modeling of spring phenology across Deciduous Forests in the Eastern United States. Global Change Biology, 22, 792-805.

Menzel, A., Sparks, T.H., Estrella, N., Koch, E., Aasa, A., Ahas, R., . . . Briede, A. (2006) European phenological response to climate change matches the warming pattern. Global Change Biology, 12, 1969-1976.

Migliavacca, M., Galvagno, M., Cremonese, E., Rossini, M., Meroni, M., Sonnentag, O., . . . Busetto, L. (2011) Using digital repeat photography and eddy covariance data to model grassland phenology and photosynthetic CO_2 uptake. Agricultural and Forest Meteorology, 151, 1325-1337.

Mitchell, T.D., & Jones, P.D. (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. International Journal of Climatology, 25, 693-712.

Mockus, J. (2012) *Bayesian approach to global optimization: theory and applications*, Springer Science & Business Media.

Murray, M., Cannell, M., & Smith, R. (1989) Date of budburst of fifteen tree species in Britain following climatic warming. Journal of Applied Ecology, 26, 693-700.

Myking, T., & Heide, O. (1995) Dormancy release and chilling requirement of buds of latitudinal ecotypes of Betula pendula and B. pubescens. Tree physiology, 15, 697-704.

Myneni, R.B., & Hall, F.G. (1995) The interpretation of spectral vegetation indexes. Geoscience and Remote Sensing, IEEE Transactions on, 33, 481-486.

Myneni, R.B, Keeling, C., Tucker, C., Asrar, G., & Nemani, R. (1997) Increased plant growth in the northern high latitudes from 1981 to 1991. Nature, 386, 698-702.

Nash, J.E., & Sutcliffe, J.V. (1970) River flow forecasting through conceptual models part I-A discussion of principles. Journal of Hydrology, 10, 282-290.

New, M., Hulme, M., & Jones, P. (2000) Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate. Journal of Climate, 13, 2217-2238.

Peñuelas, J., Filella, I., Zhang, X., Llorens, L., Ogaya, R., Lloret, F., . . . Terradas, J. (2004) Complex spatiotemporal phenological shifts as a response to rainfall changes. New Phytologist, 161, 837-846.

Peñuelas, J., Filella, I. (2009) Phenology feedbacks on climate change. Science, 324, 887-888.

Piao, S., Fang, J., Zhou, L., Ciais, P., & Zhu, B. (2006) Variations in satellite-derived phenology in China's temperate vegetation. Global Change Biology, 12, 672-685.

Piao, S., Friedlingstein, P., Ciais, P., Viovy, N., & Demarty, J. (2007) Growing season extension and its impact on terrestrial carbon cycle in the Northern Hemisphere over the past 2 decades. Global Biogeochemical Cycles, 21, doi: 10.1029/2006GB002888.

Piao, S., Tan, J., Chen, A., Fu, Y.H., Ciais, P., Liu, Q., . . . Peñuelas, J. (2015) Leaf onset in the northern hemisphere triggered by daytime temperature. Nature Communications, 6, doi:10.1038/ncomms7911.

Piao, S.L., Liu, Z., Wang, T., Peng, S.S., Ciais, P., Huang, M.T., . . . Tans, P.P. (2017) Weakening temperature control on the interannual variations of spring carbon uptake across northern lands. Nature Climate Change, 7, 359-363

Pinzon, J.E, & Tucker, C.J. (2014) A non-stationary 1981–2012 AVHRR NDVI$_{3g}$ time series. Remote Sensing, 6, 6929-6960.

Pinzon, J.E., & Tucker, C.J. (2016) A Non-Stationary 1981-2015 AVHRR NDVI$_{3g.v1}$ Time Series: an update. Remote Sensing, in preparation.

Richardson, AD., Black, T.A., Ciais, P., Delbart, N., Friedl, M.A., Gobron, N., . . . Luyssaert, S. (2010) Influence of spring and autumn phenological transitions on forest ecosystem productivity. Philos Trans R Soc Lond B Biol Sci, 365, 3227-3246.

Richardson, AD., Anderson, R.S., Arain, M.A., Barr, A.G., Bohrer, G., Chen, G., . . . Desai, A.R. (2012) Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis. Global Change Biology, 18, 566-584.

Sarvas, R. (1974) Investigations on the annual cycle of development of forest trees: II. Autumn dormancy and winter dormancy. Communicationes Instituti Forestalis Fenniae, 84, 101.

Schwartz, M.D., Ahas, R., & Aasa, A. (2006) Onset of spring starting earlier across the Northern Hemisphere. Global Change Biology, 12, 343-351.

Sitch, S., Smith, B., Prentice, I.C., Arneth, A., Bondeau, A., Cramer, W., . . . Sykes, M.T. (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. Global Change Biology, 9, 161-185.

Verant, S., Laval, K., Polcher, J., & De, Castro. M. (2004) Sensitivity of the continental hydrological cycle to the spatial resolution over the Iberian Peninsula. Journal of Hydrometeorology, 5, 267-285.

Vitasse, Y., François, C., Delpierre, N., Dufrêne, E., Kremer, A., Chuine, I., & Delzon, S. (2011) Assessing the effects of climate change on the phenology of European temperate trees. Agricultural and Forest Meteorology, 151, 969-980.

Walther, G-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J., . . . Bairlein, F. (2002) Ecological responses to recent climate change. Nature, 416, 389-395.

Wang, X., Piao, S., Xu, X., Ciais, P., Macbean, N., Myneni, R.B., & Li, L. (2015) Has the advancing onset of spring vegetation green - up slowed down or changed abruptly over the last three decades? Global Ecology and Biogeography, 24, 621-631.

White, M.A., De Beurs, K.M., Didan, K., Inouye, D.W., Richardson, AD., Jensen, O.P., . . . Lauenroth, W.K. (2009) Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. Global Change Biology, 15, 2335-2359.

Xin, Q., Broich, M., Zhu, P., & Gong, P. (2015) Modeling grassland spring onset across the Western United States using climate variables and MODIS-derived phenology metrics. Remote Sensing of Environment, 161, 63-77.

Xu, L., & Chen, X. (2013) Regional unified model - based leaf unfolding prediction from 1960 to 2009 across northern China. Global Change Biology, 19, 1275-1284.

Yang, X., Mustard, J.F., Tang, J., & Xu, H. (2012) Regional - scale phenology modeling based on meteorological records and remote sensing observations. Journal of Geophysical Research: Biogeosciences, 117, doi: 10.1029/2012JG001977.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C., Gao, F., . . . Huete, A. (2003) Monitoring vegetation phenology using MODIS. Remote Sensing of Environment, 84, 471-475.

Zohner, C.M., Benito, B.M., Svenning, J-C., & Renner, S.S. (2016) Day length unlikely to constrain climate-driven shifts in leaf-out times of northern woody plants. Nature Climate Change, 6, 1120-1123.

Supporting Information Captions

Figure S1. Spatial distribution of the fractional cover of PFTs across the Northern Hemisphere (> 30°N). Seven main PFTs were included: (a) Temperate needleleaf evergreen (TeNE), (b) Temperate broadleaved evergreen (TeBE), (c) Temperate broadleaved deciduous (TeBD), (d) Boreal needleleaf evergreen (BoNE), (e) Boreal needleleaf deciduous (BoND), (f) Natural C3 grass (NC3), and (g) Natural C4 grass (NC4). Color indicates the proportion of a specific PFT in a 0.5° × 0.5° grid.

Figure S2. Heat plots showing the relationship between predicted and satellite-derived SOS. SOS data was randomly sampled (10%) from each PFT in the Northern Hemisphere and was excluded from the calibration of spring phenology models. a–p1: predictions were generated using six spring phenology models at the PFT scale. Diagonal lines are 1:1 lines, indicating perfect agreement between predicted and satellite-derived SOS. Colors indicate the percentage of pixels within each bin.

Figure S3. Spatial pattern of averaged SOS during the period 1982–2012 at the PFT scale. a–d was SOS inferred from satellite observations and h–p1 shows the SOS predicted by five phenology models. Pie charts beneath each model indicate the differences between predicted and satellite-derived SOS; a positive value indicates that the model produced later spring onset than observed, and vice versa.

Figure S4. Maps of RMSE (days) between satellite-derived SOS and model simulations at the PFT scale. Results are based on six spring phenology models: (a–g) Spring Warming, (h–n) Sequential, (o–u) Parallel, (v–b1) UniChill, (c1-i1) Unified, and (j1–p1) DORMPHOT. Colors indicate differences between satellite- and model-derived SOS.

Figure S5. Changes in satellite-derived and model-predicted SOS at the PFT scale during the period 1982–2012. Figures S5a–g show linear trends estimated from satellite-derived SOS, and Figures S5h–w1 show results based on six spring phenology models: (h–n) Spring Warming model, (o–u) Sequential model, (v–b1) Parallel model, (c1–i1) UniChill model, (j1–p1) Unified model, and (q1–w1) DORMPHOT model.

Figure S6. Variation of satellite-derived and model predicted SOS over the period 1982-2012. a-f shows the results across the temperate broadleaved evergreen (TeBE) based on six spring phenology models while g-l indicates the results across temperate broadleaved deciduous (TeBD).

Figure S7. Comparison between individual phenology extraction method and their ensemble mean over the period 1982-2012. Bars and error bars indicate the temporal average and standard deviation of SOS for each PFT, respectively.

Figure legends

Figure 1. Heat plots showing the relationship between predicted and satellite-derived SOS. SOS data was randomly sampled (90%) from each PFT in the Northern Hemisphere and was applied to the calibration of spring phenology models. Figure 1a–p1: predictions were generated using six spring phenology models at the PFT scale. Diagonal lines are 1:1 lines, indicating perfect agreement between predicted and satellite-derived SOS. Colors indicate the percentage of pixels in each bin area.

Figure 2. Spatial pattern of averaged SOS in the period of 1982–2012. (a) satellite-derived SOS, (b) SOS predicted by the Spring Warming model, (c) Sequential model, (d) Parallel model, (e) Unichill model, (f) Unified model, and (g) DORMPHOT model. The subplot in the bottom left of each spring phenology model shows the relationship between satellite-derived SOS and model prediction (*denotes significant correlation at $P < 0.05$). Figure 2e and i show the distribution of SOS across the latitude gradient (30°N–75°N).

Figure 3. Boxplots of root mean square error (RMSE, days) between satellite-derived SOS and model-simulated SOS (from 1982 to 2012) across seven main PFTs in the Northern Hemisphere.

Figure 4. Maps of RMSE (days) between satellite-derived SOS and model simulations. RMSEs were determined at the pixel scale during the period 1982–2012. Figures 4a–f show the spatial distribution of RMSEs across the Northern Hemisphere based on six spring phenology models including the (a) Spring Warming model, (b) Sequential model, (c) Parallel model, (d) UniChill model, (e) Unified

model, and (f) DORMPHOT model. Figure 4g is the probability of RMSEs across the Northern Hemisphere, with vertical lines showing the median of RMSEs.

Figure 5. Changes in satellite-derived and model-predicted SOS during the period from 1982 to 2012. Figures 5a and b–g indicate linear SOS trends inferred from satellite observation and model simulations, respectively. Dotted regions indicate that the trends were significant at $P < 0.05$.

Figure 6. Linear trends of satellite-derived SOS and model simulations at the PFT scale. *indicates that detected trends were statistically significant at $P < 0.05$.

Table 1. Estimates of six spring phenology models for each PFT and their performance.

Models		**Model Parameters**											**Model Evaluation$_{\text{Internal}}$**		
	PFTs	**alpha**	**beta**	**Fcrit**	**Af**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**RMSE**	**R^2**	**AIC**
a. Spring Warming	TeNE	–0.07	15.00	498.24	19.69	-	-	-	-	-	-	-	18.9	0.52*	1.24E + 0
	TeBE	–0.16	8.54	498.86	25.52	-	-	-	-	-	-	-	19.5	0.49*	1.68E + 0
	TeBD	–0.05	12.64	270.24	8.30	-	-	-	-	-	-	-	17.1	0.42*	4.03E + 0
	BoNE	–0.59	6.29	226.33	40.00	-	-	-	-	-	-	-	13.5	0.57*	1.22E + 0
	BoND	–0.65	3.60	381.57	39.99	-	-	-	-	-	-	-	8.9	0.65*	3.65E + 0
	NC3	–0.08	0.95	359.88	8.54	-	-	-	-	-	-	-	18.5	0.76*	3.58E + 0
	NC4	–0.05	32.77	495.09	30.73	-	-	-	-	-	-	-	25.8	0.29*	8.97E + 0
	PFTs	**alpha**	**beta**	**Ccrit**	**Fcrit**	**Tb**	**Td**	**Ta**	**Tc**	**Af**	**-**	**-**	**RMSE**	**R^2**	**AIC**
b. Sequential	TeNE	–0.20	–5.31	179.60	63.90	7.04	1.24	–99.38	76.63	7.31	-	-	16.0	0.55*	1.17E + 0
	TeBE	–0.35	–12.76	144.92	66.42	0.27	0.78	–99.21	83.74	36.41	-	-	17.9	0.50*	1.63E + 0
	TeBD	–0.14	-6.35	170.32	353.07	5.07	1.86	–57.10	92.72	42.81	-	-	17.5	0.41*	4.06E + 0
	BoNE	–0.14	1.84	180.57	497.71	–4.86	0.72	–81.55	98.63	37.53	-	-	12.1	0.59*	1.17E + 0
	BoND	–0.40	53.11	55.21	384.22	–7.79	2.39	–58.06	-0.94	33.08	-	-	9.1	0.64*	3.67E + 0
	NC3	–0.18	-4.46	150.49	462.42	–1.55	1.26	–63.99	90.50	38.49	-	-	17.0	0.80*	3.47E + 0
	NC4	–0.05	27.59	170.58	62.61	9.77	5.68	–29.41	89.49	20.31	-	-	26.6	0.25*	9.05E + 0
	PFTs	**alpha**	**beta**	**Ccrit**	**Fcrit**	**Tb**	**Td**	**Ta**	**Tc**	**Af**	**Kmin**	**-**	**RMSE**	**R^2**	**AIC**
c. Parallel	TeNE	–0.03	-6.94	145.13	243.68	–17.61	1.13	–59.44	4.13	26.34	0.09	-	16.7	0.57*	1.19E + 0
	TeBE	–0.10	14.59	118.10	226.64	–5.54	1.78	–73.80	0.03	11.99	0.11	-	17.3	0.50*	1.61E + 0
	TeBD	–0.65	-0.19	112.23	229.55	–12.06	4.01	–61.37	4.76	14.88	0.08	-	18.1	0.41*	4.11E + 0
	BoNE	–0.02	-11.11	139.52	154.15	–15.62	1.99	–55.54	4.33	26.99	0.05	-	12.6	0.58*	1.18E + 0
	BoND	–0.13	34.52	163.59	171.01	–8.93	0.95	–99.54	3.28	11.09	0.01	-	9.0	0.63*	3.68E + 0
	NC3	–0.02	-8.90	192.43	78.55	–9.13	3.97	–89.87	5.54	11.98	0.08	-	17.3	0.81*	3.48E + 0
	NC4	–0.11	-17.57	173.36	263.35	–10.94	0.83	–30.64	27.76	26.96	0.01	-	27.3	0.24*	9.12E + 0

d. UniChill	**PFTs**	**alpha**	**beta**	**Tb**	**Ta**	**Tc**	**Ccrit**	**Fcrit**	**-**	**-**	**-**	**-**	**RMSE**	**R^2**	**AIC**
	TeNE	–0.18	3.95	0.00	0.00	4.41	88.05	19.47	-	-	-	-	16.2	0.56*	1.18
	TeBE	–0.11	11.31	0.00	0.00	–2.67	68.74	16.26	-	-	-	-	18.3	0.45*	1.64
	TeBD	–0.08	2.26	0.01	0.00	2.45	58.94	42.52	-	-	-	-	17.3	0.41*	4.05
	BoNE	–0.38	4.60	0.66	0.02	–3.95	39.25	13.78	-	-	-	-	12.5	0.59*	1.14
	BoND	–0.97	0.24	0.39	0.01	–11.53	39.77	20.09	-	-	-	-	9.0	0.64*	3.60
	NC3	–0.39	1.86	0.01	0.00	7.50	82.99	22.29	-	-	-	-	16.8	0.80*	3.46
	NC4	–0.08	9.22	0.00	0.00	0.82	29.07	52.29	-	-	-	-	26.4	0.28*	9.02
e. Unified	**PFTs**	**alpha**	**beta**	**Tb**	**Ta**	**Tc**	**w**	**z**	**Ccrit**	**deltaP**	**-**	**-**	**RMSE**	**R^2**	**AIC**
	TeNE	–0.38	6.43	0.00	0.00	–1.93	759.34	–0.05	93.11	41	-	-	15.8	0.56*	1.17
	TeBE	–0.31	9.36	0.01	0.00	7.30	605.88	–0.06	74.56	189	-	-	17.6	0.51*	1.62
	TeBD	–0.21	5.70	0.01	0.00	4.43	235.74	–0.03	86.62	105	-	-	17.6	0.41*	4.07
	BoNE	–0.39	7.87	0.01	0.00	0.45	210.11	–0.05	80.85	22	-	-	12.3	0.61*	1.18
	BoND	–0.35	6.76	0.77	0.02	–3.52	1379.60	–0.07	77.56	129	-	-	8.7	0.66*	3.59
	NC3	–0.37	4.05	0.01	0.00	–1.99	494.70	–0.04	85.39	195	-	-	16.4	0.81*	3.43
	NC4	–0.03	2.77	0.03	0.00	4.63	110.80	–0.01	0.65	62	-	-	25.8	0.29*	8.96
f. DORMPHOT	**PFTs**	**DLcrit**	**Dcrit**	**Ccrit**	**Fcrit**	**aD**	**bD**	**aC**	**cC**	**dF**	**gT**	**hDL**	**RMSE**	**R^2**	**AIC**
	TeNE	12.75	29.43	99.68	19.45	0.05	23.17	0.01	18.41	–0.30	0.33	0.05	15.7	0.56*	1.16
	TeBE	11.01	31.94	24.38	16.32	0.04	20.42	0.02	14.59	–0.76	0.25	0.11	16.8	0.53*	1.60
	TeBD	10.47	28.42	12.49	14.88	0.03	13.36	0.01	16.13	–0.90	0.30	0.02	16.9	0.43*	4.01
	BoNE	9.69	31.21	63.20	15.81	0.91	4.91	0.01	9.70	–0.59	0.25	0.15	11.8	0.60*	1.15
	BoND	10.63	55.42	61.95	19.70	0.44	16.78	0.02	–11.01	–0.44	13.26	14.34	8.7	0.63*	3.45
	NC3	10.42	21.38	8.65	21.99	0.02	6.79	0.01	15.72	–0.77	0.30	0.04	15.7	0.82*	3.38
	NC4	10.58	33.87	49.70	6.03	0.94	19.39	0.41	0.92	–0.05	14.34	11.66	26.6	0.28*	9.04

Note: R^2 represents the percentage of variance explained by the model (*$P < 0.01$). AIC: Akaike Information Criterion; the lowest AIC indicates the optimum model for the simulation of spring phenology. NSE: Nash–Sutcliffe efficiency index; higher (positive, ranging from 0 to 1) NSE values indicate higher model reliability.

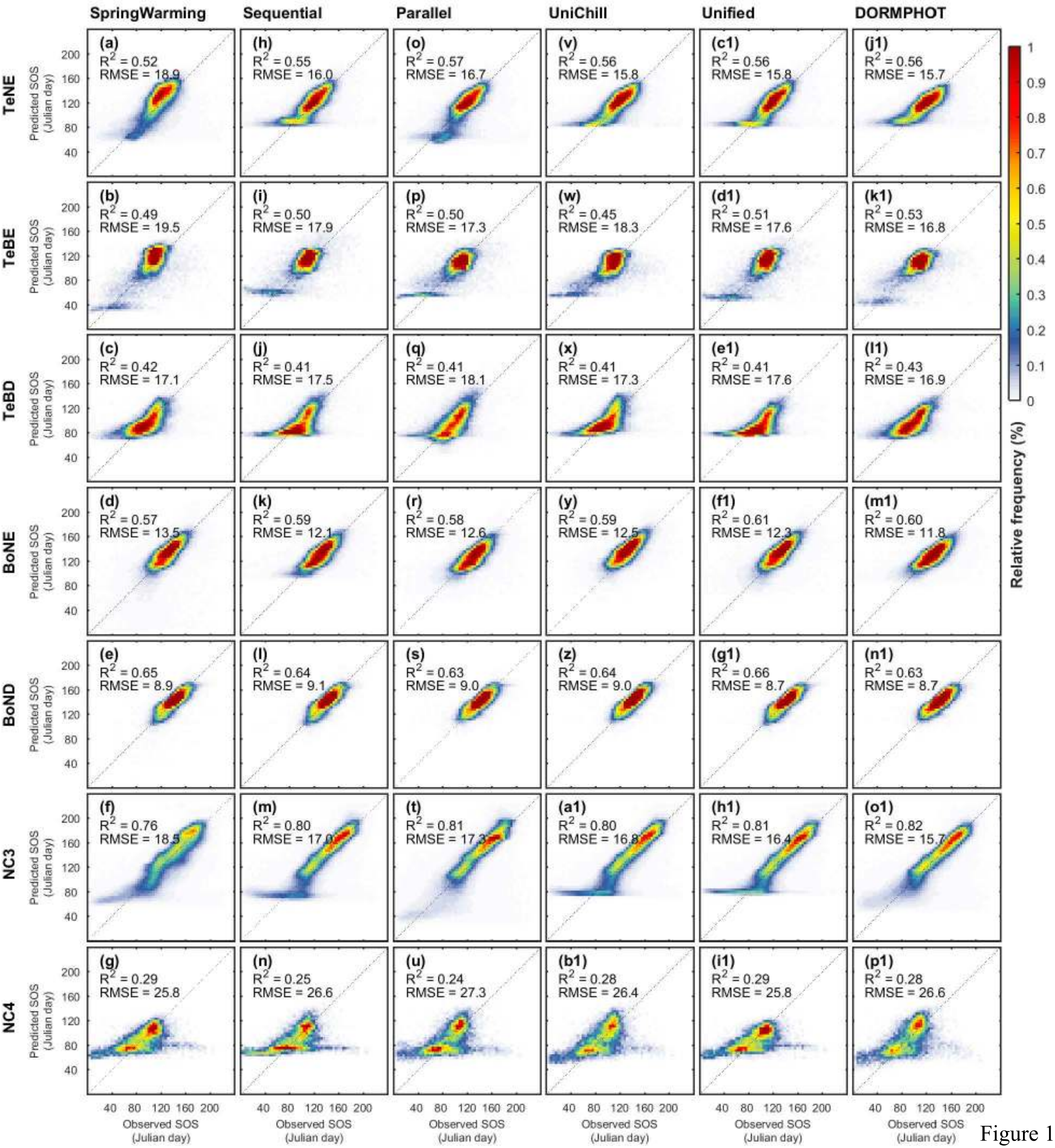
SpringWarming
Sequential
Parallel
UniChill
Unified
DORMPHOT
TeNE
TeBE
TeBD
BoNE
BoND
NC3
NC4
Predicted SOS (Julian day)
Observed SOS (Julian day)
40
80
120
160
200
Relative frequency (%)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
(a) R^2 = 0.52 RMSE = 18.9
(h) R^2 = 0.55 RMSE = 16.0
(o) R^2 = 0.57 RMSE = 16.7
(v) R^2 = 0.56 RMSE = 15.8
(c1) R^2 = 0.56 RMSE = 15.8
(j1) R^2 = 0.56 RMSE = 15.7
(b) R^2 = 0.49 RMSE = 19.5
(i) R^2 = 0.50 RMSE = 17.9
(p) R^2 = 0.50 RMSE = 17.3
(w) R^2 = 0.45 RMSE = 18.3
(d1) R^2 = 0.51 RMSE = 17.6
(k1) R^2 = 0.53 RMSE = 16.8
(c) R^2 = 0.42 RMSE = 17.1
(j) R^2 = 0.41 RMSE = 17.5
(q) R^2 = 0.41 RMSE = 18.1
(x) R^2 = 0.41 RMSE = 17.3
(e1) R^2 = 0.41 RMSE = 17.6
(l1) R^2 = 0.43 RMSE = 16.9
(d) R^2 = 0.57 RMSE = 13.5
(k) R^2 = 0.59 RMSE = 12.1
(r) R^2 = 0.58 RMSE = 12.6
(y) R^2 = 0.59 RMSE = 12.5
(f1) R^2 = 0.61 RMSE = 12.3
(m1) R^2 = 0.60 RMSE = 11.8
(e) R^2 = 0.65 RMSE = 8.9
(l) R^2 = 0.64 RMSE = 9.1
(s) R^2 = 0.63 RMSE = 9.0
(z) R^2 = 0.64 RMSE = 9.0
(g1) R^2 = 0.66 RMSE = 8.7
(n1) R^2 = 0.63 RMSE = 8.7
(f) R^2 = 0.76 RMSE = 18.5
(m) R^2 = 0.80 RMSE = 17.0
(t) R^2 = 0.81 RMSE = 17.3
(a1) R^2 = 0.80 RMSE = 16.8
(h1) R^2 = 0.81 RMSE = 16.4
(o1) R^2 = 0.82 RMSE = 15.7
(g) R^2 = 0.29 RMSE = 25.8
(n) R^2 = 0.25 RMSE = 26.6
(u) R^2 = 0.24 RMSE = 27.3
(b1) R^2 = 0.28 RMSE = 26.4
(i1) R^2 = 0.29 RMSE = 25.8
(p1) R^2 = 0.28 RMSE = 26.6

Figure 1

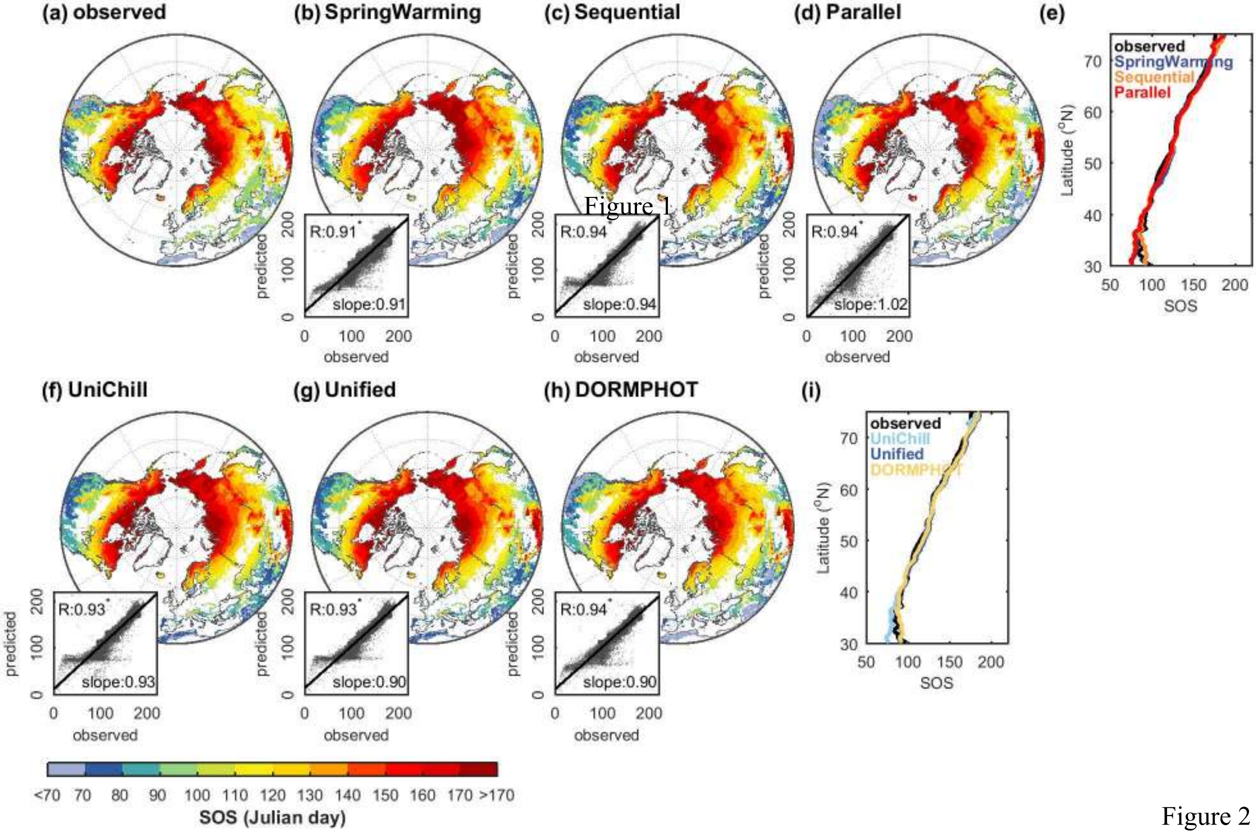

Figure 1

Figure 2

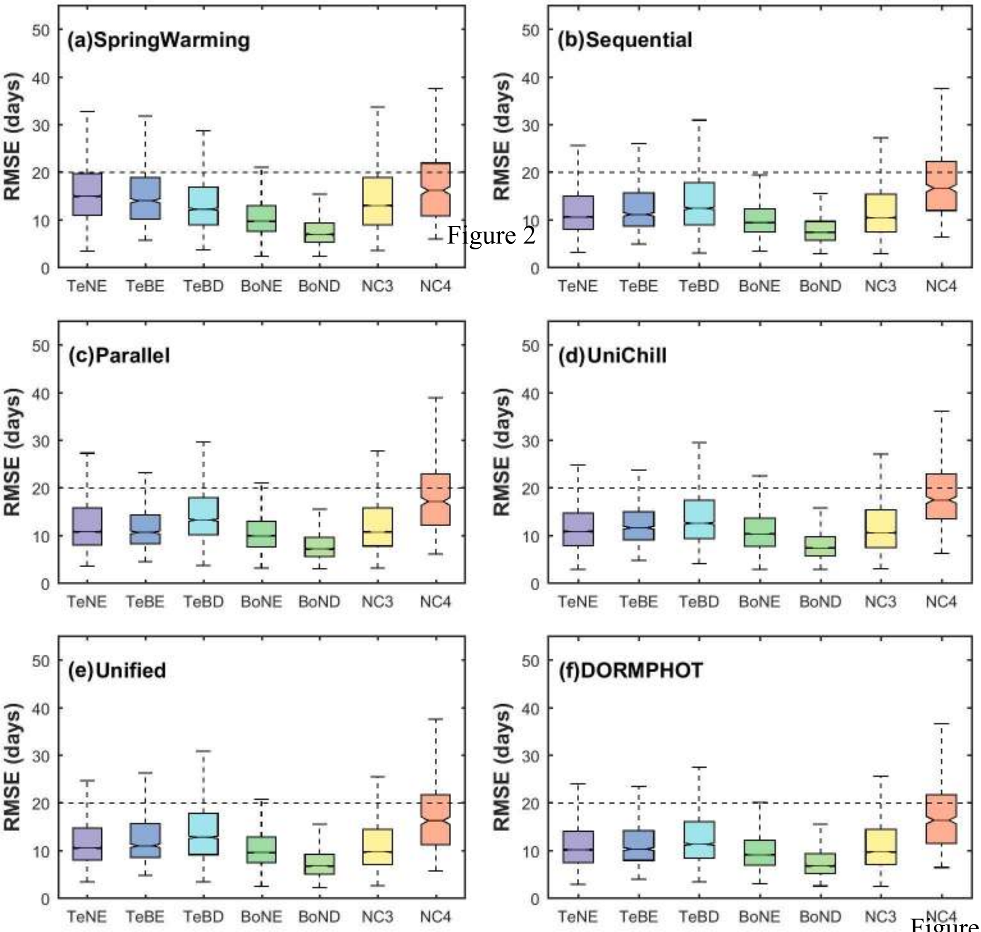

Figure 2

Figure 3

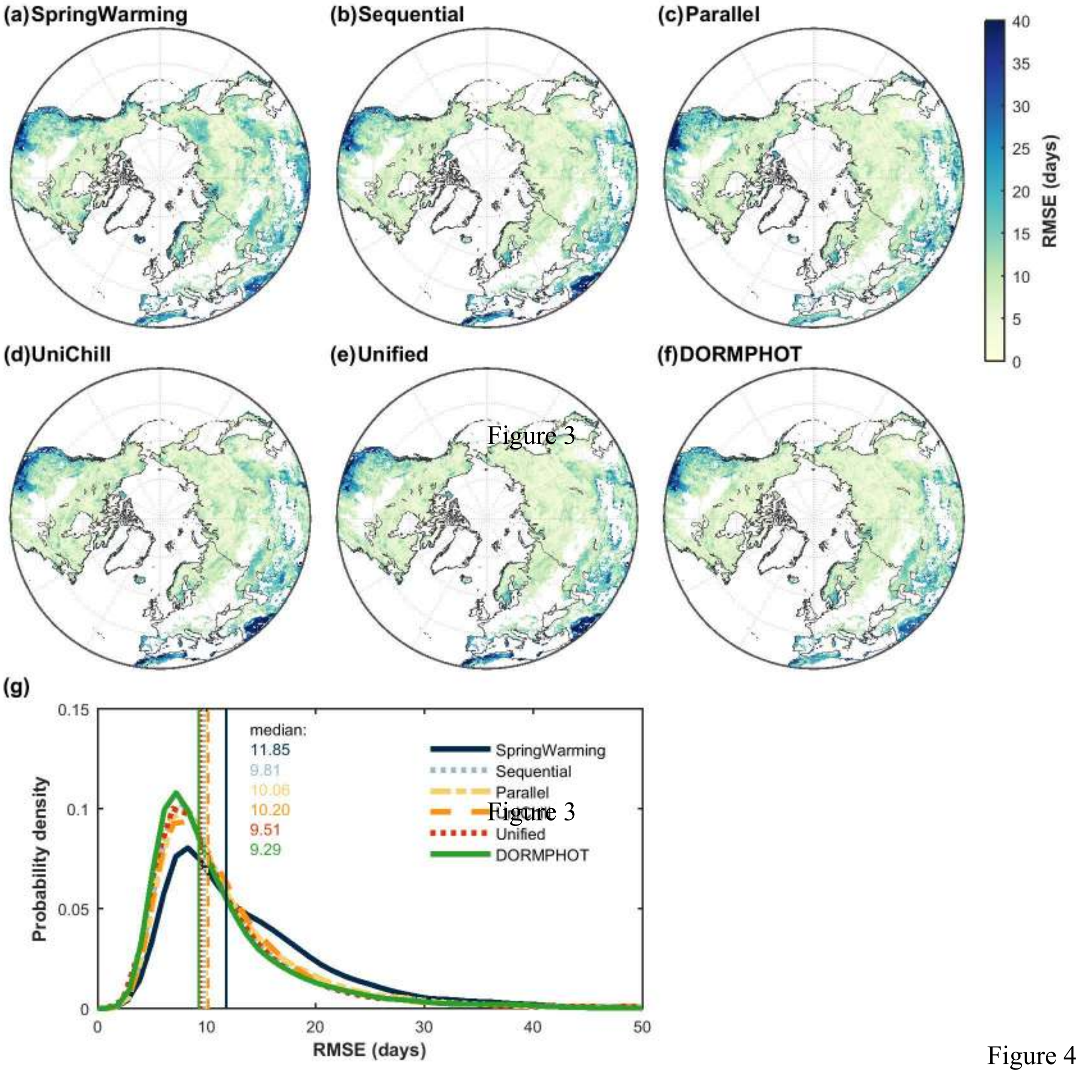

Figure 3

Figure 3

Figure 4

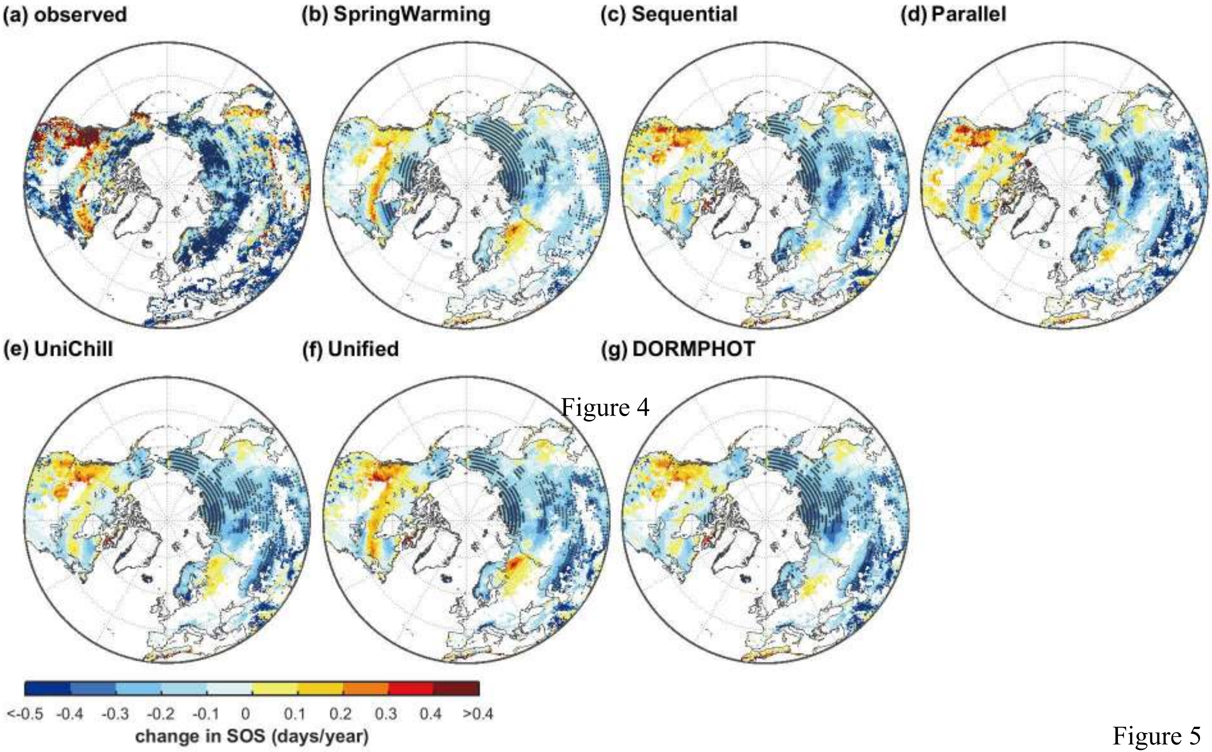

Figure 4

Figure 5

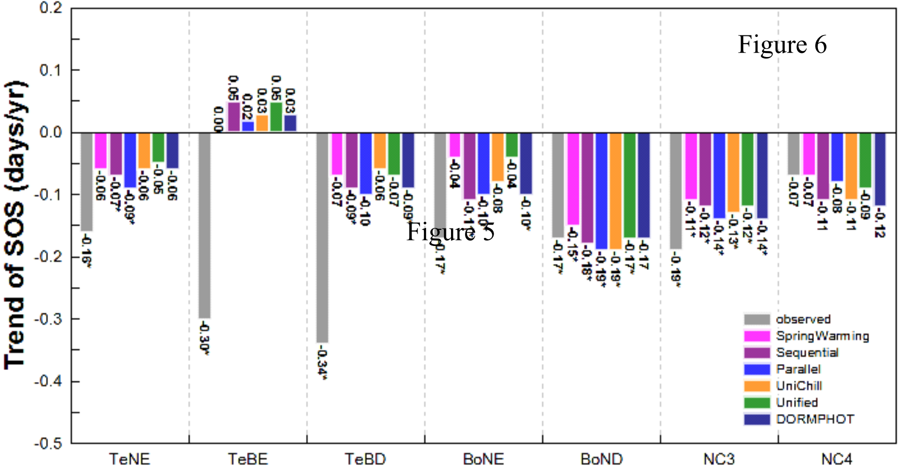

Figure 6

Figure 5

Figure S1. Spatial distribution of the fractional cover of PFTs across the Northern Hemisphere (> 30°N). Seven main PFTs were included: (a) Temperate needleleaf evergreen (TeNE), (b) Temperate broadleaved evergreen (TeBE), (c) Temperate broadleaved deciduous (TeBD), (d) Boreal needleleaf evergreen (BoNE), (e) Boreal needleleaf deciduous (BoND), (f) Natural C3 grass (NC3), and (g) Natural C4 grass (NC4). Color indicates the proportion of a specific PFT in a 0.5° × 0.5° grid.

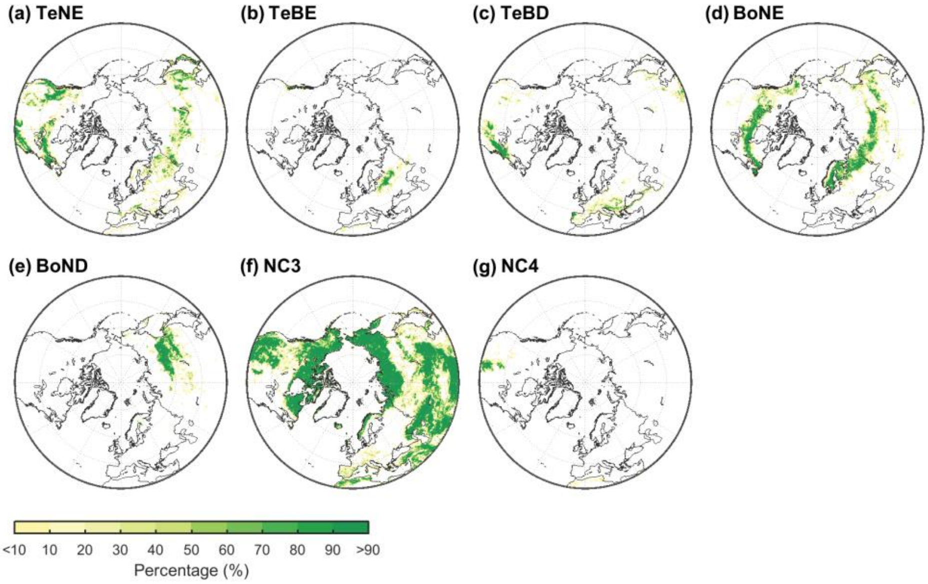

Figure S2. Heat plots showing the relationship between predicted and satellite-derived SOS. SOS data was randomly sampled (10%) from each PFT in the Northern Hemisphere and was excluded from the calibration of spring phenology models. a–p1: predictions were generated using six spring phenology models at the PFT scale. Diagonal lines are 1:1 lines, indicating perfect agreement between predicted and satellite-derived SOS. Colors indicate the percentage of pixels within each bin.

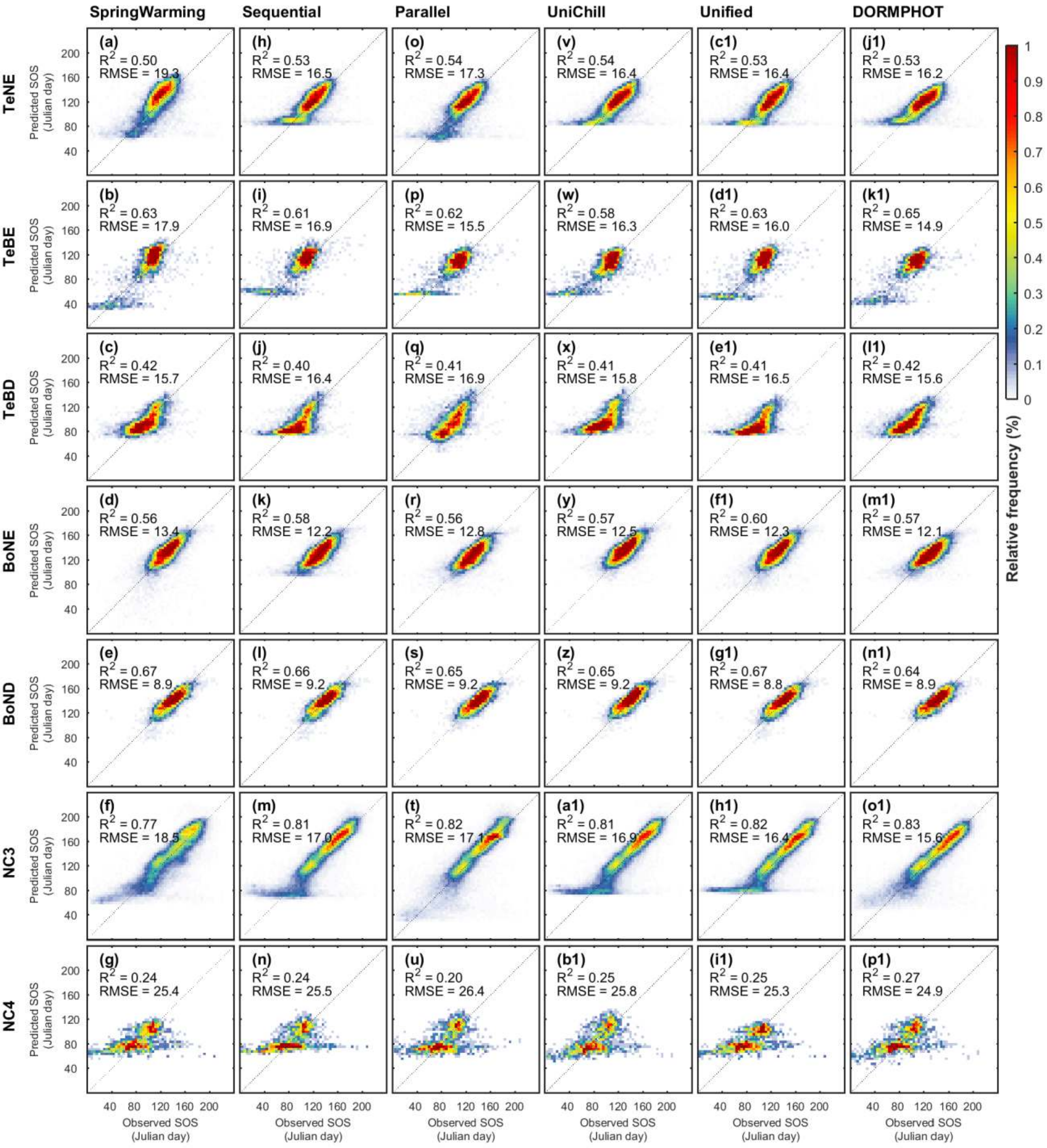

Figure S3. Spatial pattern of averaged SOS during the period 1982–2012 at the PFT scale. a–d was SOS inferred from satellite observations and h–p1 shows the SOS predicted by five phenology models. Pie charts beneath each model indicate the differences between predicted and satellite-derived SOS; a positive value indicates that the model produced later spring onset than observed, and vice versa.

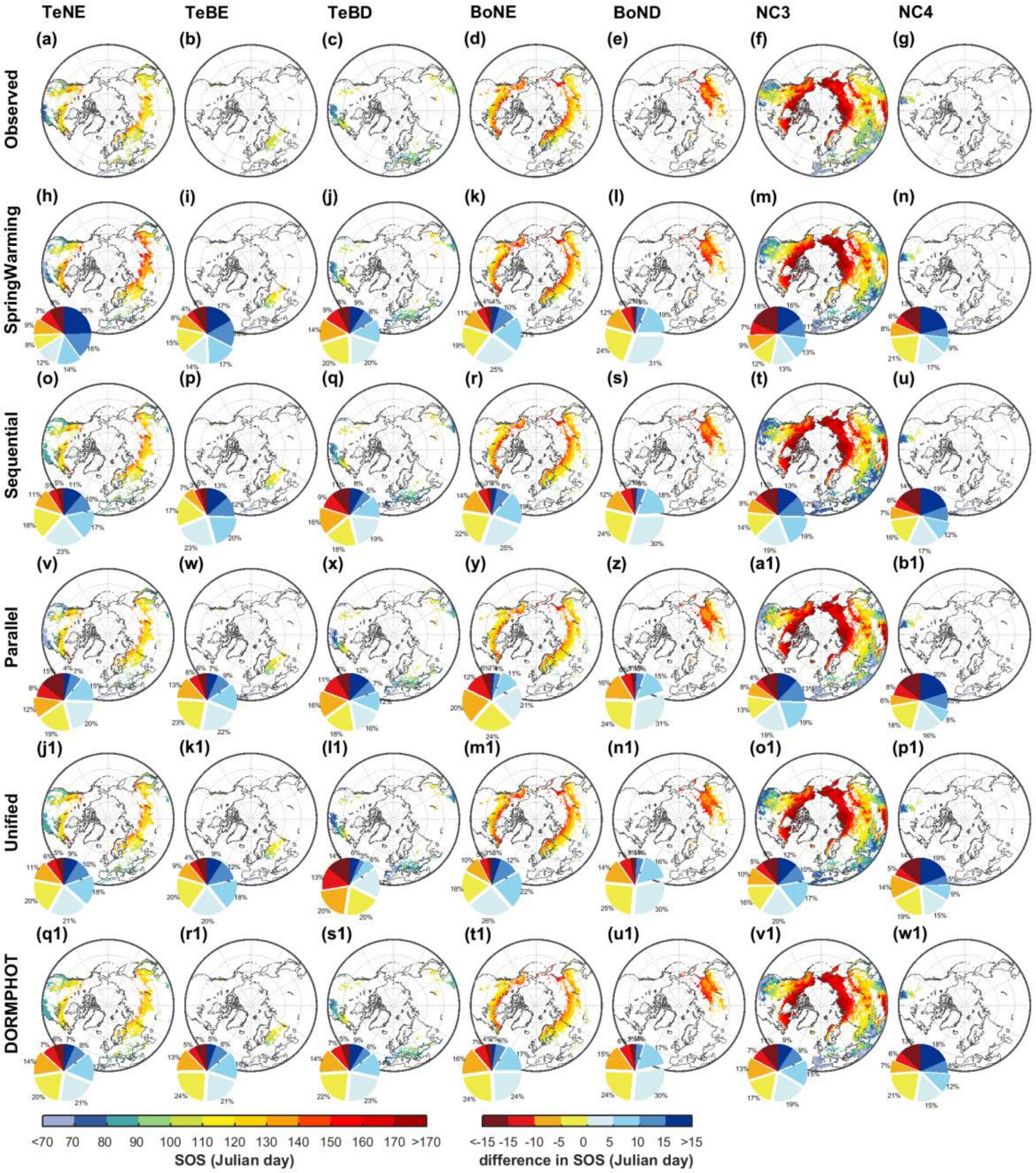

Figure S4. Maps of RMSE (days) between satellite-derived SOS and model simulations at the PFT scale. Results are based on six spring phenology models: (a–g) Spring Warming, (h–n) Sequential, (o–u) Parallel, (v–b1) UniChill, (c1-i1) Unified, and (j1–p1) DORMPHOT. Colors indicate differences between satellite- and model-derived SOS.

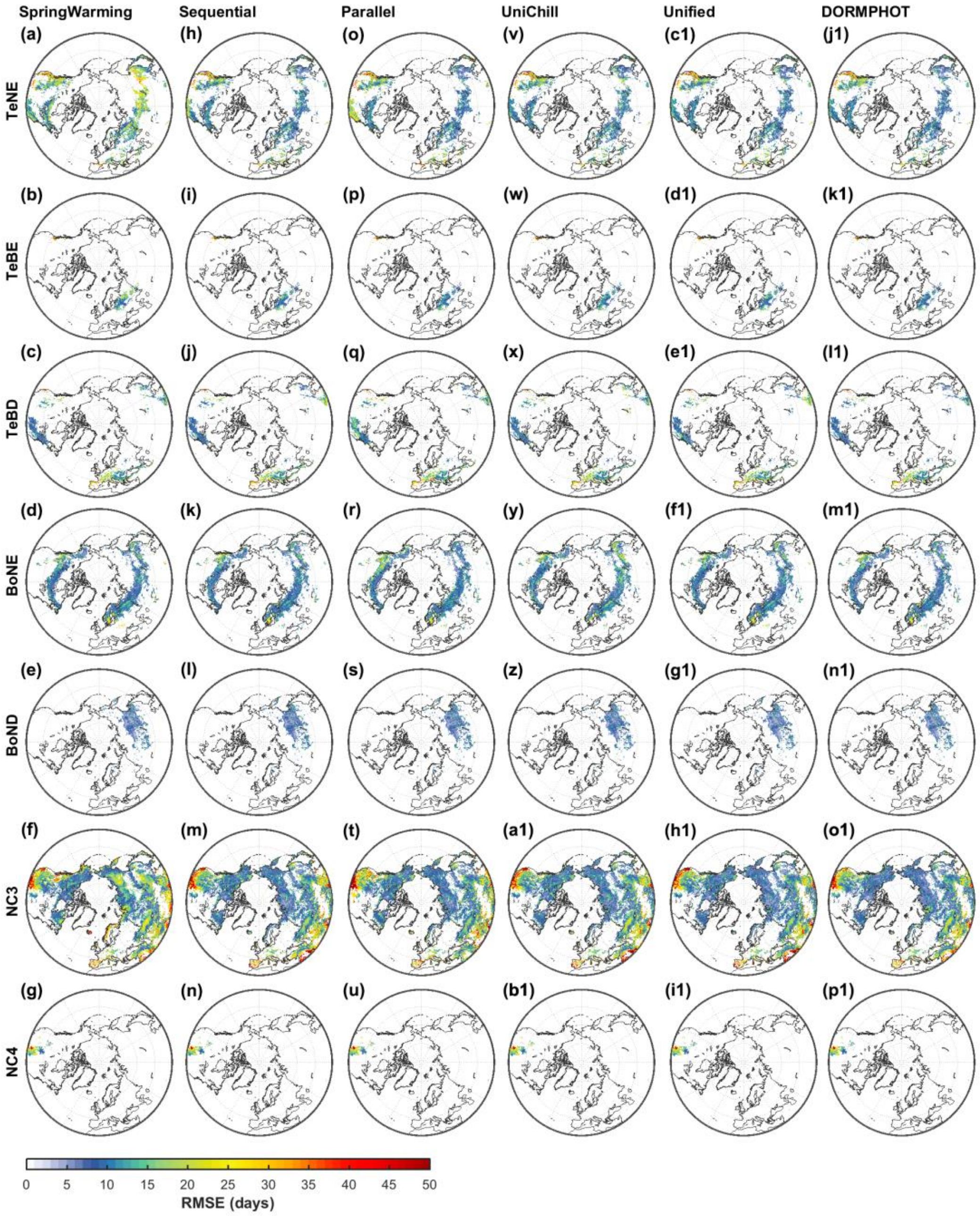

Figure S5. Changes in satellite-derived and model-predicted SOS at the PFT scale during the period 1982–2012. Figures S5a–g show linear trends estimated from satellite-derived SOS, and Figures S5h–w1 show results based on six spring phenology models: (h–n) Spring Warming model, (o–u) Sequential model, (v–b1) Parallel model, (c1–i1) UniChill model, (j1–p1) Unified model, and (q1–w1) DORMPHOT model.

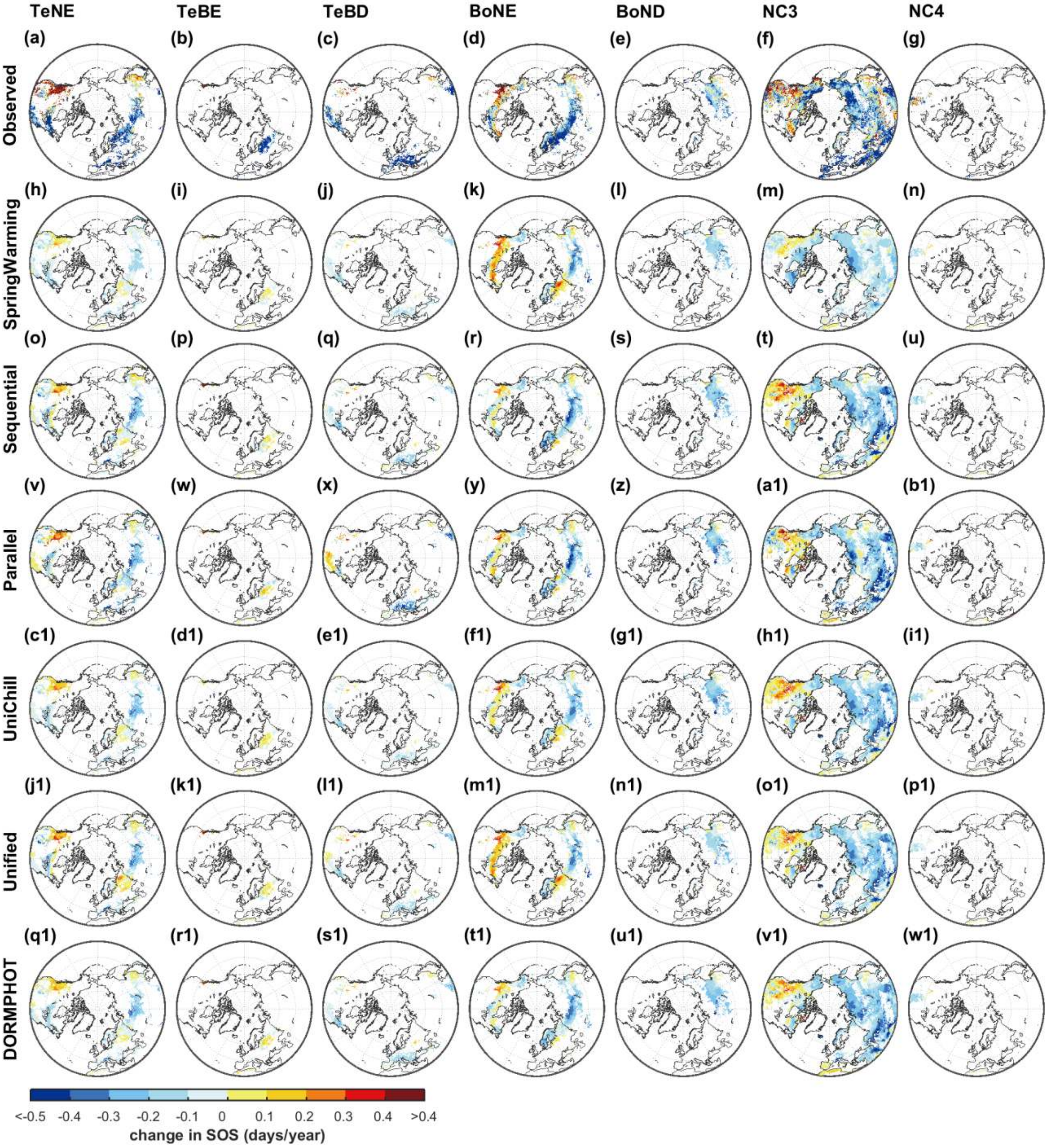

Figure S6. Variation of satellite-derived and model predicted SOS over the period 1982-2012. a-f shows the results across the temperate broadleaved evergreen (TeBE) based on six spring phenology models while g-l indicates the results across temperate broadleaved deciduous (TeBD).

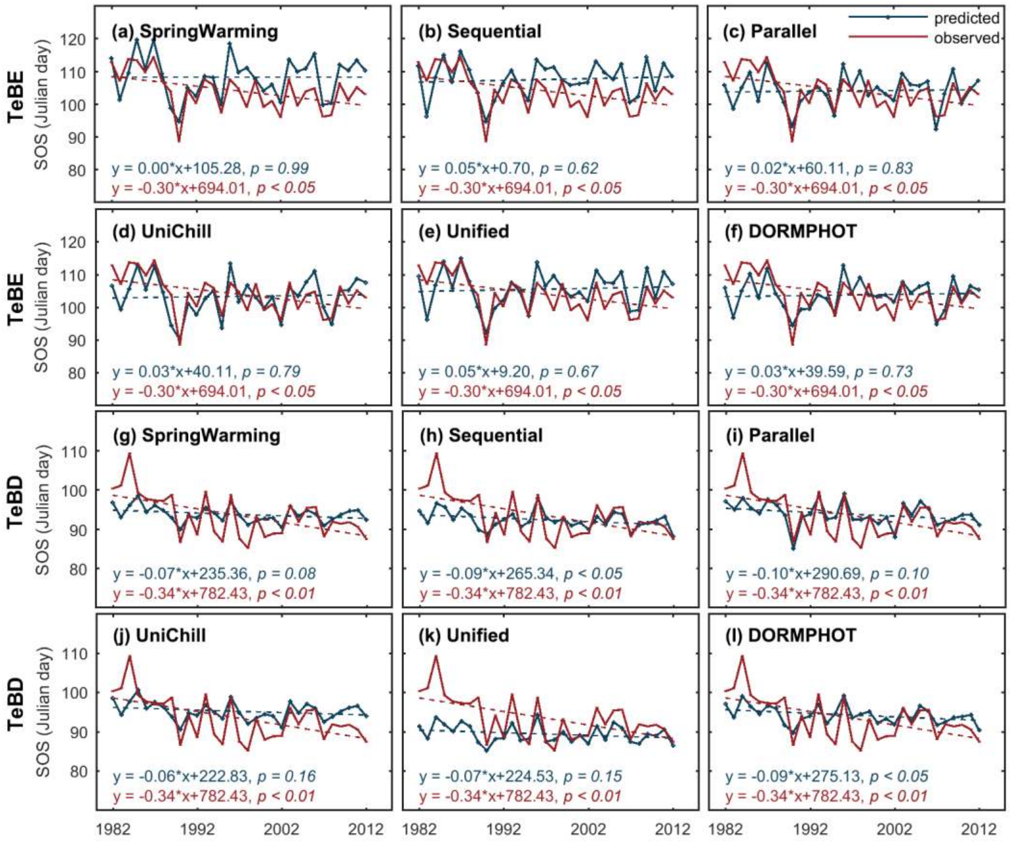

Figure S7. Comparison between individual phenology extraction method and their ensemble mean over the period 1982-2012. Bars and error bars indicate the temporal average and standard deviation of SOS for each PFT, respectively.