MODELING THE PHENOLOGICAL RESPONSE TO CLIMATE CHANGE AND ITS IMPACT ON CARBON CYCLE IN NORTHEASTERN U.S. FORESTS

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

HONG XU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

TRACY E. TWINE

MARCH 2015

Acknowledgements

On the journey toward the completion of my research and dissertation, many people have helped me in various ways. The first person I owe thanks to is my advisor, Dr. Tracy Twine. Without her guidance, support and encouragement, it would have been impossible for me to progress smoothly. She has given me not only plenty of freedom to explore what interest me, but also insightful feedbacks to the work that I have done. I feel very lucky to have a mentor who always believes in me. Second, I would like to thank my friend Dr. Xi Yang, who has the same research interest as I do. Our wonderful discussion has inspired many parts of my dissertation. I would also like to extend my sincere thanks to my committee members Dr. Marvin Bauer, Dr. John Baker, Dr. Peter Snyder and Dr. Joseph Knight, for their participation in my preliminary exam and/or dissertation committee and their valued feedbacks.

Studying abroad can be difficult for most people. But great friendship has made it much easier. I thank my dear friends who have been around sharing my happiness and sadness. Ming Chen, Zewei Song, Lu Hu, Yurong Luan, Jian Sun, Lei Wang, Mu Li, Ding Fei, Yan Wang, Rong Cao, Ming Zhao, Xijia Zhang, Shuang Wang, Shiyang Su, Wei Zhang, Bo Jiang, Ruoshui Zhang, Tuo Chen, Yulin Ye, Qianhui Shi and Sheng Tian. I particularly thank my “Unofficial Dad” James Weiske, who treated me as family and introduced the beauty of Minnesota to me.

I am also grateful to my father Yaling Xu, and my mother Jialan Wang, whom I owe everything I am today. Last but not least, none of this would have been possible without the love and patience of my wife Qianqian Tang.

Dedication

This dissertation is dedicated to my parents, Yaling Xu and Jialan Wang.

Abstract

By controlling the timing of leaf activities, vegetation phenology plays an important role in regulating photosynthesis and other ecosystem processes. As driven by environmental variables, vegetation phenology has been shifting in response to climate change. The shift in vegetation phenology, in turn, exerts various feedbacks to affect the climate system. The magnitude of phenological change and the feedbacks has yet been well understood. The goal of this dissertation is to use phenological model with remote sensing and climate data to quantify historical and future trends in leaf onset and offset in northeastern U.S. forests, and use a dynamic ecosystem model, Agro-IBIS, to quantify the impact of phenological change on terrestrial carbon balance. This dissertation has three major parts. First, six phenological metrics based on remotely sensed vegetation index were evaluated with ground phenological observation in Agro-IBIS. Second, a modified phenological metric was used to parameterize a set of phenological models at regional scale; one model for each of leaf onset and offset were selected to examine historical trends; Agro-IBIS simulations were run to quantify the impact of phenological trends on ecosystem productivities. Finally, downscaled climate projections from global climate models under two emission scenarios were used to drive phenological models to predict the trends in leaf onset and offset in the 21st century; and the impact of photoperiod on leaf onset were particularly examined. The results of this study suggest that remotely sensed phenological metrics can be used to improve phenological models with evaluation and adjustment; advancement of leaf onset and delay of leaf offset in the past have increased productivities and could potentially mitigate the warming

temperature in the future; lack of physiological understanding of the driving factors of phenology such as photoperiod could result in large uncertainties in phenological projections.

Table of Contents

Acknowledgements i
Abstract iii
List of Tables vii
List of Figures viii
Chapter 1 General Introduction 1
Chapter 2: Evaluating Remotely Sensed Phenological Metrics In Agro-IBIS Dynamic Ecosystem Model 4
2.1 Introduction 4
2.2 Method and Material 8
2.2.1 Agro-IBIS model description 8
2.2.2 Evaluation of ground phenology observations 10
2.2.3 Evaluation of remotely sensed phenological metrics 12
2.2.4 Evaluation of the propagation of bias in phenology 16
2.2.5 Errors in simulated productivities caused by biases in phenology 19
2.3 Results 20
2.3.1 Ground phenology reference 20
2.3.2 Remotely sensed phenological metrics 22
2.3.3 Propagation of bias in phenology 24
2.3.4 Impact of bias in phenology on simulated productivities 27
2.4 Discussion 30
2.4.1 Ground phenology reference and Agro-IBIS 30
2.4.2 Remotely sensed phenology 31
2.4.3 Modeled phenology 34
2.4.4 Impact of phenology on simulated productivities 35
2.5 Conclusion 37
Chapter 3: Modeling historical trends in the phenology of northeastern U.S. forests and evaluating their impact on the carbon cycle 39
3.1 Introduction 39
3.1.1 Historical changes in vegetation phenology 39
3.1.2 Phenological control on terrestrial carbon balance 41
3.1.3 Phenological models 43
3.1.4 Goal of this study 47
3.2 Method and Material 47
3.2.1 Remote sensed phenological metrics 47
3.2.2 Parameterization of phenological models at regional scale 50
3.2.3 Retrospective analysis of phenology 53
3.2.4 Agro-IBIS simulations 54
3.3 Results 55
3.3.1 Remotely sensed phenological metrics 55
3.3.2 Phenological models 58
3.3.3 Historical change in leaf onset and offset 61
3.3.4 Impact of phenological changes on terrestrial carbon cycle 64

3.4 Discussion 68
3.4.1 Historical change in phenology 68
3.4.2 Impact of phenological change on the carbon cycle 71
3.4.3 Limitations 73
3.5 Conclusion 75
Chapter 4 Modeling The Phenological Response To Future Climate Change in Northeastern U.S. Forests 77
4.1 Introduction 77
4.2 Materials and Method 79
4.3 Results 82
4.3.1 Historical phenology simulated with CMIP5 data 82
4.3.2 Phenological changes in the 21st century 86
4.3.3 Impact of photoperiod on leaf onset 92
4.4 Discussion 104
4.5 Conclusion 108
Chapter 5 General Conclusion 110
Bibliography 114
Appendix 124

List of Tables

Table 2.1 Description of the methods for retrieving phenology 13
Table 2.2 Mean percentage error between simulated and observed LAI for different simulations 20
Table 2.3 Performance of remotely sensed onset..... 23
Table 2.4 Performance of remotely sensed offset 24
Table 2.5 Performance of modeled leaf onset..... 25
Table 2.6 Performance of modeled leaf offset 27
Table 3.1 Description of the phenology index at Hubbard Brook Experimental Forest.. 50
Table 3.2 Summary of leaf onset models..... 52
Table 3.3 Summary of leaf offset models 53
Table 3.4 Performance of different thresholds for retrieving leaf onset 56
Table 3.5 Performance of different thresholds for retrieving leaf offset 56
Table 3.6 Regional statistics of RMSD and AIC_c for different phenological models 60
Table 4.1 Summary of phenological models..... 80
Table 4.2 Summary of CMIP5 models 81
Table 4.3 Regional average absolute difference (days) in 20-year average phenological dates between simulations with CMIP5 and ZedX data 86
Table 4.4 Regional statistics of RMSD and AIC_c for leaf onset models 95

List of Figures

Figure 2.1 Simulated and observed LAI 21
Figure 2.2 Simulated and observed annual carbon cycle component 21
Figure 2.3 Ground observed and remotely sensed leaf onset and offset 23
Figure 2.4 Ground observed and simulated leaf onset and offset. 26
Figure 2.5 Simulated annual GPP and NEP from Dynamic Onset experiment 27
Figure 2.6 Simulated annual GPP and NEP from Dynamic Offset experiment 28
Figure 2.7 Relationship between errors in phenology and errors in simulated productivities 29
Figure 3.1 Remotely sensed vs. observed leaf onset and offset 56
Figure 3.2 Average leaf onset and offset from remote sensing and the spatial variability for 2000-2007 57
Figure 3.3 RMSD between modeled and remotely sensed phenology for 2000-2007 59
Figure 3.4 Long-term average phenological dates over 1958-2007 62
Figure 3.5 Linear trends in leaf onset and offset 62
Figure 3.6 Trends in average air temperature 63
Figure 3.7 Trends in the difference in simulated productivities between the ONSET and STATIC run 65
Figure 3.8 Trends in the difference in simulated productivities between the OFFSET and STATIC run 67
Figure 3.9 Relationship between productivities and phenology 68
Figure 4.1 Difference (CMIP5-ZedX) in 20-year (1986-2005) average leaf onset simulated using SW 83
Figure 4.2 Difference (CMIP5-ZedX) in 20-year (1986-2005) average leaf onset simulated using SW2 84
Figure 4.3 Difference (CMIP5-ZedX) in 20-year (1986-2005) average leaf offset simulated using DLP 85
Figure 4.4 Total days of change in leaf onset over 2006-2100 simulated using SW under RCP4.5 87
Figure 4.5 Box-and-whisker plot of the changes in leaf onset simulated using SW under RCP4.5 88
Figure 4.6 Total days of change in leaf onset over 2006-2100 simulated using SW2 under RCP4.5 89
Figure 4.7 Box-and-whisker plot of the changes in leaf onset simulated using SW2 under RCP4.5 90
Figure 4.8 Total days of change in leaf offset over 2006-2100 simulated using DLP under RCP4.5 91
Figure 4.9 Box-and-whisker plot of the changes in leaf offset simulated using DLP under RCP4.5 92
Figure 4.10 Total days of change in leaf onset over 2006-2100 simulated using SW under RCP8.5 93

Figure 4.11 Total days of change in leaf onset over 2006-2100 simulated using SW2 under RCP8.5 94
Figure 4.12 Box-and-whisker plot of the changes in leaf onset simulated using SW under RCP8.5 95
Figure 4.13 Box-and-whisker plot of the changes in leaf onset simulated using SW2 under RCP8.5 95
Figure 4.14 Total days of change in leaf offset over 2006-2100 simulated using DLP under RCP8.5 96
Figure 4.15 Box-and-whisker plot of the changes in leaf offset simulated using DLP under RCP8.5 97
Figure 4.16 Difference (SW-SW2) in the change in leaf onset over 2006-2100 under RCP4.5 98
Figure 4.17 Difference (SW-SW2) in the change in leaf onset over 2006-2100 under RCP8.5 99
Figure 4.18 Box-and-whisker plot of the difference (SW-SW2) in leaf onset change under RCP4.5 100
Figure 4.19 Box-and-whisker plot of the difference (SW-SW2) in leaf onset change under RCP8.5 100
Figure 4.20 Difference (SWb-SW) in the change in leaf onset over 2006-2100 under RCP4.5 101
Figure 4.21 Difference (SWb-SW) in the change in leaf onset over 2006-2100 under RCP8.5 102
Figure 4.22 Box-and-whisker plot of the difference (SWb-SW) in leaf onset change under RCP4.5 103
Figure 4.23 Box-and-whisker plot of the difference (SWb-SW) in leaf onset change under RCP8.5 103

Chapter 1 General Introduction

Phenology is the study of the timing of recurrent biological events and causes of their temporal change regarding biotic and abiotic forces (Lieth, 1974). Vegetation phenology (i.e., the timing of plant growth stages) is a robust indicator of climate variation as it is driven by environmental factors such as temperature, precipitation and photoperiod (Rosenzweig et al., 2007). The timing of leaf onset and offset are the two most important phenological phases (i.e., phenophases) in terrestrial ecosystem as they determine the length of growing season, thus regulating photosynthesis and other ecosystem processes. Long-term phenological observations have provided strong evidence that climate change has caused shifts in vegetation phenology (e.g., earlier arrival of budbreak and later leaf fall) during the recent decades, resulting in a longer vegetated season (Chmielewski and Rotzer, 2001; Menzel and Fabian, 1999; Schwartz et al., 2006). These shifts have important implications for ecosystem functions and biosphere-atmosphere interactions. Shifting vegetation phenology may exert feedbacks to the climate system through the effects of biogeochemical processes (e.g., exchange of carbon dioxide, production of biogenic volatile organic compounds, etc.) and biophysical properties (e.g., seasonal variation in albedo) (Peñuelas et al., 2009; Richardson et al., 2013). The magnitude of these feedbacks is still uncertain. In general, extended growing season is expected to mitigate global warming by increasing the terrestrial carbon storage, because extra days are available for carbon assimilation. Efforts have been made to examine the relationship between net ecosystem production (NEP) and vegetation growing season length (GSL). Eddy covariance measurements across sites show that NEP

is positively correlated to carbon uptake period (CUP), which is an alternative measure of GSL defined as the number of days per year when the ecosystem is net carbon sink (Baldocchi, 2008; Churkina et al., 2005). However, some studies have indicated contradictory results when using other definitions of growing season (e.g., canopy duration), which might be caused by different reasons (Dunn et al., 2007; Hu et al., 2010; White and Nemani, 2003). For example, additional carbon gain due to earlier spring could be offset by the enhanced respiration in autumn; earlier spring could also increase transpiration rates, leading to higher moisture stress in summer and reducing the productivity (Richardson et al., 2010). Therefore, quantification of the phenological control on terrestrial carbon balance at regional and landscape scale is necessary for reducing the uncertainties. Dynamic Ecosystem Model (DEM) is a unique tool that can be used to conduct such quantification at a single site or over a large region (Piao et al., 2007; White et al., 1999). However, phenology is usually the most empirical component in DEMs, and not capable of predicting the phenological dates accurately (Richardson et al., 2012). This will almost certainly lead to significant errors in modeling seasonal evolution of leaf area index (LAI) and in turn the productivity, particularly for deciduous forests (Richardson et al., 2012; White et al., 1999). Moreover, in order to understand the phenological feedbacks to the climate system and make more accurate climate projection, it is important to predict how vegetation phenology will respond to future climate change.

In this dissertation, remote sensed phenological metrics were used to parameterize phenological models in the context of a DEM, Agro-IBIS (the Integrated Biosphere Simulator, agricultural version). The phenological models were used with a high-

resolution climate dataset to reconstruct the historical phenology time series of the northeastern U.S. forest. Then a series of Agro-IBIS simulations were conducted to quantify the impact of phenological change on terrestrial carbon balance. Finally, climate projections under different scenarios were used to predict the change in phenology in the 21st century. This dissertation is organized with the first and last chapter focusing on the general introduction and conclusion. Chapter 2 describes a comprehensive evaluation of existing algorithms that can be used to retrieve vegetation phenology in the context of Agro-IBIS model using long-term phenological observation at Harvard Forest. Chapter 3 describes the parameterization of phenological models at regional scale, the reconstruction of historical time series of leaf onset and offset, and the analysis of simulated trends in leaf onset and offset and their impact on terrestrial carbon balance. Chapter 4 describes the response of leaf onset and offset to the projected climate change, and particularly how photoperiod will potentially affect the change in leaf onset in the 21st century.

Chapter 2: Evaluating Remotely Sensed Phenological Metrics In Agro-IBIS Dynamic Ecosystem Model

2.1 Introduction

Vegetation phenology, or the timing of plant growth stages (e.g., the timing of budburst, flowering, leaf coloring), is considered a robust indicator of short-term climate variation and long-term climate trends because it is driven by environmental factors, such as temperature, precipitation and photoperiod. Vegetation phenology has received increased attention recently because evidence from ground observations as well as satellite remote sensing has shown that vegetation phenology has shifted during the past few decades (Menzel and Fabian, 1999; Myneni et al., 1997; Schwartz and Reiter, 2000; Zhou et al., 2001; Zhu et al., 2012), especially at middle and high latitudes of the Northern Hemisphere, as a result of increasing average temperature (Menzel et al., 2006; Peñuelas and Filella, 2001). On the other hand, shifts in vegetation phenology can exert strong control on the feedbacks between the biosphere and atmosphere by affecting biogeochemical processes (e.g., exchange of carbon dioxide, production of biogenic volatile organic compounds) and biophysical properties (e.g., seasonal variation in albedo) of ecosystems (Peñuelas et al., 2009; Richardson et al., 2013). Bias in vegetation phenology therefore may lead to errors in carbon and water exchange and energy budget simulated in dynamic ecosystem models (DEMs) (Richardson et al., 2012) as well as climate patterns simulated in coupled global climate models (GCMs) (Levis and Bonan, 2004).

A multi-model synthesis study has shown that vegetation phenology is poorly represented in many terrestrial biosphere models (Richardson et al., 2012), which highlighted the urgency of improving phenological models embedded in DEMs. Phenological models can potentially be improved by reducing the uncertainties that stem from model structure and model parameters, or drivers (Migliavacca et al., 2012). For example, a comprehensive comparison of existing phenological models across geographic zones may help reduce the structural uncertainties (Yang et al., 2012). Moreover, modeling studies at the regional scale demonstrated that, due to the difference in species type and composition, forests at different locations do not share common parameters, such as base temperature for growing degree day (GDD) calculation (Fisher et al., 2007; Yang et al., 2012). Thus, location-specific parameterization has the potential to reduce the uncertainty associated with model parameters. Parameterization of phenological models at a specific location requires corresponding phenological observations. As ground-based phenological observations are limited in spatial coverage and quantity, phenology derived from remote sensing becomes the only alternative when parameterization over a large continuous area is needed.

Phenology derived from remote sensing, which is referred to as land surface phenology (LSP) recently in order to distinguish it from *in situ* monitoring at species level, has long been used to examine phenological changes (de Beurs and Henebry, 2005; Liang and Schwartz, 2009; Myneni et al., 1997; Zhang et al., 2007; Zhou et al., 2001; Zhu et al., 2012) and to develop large-scale phenology models (Botta et al., 2000; White et al., 1997). Numerous remote sensing methods, such as vegetation index threshold and

curve fitting, have been developed to extract phenological metrics that describe particular timing related to leaf behaviors and photosynthetic activities (Balzter et al., 2007; de Beurs and Henebry, 2008; Fisher et al., 2006; Reed et al., 1994; Tateishi and Ebata, 2004; White et al., 1997; Zhang et al., 2003). Start of season (SOS) and end of season (EOS) (Fisher et al., 2006; Schwartz and Chen, 2002), or onset and offset (White et al., 1997) are two phenological phases (i.e., phenophases) most commonly extracted due to their importance in determining the growing season length (GSL). Some studies also derive more than two phenophases. For example, Zhang et al. (2003) extracted four phenophases including onset of greenup, maturity, senescence and dormancy. Most of the methods used to extract phenological metrics are based purely on time series of the normalized difference vegetation index (NDVI) (Reed et al., 1994; White et al., 1997) and enhanced vegetation index (EVI) (Fisher et al., 2006; Zhang et al., 2003) from various sensors (e.g., Advanced Very High Resolution Radiometer (AVHRR), Moderate-Resolution Imaging Spectroradiometer (MODIS)).

Although phenological metrics derived using different methods share the same name (e.g., SOS), they could actually represent different phenological stages (e.g., the timing when vegetation starts to green up, the timing when vegetation grows the fastest). An intercomparison of SOS retrieved using ten satellite methods shows that the difference between individual methods can reach as large as two months; and two methods were more closely related to ground observations than other methods (White et al., 2009). Although validation against ground observations has been conducted for remotely sensed phenological metrics in many studies (Fisher et al., 2007; Schwartz et al., 2002; White et

al., 2009; Yang et al., 2012; Zhu et al., 2012), the validation process is not standardized because the phenology-monitoring method usually varies among sites, and even the same dataset can be processed differently. More importantly, remotely sensed phenological metrics have not been evaluated in the context of DEMs. In order to improve the accuracy of carbon and water budgets derived from DEMs, there is still a need to define and test phenology transition periods as estimated by satellite sensors (Ahl et al., 2006). Many issues therefore need to be addressed to determine whether a phenological metric can be used as prescribed phenology in a DEM or to parameterize the embedded phenological model. For example, it should be ensured that the choice of phenology references from available ground observations, against which the remotely sensed phenological metrics are evaluated, represent the phenology requirements in a DEM. Otherwise, even if the remotely sensed phenological metrics are able to capture some ground phenological metrics that is selected based on the needs of certain applications, they may not be the appropriate variable to be used in a DEM. Remotely sensed phenological metrics depend not only on the method, but also the data source. When the remote sensing data source changes (i.e., from AVHRR to MODIS), a given method may lose its validity due to the difference between sensors, such as spectral and spatial resolution.

In this study, phenological metrics derived using six satellite methods for temperate deciduous trees were evaluated in the context of a DEM, Agro-IBIS (Foley et al., 1996; Kucharik, 2003; Kucharik et al., 2000), using the long-term phenological observations (O'Keefe) and measurements at the Harvard Forest AmeriFlux site

(http://ameriflux.ornl.gov/fullsiteinfo.php?sid=50). This study aims to establish a systematic evaluation process that can be used for the parameterization of phenology models embedded in DEMs. First, the reference phenological metrics were identified from ground observations according to the definition of phenology in the Agro-IBIS model, and used as prescribed phenology to assess how well Agro-IBIS captures the seasonal evolution of LAI and carbon cycle components. Second, phenological metrics derived from remote sensing data were compared with the ground reference. Then all phenological metrics were used to parameterize the phenology models to examine the propagation of errors during the parameterization and modeling process. Finally, the modeled phenology was used in Agro-IBIS to evaluate the sensitivity of simulated carbon cycle to phenology.

2.2 Method and Material

2.2.1 Agro-IBIS model description

Agro-IBIS is an improved version of the IBIS DEM (Foley et al., 1996; Kucharik et al., 2000), with the capability to represent both natural and managed ecosystems (Kucharik, 2003). The model was developed to simulate the rapid biophysical processes and long-term ecosystem dynamics in response to environmental drivers. It has been evaluated within forests at local and regional scales (Kucharik et al., 2006; Twine and Kucharik, 2008), and has been used for many applications such as the quantification of trends in net primary productivity in the 20th century (Twine and Kucharik, 2009) and climate regulation services of ecosystems throughout the Western Hemisphere (Anderson-Teixeira et al., 2012). The model is designed with a hierarchical conceptual

framework, and includes several sub-models (e.g., land surface module, vegetation dynamic module, soil biogeochemistry module) that are capable of simulating vegetation canopy physics, vegetation phenology, soil physics and hydrology, and ecosystem biogeochemistry.

Agro-IBIS has two critical phenophases for natural vegetation—leaf onset and leaf offset. For temperate deciduous trees, leaf onset and offset are defined as the date when LAI starts to increase from a minimum value, and the date when LAI starts to decrease from the peak value, respectively. The model originally used a simple scheme in which leaf onset and offset were both triggered by a critical temperature threshold (Foley et al., 1996). Currently, the phenology model is modified from the algorithm developed by White et al. (1997), which is based on GDD for leaf onset and the combination of photoperiod and temperature threshold for leaf offset. An evaluation study at three AmeriFlux sites showed that both schemes had poor performance in representing the phenology at the individual site level; simulated leaf onset dates were generally earlier than the observations with biases up to seven weeks, which led to large errors in canopy structure, such as canopy height and maximum LAI, and in turn the carbon and water exchange (Kucharik et al., 2006). Evaluation at the regional scale showed relatively good performance in capturing the LAI evolution in the northern portion of U.S. eastern deciduous forest; however, earlier onsets were also found in the southern portion, which might be a result of the single threshold of GDD used in the model (Twine and Kucharik, 2008). While the regional evaluation supports the argument that parameters of phenological models may vary with geographic location (Fisher et al., 2007), the local

evaluation implies that the applicability of parameters may change with spatial scale (e.g., from continental scale to site scale).

2.2.2 Evaluation of ground phenology observations

Harvard Forest is a mixed forest dominated by red maple (*Acer rubrum*) and red oak (*Quercus rubra*), both of which are cold-deciduous trees. Harvard Forest is one of few sites that report continual phenology observations for a relatively long period. Spring phenology has been observed since 1990 for 33 species (reduced to nine after 2002). Autumn phenology observations started in 1991 and were reduced to 14 species in 2002 (O'Keefe). Spring phenology is recorded as three metrics—percentage of buds on the tree that have broken open (BBRK), percentage of leaves on the tree that are at least 75% of their total size (L75), and percentage of leaves on the tree that are greater or equal to 95% of their final size (L95). Autumn phenology is recorded as the percentage of leaves remaining on the tree that have changed color (LCOLOR), and the percentage of leaves that have fallen (LFALL).

For Agro-IBIS runs, observations from the dominant species (i.e., red maple and red oak) were used to characterize the temperate deciduous tree plant functional type (PFT). Following Yang et al. (2012), each metric was fitted for each individual tree sample (multiple tree samples are observed for each species) using a logistic function. The Day of Year (DOY) when the fitted metrics reached certain amplitude between minimum and maximum at an interval of 10% from 10% to 90% was first calculated (e.g., DOY when 20% of buds have broken denoted by BBRK20, DOY when 30% of leaves have changed

color denoted by LCOLOR30). Then the average DOY of five red maple individuals and four red oak individuals was used to represent the phenology of the site.

A series of Agro-IBIS simulations were run at the Harvard Forest AmeriFlux site (42.5378˚N, 72.1715˚W) to determine which observation-based phenological metrics best represent the leaf onset and offset, and how well the model simulates carbon exchange with those metrics. Simulations were conducted with spring onset and autumn offset prescribed as each combination of the observation-based phenological metrics (e.g., BBRK20 as the onset and LCOLOR20 as the offset). The model was driven with a high-resolution (5 minute latitude/longitude grid, ~ 9km on a side) historical climate dataset created by ZedX Inc., which contains daily values of the six variables required by the Agro-IBIS model—maximum and minimum air temperature, precipitation, incoming shortwave radiation, relative humidity and wind speed, over the conterminous U.S. for the period 1948-2007. More detailed information about the dataset can be found in Motew and Kucharik (2013). Data from the grid cell containing the Harvard Forest AmeriFlux site were used to drive Agro-IBIS. The area where phenology is observed (42.53˚N-42.54˚N, 72.18˚W-72.19˚W) is approximately 1km away from the Harvard Forest AmeriFlux site; however, both the phenology observation and the AmeriFlux site are located within the same grid cell of the climate dataset. It is therefore assumed that there was no variability in phenology within the grid cell. For each simulation, a soil spin-up was conducted so that soil carbon reached near equilibrium. Then the model was run over the period 1948-2007 with phenology simulated using the embedded phenology module for 1948-1990 and prescribed for 1991-2007. The LAI simulated in Agro-IBIS

was compared with the deciduous overstory LAI (i.e., LAI of deciduous canopy without the effect of stems, calculated as the overall LAI measured minus the lowest LAI value during the time when no leaves exist) at the Harvard Forest AmeriFlux site. LAI measurements were taken in 1998, 1999, 2005, 2006, 2007 and 2008. Data from 2005 and 2008 were not included because the measurement records are too few in 2005 (only four) and 2008 is beyond the time period of the climate dataset. The mean percentage error (MPE) between simulated and observed LAI for each simulation was calculated to measure the performance of the model. The combination of spring onset and autumn offset metrics, with which the Agro-IBIS model had the best performance in simulating the LAI (i.e., the lowest MPE), was chosen as the ground reference to evaluate the remotely sensed phenological metrics. We also compared simulated annual average gross primary productivity (GPP), ecosystem respiration (Re) and net ecosystem production (NEP) with the gap-filled (Version 7, Level 2) eddy covariance measurements (Urbanski et al., 2007).

2.2.3 Evaluation of remotely sensed phenological metrics

Six VI-based methods were used to extract onset and offset dates. A brief description of each method is listed in Table 2.1. Although some previously published methods used EVI (e.g., Zhang et al., 2003), and other methods used NDVI (e.g., White et al., 1997), All methods were tested with both VIs. NDVI and EVI were calculated following equations (2.1) and (2.2), respectively, using the 8-Day 500m MODIS surface reflectance product (code: MOD09A1) acquired from the NASA LPDAAC (https://lpdaac.usgs.gov/).

Table 2.1 Description of the methods for retrieving phenology

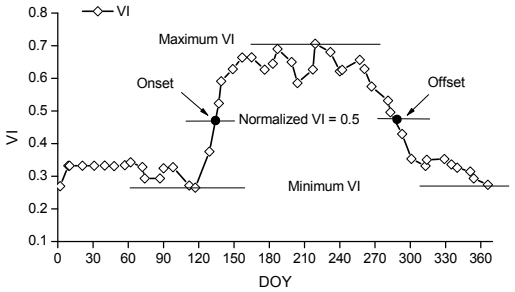
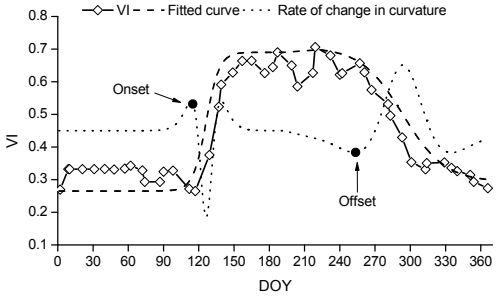
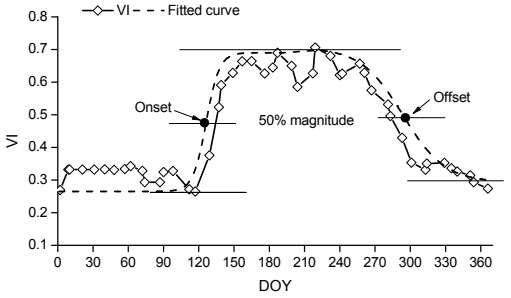
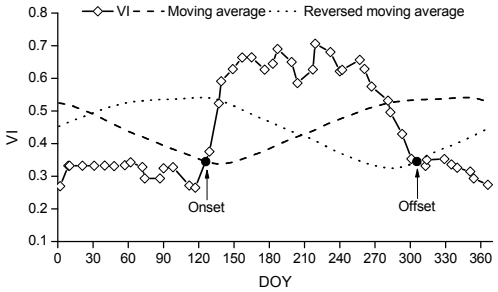
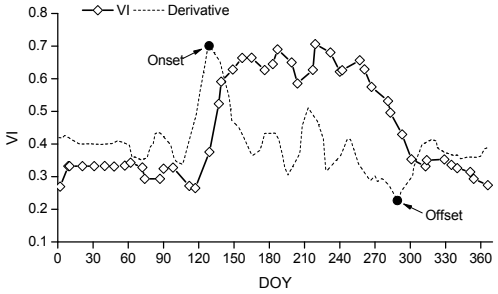
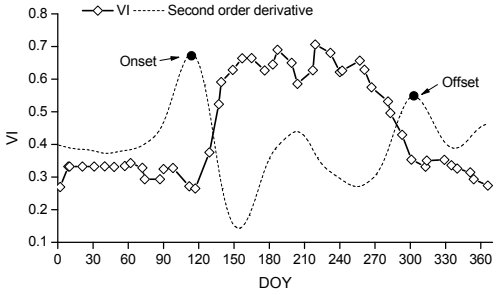
Abbreviation	Description	Example	Reference
MIDPOINT	VI is normalized to a range of 0-1 using following equation. Onset is defined as the DOY when normalized VI exceeds 0.5 in the spring. Offset is defined as the DOY when normalized VI decreases below 0.5 in the autumn. $$VI_N = \frac{VI - VI_{\min}}{VI_{\max} - VI_{\min}}$$		White et al., 1997
LOGISTIC1	VI time series is fitted using logistic function. Then the rate of change in curvature of fitted function is calculated. Onset is defined as the DOY when the rate of change in curvature reaches the first local maximum in the spring. Offset is defined as the DOY when the rate of change in curvature reaches the first local minimum in the autumn.		Zhang et al., 2003
LOGISTIC2	VI time series is fitted using logistic function. Onset is defined as the DOY when fitted VI exceeds 50% amplitude between the minimum and maximum in the spring. Offset is defined as the DOY when fitted VI decreases below 50% amplitude between the minimum and maximum in the autumn.		Fisher et al., 2006

Table 2.1 Cont.

Abbreviation	Description	Example	Reference
MOVING	A new VI curve is established from moving average models with an introduced time lag of 225-days (equivalent to the fifteen 15-day composite used in White et al., 2009). Onset is defined as the DOY when the original VI time series crosses the moving-average curve. Offset is defined the same way as onset with the VI time series reversed.		Reed et al., 1994
DERIVATIVE	The derivative of VI time series is derived by calculating the change in VI with a 20-day moving window (equivalent to the 3 consecutive 10-day composite used in Tateishi and Ebata, 2004). Onset is defined as the DOY when the maximal increase in VI is reached. Offset is defined as the DOY when the maximal decrease in VI is reached.		Tateishi & Ebata, 2004
CAMELBACK	A moving window of 50 days (equivalent to the 5 10-day composite used in Balzter et al., 2007) is passed over the VI time series. The slope of the regression of the VI against time within every window is calculated to establish the first order derivative time series. Then the second order derivative is calculated using the same process and window. Onset is defined as the DOY when the second derivative time series reaches a local maximum and the slope is positive. Offset is determined at the time where the second order derivative reaches a local maximum and the slope is negative.		Balzter et al., 2007

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (2.1)$$

$$EVI = 2.5 \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + 6\rho_{RED} - 7.5\rho_{BLUE} + 1} \quad (2.2)$$

In equation 2.1 and 2.2, ρ_{NIR} is the near-infrared band reflectance, ρ_{RED} is the red band reflectance, and ρ_{BLUE} is the blue band reflectance. Because phenology derived from different data products could produce different results even if the same method were applied (Ahl et al., 2006), phenology derived from MOD09A1 product was compared with that derived from the 16-Day vegetation indices product (code: MOD13A1; Figure A2.1 and Table A2.1 and A2.2 in Appendix) and 8-Day nadir BRDF-adjusted reflectance (NBAR) product (code: MCD43A4; Figure A2.2 and Table A2.3 and A2.4 in Appendix), both also having a spatial resolution of 500m. MOD09A1 showed the best performance (i.e., with inter-annual variability, which is consistent with Ahl et al. (2006)), therefore only results from MOD09A1 were discussed for the remainder of this chapter. In order to show how different data sources may affect the results, the land surface dynamic product (code: MCD12Q2) was included in the comparison, because it was derived using one of the methods evaluated here (i.e., LOGISTIC1, Table 2.1), but with NBAR EVI as input.

The quadratic model (de Beurs and Henebry, 2008), which involves temperature, was not considered in this study, because it is fundamentally different from the nonlinear fitting methods purely based on VI, and it sometimes fails to capture the offset. Methods based on arbitrary thresholds such as NDVI0.2 and NDVI0.3 used in White et al. (2009) (i.e., 0.2 and 0.3 were used as NDVI threshold to determine onset) were also excluded, because the difference between sensors could yield large discrepancies in the range of

VIs. Preliminary investigation showed that the MODIS NDVI at the study site is sometimes larger than 0.3 over the entire course of a year, which makes it impossible to determine the phenological dates.

In order to derive the phenological metrics, an algorithm based on the Savitzky-Golay Filter was first applied with band quality files and state flags to smooth the VI time series (Chen et al., 2004). Then the reconstructed VI time series were interpolated to daily values using a linear model. Dates of onset and offset for 2000-2010 were derived using each method for the five MODIS pixels that are encompassed in the phenology-observation area. The phenological dates averaged over the five pixels were compared with the metrics based on ground observations selected in section 2.2.2. The performance of each metric was evaluated using the root mean square deviation (RMSD, equation (2.3)) and Spearman's rank correlation coefficient (ρ). RMSD describes how close remotely sensed phenological metrics are to ground observations, while the correlation coefficient describes how well the remotely sensed phenological metrics capture the inter-annual variability. In equation 2.3, DOY_{obs} is observed phenology and DOY_{pred} is remotely sensed or simulated phenology.

$$RMSD = \sqrt{\frac{1}{N}\sum\left(DOY_{obs} - DOY_{pred}\right)^2} \quad (2.3)$$

2.2.4 Evaluation of the propagation of bias in phenology

The onset and offset dates derived using different methods with satellite data in section 2.2.3 were used to parameterize an onset model (Sarvas, 1974) (referred to as the 'Sequential' model hereafter) and an offset model (Delpierre et al., 2009) (referred to as

the ‘Delpierre’ model hereafter). The ‘Sequential’ model assumes that leaf onset is triggered when a critical GDD threshold is exceeded after a chilling requirement is fulfilled (equation (2.4)). The ‘Delpierre’ model assumes that both temperature and photoperiod control the senescence process (equation (2.5)). We chose the ‘Delpierre’ model because it has been proven to have relatively good performance (Yang et al., 2012). Although several model structures are available for the onset, only ‘Sequential’ model was used here as an example to show the propagation of bias in phenology, because it has moderate complexity in terms of parameter number. Tests of the ‘Spring Warming’ (Hunter and Lechowicz, 1992) and ‘Parallel’ models (Kramer, 1994) did not change the conclusion.

$$\begin{cases} S_c = \sum_{t_0}^{t_h} R_c(x_t) \text{ if } T_{chill} > x_t \text{ then } R_c = 1 \text{ else } R_c = 0 \\ \text{When } S_c \geq C_{total} \text{ heat accumulation starts} \\ S_f = \sum_{t_h}^{t_h} R_f(x_t) \text{ if } T_{base} \geq x_t \text{ then } R_f = 0 \text{ else } R_c = x_t - T_{base} \\ \text{When } S_f \geq F^* \text{ onset is triggered} \end{cases} \quad (2.4)$$

$$\begin{cases} \text{If } P(t) \leq P_{start} \text{ and } x_t \leq T_{chill}, \text{ then } S_{sen} = \sum_{t_1}^{t_s} R_{sen}(x_t) \\ \text{Where } R_{sen}(x_t) = [T_{chill} - x_t]^x \times [P(t) / P_{start}]^y \\ \text{When } S_{sen} \geq Y_{crit} \text{ offset is triggered} \end{cases} \quad (2.5)$$

In equation 2.4 and 2.5, S_f is the accumulated heat forcing units (degree-day); R_f is the rate of heat forcing (degree-day); S_c is the accumulated chilling units (day); R_c is the rate of chilling (day); x_t is the temperature at time t; T_{base}, is the base temperature (˚C) required by the heat accumulation process; T_{chill} is base temperature (˚C) required by

chilling accumulation process; t_0 is the starting date of accumulation (DOY); t_b is the date of onset (DOY); t_h is the date when the chilling accumulation is completed (DOY); t_s is the date of offset (DOY); F^* is the critical threshold of heating process (onset) (degree-day); C_{total} is the critical threshold of the chilling process (day); S_{sen} is the accumulated forcing units for senescence (°C hour hour^{-1}), R_{sen} is the rate of forcing (°C hour hour^{-1}/day); Y_{crit} is the critical threshold of the senescence process (°C hour hour^{-1}); P_{start} is the photoperiod threshold for senescence process (hour); $P(t)$ is the photoperiod for day t; x, y are parameters for the 'Delpierre' model.

Data used to drive the phenology models include daily temperature from the ZedX dataset and photoperiod, which is calculated as a function of latitude and DOY (Campbell and Norman, 1998). A simple genetic algorithm written in Interactive Data Language (http://www.ncnr.nist.gov/staff/dimeo/idl_programs.html) was applied to optimize the model parameters by minimizing the RMSD between the modeled and remotely sensed phenological dates. Convergence is achieved when RMSD can no longer be reduced or the 100th generation of parameters is reached. All the parameters of the phenological models were optimized (i.e., T_{chill}, C_{total}, T_{base}, and F^* for the 'Sequential' model; P_{start}, T_{chill}, x, y, and Y_{crit} for the 'Delpierre' model). Data from the period 2000-2007, which is the overlap between the references (i.e., remotely sensed phenology available since 2000) and the driving data (i.e., ZedX data, available for 1948-2007), were used for the optimization. The onset and offset dates for 1991-2007 were simulated using the 'Sequential' and 'Delpierre' models, respectively, with parameters optimized using each remotely sensed phenological metric as reference (i.e., 6 methods x 2 VIs = 12 sets of

parameters). As a test of model improvement, the onset and offset dates were also simulated using the default Agro-IBIS phenology algorithm. Then these dates were compared with ground observations in the same manner as the evaluation of remotely sensed phenology.

2.2.5 Errors in simulated productivities caused by biases in phenology

Bias in phenology is known to cause errors in ecosystem processes simulated in DEMs (Richardson et al., 2012). In this study, a series of Agro-IBIS simulations were conducted to examine the sensitivity of simulated GPP and NEP to phenology. First, a control simulation was run with both onset and offset prescribed as the observed phenological metrics (section 2.2.2). Then two sets of experimental simulations were run. In one experiment (Dynamic Onset), offset was prescribed with observations and onset was predicted using the phenological model parameterized with the six remotely sensed phenological metrics (section 2.2.4). In the second experiment (Dynamic Offset), onset was prescribed with observations and offset was predicted. Then GPP and NEP simulated in both model runs were compared. Because all the parameters, settings, and driving data are identical except the phenology, the difference in simulated GPP and NEP can be attributed to differences in phenology. For example, the difference between Dynamic Onset and the control can be attributed to the difference in the date of onset. A regression analysis was conducted to evaluate the relationship between the difference in simulated GPP and NEP and the differences in phenology.

2.3 Results

2.3.1 Ground phenology reference

It is found that Agro-IBIS had the best performance in capturing the seasonal evolution of LAI (i.e., smallest MPE between simulated and observed LAI; Table 2.2) with the onset prescribed as BBRK30 (i.e., the DOY when 30% of the buds have broken) and the offset prescribed as LCOLOR20 (i.e., the DOY when 20% of the leaves have changed color) (Fig. 2.1). These two metrics based on ground observation represent well the beginning of the increase in LAI in the spring and the decrease in the autumn. The rate of increase in simulated LAI is slightly lower than the observation in spring, while the simulated LAI generally decreases faster than the observation in autumn, particularly in 2007. The simulation shows that the model captured the peak value of LAI.

Table 2.2 Mean percentage error between simulated and observed LAI for different simulations

(%)	BBRK10	BBRK20	BBRK30	BBRK40	BBRK50	BBRK60	BBRK70	BBRK80	BBRK90
LCOLOR10	10.46	10.69	10.51	10.79	10.92	11.04	11.03	11.31	11.45
LCOLOR20	9.09	9.09	**8.86**	9.10	9.19	9.30	9.23	9.45	9.48
LCOLOR30	9.45	9.53	9.28	9.51	9.61	9.72	9.63	9.83	9.84
LCOLOR40	10.22	10.26	9.98	10.15	10.22	10.32	10.20	10.38	10.36
LCOLOR50	11.38	11.40	11.11	11.30	11.35	11.45	11.31	11.46	11.41
LCOLOR60	12.63	12.65	12.35	12.54	12.59	12.71	12.55	12.68	12.60
LCOLOR70	14.21	14.22	13.91	14.10	14.15	14.26	14.10	14.21	14.11
LCOLOR80	15.83	15.80	15.48	15.66	15.72	15.83	15.66	15.76	15.65
LCOLOR90	17.98	17.92	17.59	17.76	17.82	17.93	17.75	17.85	17.71

The Agro-IBIS model performed well in simulating annual GPP, Re and NEP (Fig. 2.2). The magnitudes of these variables were reproduced well compared with the eddy covariance measurements. The multi-year average GPP, Re and NEP observations are 1428.2 gC/m2, 1135.1 gC/m2, and 293.1 gC/m2, respectively; while for the Agro-IBIS simulation, they are 1420.1 gC/m2, 1196.7 gC/m2, and 223.4 gC/m2. The model captured the inter-annual variability in GPP relatively well with a correlation of 0.5. In

contrast, the model did not capture the inter-annual variability in Re and NEP, particularly for the last four years of the simulation.

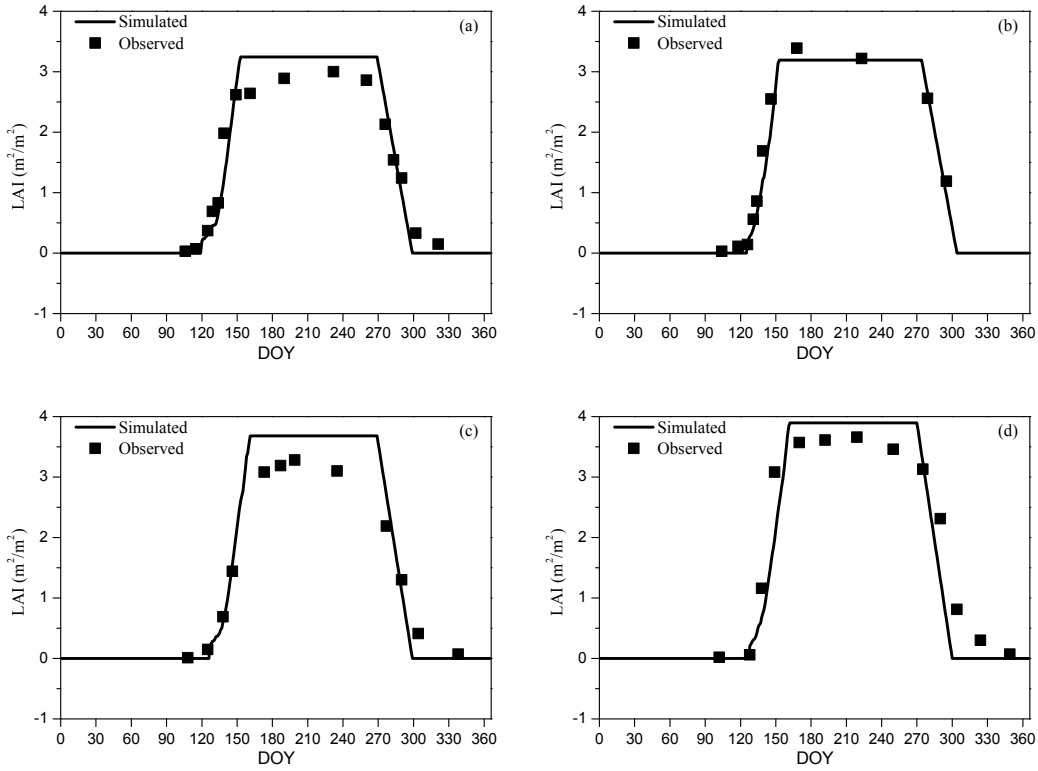


Figure 2.1 Simulated and observed LAI
1998 (RMSD=0.33) (a), 1999 (RMSD=0.24) (b), 2006 (RMSD=0.37) (c), and 2007 (RMSD=0.53) (d). Simulated LAI was computed with the vegetation dynamics module of Agro-IBIS. Observed LAI was measured using a LAI2000 sensor at Harvard Forest AmeriFlux site.

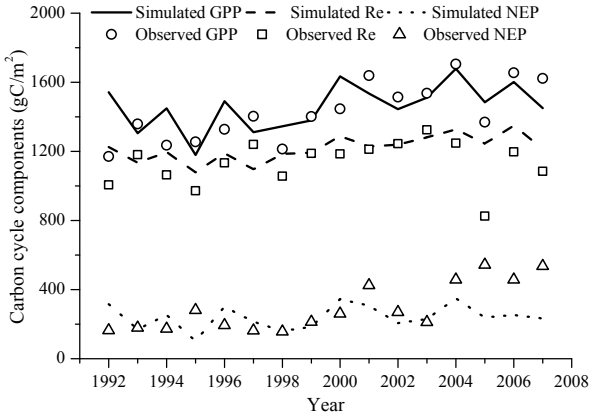


Figure 2.2 Simulated and observed annual carbon cycle component
GPP (RMSD=148.3 gC/m2), Re (RMSD=137.3 gC/m2) and NEP (RMSD=157.7 gC/m2).

2.3.2 Remotely sensed phenological metrics

Figure 2.3 shows the BBRK30 and LCOLOR20 for 1991-2010 along with the remotely sensed onset and offset derived using satellite methods for 2000-2010. At Harvard Forest, BBRK30 varies in the range of DOY112 to DOY135 with an average date of DOY125. In each year, BBRK30 also varies across species and individual trees. The standard deviation fell in the range of 1.1-9.0 days. When NDVI was used to retrieve the leaf onset, LOGISTIC1 and CAMELBACK produced earlier dates than BBRK30 (Fig. 2.3a) with an RMSD of 36.1 and 16.5 days (Table 2.3), respectively. The onset dates retrieved using LOGISTIC2 and MOVING varied around BBRK30 showing the smallest RMSD (less than a week) and relatively high ρ (Fig. 2.3a, Table 2.3). MIDPOINT and DERIVATIVE generally produced onset dates that are later than BBRK30. MOVING had the best performance capturing the inter-annual variability with the highest ρ of 0.54 (Table 2.3), whereas the ρ of LOGISTIC1, CAMELBACK and DERIVATIVE were relatively low (Table 2.3). When EVI was used to retrieve the leaf onset, LOGISTIC1 and CAMELBACK still produced earlier dates although they were closer to BBRK30, whereas LOGISTIC2, MIDPOINT and DERIVATIVE produced later dates (Fig. 2.3b). The RMSD of LOGISTIC2, MIDPOINT and MOVING were larger than those when NDVI was used (Table 2.3). For all the methods, the onset dates derived from EVI have better correlation with BBRK30 than those derived from NDVI (Table 2.3), suggesting that the inter-annual variability was better captured with EVI.

The average date of LCOLOR20 at Harvard Forest is DOY273 over the period of 1991-2010 with relatively small inter-annual variability (Fig. 2.3c). However, in a

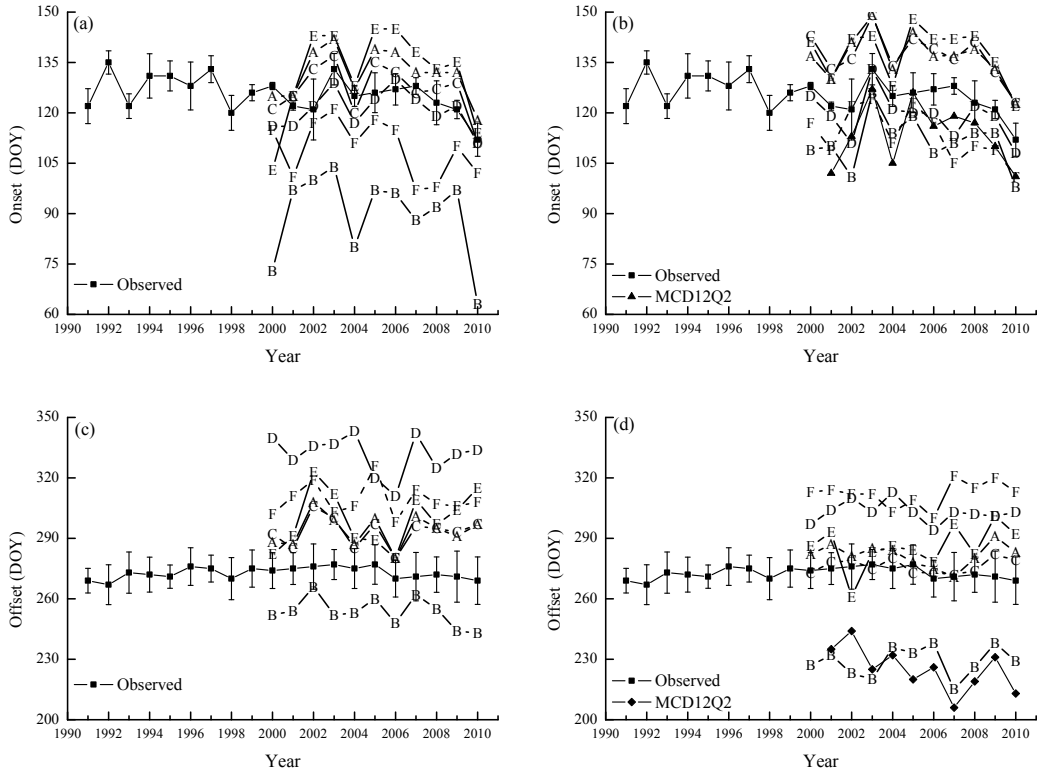


Figure 2.3 Ground observed and remotely sensed leaf onset and offset
Onset from NDVI (a), onset from EVI (b), offset from NDVI (c), and offset from EVI (d). Error bars indicate the standard deviation of observation. Methods used to retrieve phenology are as follows: (A) MIDPOINT, (B) LOGISTIC1, (C) LOGISTIC2, (D) MOVING, (E) DERIVATIVE, and (F) CAMELBACK.

Table 2.3 Performance of remotely sensed onset

Leaf	NDVI		EVI	
Onset	RMSD	ρ	RMSD	ρ
LOGISTIC1	36.1	0.06	13.8	0.44
LOGISTIC2	6.3	0.31	13.3	0.80
MIDPOINT	9.2	0.42	13.1	0.54
MOVING	5.3	0.54	6.5	0.68
DERIVATIVE	14.5	0.28	14.7	0.51
CAMELBACK	16.5	0.30	12.2	0.48

The RMSD and ρ between the onset from MCD12Q2 product and observations are 11.2 days and 0.79, respectively.

particular year, the difference between individual trees is larger than that for BBRK30 (the standard deviation ranges from 5.8 days to 12.7 days). When NDVI was used to retrieve the offset, all the methods produced later dates than LCOLOR20 except for LOGISTIC1 (Fig. 2.3c). The discrepancy was large with the RMSD ranging from 20.6 to

59.1 days (Table 2.4). The correlations between remotely sensed offsets and LCOLOR20 were weak (Table 2.4). A similar pattern was found for the offsets retrieved using EVI (Fig. 2.3d) except that the later dates were closer to LCOLOR20 (i.e., smaller RMSD) whereas the earlier dates (i.e., offset derived using LOGISTIC1) were farther. LOGISTIC2 had relatively good performance as the offsets fell within one standard deviation of the ground observation for most years. LOGISTIC1, DERIVATIVE and CAMELBACK were negatively correlated with LCOLOR20, whereas MOVING showed relatively high ρ (Table 2.4).

Table 2.4 Performance of remotely sensed offset

Leaf	NDVI		EVI	
Offset	RMSD	ρ	RMSD	ρ
LOGISTIC1	20.6	0.53	45.3	-0.25
LOGISTIC2	21.1	0.38	5.2	0.02
MIDPOINT	21.8	0.32	9.8	0.17
MOVING	59.1	0.13	30.0	0.51
DERIVATIVE	29.4	0.03	17.0	-0.38
CAMELBACK	36.4	0.30	39.3	-0.33

The RMSD and ρ between the offset from MCD12Q2 product and observations are 49.1 days and 0.37, respectively.

Moreover, the phenology from the MCD12Q2 product showed similar biases as LOGISTIC1 with EVI calculated using MOD09A1. The RMSD and ρ between MCD12Q2 and the observations are 11.2 days and 0.79 for the onset (13.8 days and 0.44 for LOGISTIC1 with EVI), and 49.1 days and 0.37 for the offset (45.3 days and -0.25 for LOGISTIC1 with EVI)

2.3.3 Propagation of bias in phenology

With different remotely sensed phenological metrics used as reference, the parameters of phenological models showed different capabilities of being optimized. For the ‘Sequential’ model, the RMSD between modeled and remotely sensed onset was

minimized to a range of 4.0-10.2 days. Specific RMSD depended on the combination of method and VI used to retrieve the onset. The RMSD from the 'Delpierre' model ranged from 3.4 to 11.8 days. The modeled phenology generally showed a similar pattern of bias as the remotely sensed phenology used for parameterization in terms of whether it is earlier or later than the ground observation (Fig. 2.4). When the onsets derived from NDVI were used as reference to parameterize the 'Sequential' model, LOGISTIC1, LOGISTIC2, MOVING and DERIVATIVE showed a smaller RMSD (Table 2.5) than that between the remotely sensed onset and BBRK30 (Table 2.3), suggesting that the modeled onsets were closer to the ground observation. The correlation was increased for all the methods except CAMELBACK. When the onsets derived from EVI were used for parameterization, LOGISTIC1, LOGISTIC2, MIDPOINT and DERIVATIVE showed a slightly reduced RMSD; and LOGISTIC1， LOGISTIC1 and CAMELBACK showed a decrease in the correlation coefficient.

Table 2.5 Performance of modeled leaf onset

Leaf	NDVI		EVI	
Onset	RMSD	ρ	RMSD	ρ
LOGISTIC1	34.5	0.20	12.2	0.31
LOGISTIC2	5.7	0.63	12.5	0.68
MIDPOINT	10.2	0.72	12.5	0.68
MOVING	5.0	0.67	9.4	0.67
DERIVATIVE	11.3	0.65	13.5	0.80
CAMELBACK	17.4	0.30	13.3	0.32

The RMSD and ρ between the onset simulated using the original Agro-IBIS algorithm and observations are 10.9 days and 0.43, respectively.

The modeled leaf offsets showed smaller RMSD and higher correlation with the ground observation when the offsets derived from NDVI were used as reference to parameterize the 'Delpierre' model, regardless of methods (Table 2.4, 2.6). In contrast, when the offsets derived from EVI were used as reference, increased RMSD was only

found for LOGISTIC1. MIDPOINT and MOVING showed lowered correlation coefficient, while the other methods showed higher correlations coefficient.

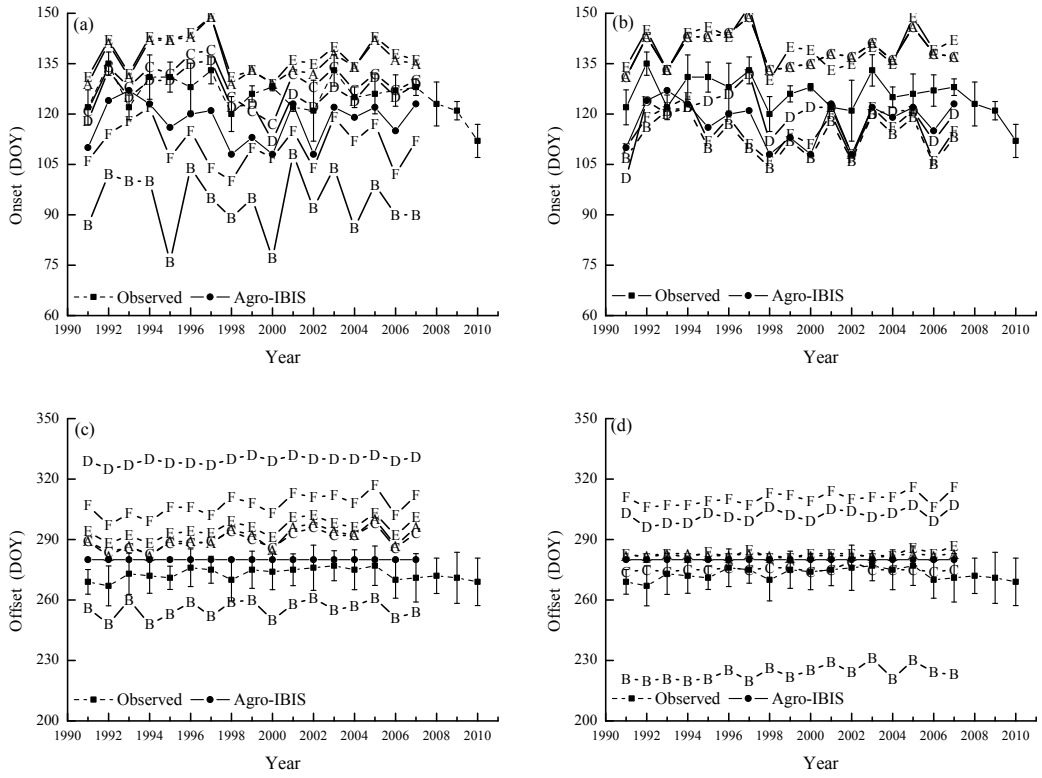


Figure 2.4 Ground observed and simulated leaf onset and offset.
Error bars indicate the standard deviation of observation. 'Agro-IBIS' is the leaf onset or offset simulated using the original Agro-IBIS algorithm and parameters. Onset was simulated using the 'Sequential' model and the parameters were optimized with onset derived from NDVI (a), and EVI (b); offset was simulated using the 'Delpierre' model and parameters were optimized with offset derived from NDVI (c), and EVI (d). Methods used to derive the onset and offset are as follows: (A) MIDPOINT, (B) LOGISTIC1, (C) LOGISTIC2, (D) MOVING, (E) DERIVATIVE, and (F) CAMELBACK.

Moreover, the leaf onset simulated using the Agro-IBIS algorithm was generally earlier than the ground observation (Fig. 2.4) with an RMSD of 10.9 days. The correlation coefficient was 0.43, suggesting the inter-annual variability is not represented well. The leaf offset simulated using the Agro-IBIS algorithm was constant at DOY280 during the simulation period, which implies that the offset is only controlled by photoperiod, even though low temperature is also considered in the algorithm.

Table 2.6 Performance of modeled leaf offset

Leaf Offset	NDVI		EVI	
	RMSD	ρ	RMSD	ρ
LOGISTIC1	18.1	0.53	49.5	0.50
LOGISTIC2	17.1	0.47	3.4	0.40
MIDPOINT	18.1	0.50	9.6	0.04
MOVING	56.3	0.39	29.0	0.23
DERIVATIVE	22.6	0.52	10.3	0.02
CAMELBACK	34.1	0.50	37.3	0.35

The RMSD between the offset simulated using the original Agro-IBIS algorithm and observations is 7.5 days. Correlation is not available because the modeled offset is constant.

2.3.4 Impact of bias in phenology on simulated productivities

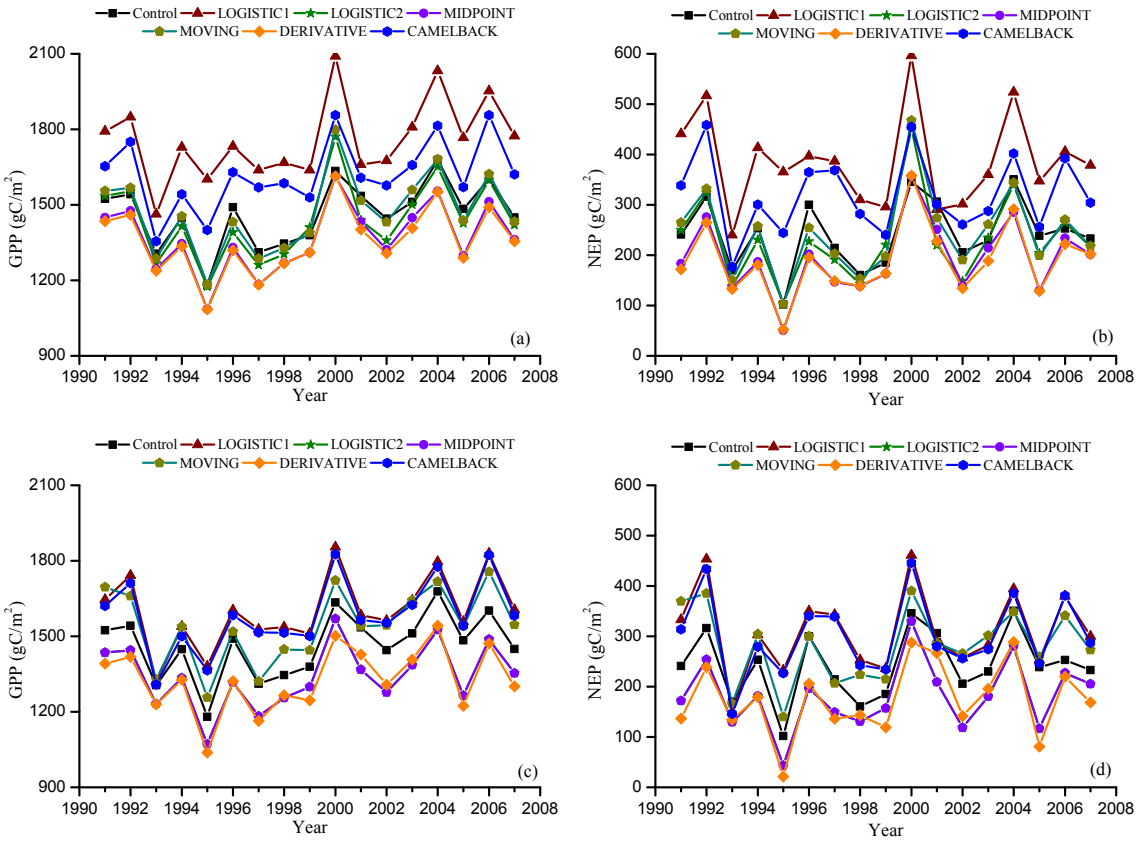


Figure 2.5 Simulated annual GPP and NEP from Dynamic Onset experiment
Leaf onset dates were simulated using the 'Sequential' model with the parameters optimized against remotely sensed onset using NDVI (a)(b), and EVI (c)(d).

Figures 2.5 and 2.6 show the GPP and NEP simulated in the Dynamic Onset and Dynamic Offset runs, respectively. The GPP and NEP from experimental simulations have similar inter-annual variability as the control; however, their magnitude was overall

increased or decreased compared with the control. This corresponds to the overall advanced or delayed phenology in the experimental simulations because the environmental conditions are the same. In general, higher GPP and NEP were found for earlier onset and later offset, mainly due to the extra days of photosynthesis.

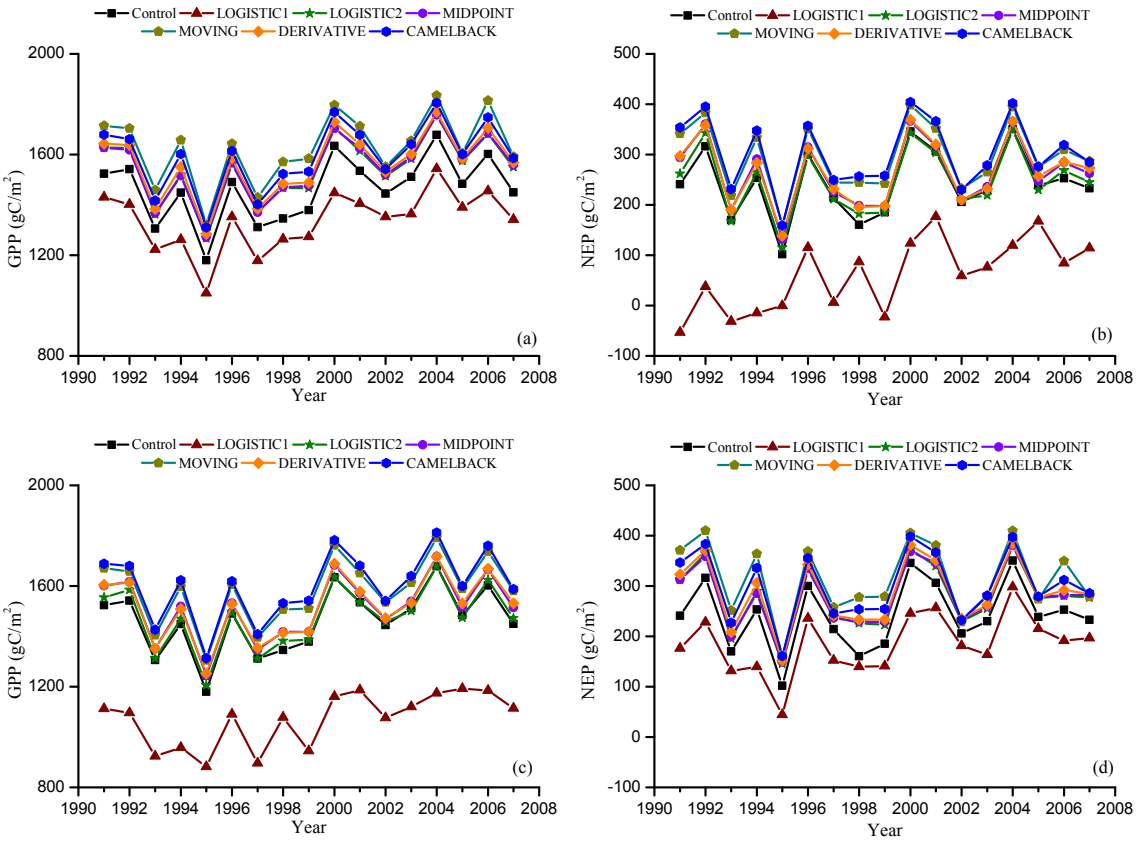


Figure 2.6 Simulated annual GPP and NEP from Dynamic Offset experiment
Leaf offset dates were simulated using the 'Delpierre' model with the parameters optimized against remotely sensed offset using NDVI (a)(b), and EVI (c)(d).

The regression analysis showed a strong negative linear correlation between the bias in the onset (i.e., difference between modeled onset and observed onset) and the errors in simulated productivities (i.e., difference in simulated GPP and NPP between the Dynamic Onset experiment and the control) (Fig. 2.7a and 2.7b). The slopes for GPP (R^2=0.98, P<0.01) and NEP (R^2=0.93, P<0.01) were -9.48 and -5.02, respectively, indicating that a

one-day bias in the leaf onset would result in an error of 9.48 g C m^{-2} yr^{-1} in GPP and 5.02 g C m^{-2} yr^{-1} in NEP.

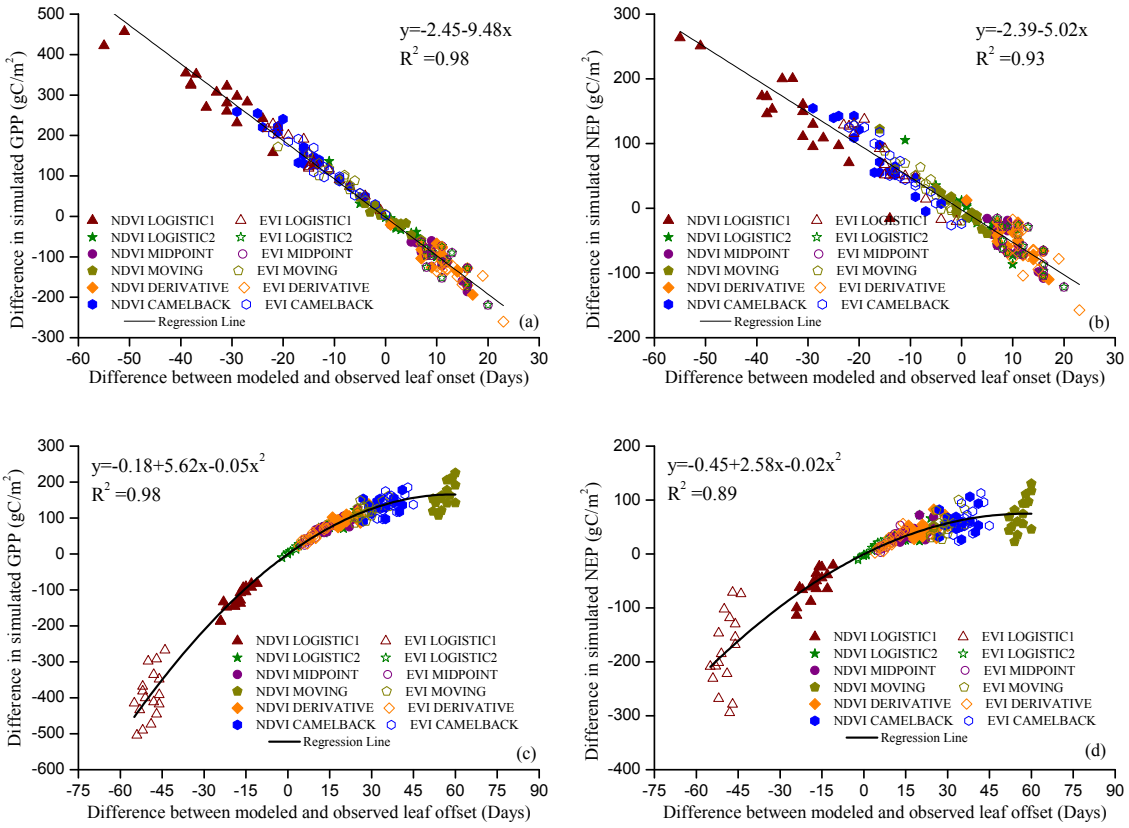


Figure 2.7 Relationship between errors in phenology and errors in simulated productivities Leaf onset vs. GPP (a), leaf onset vs. NEP (b), leaf offset vs. GPP (c) and leaf offset vs. NEP (d). Regressions were conducted using all data in the same category of simulations. P-Value for all regressions is less than 0.01. Symbols indicate the phenology method used. For example, 'NDVI LOGISTIC1' in (a) indicates that the leaf onset dates were modeled using the 'Sequential' model and the parameters were optimized with onset dates derived from NDVI using method 'LOGISTIC1'.

The difference in simulated GPP and NEP between the Dynamic Offset and the control can be represented as a quadratic function of the difference between modeled and observed leaf offset (Fig. 2.7c and 2.7d). As the coefficients of the quadratic term were small (-0.05 for GPP and -0.02 for NEP), the relationship is approximately linear when the bias in offset is small (e.g., when the bias is less than 10 days). The quadratic

relationship also implies that the magnitude of error in productivities resulting from a negative bias (i.e., earlier offset) is larger than that resulting from a positive bias (i.e., later offset). Moreover, the correlation between the errors in simulated GPP and the bias in phenology is slightly stronger than that between the errors in NEP and the bias in phenology. Since NEP is the difference between GPP and Re, this implies that the Re is not as strongly controlled by the phenology as the photosynthesis.

2.4 Discussion

2.4.1 Ground phenology reference and Agro-IBIS

The choice of ground phenology reference was based on the model performance in simulating the evolution of LAI. This method allowed us to choose the ground phenology reference quantitatively. The differences in MPE resulting from different combinations of BBRK and LCOLOR are relatively small (Table 2.1). This is due to the small difference between different phenological levels, which were derived from interpolation. On average, it takes 4.7 days for the buds to break from 10% to 90%. Therefore, in some years, the adjacent two phenological levels (e.g., BBRK20 and BBRK30) could be the same. As it takes longer for the leaves to change color (e.g., averaging 9.5 days from 10% to 50% coloring), the difference resulting from different LCOLOR levels is larger (Table 2.1). Although uncertainties remain in the chosen ground phenology reference, they were confined to a small range based on the choice of metrics. The uncertainties in the offset are larger than those in the onset due to the larger variability in the offset across species and individual trees.

Agro-IBIS captured the magnitude of productivity variables when compared with

observations, although uncertainties exist because of model limitations. For example, the slower increase in LAI in spring compared with the observations (Fig. 2.1) results from a small underestimation in net primary productivity (NPP), a component of simulated LAI. The faster decrease in simulated LAI in autumn might be because of the relatively simple scheme used in Agro-IBIS. Once the offset is triggered, the LAI decreases to a minimum linearly in a 30-day period. However, as the canopy photosynthesis is scaled using LAI in the model, the faster decrease in LAI can partly correct the errors caused by not taking into account the effect of leaf age and coloring.

The discrepancies between the simulated carbon cycle components and the flux measurements can be explained by the following possible reasons: (1) the grid cell was simulated as temperate deciduous forest so that only one PFT existed (Kucharik et al., 2006); (2) the footprints of other PFTs such as evergreen trees and understory shrubs were not simulated although they are likely small; (3) the meteorological data used to drive the model represents the average condition of a grid cell, which could be slightly different from the real condition at the site; and (4) there might also be uncertainties in the flux measurements and the post-processing such as gap-filling (Moffat et al., 2007).

2.4.2 Remotely sensed phenology

With the remote sensing product chosen in this study, including the products shown in the supporting material, most of the six remote sensing methods in this study show relatively poor performance compared with the ground-observed phenology, regardless of which VI is used. The discrepancy can be explained by several possible reasons, within which the difference in definition might be the most contributing. For example, the onset

retrieved using DERIVATIVE is later than BBRK30 (Fig. 2.3a) because the maximum of the first derivative of VI time series represents the time when VI increases the fastest, which usually responds to the period of fast leaf expansion after all buds have broken. LOGISTIC2 and MIDPOINT arbitrarily define the onset as the time when 50% of the amplitude between the minimum and maximum of either fitted or normalized VI time series is reached, which is expected to be later than the time when buds break. LOGISTIC1 and CAMELBACK are both based on the second derivative of VI time series. The local maximum of the second derivative tends to capture the subtle change in the VI, which is too sensitive to the growth of understory that occurs earlier than the development of canopy (Fisher et al., 2006; Richardson and O'Keefe, 2009). The onset retrieved using LOGISTIC1 can also be affected by the curve fitting, because the maximum rate of change in the curvature is determined by the shape of fitted curve (i.e., the parameters of logistic function), and the shape is controlled by how the VI changes when it starts increasing as well as when it reaches the peak.

For the offset, MOVING and CAMELBACK use a process symmetrical to the onset. Because the onset derived using these methods well represents the period when LAI starts to increase from the minimum value, the offset tends to represent the period when LAI drops to the minimum value, which is much later than when LCOLOR20 is reached. LOGISTIC2, MIDPOINT and DERIVATIVE also produce dates later than LCOLOR20 as they tend to capture the period when VI drops the fastest, which usually corresponds to the fast change in leaf color and decrease in LAI rather than the beginning of offset. Although LOGISTIC1 is trying to capture the period when VI starts to decrease, there is

still a difference between the offset derived using LOGISTIC1 and that based on ground observation, because the VI does not change synchronously with LAI. Moreover, similar to the onset, the offset derived using LOGISTIC1 can be affected by the shape of the fitted curve, which is controlled by the change in VI around the period when VI drops to the minimum. Regardless of remote sensing method, offset is later when using NDVI than with EVI. This is likely because NDVI tends to saturate when the LAI is high (3 or larger for a pure forest pixel) (Birky, 2001; Lüdeke et al., 1991) so that it is not as sensitive as EVI to the drop in LAI. Generally, EVI is more responsive to the canopy structural variation, such as LAI, and NDVI is more sensitive to chlorophyll (Huete et al., 2002).

There is a difference between the phenological dates retrieved using LOGISTIC1 with EVI, and those from MCD12Q2, which was produced using LOGISTIC1 as well with NBAR EVI. This highlights the fact that the phenology retrieved can also be affected by factors other than the method, such as the choice of data source (see Appendix) (Ahl et al., 2006) and data processing. On the other hand, the similar patterns shown by the two results suggest that the general features of a certain method are relatively independent of data source. Another issue related to the satellite methods is the parameterization. The width of the moving window used in MOVING, DERIVATIVE and CAMELBACK as well as the 50% amplitude used in LOGISTIC2 and MIDPOINT can be considered as parameters. Because those parameters were chosen based on the input data used when developing the method, when the data source changes, they may no longer be optimal and can contribute to the discrepancies between the remotely sensed

phenology and the ground observation. Preliminary investigation suggests that, by adjusting parameters, it is possible to reduce the discrepancy between remotely sensed phenology and the ground phenology reference at a specific site (see Figure A2.3 in Appendix). However, these parameters must be evaluated at other locations. Currently, long-term ground observations of phenology are limited, and the phenology is usually recorded in different ways at different locations (e.g., Harvard Forest vs. Hubbard Brook Experimental Forest), which causes evaluation to be confounded. Recently, digital cameras have been widely installed to observe vegetation phenology. Since the phenological information gathered from digital camera can be standardized, it can potentially be used to validate the phenology retrieved from satellite imagery at multiple locations, and help resolve the issue of scale difference between ground observations and satellite imagery (i.e., individual trees vs. pixels) (Sonnentag et al., 2012).

2.4.3 Modeled phenology

Differences in RMSD and correlation coefficient between the modeled and observed phenology have been found in comparison with those between the corresponding remotely sensed phenology and ground observation. However, there is no evident pattern of whether the RMSD and correlation coefficient would increase or decrease, which suggests that the discrepancy between modeled and remotely sensed phenology may either add to or offset the discrepancy between remotely sensed phenology and ground observations. Overall, the magnitude of differences in the RMSD is small (usually less than 1 day for the onset, and less than 5 days for the offset). Thus, the RMSD between modeled and observed phenology is still on the same order as the RMSD between the

phenology used for parameterization and observed on the ground. In other words, the bias in remotely sensed phenology is generally maintained by the modeled phenology. The magnitude of changes in correlation varies in a relatively wide range (Table 2.3 vs. Table 2.5 and Table 2.4 vs. Table 2.6), because the correlation is not considered in the cost function of the genetic algorithm. In most cases, correlation became lower suggesting the capability of capturing the inter-annual variability is weakened after the modeling process. Even though correlation became higher in some cases, it does not necessarily indicate that the capability of capturing the inter-annual variability has been improved. If the modeling period were extended, the correlation might further change. This highlights the fact that there is still a need to evaluate whether the phenological models and the optimized parameters can properly capture the changes in phenology in response to the changing climate, even if the RMSD is minimized. One possible solution is to maximize the correlation between modeled phenology and the reference and minimize the RMSD during the optimization process, which requires developing a cost function that incorporates the both metrics.

2.4.4 Impact of phenology on simulated productivities

The analysis indicates that errors in simulated GPP and NEP result from the bias in simulated phenology (Fig. 2.5 and 2.6). Although the sign of NEP did not change due to the bias in phenology, as the study site is a relatively large carbon sink, this might not be the case at other locations that are closer to carbon neutral. The relationship between the errors in simulated GPP and NEP and the bias in phenology also has implications on the impact of phenological shifts on carbon assimilation. The linear relationship between

errors in GPP and NEP and the bias in leaf onset means that a one-day advance in leaf onset would result in an increase of 9.48 g C m^{-2} yr^{-1} in GPP and 5.02 g C m^{-2} yr^{-1} in NEP (Fig. 2.7a and 2.7b). The quadratic relationship between errors in GPP and NEP and the bias in leaf offset (Fig. 2.7a, 2.7b) suggests that delayed leaf offset leads to higher GPP and NEP. However, the marginal increase in GPP and NEP declines with the days of delay. This might be because the environmental condition becomes less and less favorable for carbon uptake late in a year.

Because vegetation growing season length (GSL) is determined by leaf onset and offset, productivity is usually correlated with GSL. Several studies have shown significant control of GSL on the productivities. For, example, a modeling study found that an extension of one day in GSL would result in an increase of 9.8±2.6 gC m^{-2} yr^{-1} in GPP for temperate deciduous broadleaf forest (Piao et al., 2007); carbon flux measurements showed an increase of 5.57 m^{-2} yr^{-1} in NEP with a one day extension of carbon uptake period (i.e., number of days when it is a net carbon sink), which is an alternative definition of GSL (Baldocchi, 2008). These relationships are similar to the impact of leaf onset on the productivities estimated from simulations in this study. This might be because the variation in GSL in those studies is dominated by the variation in leaf onset (Piao et al., 2007) as the regression analysis indicates that the leaf offset controls the productivities in a different way than leaf onset, and its impact is generally smaller. Thus, it might be better to examine the individual impact of leaf onset and offset instead of the overall impact of GSL in future studies.

2.5 Conclusion

In this study, long-term phenological observations were used along with LAI and carbon flux measurements made at ground level to evaluate remotely sensed phenological metrics and simulated phenology, GPP, and NEP using the Agro-IBIS dynamic ecosystem model. The results show that the phenological phases that the evaluated methods represent have relatively large discrepancies compared with the ground phenology reference chosen according to the definition in Agro-IBIS. With the input data presented (i.e., MOD09A1), only two methods for leaf onset and one method for leaf offset show a bias of less than a week. The discrepancies can be attributed to the definition, parameters used for a certain method as well as the input data. When the remotely sensed phenological metrics are used to parameterize phenological models, the bias is generally maintained in the modeled phenology, while the capability of capturing the inter-annual variability was entirely changed. According to simulations in this study, large errors in simulated GPP and NEP resulted from the bias in phenology, and had different relationships with leaf onset and offset. The same magnitude of bias in leaf onset led to larger errors in simulated productivity than in leaf offset, which implies that changes in leaf onset have a larger impact on carbon assimilation than leaf offset. The results suggest that, for the purpose of directly specifying the phenology in DEMs or parameterizing the embedded phenological models, the method used to retrieve phenology should be developed or adjusted based on the definition of phenophases in DEMs as well as the characteristics of satellite data, such as temporal and spatial resolution. Validation should also be conducted at different locations, for which

phenological observations based on digital cameras have the potential to provide standardized information. The results do not imply that those methods with poor performance in this study are invalid. They could still be useful for other applications as they represent different phenophases. Because it is important for phenological models to capture the inter-annual variability and in turn the long-term trend in phenology as well as the absolute timing, metrics that quantify such capability might need to be included in the process of optimizing model parameters. The different control of leaf onset and offset suggests that GSL might not be the best metric to represent the impact of phenology. Moreover, the relationship between the errors in simulated productivities and bias in phenology might not be able to represent other locations, as the environmental conditions vary. To quantitatively understand the phenological control on carbon assimilation and other ecosystem processes, there is still a need to conduct simulations at regional or larger scales with well-represented phenology.

Chapter 3: Modeling historical trends in the phenology of northeastern U.S. forests and evaluating their impact on the carbon cycle

3.1 Introduction

Vegetation phenology responds to both short-term climate variation and long-term climate change as driven by environmental factors including temperature, precipitation, and photoperiod (or day length). Vegetation phenology also controls many feedbacks to the climate system by influencing biogeochemical processes (e.g., exchange of carbon dioxide, production of biogenic volatile organic compounds) and biophysical properties (e.g., seasonal variation in albedo, evapotranspiration) of ecosystems (Peñuelas et al., 2009; Richardson et al., 2013). The important role of vegetation phenology as both indicator and regulator of global change has been increasingly acknowledged. Although efforts have been made to examine the historical change in vegetation phenology as well as the feedbacks to the climate system, there are still large uncertainties due to limited phenological observations and uncertainties in phenological models.

3.1.1 Historical changes in vegetation phenology

Phenological observations show that vegetation phenology has shifted during the past several decades in response to climate change. Advanced spring phenology (e.g. leaf unfolding) is widely reported in the middle and higher latitudes. For example, spring events have advanced 5~8 days in Europe and North America in the last few decades of

the 20th century (Chmielewski and Rotzer, 2001; Menzel and Fabian, 1999; Schwartz and Reiter, 2000). At the global scale, two meta-analyses showed that spring phenology of wild plants and animals has become earlier at a rate of 2.3 days per decade (Parmesan and Yohe, 2003) and 5.1 days per days per decade (Root et al., 2003), respectively. The difference in rate might be due to the studies included and methods used in their analyses.

Compared with ground phenological observation, satellite remote sensing has advantages in monitoring the seasonal change and inter-annual variability of vegetation properties over large spatial extent. Satellite imagery from different sensors, such as the Advanced Very High Resolution Radiometer (AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS) has been widely used in phenological studies. Various approaches have been developed to capture the phenophases of vegetation (Myneni et al., 1997; Reed et al., 1994; Zhang et al., 2003). While some studies showed no significant trends (Zhu et al., 2012), most of the phenological assessments based on remote sensing have detected an earlier spring onset of vegetation greenness. For example, an advance of eight days in the start of the vegetation growing season was found for 1981-1991 in the northern high latitudes (Myneni et al., 1997). An advance of eight days in North America and six days in Eurasia was also found over 1981-1999 (Zhou et al., 2001). Another study showed an earlier trend of 5.4 days per decade for spring greening in Europe over the period of 1982-2001 (Stockli and Vidale, 2004). The earlier trends in spring phenology were well correlated with the increasing temperature (Chmielewski and Rotzer, 2001; Menzel and Fabian, 1999; Menzel et al., 2006; Peñuelas and Filella, 2001).

Compared to spring phenology, the shifts in autumn phenology have been observed but appear to show less robust patterns. Menzel and Fabian (1999) found a delay of 4.8 days for leaf coloring in Europe over 1959-1993, while Chmielewski and Rotzer (2001) only detected a slight delay for the period 1969-1998. A study of four deciduous tree species in Germany showed that leaf coloring in some areas exhibited acceleration whereas others showed delay, which might be because leaf coloring is triggered by a set of factors other than temperature alone (Estrella and Menzel, 2006).

In contrast to ground observations, results from satellite observations consistently found that the start of autumn was delayed in the late 20th century. An analysis across the entire northern high latitudes found a prolongation of 5 days per decade for the declining phase of vegetation growing season over (Myneni et al., 1997), while studies of North America have found delays of 2.2 days per decade (Zhou et al., 2001) and 5.5 days per decade (Zhu et al., 2012), and in Europe of 9.6 days per decade (Stockli and Vidale, 2004), and 5.5 days per decade in North America over 1982-2006 (Zhu et al., 2012).

3.1.2 Phenological control on terrestrial carbon balance

Due to the changes in vegetation phenology, the vegetation growing season (i.e., the period between bud burst and leaf fall) has been lengthened by 10–20 days in the last few decades of the 20th century, which is mostly characterized by an earlier spring onset (Linderholm, 2006; Menzel and Fabian, 1999). Among the feedbacks of the changes in vegetation phenology to the climate system (Peñuelas et al., 2009; Richardson et al., 2013), the influence on the terrestrial carbon balance has received the most concern. The relationship between net ecosystem production (NEP) and vegetation growing season

length (GSL) has been widely examined. Eddy covariance measurements across sites generally support the assumption that a longer growing season increases the NEP because of extra days for carbon assimilation. For example, Baldocchi (2008) found that annual NEP increases by 5.57 g C m^{-2} with a one day extension in carbon uptake period (CUP, i.e., the number of days when the ecosystem is a net carbon sink) for deciduous broadleaf forests. Churkina et al. (2005) also found strong linear correlations between NEP and CUP based on data from FLUXNET sites (5.77 g C m^{-2} per day for deciduous broadleaf forests, 3.37 g C m^{-2} per day for evergreen conifer forests, and 7.91 g C m^{-2} per day for herbaceous vegetation).

Modeling studies have confirmed the positive correlation between NEP and CUP, but not a significant correlation between NEP and GSL (Piao et al., 2007; White and Nemani, 2003). This might be explained by the effect of enhanced soil carbon decomposition induced by warming, which offsets the increased gross primary productivity (GPP) resulting from a prolonged growing season (Piao et al., 2007). Analysis based on flux measurements at 21 FLUXNET sites has shown the same mechanism, except that the increase in GPP is greater than the concurrent increase in ecosystem respiration (RE), leaving an overall positive effect of prolonged vegetation growing season on NEP (Richardson et al., 2010). A 10-year observation at the Morgan-Monroe State Forest AmeriFlux site has shown an increasing trend in annual NEP of about 10 g C m^{-2} yr-1 (Dragoni et al., 2011). Half of the trend can be explained by a longer vegetated season, and the other half can be explained by the decrease in ecosystem

respiration during the winter due to a negative trend in air and soil temperature (Dragoni et al., 2011).

In some cases, NEP is negatively correlated with GSL. For example, Hu et al. (2010) found a negative relationship between GSL and NEP in a subalpine forest, which is due to the dependence of forest carbon uptake on snow melt water, and a longer GSL usually results from a shallower snowpack. These studies suggest that phenological control on the carbon balance is complicated. It is the combination of changes in phenology and environmental drivers rather than GSL alone that determines the change in NEP. Therefore, although CUP is controlled by GSL (Churkina et al., 2005), it may be a better indicator of NEP as it already includes the effect of environmental conditions. Furthermore, as environmental conditions play a role in the phenological control on the carbon balance, the relationship between NEP and GSL derived at one location may not be representative for other locations. Therefore, it is important to separate the effect of phenological shifts from that of other environmental changes, as well as to examine the dependence of the relationship between NEP and GSL on environmental conditions. However, these studies are rare and usually conducted at a site level (White et al., 1999). There is still a need to quantify the phenological control on terrestrial carbon balance and its spatial variability at regional to continental scales.

3.1.3 Phenological models

Dynamic ecosystem models (DEMs) can be used to quantify the phenological control on various feedbacks at local and regional scales, because phenological models are embedded in DEMs to predict the timing of leaf onset and offset and simulate the

seasonal evolution of LAI. The accuracy of phenological models embedded in DEMs is therefore critical in such application. Keenan et al. (2012) demonstrated that failure to accurately predict phenology is a key reason why many DEMs are unable to predict inter-annual variability in either GPP or NEP.

Phenological models embedded in DEMs may vary in their structure, depending on the assumption of the driving environmental factors and the mechanisms behind. It is widely accepted that leaf onset in temperate and boreal forests is mainly driven by temperature (Cannell and Smith, 1983; Peñuelas and Filella, 2001; Polgar and Primack, 2011). For example, in the CLASS (Canadian Land Surface Scheme) model, the breaking of dormancy is triggered when both air temperature and the temperature of the top soil layer first exceed 0ºC (Verseghy et al., 1993). Both the original version of IBIS (Foley et al., 1996) and LPJ (Lund-Potsdam-Jena Dynamic Global Vegetation Model) (Sitch et al., 2003) use 5°C as the threshold.

Another category of leaf onset models is based on thermal time. Leaf onset is triggered when accumulated growing degree-days (GDD) reaches a critical value (often called critical forcing temperature). Some models such as 'Sequential' (Sarvas, 1974), 'Parallel' (Kramer, 1994; Landsberg, 1974), 'Alternating' (Murray et al., 1989) and 'UniChill' (Chuine, 2000) require a certain amount of chilling days, defined as the number of days following a predetermined date (usually 1 November in the Northern Hemisphere) with daily temperature below a particular threshold (e.g., 2°C), to initiate the GDD accumulation process, while Spring Warming model (Hunter and Lechowicz, 1992) implicitly assumes that winter chilling requirement is always fulfilled, so that the

GDD accumulation process starts at a predetermined date. The 'Alternating' model has been adopted in a global phenology scheme (Botta et al., 2000), Hybrid model (Friend et al., 1997; Friend and White, 2000) and PnET model (Chiang and Brown, 2007). The model of White et al. (1997), which has also been adopted widely (Kim and Wang, 2005; Kucharik et al., 2006; White et al., 2000), is also based on thermal time, while soil temperature is used to calculate GDD and combined with radiation information to trigger leaf onset. Other factors such as photoperiod (Partanen et al., 1998), precipitation and nitrogen deposition (Cleland et al., 2007) are considered to have minimal effect on leaf onset at ecosystem level. However, recent studies suggest that it is necessary to constrain the leaf onset model with photoperiod, because an advance is implicitly included in the common 'Spring Warming' model, which is problematic in predicting the budburst in the future (Blümel and Chmielewski, 2012; Körner and Basler, 2010). Moreover, Arora and Boer (2005) developed a new approach in CTEM (Canadian Terrestrial Ecosystem Model) to predict leaf onset, which is triggered by the environmental conditions favorable for carbon gain.

Temperature and photoperiod are considered to be the most important factors that regulate the timing of leaf offset. Temperature threshold alone has been used to trigger the date of leaf offset in some models. For example, in CLASS model, leaf offset occurs when air temperature drops below 0ºC (Verseghy et al., 1993). By contrast, the model of Kaduk and Helmann (1996), original IBIS (Foley et al., 1996) and LPJ (Sitch et al., 2003) all use 5ºC as the threshold. IBIS2.0 triggers the leaf offset when 10-day running average temperature drops below 0ºC, or 5ºC warmer than the coldest monthly temperature

(Kucharik et al., 2000). A photoperiod threshold method is employed in the Hybrid model (Friend et al., 1997). The autumn photoperiod required for leaf senescence is empirically calculated depending on latitude. The photoperiod alone without interacting with temperature cannot explain the inter-annual variability of leaf senescence because the photoperiod of a particular day is the same every year (Chuine et al., 2003). Both photoperiod and temperature are considered in the continental phenology model (White et al., 1997), which is adopted by many DEMs such as BIOME-BGC (White et al., 2000) and CTEM (Arora and Boer, 2005). The leaf offset is triggered using a photoperiod threshold with upper and lower temperature limits. Delpierre et al. (2009) found that a combination of temperature and photoperiod was sufficient to predict the senescence date of Sessille Oak and European beech. Vitasse et al. (2011) extended this method to four dominant species in European temperate forests and found that the senescence model has good predictability for three species.

According to a study that evaluated 14 ecosystem models (Richardson et al., 2012), the phenological models embedded tend to substantially overestimate the length of the growing season in temperate deciduous forests, with spring onset coming too early, and autumn dormancy too late, which would introduce large bias in simulated productivities. In order to improve the overall performance of ecosystem models, additional work is still needed to improve current phenological models or develop new models (Richardson et al., 2013).

3.1.4 Goal of this study

To some extent, the detected trends in vegetation phenology depend on the phenological metrics used in the analysis (e.g., leaf unfolding vs. flowering, NDVI threshold vs. rate of change in NDVI). Therefore, when examining the change in vegetation phenology, it is important to explicitly define the phenophases in discussion. This study aims to use improved phenological models to examine the historical trends in the two key phenophases (i.e., leaf onset and offset) required in a DEM, Agro-IBIS, over Northeastern U.S. deciduous forests, and use Agro-IBIS to quantify the impact of phenological changes on the terrestrial carbon cycle and the relationship between productivities (i.e., GPP and NEP) and phenology. First, remotely sensed leaf onset and offset were used as reference to parameterize a set of phenological models on a grid cell basis. One leaf onset model and offset model were chosen to reconstruct historical time series with a high-resolution climate dataset. Then the trends in reconstructed leaf onset and offset were analyzed for two periods. Finally, a series of Agro-IBIS simulations were run with leaf onset and offset simulated using the chosen phenological models, to quantitatively separate the impact of changes in onset and offset on GPP and NEP.

3.2 Method and Material

3.2.1 Remote sensed phenological metrics

According to the analysis in Chapter 2 (Xu et al., 2014), although existing remote sensing methods have relatively large biases in capturing the leaf onset and offset dates required in Agro-IBIS, it is possible to reduce the bias by adjusting the parameters in those methods. In this study, LOGISTIC2 (Fisher et al., 2006) with enhanced vegetation

index (EVI) was modified to retrieve both leaf onset and offset for following reasons. First, the sigmoid curve of logistic function relatively well characterizes the seasonal behaviors of deciduous forests; second, LOGISTIC2 with EVI has relatively good performance in capturing the inter-annual variability in leaf onset with moderate bias and also has relatively small bias in capturing the leaf offset; third, the parameter (i.e., the relative threshold used to determine leaf onset and offset) can be easily adjusted. Originally, LOGISTIC2 method fits the EVI time series to logistic functions, and determines the leaf onset as the date when the fitted EVI exceeds 50% of the amplitude between minimum and maximum in the spring, and the leaf offset as the date when the fitted EVI drops to 50% amplitude in the autumn (Fisher et al., 2006). In this study, the threshold of 50% was changed to other values between 10% and 90% with an interval of 5% (e.g., 10%, 15%, 20%). One threshold was selected for leaf onset and offset, respectively, based on their performance (measured as RMSD between remotely sensed and observed leaf onset and offset) at both the Harvard Forest and the Hubbard Brook Experimental Forest (http://www.hubbardbrook.org).

At the Harvard Forest, EVI time series were first calculated using the 8-Day 500m MODIS surface reflectance product (code: MOD09A1, https://lpdaac.usgs.gov/) and smoothed based on the Savitzky-Golay Filter with band quality files and state flags (Chen et al., 2004). Then LOGISTIC2 method was applied with aforementioned thresholds to retrieve leaf onset and offset for 2000-2010. The average date over the five MODIS pixels encompassed in the phenology-observation area was used as the site value to compare with observations that represent the requirement of phenology in Agro-IBIS

(i.e., BBRK30 for leaf onset and LCOLOR20 for leaf offset, see Chapter 2 or Xu et al., 2014).

The Hubbard Brook Experimental Forest is located within the White Mountain National Forest in central New Hampshire. The forest is dominated by sugar maple (*Acer saccharum* Marsh.), American beech (*Fagus grandifolia* Ehrh.), and yellow birch (*Betula alleghaniensis* Britt.). Phenology is routinely recorded at nine different areas. Average phenological dates across species at locations 1 and 6 were used for validation because the locations of observations within these two areas could be linked to a few MODIS pixels, thus decreasing errors related to averaging over a large region. Phenology at the Hubbard Brook Experimental Forest was quantified with an index ranging from 0 to 4 for both spring and autumn, which is different from the metrics used in the Harvard Forest (e.g., BBRK, LCOLOR). Table 3.1 describes the plant development stages corresponding to different index values. Observations were converted using a method similar to Richardson et al. (2006) in order to make the data comparable with the metrics used in Harvard Forest. The phenology index time series was fitted to a sigmoid logistic function with maximum set to 4 and minimum set to 0. Based on the phenology index (Table 3.1), leaf onset was determined as the date when fitted phenology index exceeds 1.2, and leaf offset was determined as the date when fitted phenology index drop to 3. Remotely sensed phenological metrics at Hubbard Brook were derived in the same manner as at Harvard Forest. The average date across the six MODIS pixels encompassed within locations 1 and 6 was used as the site value.

Table 3.1 Description of the phenology index at Hubbard Brook Experimental Forest

Season	Index	Description
Spring	0	All leaves fallen except remnants on beech winter condition
	1	Bud swelling noticeable
	2	Small leaves or flowers visible, initial stages of leaf expansion, leaves about 1 cm long
	3	Leaves 1/2 of final length, leaves 5 cm long, leaves obscure 1/2 of sky as seen through crowns
	3.5	Leaves 3/4 expanded, sky mostly obscured through crown, crowns not yet in summer condition
	4	Canopy appears in summer condition, leaves fully expanded, little sky visible through crowns
Fall	4	Canopy appears in summer condition, only scattered leaves or branches have any color change
	3	Many leaves have noticeable reddening or yellowing, much green still present
	2	Most leaves yellow or red few, leaves have fallen, leaves still obscure most of sky as seen through crown
	1	No more green in canopy, half of leaves have fallen, leaves still obscure half of sky as seen through crown
	0.5	Most leaves fallen
	0	All leaves fallen except remnants on beech winter condition

3.2.2 Parameterization of phenological models at regional scale

Because this study only focused on the phenology of deciduous trees, the study region selected is northeastern U.S. encompassing 35° to 45° latitude and -85° to -70° longitude, where most of the deciduous forests in U.S. are located. As Fisher et al. (2007) demonstrated that parameters of phenological models might vary with geographic location and species composition, a possible way to improve the overall performance of phenological models over a particular region is to conduct location-specific parameterization, which is supported by multi-sites modeling work (Yang et al., 2012). In this study, remotely sensed phenological metrics were used as reference to conduct location-specific (i.e., grid cell by grid cell) parameterization of different phenological models over the study region. First, leaf onset and offset were retrieved using the modified LOGISTIC2 method with EVI at the original resolution of MOD09A1 (i.e.,

500m) for 2000-2007, which is the temporal overlap with the historical climate dataset created by ZedX Inc. (Motew and Kucharik, 2013). Only the pixels, which were classified as deciduous broadleaf forest or mixed forest in either 2000 or 2007, were considered based on the MODIS land cover type product (code: MCD12Q1) (Friedl et al., 2002; Friedl et al., 2010), which has the same spatial resolution as MOD09A1. In order to match the historical climate dataset spatially, retrieved phenological dates were aggregated into a latitude-longitude grid by averaging the value of pixels encompassed in each 5 minute by 5 minute grid cell. Then, in each grid cell, a simple genetic algorithm was applied to optimize the parameters of various phenological models by minimizing the RMSD between the modeled and remotely sensed phenology over the period of 2000-2007.

For leaf onset, the 'temperature threshold' model (Foley et al., 1996; Sitch et al., 2003), widely tested 'Spring Warming' model (Hunter and Lechowicz, 1992), 'Sequential' model (Sarvas, 1974) and 'Parallel' model (Kramer, 1994), as well as modified 'Spring Warming' model with explicit photoperiod term (denoted as 'Spring Warming II') (Blümel and Chmielewski, 2012), were optimized. For leaf offset, the 'Temperature threshold' model (Foley et al., 1996; Sitch et al., 2003), the 'White' model (White et al., 1997) currently adopted in Agro-IBIS (Twine and Kucharik, 2008), and the 'Delpierre' model (Delpierre et al., 2009), were selected. For the 'White' model, the model structure was taken, but air temperature was used instead of soil temperature. The parameters of each model were all optimized in their preset ranges. The structures and

range of parameters for leaf onset and offset are summarized in Table 3.2 and Table 3.3, respectively.

Table 3.2 Summary of leaf onset models

Model Name	Equation	Parameter Range
Temperature threshold (TT)	When $T_{avg10} \geq T^*$, leaf onset occurs, where T_{avg10} is 10-day running average temperature	T^* : [-10:10]
Spring Warming (SW)	$S_f = \sum_{t_0}^{t_y} R_f(x_t)$ if $T_b > x_t$ then $R_f = 0$ else $R_f = x_t - T_b$ when $S_f \geq F^*$ leaf onset occurs	T_b : [0:10] t_0 : [1:100] F^* : [100:1500]
Spring Warming II (SW2)	$S_f = \sum_{t_0}^{t_y} R_f(x_t)$ when $S_f \geq F^*$ leaf onset occurs If $T_b > x_t$ then $R_f = 0$ else $R_f = (x_t - T_b)[P(t)/10]^a$	T_b : [0:10] t_0 : [1:100] F^* : [100:1500] a : [0:5]
Sequential (SEQ)	$S_c = \sum_{t_0}^{t_h} R_c(x_t)$ If $T_c > x_t$ then $R_c = 1$ else $R_c = 0$ when $S_c \geq C_{total}$ heat accumulation starts $S_f = \sum_{t_h}^{t_y} R_f(x_t)$ when $S_f \geq F^*$ leaf onset occurs If $T_b > x_t$ then $R_f = 0$ else $R_f = x_t - T_b$	T_c : [-10:10] C_{total} : [1:150] T_b : [0:10] F^* : [100:1500]
Parallel (PAR)	$S_c = \sum_{t_0}^{t_h} R_c(x_t)$ if $T_c > x_t$ then $R_c = 1$ else $R_c = 0$ $S_f = \sum_{t_0}^{t_h} K_m \cdot R_f(x_t)$ when $S_f \geq F^*$ leaf onset occurs If $S_c / C_{total} > 1$ then $K_m = 1$ else $K_m = S_c / C_{total}$ If $T_b > x_t$ then $R_f = 0$ else $R_f = x_t - T_b$	T_c : [-10:10] C_{total} : [1:150] T_b : [0:10] F^* : [100:1500]

x_t (˚C) is daily average temperature; T_{avg10} (˚C) is 10-day running average temperature; *P(t)* (hour) is the photoperiod for day *t*; T^* (˚C) is the critical temperature threshold for leaf onset; T_b (˚C) is the base temperature for heat accumulation; T_c is the base temperature for chilling accumulation; C_{total} (days) is the critical threshold of chilling process; R_c (day) is the rate of chilling accumulation; S_c (day) is the status of chilling accumulation; R_f (degree-day) is the rate of heat accumulation; S_f (degree-day) is the status of heat accumulation; K_m is adjusting factor for heat accumulation based on chilling status; F^* (degree-day) is the critical threshold for heat accumulation; a is empirical parameter; t_0 (day of year) is the starting date of accumulation; t_h (day of year) is the date when chilling requirement is fulfilled; t_y (day of year) is the date of leaf onset

Table 3.3 Summary of leaf offset models

Model Name	Equation	Parameter Range
Temperature threshold (TT)	When $T_{avg10} \le T^*$, leaf onset occurs	T^* : [-15:20]
White Model (WHT)	When $P(t) \le P^*$ and $T_{avg10} \le T_{high}$, or $T_{avg10} \le T_{low}$ leaf offset occurs	P^* : [10:16] T_{high} : [5:30] T_{low} : [-10:10]
Delpierre Model (DLP)	If $P(t) \le P_{start}$ and $x_t \le T_c$, then $S_{sen} = \sum_{t_0}^{t_s} R_{sen}(x_t)$ where $R_{sen}(x_t) = [T_c - x_t]^x \times [P(t) / P_{start}]^y$ when $S_{sen} \ge Y_{crit}$ leaf offset occurs	P_{start} : [10:16] T_c : [1:100] x & y : [0:5] Y_{crit} : [0:6000]

x_t (˚C) is daily average temperature; T_{avg10} (˚C) is 10-day running average temperature; $P(t)$ (hour) is the photoperiod for day t; T^* (˚C) is the critical temperature threshold for leaf offset; T_b (˚C) is the base temperature for heat accumulation; T_c is the base temperature for chilling accumulation; T_{high} (˚C) is the high temperature limit for leaf offset; T_{low} (˚C) is the low temperature limit for leaf offset; P_{start} (hour) is the critical photoperiod threshold for leaf offset; R_{sen} (˚C hour hour^{-1} day^{-1}) is the rate of forcing for leaf offset; S_{sen} (˚C hour hour^{-1}) is the accumulated forcing units for leaf offset; Y_{crit} is the critical forcing threshold of leaf offset; t_0 (day of year) is the starting date of accumulation; t_s (day of year) is the date of leaf senescence

3.2.3 Retrospective analysis of phenology

Once the parameters of phenological models were optimized grid cell by grid cell over the study region, one leaf onset model and one offset model were selected based on their overall performance to conduct retrospective modeling. As some studies suggested that Akaike's Information Criterion (AIC) should be used in model selection to consider the trade-off between the goodness-of-fit (model explanatory power) and the model complexity (number of parameters) (Migliavacca et al., 2012; Richardson et al., 2013), the small sample corrected criterion, AICc (Burnham and Anderson, 2002) were also calculated for all the grid cells using equation 3.1,

$$AIC_C = n \log \sigma^2 + 2p + \frac{2p(p+1)}{n-p-1} \quad (3.1)$$

where n is the number of samples, σ is the RMSD, and p is the number of parameters.

Selected models were run with the ZedX dataset for 1958-2007 to reconstruct the historical time series of leaf onset and offset. Then the temporal trends were determined in each grid cell for two periods (1958-2007 and 1983-2007) using linear regression, and the statistical significance of the trends was determined using the t-test.

3.2.4 Agro-IBIS simulations

In order to quantify the impact of phenological changes on GPP and NEP, two sets of Agro-IBIS simulations (one set of simulations for 1958-2007 and one for 1983-2007) were conducted using an experimental design similar to White et al. (1999). For each set of simulations, a control simulation (STATIC) was run with leaf onset and offset fixed as the long-term means over the analysis period (e.g., 50-year average leaf onset date for 1958-2007). In other words, both leaf onset and offset date were constant over the entire simulation period. Xu et al. (2014) demonstrated that it is necessary to separately account for changes in leaf onset and leaf offset, rather than accounting for the change in vegetation growing season length (GSL) alone (White et al., 1999), because of their different effects on carbon uptake. Therefore, two experimental simulations were run. A Dynamic onset (ONSET) simulation was run with leaf offset fixed as the long-term mean and leaf onset set as that simulated using the parameterized model selected in section 3.2.3. Similarly, a Dynamic offset (OFFSET) simulation was run with leaf onset fixed as the long-term mean and leaf offset set as that simulated using the parameterized model selected in section3.2.3.

Results were analyzed in two ways. First, the linear trends of the difference in simulated productivities (i.e., GPP and NEP) between experimental and control

simulations were calculated for both simulation periods. Because all the parameters, settings, and driving data were identical, the difference in simulated productivities can be attributed to the differences in phenology. For example, the trends in the difference in simulated productivities between the ONSET and the STATIC reflect the impact of leaf onset. Second, a linear regression analysis similar to that performed in Chapter 2 was conducted to predict the relationship between the difference in simulated productivities and the variation in phenology. For example, the difference in simulated GPP between the ONSET and the STATIC simulation was regressed on the leaf onset anomalies. The regression slopes reflect the impact of phenology on GPP.

3.3 Results

3.3.1 Remotely sensed phenological metrics

Table 3.4 and 3.5 show the performance of different thresholds used in LOGISTIC2 method in retrieving leaf onset and offset, respectively. At the Harvard Forest, the RMSD between remotely sensed and ground-observed leaf onset was minimized when 25% used as threshold (i.e., the amplitude between the minimum and maximum of fitted EVI time series that determines leaf onset); while at the Hubbard Brook Experiment Forest, the RMSD was minimized when 35% was used as threshold. 30% was finally selected as the threshold to retrieve leaf onset over the study domain, because the RMSD was relatively small (slightly larger than the minimum RMSD) at both sites when it was used. As the RMSD between remotely sensed and ground-observed leaf offset was minimized at both sites when 55% was used as threshold, it was selected to retrieve leaf offset over the study domain. The inter-annual variability in leaf onset was also captured by the remote

sensing method at both sites (Fig. 3.1), resulting in correlation coefficients of 0.89 and 0.64 at the Harvard Forest and the Hubbard Brook Experimental Forest, respectively. However, the inter-annual variability in remotely sensed offset did not show good agreement with that in the observation at both sites.

Table 3.4 Performance of different thresholds for retrieving leaf onset
(RMSD between remotely sensed and observed leaf onset in days)

	10%	15%	20%	25%	30%	35%	40%	45%	50%
Harvard Forest	12.4	7.4	4.8	3.4	4.8	6.8	9.0	11.2	13.3
Hubbard Brook	16.7	12.2	9.0	6.6	5.0	4.6	5.2	6.6	8.0

Table 3.5 Performance of different thresholds for retrieving leaf offset
(RMSD between remotely sensed and observed leaf offset in days)

	35%	40%	45%	50%	55%	60%	65%	70%	75%
Harvard Forest	15.2	11.8	8.3	5.2	4.6	6.8	10.6	15.1	19.7
Hubbard Brook	13.4	10.6	8.0	6.0	5.3	6.1	8.4	11.2	14.7

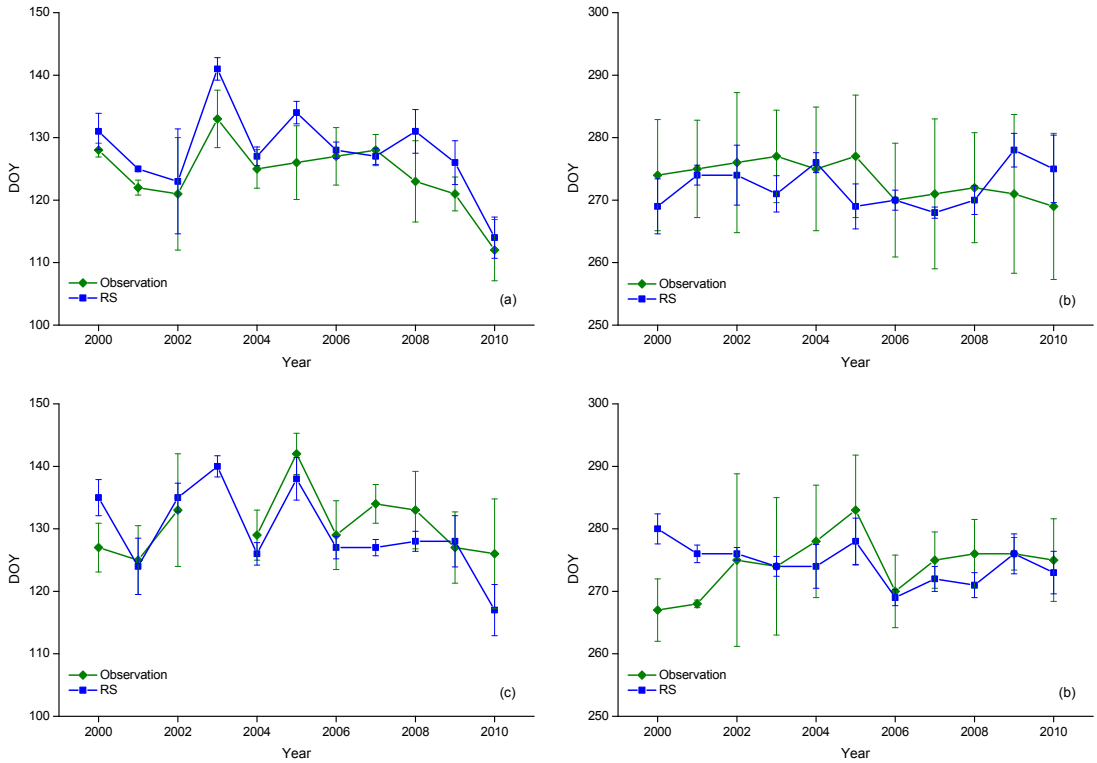


Figure 3.1 Remotely sensed vs. observed leaf onset and offset
(a) Onset at Harvard Forest (HF); (b) offset at HF; (c) onset at Hubbard Brook Experimental Forest (HB); and (d) offset at HB. Error bars indicate the standard deviation.

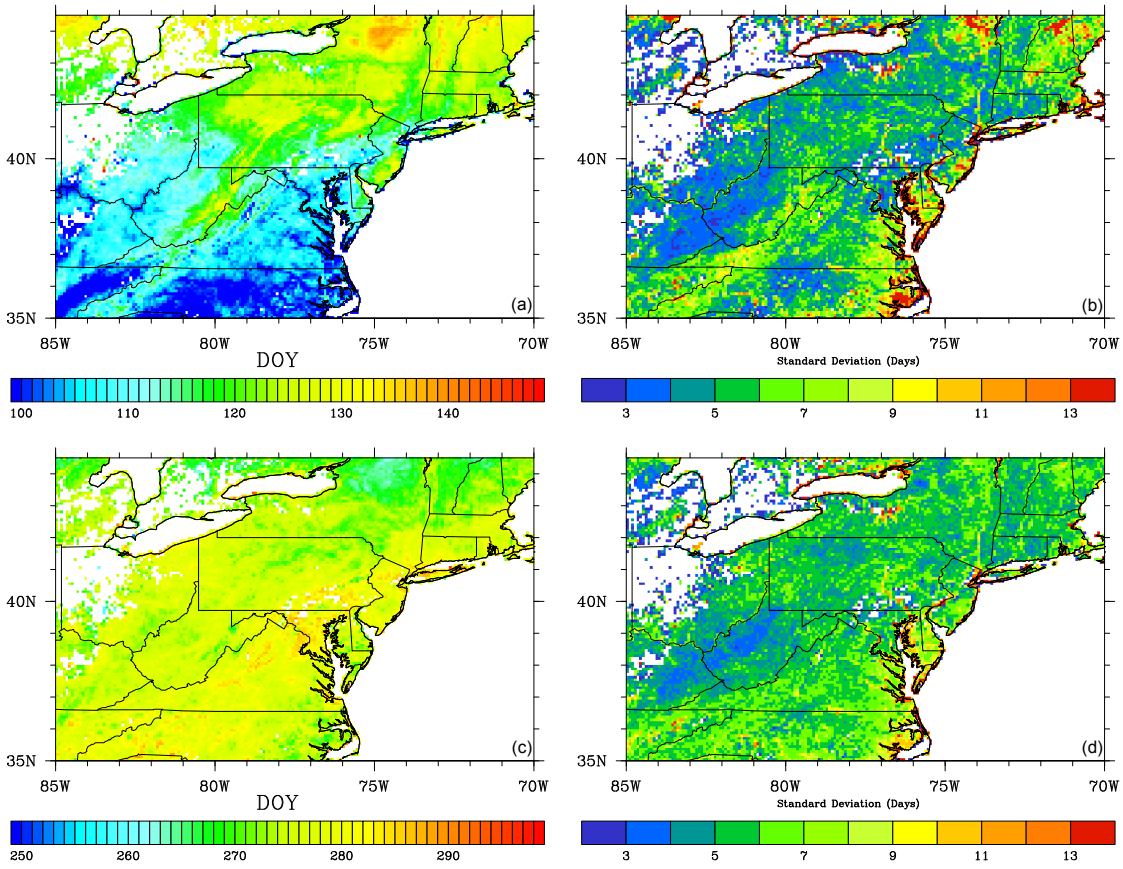


Figure 3.2 Average leaf onset and offset from remote sensing and the spatial variability for 2000-2007 (a) leaf onset; (b) standard deviation of leaf onset; (c) leaf offset; (d) standard deviation of leaf offset. Some grid cells are shown with no value (e.g., west Ohio), because no MODIS pixels located in those grid cells were classified as deciduous or mixed forest.

Figure 3.2 shows the 8-year (i.e., 2000-2007) average date of leaf onset and offset from MOD09A1 after spatially aggregated to a 5-minute grid, as well as their average spatial variability (i.e., the standard deviation of phenological dates across MODIS pixels in each grid cell) over the same time period. Remotely sensed leaf onset shows an apparent latitudinal gradient from day of year (DOY) ~90 in the south to DOY ~145 in the north. Leaf onset also shows an altitudinal gradient with later onset appearing at higher elevation, which can be seen over the area of Appalachian Mountains (Fig 3.2a).

Leaf offset in the north part of the study area are generally earlier than offset in the south, although the gradient is small (Fig 3.2c). These patterns are consistent with Hopkin's Law (Fitzjarrald et al., 2001; Hopkins, 1938), which suggests that the modified LOGISTIC2 method validated at site level is also effective at larger scales. For both leaf onset and offset, the 8-year average standard deviation of most grid cells was between 4 and 9 days (Fig 3.2c and 3.2d), suggesting that the spatial variability was relatively small. Grid cells with high spatial variability (e.g., standard deviation above two weeks) were generally found along the East Coast and the boundary of the Great Lakes.

3.3.2 Phenological models

Figure 3.3 shows the RMSD between modeled and remotely sensed phenology after the parameters were optimized using the genetic algorithm. The regional statistics of RMSD and AICc are summarized in Table 3.6. For leaf onset, when measured by RMSD, SW2 has the best performance by producing the lowest regional average value (Table 3.6). The spatial patterns of the RMSD from SW and SW2 are almost identical with most grid cells having an RMSD around 3 days (Fig. 3.3b and 3.3c). The SEQ and PAR show a little higher RMSD than SW and SW2 with regional average RMSD slightly lower than 4 days (Fig 3.3d and 3.3e). This is consistent with the findings that simple spring warming model is generally better than models considering chilling requirement in the similar region (Migliavacca et al., 2012; Yang et al., 2012), and the spring warming model with photoperiod term (SW2) has slightly better performance than the traditional spring warming model (SW)(Blümel and Chmielewski, 2012). The TT model shows the highest average RMSD with the value of most grid cells greater than 7 days (Fig. 3.3).

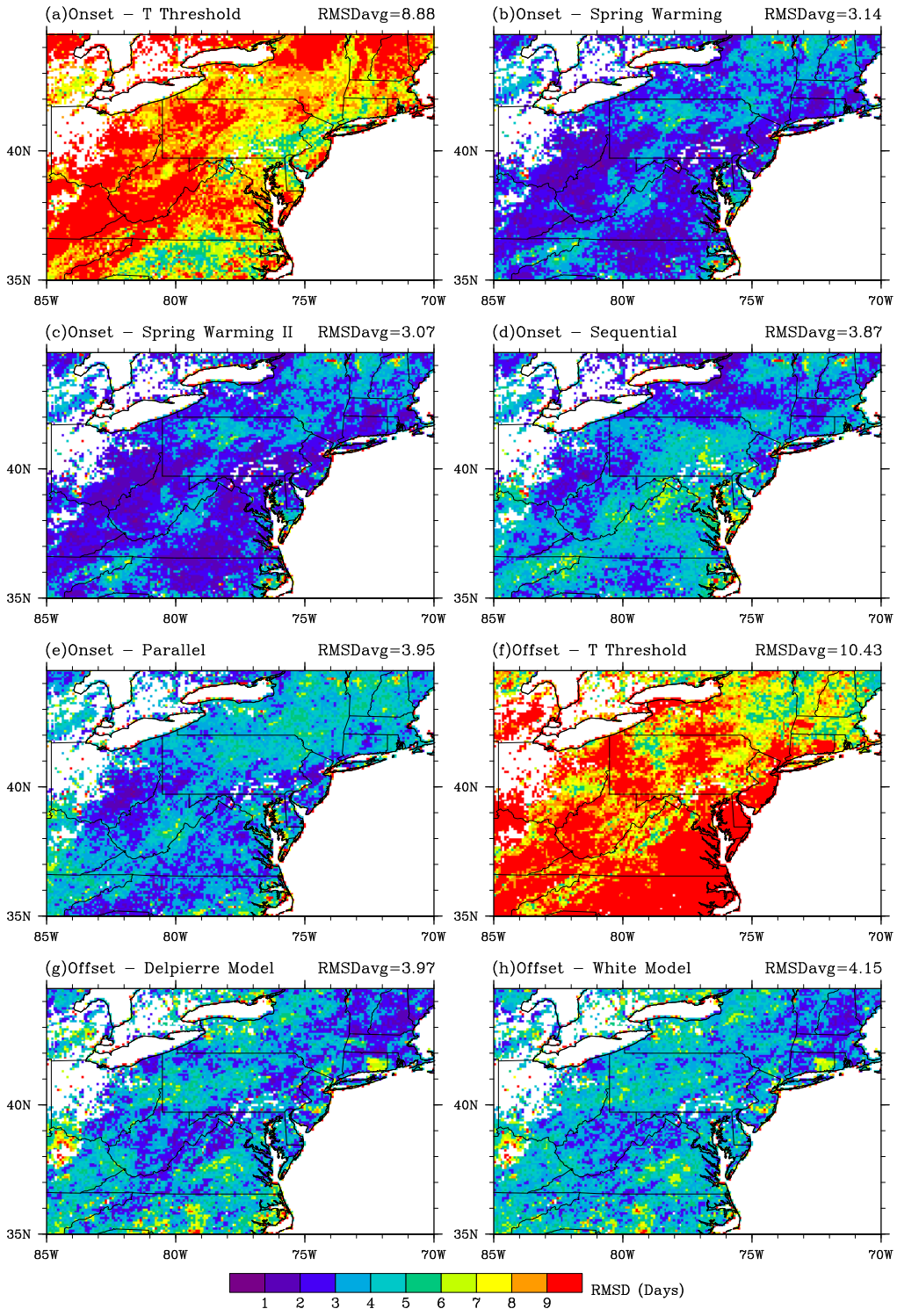


Figure 3.3 RMSD between modeled and remotely sensed phenology for 2000-2007 (a) leaf onset-TT; (b) leaf onset-SW; (c) leaf onset-SW2; (d) leaf onset-SEQ; (e) leaf onset-PAR; (f) leaf offset-TT; (g) leaf offset-WHT; and (h) leaf offset-DLP

Table 3.6 Regional statistics of RMSD and AIC_c for different phenological models

Model	Regional Average RMSD	Standard Deviation of RMSD	Regional Average AIC_c	Standard Deviation of AIC_c
Onset – TT	8.88	2.26	37.12	3.91
Onset – SW	3.14	1.92	28.58	6.91
Onset – SW2	3.07	1.88	37.54	6.98
Onset – SEQ	3.88	1.89	41.69	6.14
Onset – PAR	3.95	1.91	42.09	5.91
Offset – TT	10.43	4.73	38.89	6.15
Offset – WHT	4.15	1.56	33.86	5.31
Offset – DLP	3.97	1.69	60.84	6.24

However, when considering the number of parameters, SW showed the lowest AICc score (Table 3.6), because it has similar RMSD as SW2 but one fewer parameter (Table 3.2). As stated in Migliavacca et al. (2012), if the difference in AICc score between two models is less than 2, the two models are approximately equivalent; if the difference is larger than 6, the model with lower AICc score is about 20 more likely to be the true model. According to these criteria, the TT model is equivalent to SW2 and outperforms SEQ and PAR, which is not reasonable because the TT model produces much larger errors (i.e., RMSD). AICc score has overemphasized the simplicity of the TT model (i.e., only one parameter. Therefore, before applying AIC criteria to avoid over-fitting, traditional measures of model performance (e.g., mean absolute error, RMSD) may need to be used to eliminate models with large errors. Moreover, according to the AIC_c criteria, SW is more likely to be the true model than SW2, which shows a slightly lower RMSD but is penalized by its complexity (i.e., one more parameter than SW). However, the explicit photoperiod term applied to correct the forcing function in SW2 (Table 3.2) may be a mechanism that limits the change in leaf onset in the future (Blümel and Chmielewski, 2012). In fact, as a preliminary investigation suggests that historical trends

in leaf onsets simulated using SW2 and SW are not significantly different (see Figure A3.1 in Appendix), results from SW2 are presented in a later retrospective analysis.

For leaf offset, DLP shows the lowest regional average RMSD (Table 3.6). However, it was penalized by the number of parameters (five) so that its AICc score is high (Table 3.6). WHT model shows a spatial pattern of RMSD similar to DLP model (Fig 3.3g and 3.3h). The TT model shows the highest regional average RMSD (Table 3.6) with most grid cells higher than 8 days (Fig 3.3f). Similar to the leaf onset, the effect of TT's simplicity is dominant so that its AIC_c score is lower than DLP's. As the structure of WHT model (Table 3.3) suggests that the leaf offset only responds to photoperiod except in those years with extremely low temperature, it is found that the leaf offset predicted using WHT remained constant every year in most grid cells. Since DLP shows overall lower RMSD and it better represents the inter-annual variability, it was finally selected to conduct the retrospective analysis.

3.3.3 Historical change in leaf onset and offset

The SW2 and DLP model were run with the ZedX climate dataset to simulate the date of leaf onset and offset, respectively, for the period 1958-2007. The 50-year average leaf onset and offset showed spatial patterns (Fig. 3.4) similar to those derived from remote sensing (Fig. 3.2).

Model simulation results show a negative (advanced) trend in leaf onset over most of the study area for both the 1958-2007 and 1983-2007 time periods (Fig. 3.5a and 3.5b). There are also delayed trends in the Appalachian Mountain region (Fig 3.5a and 3.5b), which are associated with a decreasing trend in average spring temperature over both

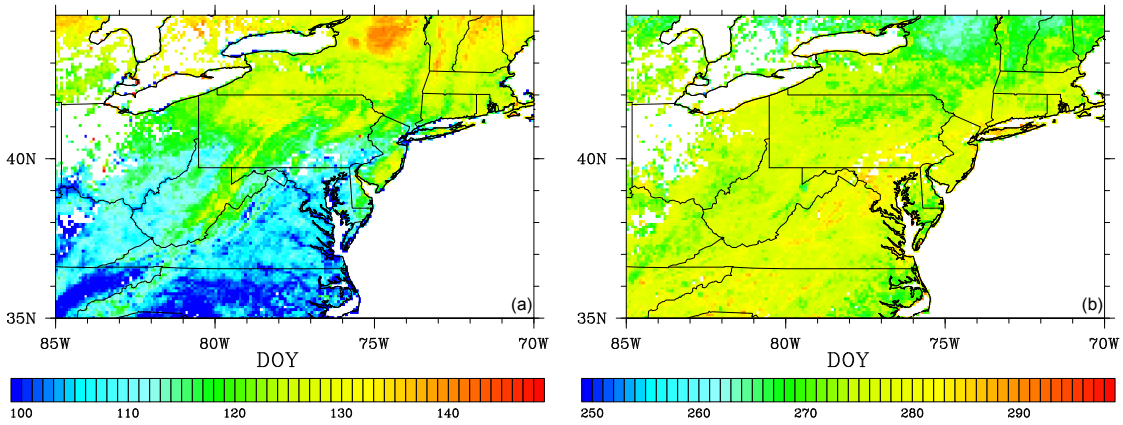


Figure 3.4 Long-term average phenological dates over 1958-2007
(a) leaf onset simulated using SW2; (b) leaf offset simulated using DLP

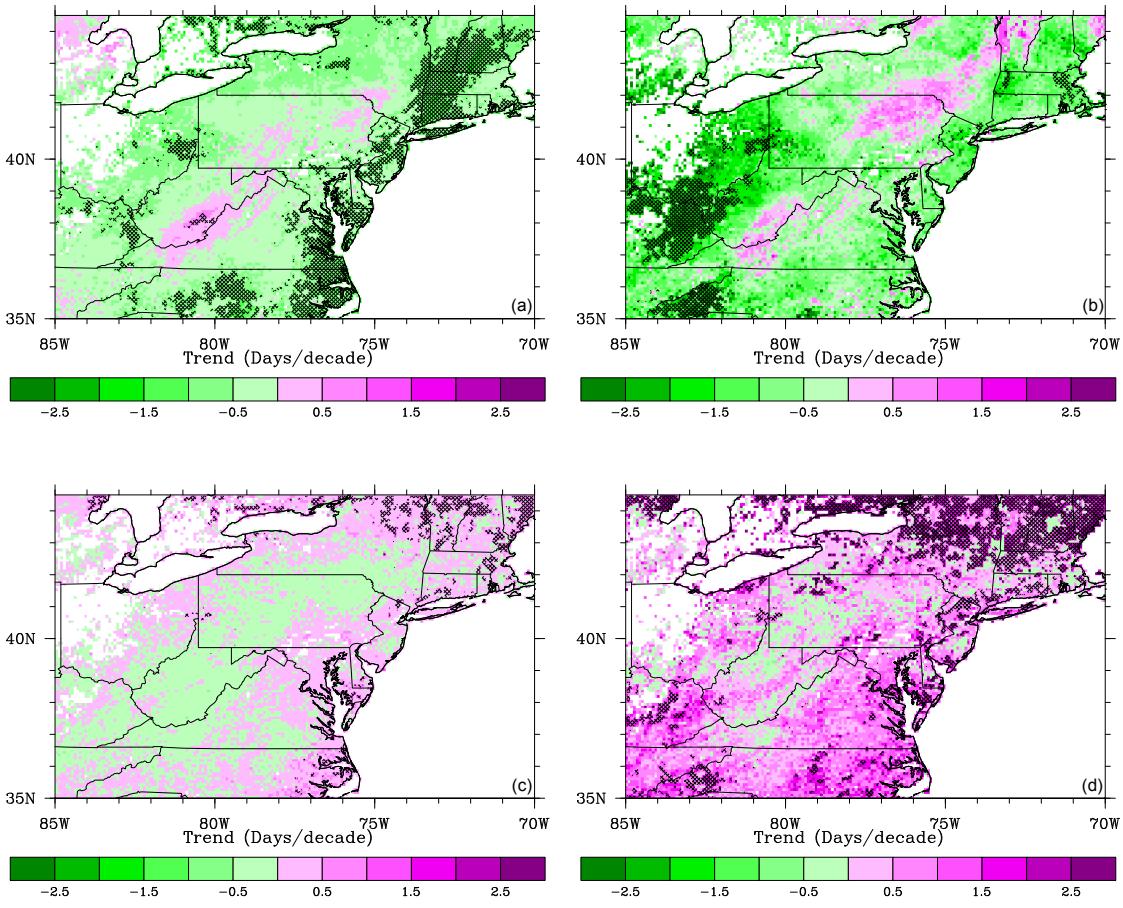


Figure 3.5 Linear trends in leaf onset and offset
(a) leaf onset: 1958-2007; (b) leaf onset: 1983-2007; (c) leaf offset: 1958-2007; (d) leaf offset: 1983-2007
Hatching indicates that the trends are statistically significant at P<0.10

time periods (Fig. 3.6a and 3.6b). While 37.2% of the grid cells show a trend between -0.5 days decade^{-1} (i.e., 2.5 days earlier in 50 years) and 0, 42.9% of the grid cells show a trend ranging from -1 days decade^{-1} (i.e., 5 days earlier in 50 years) to -0.5 days decade^{-1}. Negative trends significant at P<0.1 (~24.4% of grid cells) are found along the East Coast and in New England (Fig. 3.5a). For 1983-2007, the magnitude of advanced trends in leaf onset is slightly larger than those for 1958-2007. Grid cells that show a trend less than -1 days decade^{-1} account for ~39.8% of grid cells. 10.8% of the grid cells show a trend less than -2 days decade^{-1} (i.e., 5 days earlier in 25 years). Those grid cells with relatively large negative trends are found in western North Carolina, eastern Kentucky, southern Ohio, and the western part of West Virginia (Figure 3.5b).

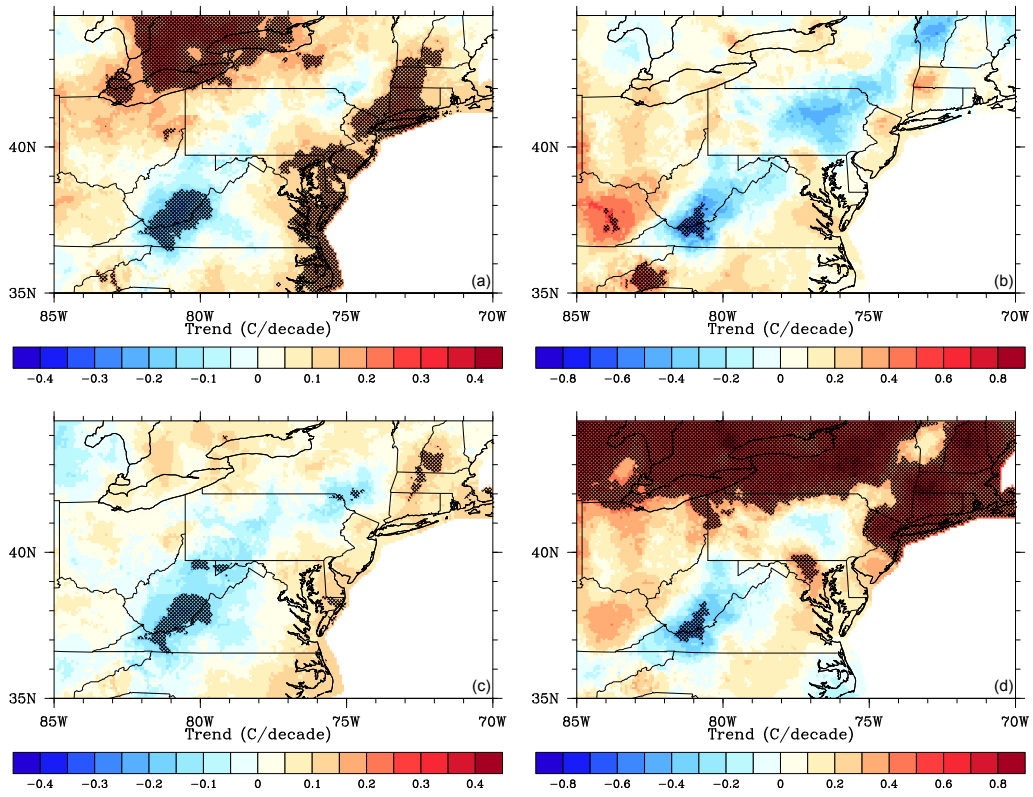


Figure 3.6 Trends in average air temperature
(a) 1958-2007 (MAM); (b) 1983-2007 (MAM); (c) 1958-2007 (SON); (d) 1983-2007 (SON)
Hatching indicates that the trends are statistically significant at P<0.01

For 1958-2007, only a small region covering parts of New York, Vermont, New Hampshire and Maine show positive (delayed) trends (Fig. 3.5c). Nearly 94% of the grid cells show very weak positive or negative trends (< +/-0.5 days decade^{-1}). This is might be because the change in autumn temperature during this period is small (Fig. 3.7c). However, no trends can be found over some areas with significant change in autumn temperature (e.g., West Virginia), suggesting that photoperiod might have a stronger control on the leaf offset than temperature over this region. For 1983-2007, there is a positive (delayed) trend for most of the study region (85.2% of the grid cells). About 7.7% of the grid cells show a trend greater than 2 days decade^{-1} (i.e., 5 days in 25 years). Grid cells with trends significant at $P<0.1$ are generally found over the northern part of the study region including Michigan, New York, Vermont, New Hampshire, Maine and Massachusetts (Figure 3.5d), which is due to the significant increasing trend in average autumn temperature over this period (Fig. 3.6d).

3.3.4 Impact of phenological changes on terrestrial carbon cycle

The simulated trends in the difference in productivities between the ONSET and STATIC runs (i.e., trend in [$GPP_{onset} - GPP_{static}$]) are positive over most of the domain during both time periods, meaning the difference between these two variables is widening with time as earlier leaf onset increases productivity. As the difference in simulated productivities between the ONSET and STATIC runs are due to the variation in leaf onset, the spatial pattern of the trends in the difference is similar to that of the leaf onset, but different in sign (Fig. 3.7). Positive trends in the difference in simulated productivities between the ONSET and STATIC run occur in the grid cells that have

negative (advanced) trends in leaf onset. The most prominent trends in GPP difference for the period of 1958-2007 are found in the New England (Fig. 3.7a), which are around 12 g C m^{-2} $decade^{-1}$ (i.e., an increase of 60 g C m^{-2} in 50 years). Some significant (at P<0.10) trends can also be found in northern New York, eastern Ohio, and central North Carolina (Fig. 3.7a). For 1983-2007, most of the significant trends are found in the area of eastern Kentucky, southern Ohio and western part of West Virginia. The trends are around 25 g C m^{-2} $decade^{-1}$ and up to 40 g C m^{-2} $decade^{-1}$. With respect to the difference in simulated NEP, although there is a positive trend over most of the study region for

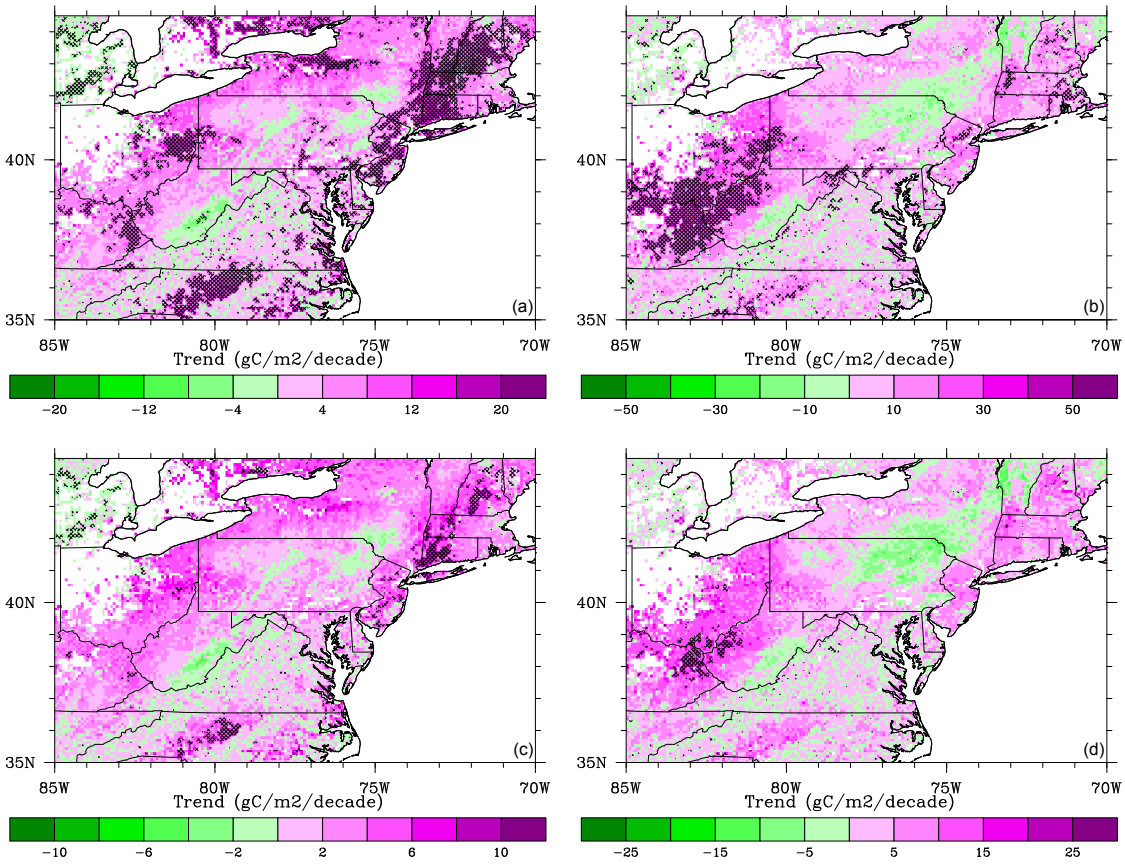


Figure 3.7 Trends in the difference in simulated productivities between the ONSET and STATIC run (a) Trends in GPP_{ONSET}-GPP_{STATIC}: 1958-2007; (b) Trends in GPP_{ONSET}-GPP_{STATIC}: 1983-2007; (c) Trends in NEP_{ONSET}-NEP_{STATIC}: 1958-2007; (d) Trends in NEP_{ONSET}-NEP_{STATIC}: 1983-2007 Hatching indicates that the trends are statistically significant at P<0.10

both periods (Fig 3.7c and 3.7d), the trends are relatively small and only statistically significant at $P<0.10$ for a few grid cells.

As there is generally no significant trend in leaf offset for the period of 1958-2007 (Fig. 3.5c), no significant trends in the difference in simulated productivities between the OFFSET and the STATIC run are found, except for a few grid cells in the north (Fig. 3.8a and 3.8c). For 1983-2007, positive trends in the difference in simulated GPP and NEP between the OFFSET and the STATIC run can be found over most of the study region (Fig. 3.8b and 3.8d) in response to the delayed leaf offset. However, those trends are relatively small (mostly under 10 g C m^{-2} $decade^{-1}$ for GPP and under 5 g C m^{-2} $decade^{-1}$ for NEP) and only statistically significant at $P<0.1$ over some parts of New York, Vermont, New Hampshire and Maine (Fig. 3.8b and 3.8d).

Most of the regressions between the difference in simulated productivities (between the experimental and control simulation) and the anomalies in phenology are statistically significant at $P<0.01$, suggesting that the productivities are sensitive to phenology (Fig. 3.9). Generally, GPP and NEP are negatively correlated to leaf onset (Fig. 3.9a and 3.9b), indicating that productivities increase as the leaf onset date moves earlier. Most of the slopes (~81.7% of grid cells) for the regression of GPP on leaf onset are in the range of -11 g C m^{-2} day^{-1} to -9 g C m^{-2} day^{-1} with an average of -10.10 g C m^{-2} day^{-1} (meaning that for every day that onset advances, an extra 10.10 g C m^{-2} is predicted to be assimilated by the vegetation), while for the regression of NEP on leaf onset, 89.1% of grid cells have a slope ranging from -8 g C m^{-2} day^{-1} to -6 g C m^{-2} day^{-1} with an average of -7.24 g C m^{-2} day^{-1}. The positive slopes of regressions between productivities

and leaf offset (Fig. 3.9c and 3.9d) indicate that productivities increase as the leaf offset moves later. However, the magnitude of the slopes is generally smaller than the regression on leaf onset, which suggests that the productivities are more sensitive to changes in leaf onset than changes in leaf offset. For the regression of GPP on leaf offset, most of the slopes (91.1%) fall in the range of 5 g C m^{-2} day^{-1} to 7 g C m^{-2} day^{-1} with an average of 5.76 g C m^{-2} day^{-1}, while for the regression of NEP on leaf offset, 81.2% of grid cells have a slope ranging from 3.5 g C m^{-2} day^{-1} to 4.5 g C m^{-2} day^{-1} with an average of 3.89 g C m^{-2} day^{-1}. There is also a tendency for grid cells at lower latitude to have greater slope (Fig. 3.9c and 3.9d).

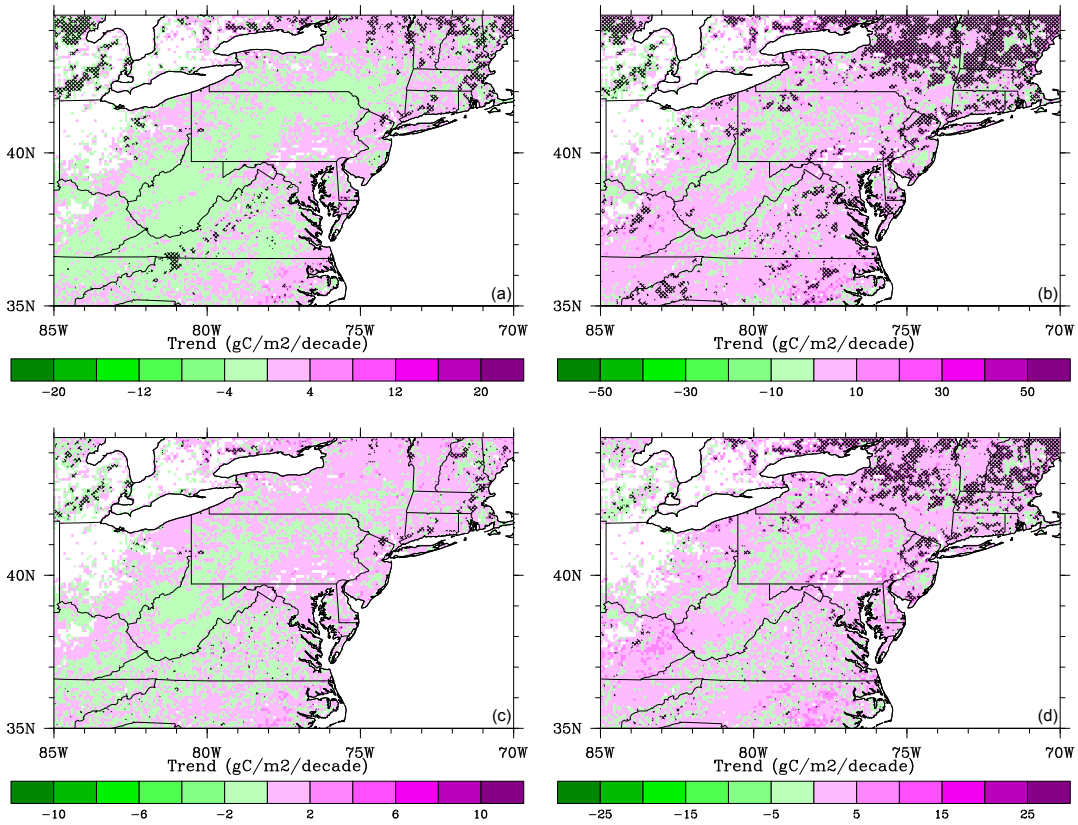


Figure 3.8 Trends in the difference in simulated productivities between the OFFSET and STATIC run (a) Trends in GPP_{OFFSET}-GPP_{STATIC}: 1958-2007; (b) Trends in GPP_{OFFSET}-GPP_{STATIC}: 1983-2007; (c) Trends in NEP_{OFFSET}-NEP_{STATIC}: 1958-2007; (d) Trends in NEP_{OFFSET}-NEP_{STATIC}: 1983-2007 Hatching indicates that the trends are statistically significant at P<0.10

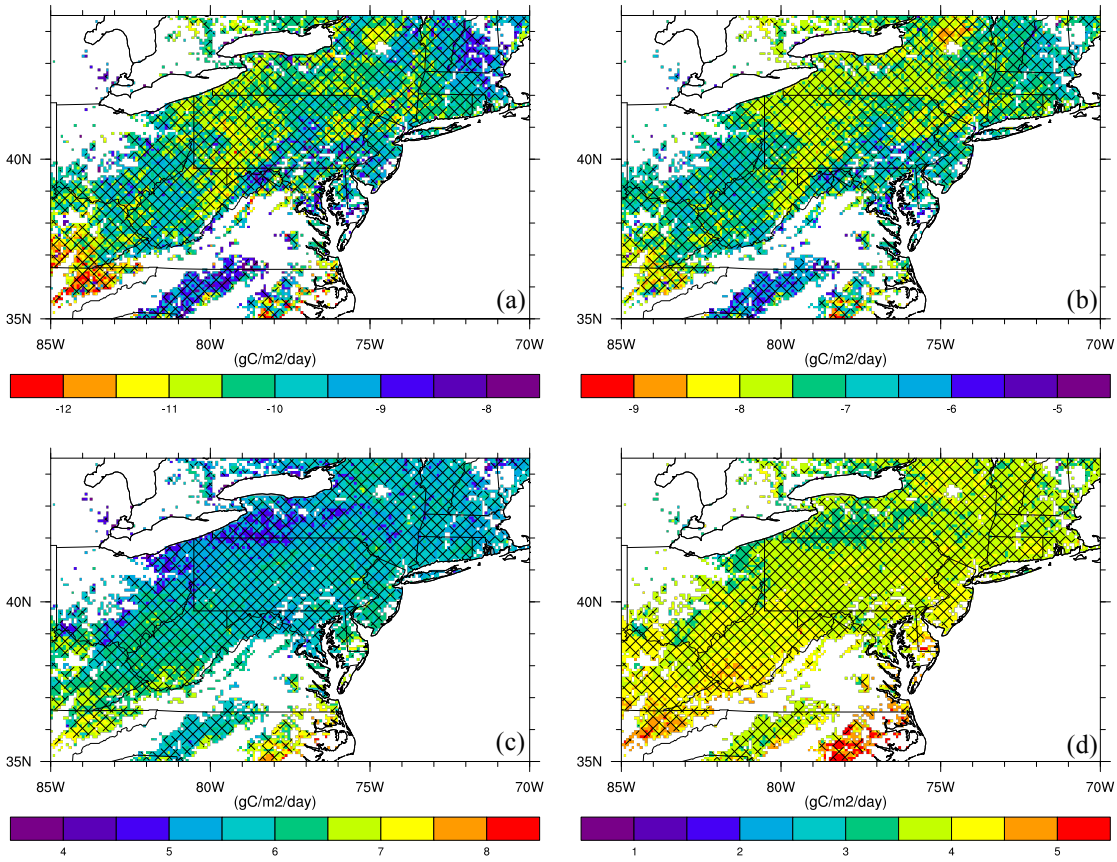


Figure 3.9 Relationship between productivities and phenology
(a) GPP vs. leaf onset; (b) NEP vs. leaf onset; (c) GPP vs. leaf offset; (d) NEP vs. leaf offset
Hatching indicates that the regressions are statistically significant at $P<0.01$. The slopes of regression are not shown for some grid cells in the southern part (most of Virginia, part of Tennessee and North Carolina) and the northeastern corner (part of New Hampshire and Maine) of the study region, because the dominant plant functional type in those grid cells were determined as evergreen trees in Agro-IBIS, which do not respond to the variation in leaf onset and offset and result in slopes around 0.

3.4 Discussion

3.4.1 Historical change in phenology

In this study, long-term time series of leaf onset and offset were reconstructed at a regional scale using the phenological models parameterized grid cell by grid cell with a gridded climate dataset. In general, the results show advances in leaf onset and delays in leaf offset, which is consistent with the findings of previous studies for North America (Dragoni and Rahman, 2012; Myneni et al., 1997; Richardson et al., 2006; Schwartz et

al., 2006; Yang et al., 2012; Zhang et al., 2014; Zhang et al., 2007; Zhou et al., 2001; Zhu et al., 2012). The magnitude of the trends detected in this study varies in space. For leaf onset, most of the advanced trends during 1958-2007 vary around ~1 day decade^{-1} (Fig. 3.5a), which is slightly smaller than the 1.2 days decade^{-1} in a temperature-driven spring index based on lilac records during 1956-2003 (Schwartz et al., 2006). The advanced trends in New England range from 0.4 to 1.7 days decade^{-1}, which is comparable with the 1.4±0.9 days per decade during 1960-2010 reported in the modeling study for New England (Yang et al., 2012) and the 1.6±0.4 days per decade during 1957-2004 at Hubbard Brook Experimental Forest (Richardson et al., 2006). For 1983-2007, the results show larger advances in leaf onset and the spatial pattern of trends is different from that for 1958-2007, suggesting that the trends in phenology depend on the period of analysis (Zhu et al., 2012). For the most recent few decades, there are discrepancies in detected trends in spring phenology among different studies. For example, based on NDVI threshold, Zhou et al. (2001) found an advanced trend of 4.2 days per decade in spring greening in North America for the period of 1982-1999. However, studies based on both remote sensing (Zhu et al., 2012) and modeling (Piao et al., 2007) detected no significant trend in spring greenup onset for 1982-2006 and 1980-2002, respectively. The discrepancies can be attributed to several factors. First, the same phenological term used in different studies could have different biophysical meanings, especially for those based on satellite imagery (White et al., 2009; Xu et al., 2014; Zhang et al., 2014). The leaf onset examined in this study was calibrated to the Agro-IBIS model, which uses a slightly different definition of onset of greenup used in previous studies. Second, with

respect to phenology retrieved using satellite data, even with the same method, inconsistency could stem from different data sources (e.g., AVHRR vs. MODIS) and vegetation indices (Zhang et al., 2014). Moreover, the difference in spatial scales and coverage of analysis could also lead to discrepancies in the detected trends (Yang et al., 2012).

For autumn phenology, previous studies focused on the onset of dormancy (i.e., the date when vegetation activities have ceased or leaves are completely removed from trees) (Piao et al., 2007; Zhou et al., 2001; Zhu et al., 2012). In contrast, the leaf offset examined in this study is defined as the date when LAI starts decreasing to meet the needs of the Agro-IBIS model, which is similar to the end of season (EOS) used in Dragoni and Rahman (2012) in terms of the retrieving method based on remote sensing. Despite the difference in definition, comparison with other results can still improve the understanding of the autumn phenology, because the change in the onset of dormancy is the combination of change in leaf offset and the change in the period the plant takes to shed leaves. For example, even if there is no change in leaf offset, delayed dormancy could still be detected if the period for leaves to color and drop were shortened. Delayed start of dormancy was widely found in North America for the recent few decades. Based on NDVI, Zhou et al. (2001) and Zhu et al. (2012) found a trend of 2.1 days per decade for 1982-1999 and 5.5 days per decade for 1982-2006, respectively. Model simulation showed a trend of 2.8 days per decade for 1980-2002 (Piao et al., 2007). Although results of this analysis also show delayed trends in leaf offset for 1983-2007 over most of the study region, the magnitude of trends are generally smaller than those reported in

previous studies. However, larger and significant trends found over the northern part (Fig. 3.5d) show good agreement with the pattern of the trends in end of season detected over the same region for the period of 1989-2008 (Dragoni and Rahman, 2012). Moreover, this is consistent with the fact that most significant trends in the onset of dormancy were found at relatively high latitude (Piao et al., 2007; Zhu et al., 2012).

Overall, this analysis suggests that the long-term change in leaf onset and offset are relatively small in the study region. The change becomes more pronounced during the more recent decades corresponding to an enhanced warming over this region. The difference between this and other analyses is likely due to the different phenological stages examined and spatial coverage. In order to fully understand the change in vegetation activities during the spring and autumn transitional period, additional work is still needed, especially for leaf senescence (Richardson et al., 2013).

3.4.2 Impact of phenological change on the carbon cycle

As terrestrial carbon balance is closely linked to the atmospheric greenhouse gas concentration, the impact on terrestrial carbon balance is one of the most important phenological feedbacks to the climate system (Peñuelas et al., 2009; Richardson et al., 2013). By comparing results from simulations with dynamic and static phenology, the historical change in GPP and NEP caused by the separate changes in leaf onset and offset was quantified. This analysis shows that increased (decreased) GPP and NEP were associated with earlier (later) leaf onset and later (earlier) leaf offset (Fig. 3.7 and 3.8). While the magnitude of historical change in GPP and NEP caused by phenological change is relatively small over the study region due to the relatively small change in

phenology, there is a strong correlation between the productivities and phenology (Fig. 3.9). It is generally accepted that a prolonged vegetation growing season (due to either earlier leaf onset or later leaf offset) increases GPP as a result of allowing more time for photosynthesis to occur (Piao et al., 2007). However, while some studies suggest that a prolonged vegetation growing season does not necessarily increase NEP because increased GPP would be offset by enhanced ecosystem respiration (Piao et al., 2007; White and Nemani, 2003), this analysis supports the hypothesis that only part of the increase in GPP due to prolonged growing season would be offset, resulting in an overall positive effect on NEP(Dragoni et al., 2011; Richardson et al., 2010).

This analysis also reveals that the relationship between the productivities and phenology (i.e., slope of regression) varies across the study domain (Fig 3.9). In other words, a one-day extension in vegetation growing season due to earlier leaf onset could lead to different gains of gross carbon from photosynthesis as well as net carbon flux at different locations. In reality, this might be due to different species composition and environmental factors. Because Agro-IBIS simulates plant functional types rather than specific species (i.e., vegetation at different locations share the same parameters), the simulated spatial variability can be attributed to the difference in the environment conditions including soil texture and local climate. Although this analysis shows a net carbon benefit due to both earlier leaf onset and later leaf offset in the study region, which is consistent with findings of Keenan et al. (2014), the possibility that a prolonged growing season has no significant impact on net carbon flux over other regions cannot be ruled out. Moreover, unlike the estimation of Keenan et al. (2014) that productivities are

more sensitive to leaf offset, this regional-scale analysis supports the site level finding that a change in leaf onset has a greater impact on both GPP and NEP than leaf offset (i.e., the magnitude of regression slope on leaf onset is larger than leaf offset), which is probably due to the fact that the environmental condition in the spring is more favorable for vegetation growth than in the autumn (Xu et al., 2014). The regional average impact of leaf onset (i.e., 7.24 g C m^{-2} day^{-1}) and leaf offset (i.e., 3.89 g C m^{-2} day^{-1}) on NEP (i.e., 5.57 g C m^{-2} day^{-1}) is approximately the same as the impact of CUP on NEP found by synthesizing measurements at different locations (Baldocchi, 2008; Churkina et al., 2005), implying that the relationship between vegetation growing season on productivities might have represented the combined impact of leaf onset and offset. However, this analysis suggests that vegetation growing season length may not be the best indicator of the phenological control on terrestrial carbon balance, because the variation in leaf onset and offset as well as their respective impact on productivities can be different over space, and may further change with climate over time.

3.4.3 Limitations

In this study, remotely sensed phenology was used as a proxy for ground observation to parameterize phenological models. Uncertainties in the remotely sensed phenology, therefore, might have propagated to the parameters of the phenological models and on to the simulated phenology. There are several possible sources of uncertainty. First, to validate remotely sensed phenology, the ground phenological measurements at Hubbard Brook Experimental Forest were converted to match the more reliable metrics at Harvard Forest, which were chosen based on their performance in Agro-IBIS (Xu et al., 2014).

Although the thresholds of the phenology index used for conversion were selected based on the meaning of index, they were relatively arbitrary. For example, the date on which the spring phenology index at Hubbard Brook reached 1.2 may not exactly represent the date of 30% percent of budbreak at Harvard Forest. A sensitivity analysis indicated that a change of 0.1 in the spring index threshold could lead to a shift of 1-2 days in the ground phenological date. In terms of the inter-annual variability, the modified LOGISTIC2 performed well for leaf onset. However, the correlation between remotely sensed leaf offset and the ground metric was weak (Fig. 3.1). This might be because the springtime phenology is similar among species, whereas there are more pronounced cross-species differences in fall phenology (Richardson et al., 2006; Yang et al., 2012). Compared with species level phenological observations, digital imagery from "PhenoCams" (Richardson et al., 2009; Sonnentag et al., 2012) has more advantages for validating remotely sensed phenology. For example, imagery from PhenoCams has a spatial scale closer to satellite imagery; it is easier to be standardized at different locations; and the spatial coverage is increasing as more PhenoCams are installed. A disadvantage is that the records are relatively short for representing the inter-annual variability in phenology (Zhang et al., 2014). This disadvantage will be reduced with time as the records become longer.

At the regional scale, LOGISTIC2 with the thresholds determined at two sites was used to retrieve phenology over the entire study region. Uncertainties might stem from the extrapolation, because the performance of the method and thresholds may vary with location due to different species composition. A more comprehensive validation over a larger domain is still needed. Moreover, there might also be uncertainties as a result of

spatial aggregation. In order to simplify the analysis process, the aggregation from the original resolution of 500 meter to 5 minutes grid was based on MODIS land cover product (MCD12Q1). As the user's accuracy of classification for deciduous broadleaf forest and mixed forest were 75.9% and 53.1%, respectively (Friedl et al., 2010), the commission errors (i.e., pixels classified as deciduous or mixed forest could actually be other land cover type) in the land cover classification might have introduced errors in in the aggregated phenological dates. The large spatial variability in remotely sensed phenology along water bodies (Fig. 3.2) might be related to this issue. Thus, an inspection of the spectral characteristics as well as the seasonal pattern of VI can be applied in future work to improve the accuracy.

3.5 Conclusion

Based on remotely sensed leaf onset and offset, multiple phenological models were parameterized grid cell by grid cell over the study region. A leaf onset model and an offset model were selected to reconstruct historical time series of leaf onset and offset, respectively. The trends in leaf onset and offset were quantified for the two time periods of 1958-2007 and 1983-2007. By conducting a series of Agro-IBIS simulations with the reconstructed phenology, the impact of phenological change on GPP and NEP was quantified. This analysis suggests that the modified remote sensing method based on logistic curve-fitting has relatively good performance in capturing the timing of leaf onset and offset as well as the inter-annual variability in leaf onset. The inter-annual variability in leaf offset is poorly represented at the two validation sites possibly because of the large number of different species. This underlines the scale issue of validating remotely sensed

phenology with species-level observation, and highlights the need of validation over large spatial coverage using observation with a spatial scale closer to satellite imagery, such as "PhenoCams" (Richardson et al., 2009; Sonnentag et al., 2012). The evaluation of phenological models shows that, in terms of minimizing RMSD, the Spring Warming model with explicit photoperiod term and the Delpierre Model were the best for predicting leaf onset and offset, respectively. It also suggests that, before applying the Akaike's Information Criterion to avoid over-fitting, model candidates should be screened by examining the traditional measures of error and investigating the key mechanism represented in the models. The trends in phenology are relatively small during the period of 1958-2007, especially for leaf offset. The trends in both leaf onset and offset have increased during the more recent period of 1983-2007 and the spatial pattern has changed as well corresponding to the pattern of changing temperature. As a result of small historical trend in phenology, the actual change in productivities due to phenological change is relatively small. On the other hand, the strong correlation between the change in productivities and the change in phenology provide implication that, should the phenology further changes in the future with rising temperature, there could be a further enhancement in carbon sequestration, which would in turn mitigate the rising temperature. However, the relationships between phenology and the productivities may themselves change in the future. Therefore, simulations that predict how phenology and productivities will respond to future climate change are still needed.

Chapter 4 Modeling The Phenological Response To Future Climate Change in Northeastern U.S. Forests

4.1 Introduction

As vegetation phenology is mainly driven by environmental factors such as temperature, precipitation and photoperiod, it is considered as a simple and important indicator of climate change (Rosenzweig et al., 2007). Shifts in vegetation phenology in the recent few decades have been reported by numerous studies using either ground observations (Keenan et al., 2014; Menzel and Fabian, 1999; Schwartz and Reiter, 2000) or remote sensing(Myneni et al., 1997; Zhang et al., 2014; Zhou et al., 2001; Zhu et al., 2012). Variation in vegetation phenology, particularly leaf onset and offset, in turn exerts various feedbacks to the climate system by affecting biophysical properties (e.g., albedo) and biogeochemical processes (e.g., carbon cycling) (Peñuelas et al., 2009; Richardson et al., 2013). Inaccurate representation of vegetation phenology may lead to errors in the simulation of carbon, water and energy exchange in dynamic ecosystem models (Richardson et al., 2012) as well as coupled climate models (Levis and Bonan, 2004). In order to improve the representation of phenological feedbacks in coupled climate models, it is important to understand how vegetation phenology alone will respond to future climate change. However, efforts made on quantification of phenological response to projected climate are relatively rare. A phenological forecasting study suggested that the uncertainty in phenological projection could be classified into three categories: uncertainty due to model structure, model parameter and model (Migliavacca et al.,

2012). Among the three sources, parameter uncertainty is relatively small because of the optimization process. The structure uncertainty, which mainly stems from different mechanisms considered and assumptions made to build phenological models, is larger. For example, some studies suggest that spring phenology such as bud-burst and flowering of most plants are predominantly driven by air temperature(Chuine et al., 2010; Morin et al., 2009; Vitasse et al., 2011). Körner and Basler (2010) suggested that the advance in spring phenology due to future warming would be constrained by photoperiod. Blümel and Chmielewski (2012) argued that, leaf onset models without photoperiod limitation have a implied trend as temperature increase; the inclusion of a photoperiod term in a conventional spring warming model could overcome such shortcoming, and lead to physically meaningful parameterization as well as better overall performance. Moreover, as the driver uncertainty (i.e., the climate projection) is comparable with the structure uncertainty as a result of the difference in scenarios and global climate models (GCMs) (Migliavacca et al., 2012), it is also necessary to quantify the range of phenological response to projected climate across scenarios and GCMs.

This study examines the response of two key phenophases (i.e., leaf onset and offset) of the deciduous forests in northeastern U. S. to climate change in the 21st century by using downscaled climate projections from GCMs to drive phenological models with location-specific parameterization. This study aims to quantify the projected changes in leaf onset and offset and the variation across locations, GCMs and climate scenarios. Particularly, how photoperiod might affect the projected changes in leaf onset is also quantified by comparing leaf onset models with and without photoperiod limitation.

4.2 Materials and Method

Based on the performance evaluation of phenological models conducted in Chapter 3, the conventional 'Spring Warming' (SW) model based on growing degree-day (GDD) accumulation (Hunter and Lechowicz, 1992) and the modified spring warming model with explicit photoperiod term ('Spring Warming II' abbreviated as SW2) (Blümel and Chmielewski, 2012) were selected to predict leaf onset, and the 'Delpierre' (DLP) model based on both temperature and photoperiod (Delpierre et al., 2009) was selected to predict leaf offset. The structure and range of parameters for selected phenological models are summarized in Table 4.1. In each grid cell over the study area encompassing 35˚ to 45˚ latitude and -85˚ to -70˚ longitude, all the parameters of each model were optimized by using a simple genetic algorithm to minimize the root mean square deviation (RMSD) between modeled and remotely sensed phenology for 2000-2007 (Chapter 3). Although the conventional 'Spring Warming' model does not have a photoperiod term, the parameter t_0 (i.e., the start date of GDD accumulation) implicitly imposes a photoperiod limitation by setting a photoperiod threshold (photoperiod has a fixed relationship with date on certain latitude). The difference between the two leaf onset models represents the effect of the photoperiod term in 'Spring Warming II' rather than the overall effect of photoperiod. Therefore, SW and SW2 were optimized once more with t_0 fixed at January 1st (denoted as SWb and SW2b, respectively) to eliminate the implicit photoperiod limitation. The performance of SWb and SW2b were evaluated using root-mean-square-deviation (RMSD) and small sample corrected Akaike's Information Criterion (AIC_c) in the same manner as in Chapter 3. Thus, the impact of

explicit photoperiod term can be quantified by comparing the change in leaf onset projected by SW2 and SW (or SW2b and SWb), while the impact of implicit photoperiod limitation can be quantified by comparing SW2b and SW2 (or SWb and SW). Those differences also reflect the potential bias in projected leaf onset, which could propagate into the phenological feedbacks and in turn the projected climate.

Table 4.1 Summary of phenological models

Model Name	Equation	Parameter Range
Spring Warming (SW)	$S_f = \sum_{t_0}^{t_y} R_f(x_t)$ if $T_b > x_t$ then $R_f = 0$ else $R_f = x_t - T_b$ when $S_f \geq F^*$ leaf onset occurs	T_b : [0:10] t_0 : [1:100] F^* : [100:1500]
Spring Warming II (SW2)	$S_f = \sum_{t_0}^{t_y} R_f(x_t)$ when $S_f \geq F^*$ leaf onset occurs If $T_b > x_t$ then $R_f = 0$ else $R_f = (x_t - T_b)[P(t)/10]^a$	T_b : [0:10] t_0 : [1:100] F^* : [100:1500] a : [0:5]
Delpierre Model (DLP)	If $P(t) \leq P_{start}$ and $x_t \leq T_c$, then $S_{sen} = \sum_{t_0}^{t_s} R_{sen}(x_t)$ where $R_{sen}(x_t) = [T_c - x_t]^x \times [P(t)/P_{start}]^y$ when $S_{sen} \geq Y_{crit}$ leaf offset occurs	P_{start} : [10:16] T_c : [1:100] x & y : [0:5] Y_{crit} : [0:6000]

x_t (˚C) is daily average temperature; $P(t)$ (hour) is the photoperiod for day t; T_b (˚C) is the base temperature for heat accumulation; R_f (degree-day) is the rate of heat accumulation; S_f (degree-day) is the status of heat accumulation; T_c is the base temperature for chilling accumulation; P_{start} (hour) is the critical photoperiod threshold for leaf offset; R_{sen} (˚C hour hour^{-1} day^{-1}) is the rate of forcing for leaf offset; S_{sen} (˚C hour hour^{-1}) is the accumulated forcing units for leaf offset; Y_{crit} is the critical forcing threshold of leaf offset; t_0 (day of year) is the starting date of accumulation; t_y (day of year) is the date of leaf onset; t_s (day of year) is the date of leaf senescence.

The temperature data used to drive the phenological models in this study were the output from eight GCMs (Table 4.2) in the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project, phase 5 (CMIP5). Particularly, the downscaled daily maximum and minimum temperature were downloaded from the Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections archive at http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/ (Reclamation, 2013). The

downscaling method used was Bias Correction with Constructed Analogues (BCCA) (Hidalgo et al., 2008; Maurer et al., 2010). Because the parameter sets of phenological models have the same spatial resolution of 5 minutes (i.e., 1/12 degree) as the historical climate data used to drive the phenological models in the optimization process (i.e., ZedX dataset, Chapter 3), the downscaled GCM data were interpolated to a 5 minutes grid from their original spatial resolution of 1/8 degree using a bilinear method.

Table 4.2 Summary of CMIP5 models

Acronym	Full model name	Institute (Country)
BCC	BCC-CSM1.1	Beijing Climate Center, China Meteorological Administration (China)
CSIRO	CSIRO-Mk3.6.0	Commonwealth Scientific and Industrial Research Organisation in collaboration with the Queensland Climate Change Centre of Excellence (Australia)
INM	INM-CM4	Institute for Numerical Mathematics (Russia)
IPSL	IPSL-CM5A-MR	Institut Pierre-Simon Laplace (France)
MIROC	MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology (Japan)
MPIM	MPI-ESM-MR	Max Planck Institute for Meteorology (Germany)
NCAR	CCSM4	National Center for Atmospheric Research (USA)
NOAA	GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory (USA)

Although the CMIP5 data were already bias-corrected, it is still important to evaluate how they perform in simulating phenology. Therefore, data from the 'historical' experiment of CMIP5 were first used to drive the phenological models. As the GCMs do not reproduce the exact year-to-year variation of climate, the average leaf onset and offset dates over 1986-2005 simulated with CMIP5 data were compared with those simulated with ZedX data. Then, data from the two core experiments of CMIP5, RCP4.5 and RCP8.5, were used to predict leaf onset and offset in the 21st century. RCP4.5 and RCP8.5 refer to the representative concentration pathways (Moss et al., 2010; Riahi et al., 2011), where the radiative forcing in 2100 will be approximately 4.5 W m^{-2} and 8.5 W m^{-2} higher, respectively, than in the pre-industrial period. Only one ensemble member

('r1i1p1') was used for each GCM. For each scenario and GCM, the trends in simulated leaf onset and offset were determined in each grid cell for the period of 2006-2100 using linear regression, and the statistical significance of the trends was determined using the Student's t-test. The sensitivity of leaf onset (or leaf offset) to temperature (days per °C) was estimated as the ratio between the trend in leaf onset (or leaf offset) (days per decade) and the trend in spring (or autumn) temperature (°C per decade).

4.3 Results

4.3.1 Historical phenology simulated with CMIP5 data

The average leaf onset dates over 1986-2005 simulated with CMIP5 data show relatively good agreement with those simulated with ZedX data when using both SW (Fig. 4.1) and SW2 (Fig. 4.2). While relatively large positive differences in average leaf onset (i.e., later leaf onset simulated with GCMs) can be found over the Appalachian Mountain area, the absolute difference between runs with GCMs and runs with ZedX is generally less than four days (Fig. 4.1 and 4.2). The regional average of the absolute difference in leaf onset varies across GCMs in a range of 1.92 to 2.58 days when using SW, and in the range of 1.88 to 2.50 days when using SW2 (Table 4.3). The 20-year (1986-2005) average leaf offset dates simulated with CMIP5 data show a better agreement than leaf onset, with the regional average absolute difference ranging between 0.64 and 0.85 days (Table 4.3). In contrast to the leaf onset, negative differences (i.e., earlier leaf offset simulated with GCMs) are usually found over the Appalachian Mountain area (Fig. 4.3), which implies colder temperature from GCMs consistently with the later leaf onset. Overall, the average leaf onset (Fig. A4.1 and A4.2) and offset (Fig.

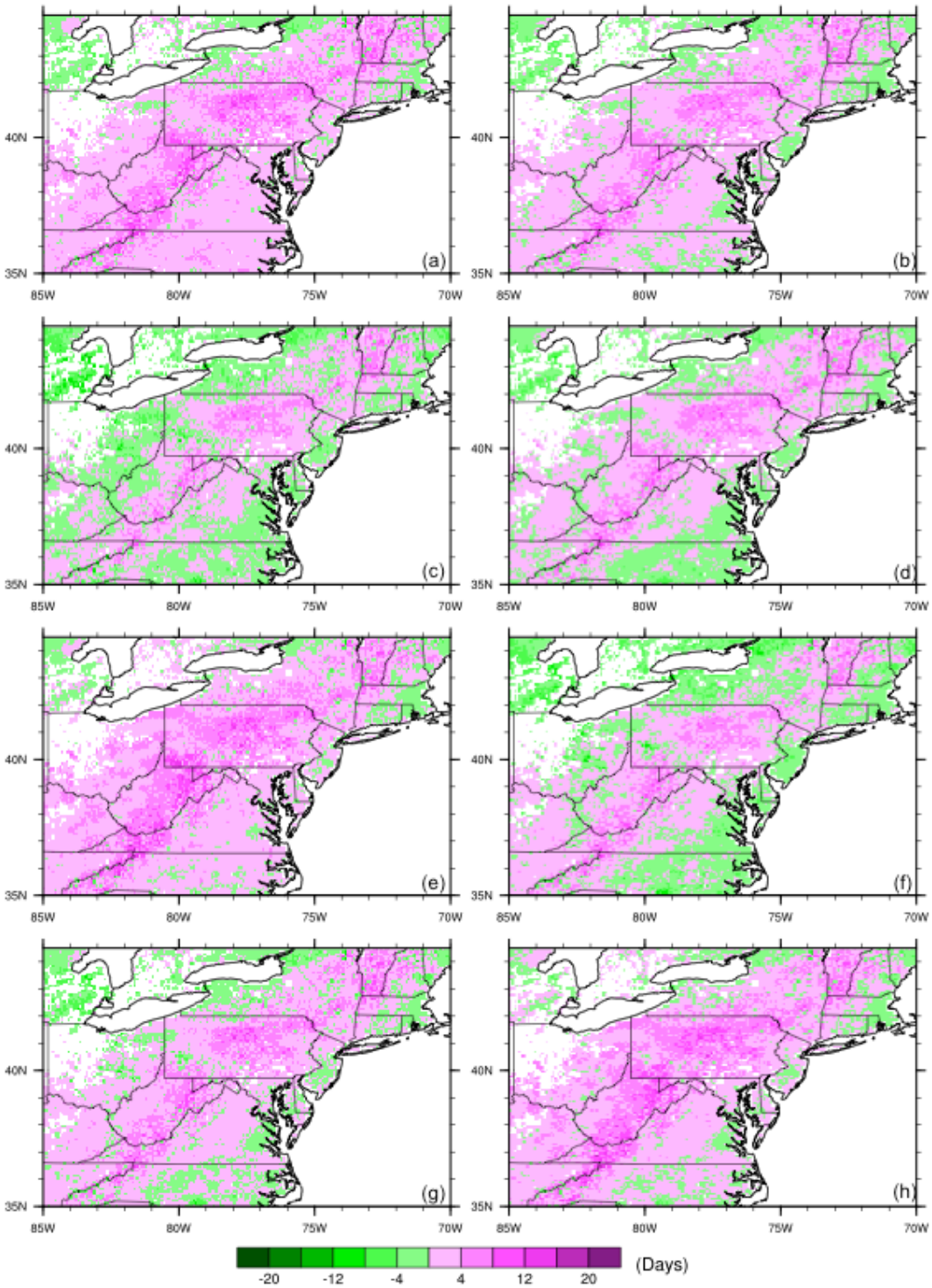


Figure 4.1 Difference (CMIP5-ZedX) in 20-year (1986-2005) average leaf onset simulated using SW Data used were the output of CMIP5 'historical' experiment from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA.

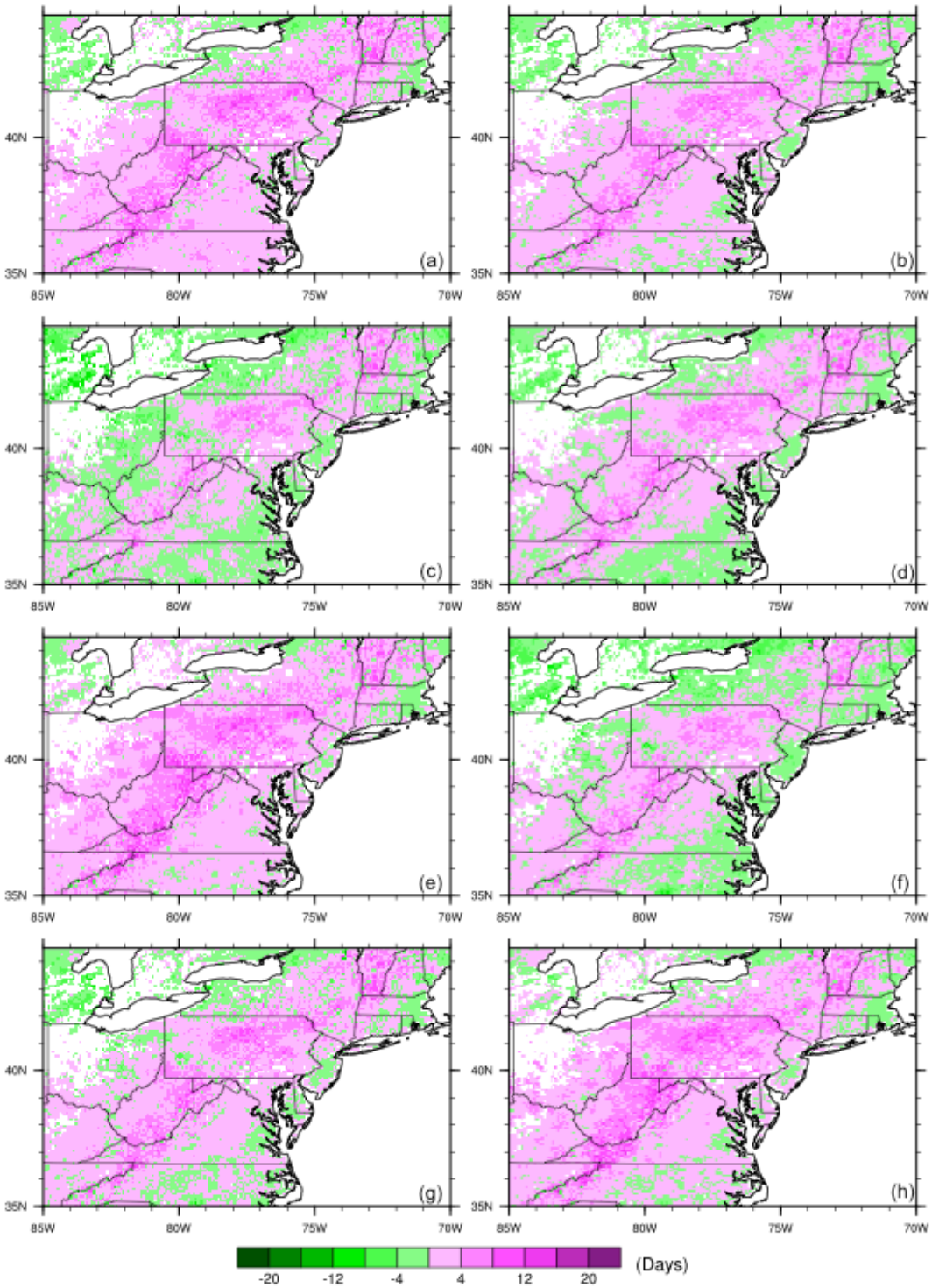


Figure 4.2 Difference (CMIP5-ZedX) in 20-year (1986-2005) average leaf onset simulated using SW2 Data used were the output of CMIP5 'historical' experiment from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA.

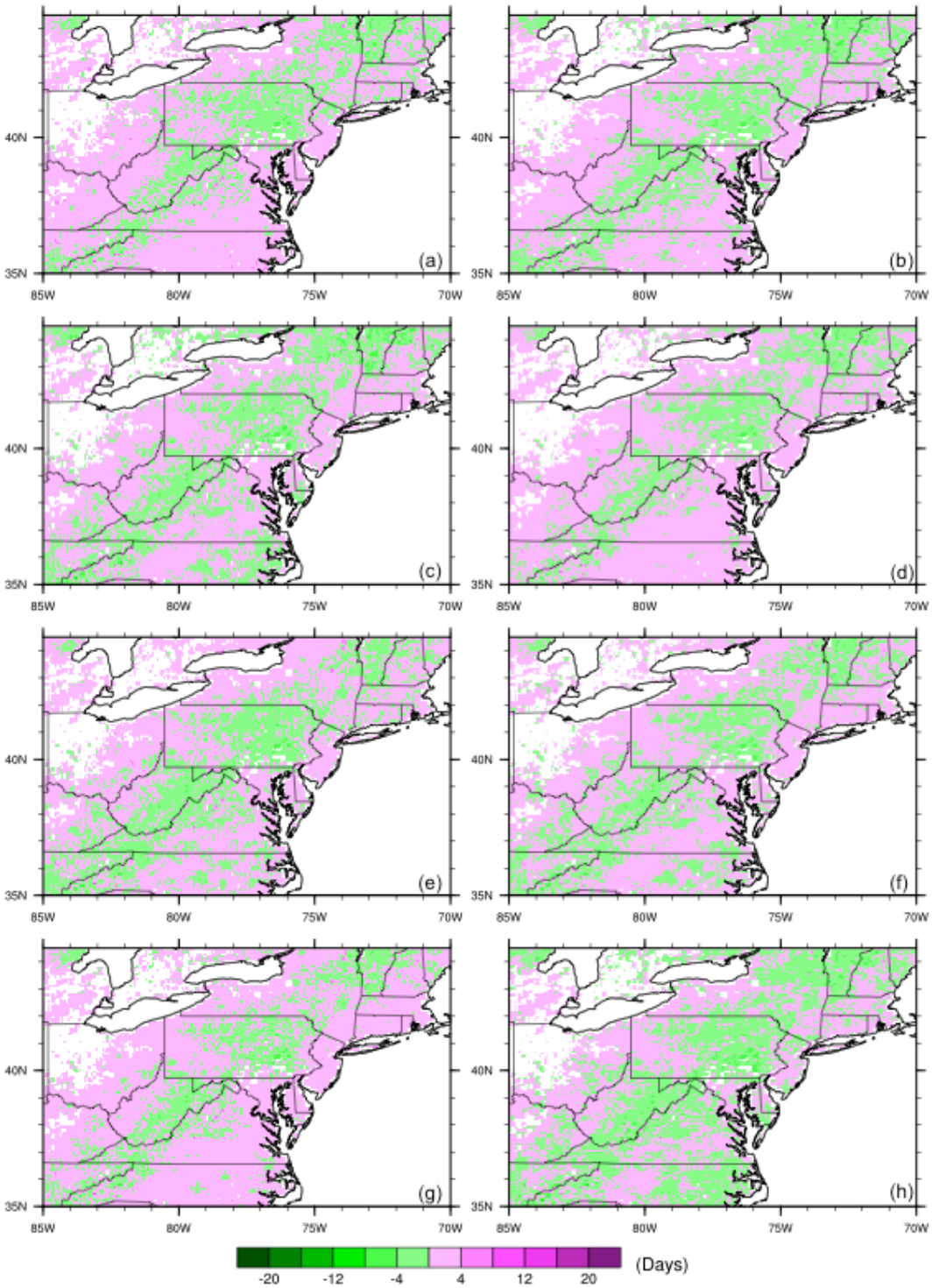


Figure 4.3 Difference (CMIP5-ZedX) in 20-year (1986-2005) average leaf offset simulated using DLP Data used were the output of CMIP5 'historical' experiment from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA.

Table 4.3 Regional average absolute difference (days) in 20-year average phenological dates between simulations with CMIP5 and ZedX data

(Days)	BCC	CSIRO	INM	IPSL	MIROC	MPI	NCAR	NOAA
Onset-SW	2.50	1.98	2.01	1.92	2.58	2.19	2.07	2.51
Onset-SW2	2.41	1.93	1.96	1.88	2.50	2.13	2.02	2.45
Offset-DLP	0.70	0.67	0.78	0.69	0.66	0.71	0.64	0.85

A4.3) simulated with CMIP5 data show spatial patterns similar to those retrieved from satellite data (Fig. 3.2), which along with the quantitative comparison suggests that there is no large bias in historical phenology simulated with CMIP5 data.

4.3.2 Phenological changes in the 21st century

Because the trends found in simulated phenology were relatively small despite of climate scenarios and GCMs, in order to be more intuitive, the trends are presented as the total days of change over 2006-2100 (i.e., linear trend multiplied by 95 years) in this chapter. Under RCP4.5 scenario, when SW model was used to simulate leaf onset, significant negative trends (i.e., earlier leaf onset) can be found over the entire study region when driven with BCC, CSIRO, MIROC and MPI (Fig. 4.4). When INM and NCAR were used to drive the model, only the northern part of the study region shows significant trends. In the contrary, only the southern part of the study region shows significant trends when driven with NOAA. The magnitude of change in leaf onset has a slight latitudinal gradient with higher latitude showing larger changes, except for the simulation with NOAA, which is consistent with the spatial patterns of the trends in projected spring temperature (Fig. A4.4). Regional average change in leaf onset varies across GCMs in a range of -8.7 to -3.8 days with the largest change found in the simulation with MIROC (Fig 4.5) due to the highest rate of warming projected by this GCM (Fig A4.5). The sensitivity of leaf onset to spring temperature is roughly -2.7 days

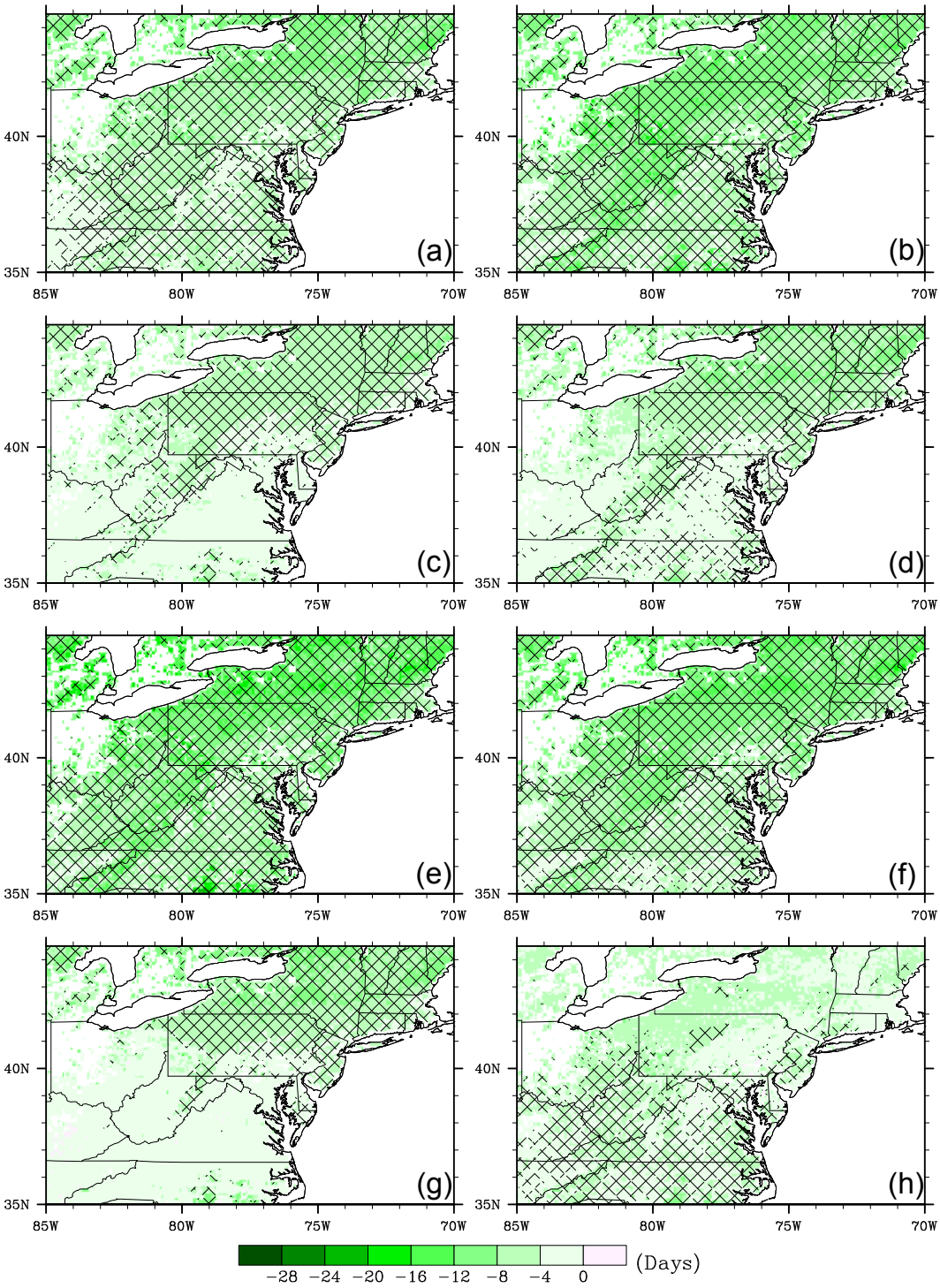


Figure 4.4 Total days of change in leaf onset over 2006-2100 simulated using SW under RCP4.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA. Hatching indicates trends are significant at $P<0.05$.

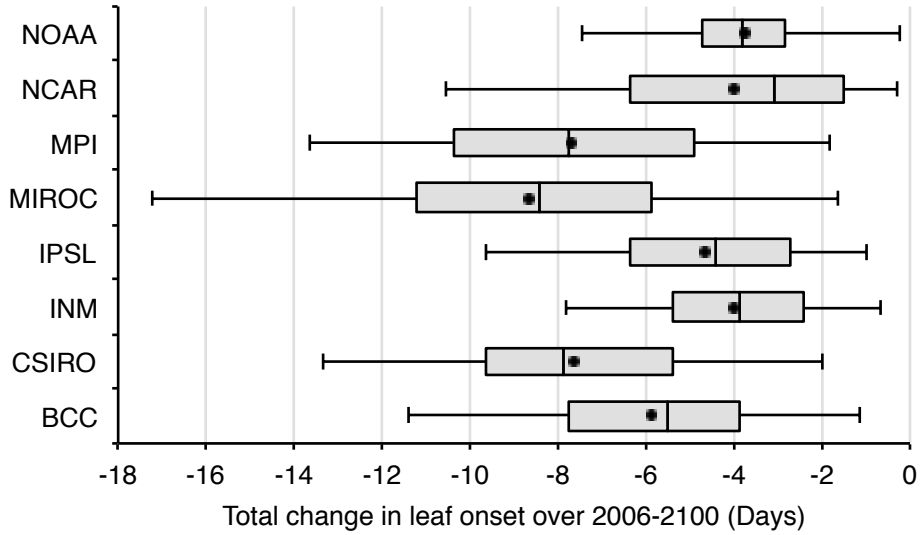


Figure 4.5 Box-and-whisker plot of the changes in leaf onset simulated using SW under RCP4.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

per °C (averaged across the region combining all GCMs). When SW2 model was used to simulate leaf onset, the long-term change is similar to the simulation using SW model in terms of spatial pattern (Fig. 4.6). The regional average change is slightly smaller than that from simulation using SW, ranging between -8.6 to -3.6 days (Fig. 4.7). The average sensitivity to spring temperature is -2.6 days per °C.

For leaf offset, significant positive trends (i.e., later leaf offset) can be found over the entire study region when driven with BCC, CSIRO, IPSL and MIRCO (Fig. 4.8), which projected relatively high rate of warming in autumn temperature (Fig. A4.6 and A4.7). Simulations with MPI, NCAR and NOAA show scattered significant positive trends across the study region, while the simulation with INM only shows positive trends in the most northern part with some insignificant negative trends found in the southern part (Fig. 4.8). Although the change in leaf offset varies spatially and sometimes exceeds 10 days in the simulation with some GCMs (e.g., BCC, IPSL, MIRCO), the regional average

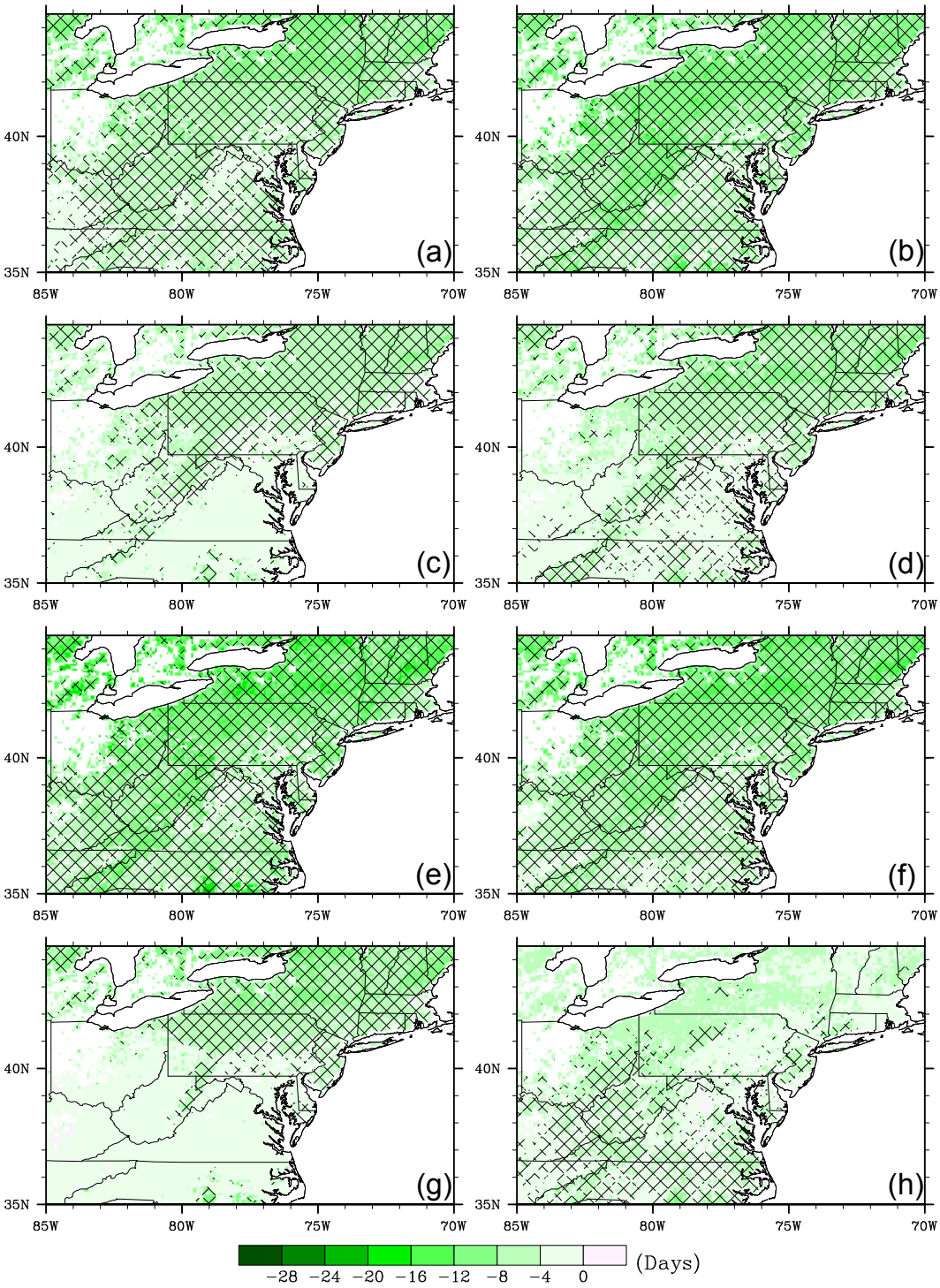


Figure 4.6 Total days of change in leaf onset over 2006-2100 simulated using SW2 under RCP4.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA. Hatching indicates trends are significant at $P<0.05$.

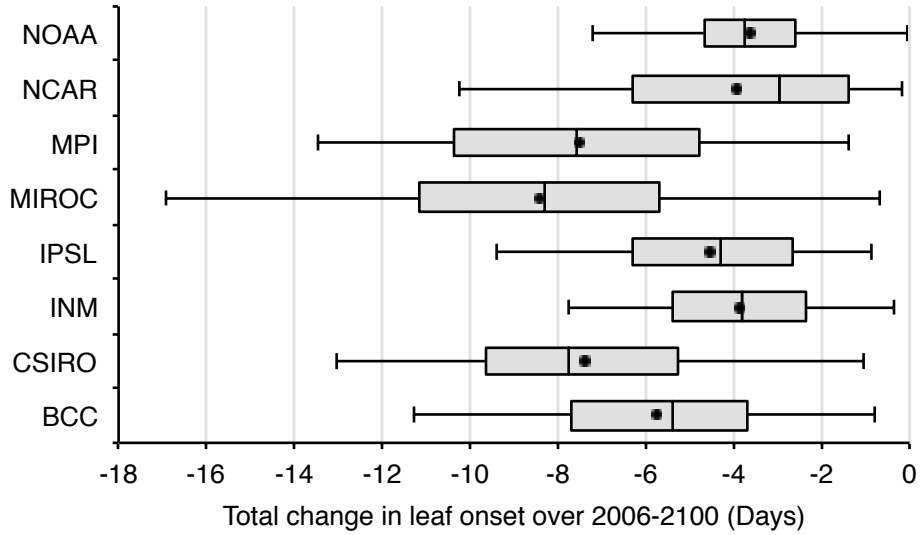


Figure 4.7 Box-and-whisker plot of the changes in leaf onset simulated using SW2 under RCP4.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

of change is fairly small, ranging between 1.2 and 4.4 days across GCMs (Fig. 4.9). The average sensitivity of leaf offset to autumn temperature is 1.4 days per °C.

Under RCP8.5 scenario, the changes in leaf onset simulated using SW and SW2 show similar spatial patterns as under RCP4.5 scenario (Fig. 4.10 and 4.11). In all runs, significant negative trends (i.e., earlier leaf onset) can be found over the entire study region with larger magnitude of change found on higher latitude in corresponding to the latitudinal gradient in the warming in spring temperature (Fig. A4.8). As a result of the higher rate of warming under RCP8.5 scenario (Fig. A4.9), the magnitude of change in leaf onset approximately doubled compared with that under RCP4.5 scenario, with the regional average change ranging between -16.1 and -9.6 days when using SW and between -15.7 and -9.3 days when using SW2 (Fig. 4.12 and 4.13), which resulted in an average sensitivity of -2.4 days per °C from SW and -2.3 days per °C from SW2. For leaf offset, no matter which GCM was used as driving data, positive trends (i.e., later leaf

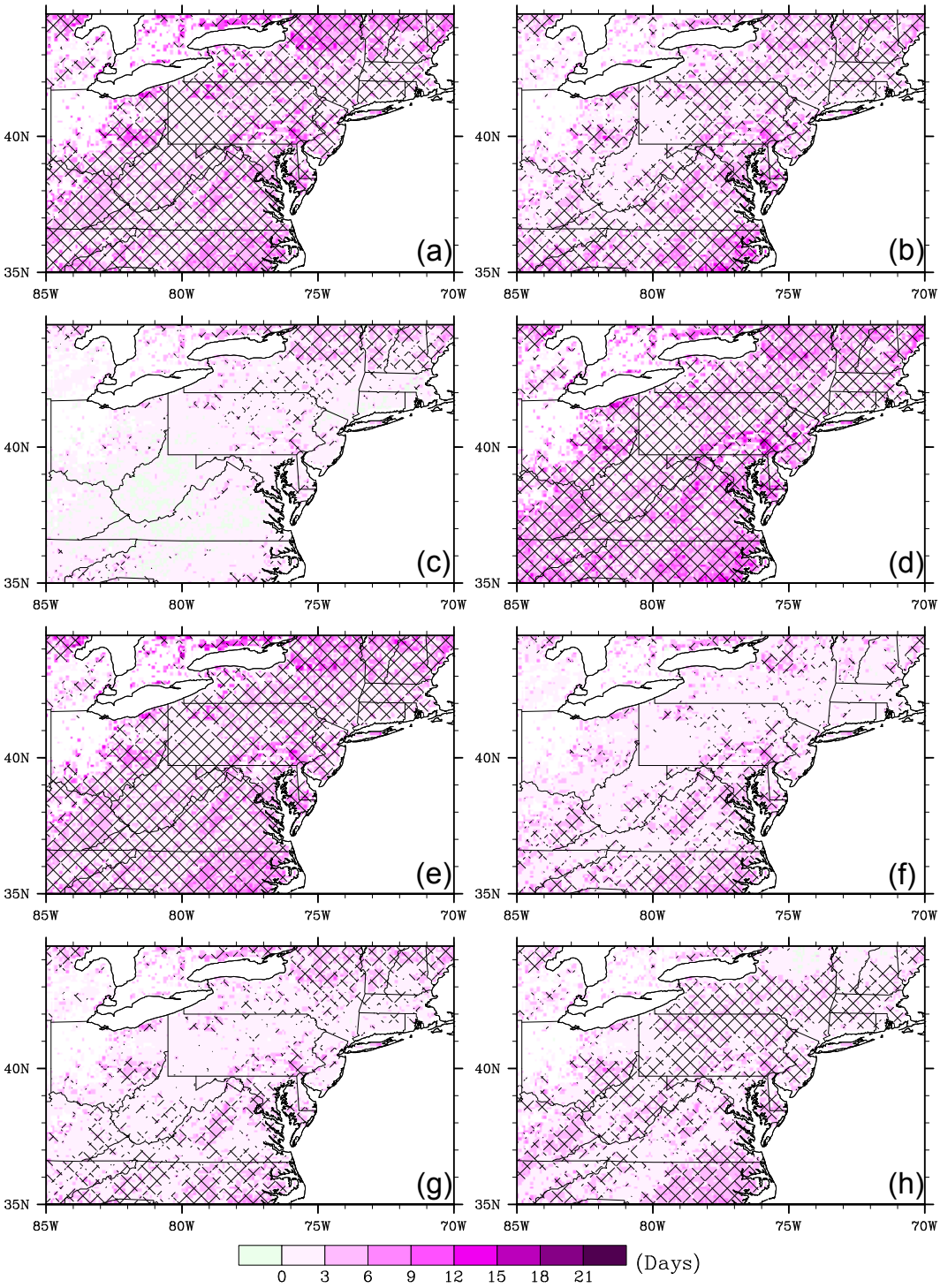


Figure 4.8 Total days of change in leaf offset over 2006-2100 simulated using DLP under RCP4.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA. Hatching indicates trends are significant at $P<0.05$.

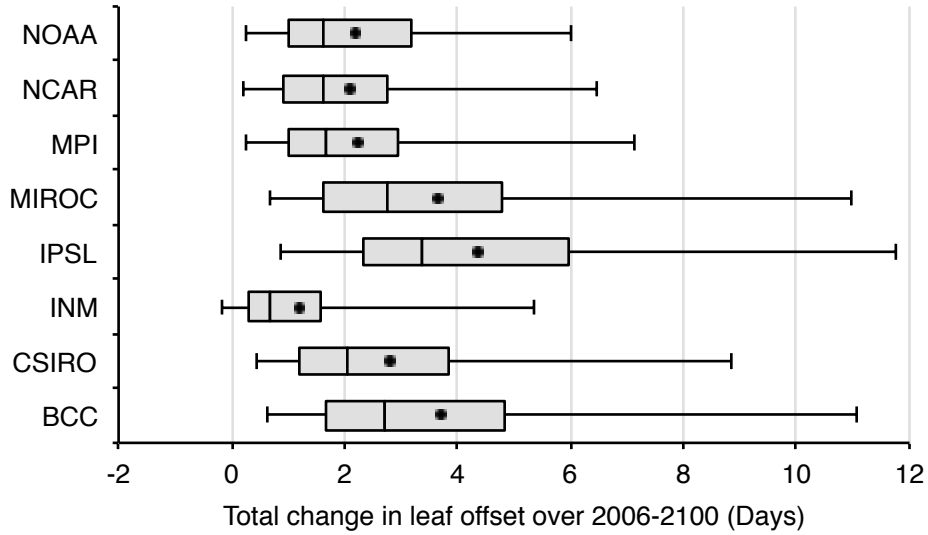


Figure 4.9 Box-and-whisker plot of the changes in leaf offset simulated using DLP under RCP4.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

offset) can be found over the entire study region in corresponding to the warming in autumn temperature (Fig. A4.10 and A4.11), with more pronounced changes appearing over the southeastern part of the study region (e.g., North Carolina) (Fig. 4.14). On regional average, the change in leaf offset varies across GCMs in a range of 5.1 to 11.1 days (Fig. 4.15), with a average sensitivity of 1.8 days per °C.

4.3.3 Impact of photoperiod on leaf onset

With t_0 optimized (i.e., the implicit photoperiod limitation imposed), the difference in the change of leaf onset between simulations using SW and SW2 is relatively small under both RCP4.5 and RCP8.5, with most grid cells showing a difference between -2 and 2 days (Fig. 4.16 and 4.17). Generally, there is no clear spatial pattern in either the sign or the magnitude of difference, except that Virginia and Maryland show some relatively large negative difference (i.e., earlier leaf onset simulated using SW). The differences in the change in leaf onset nearly cancel out across the region, resulting in a

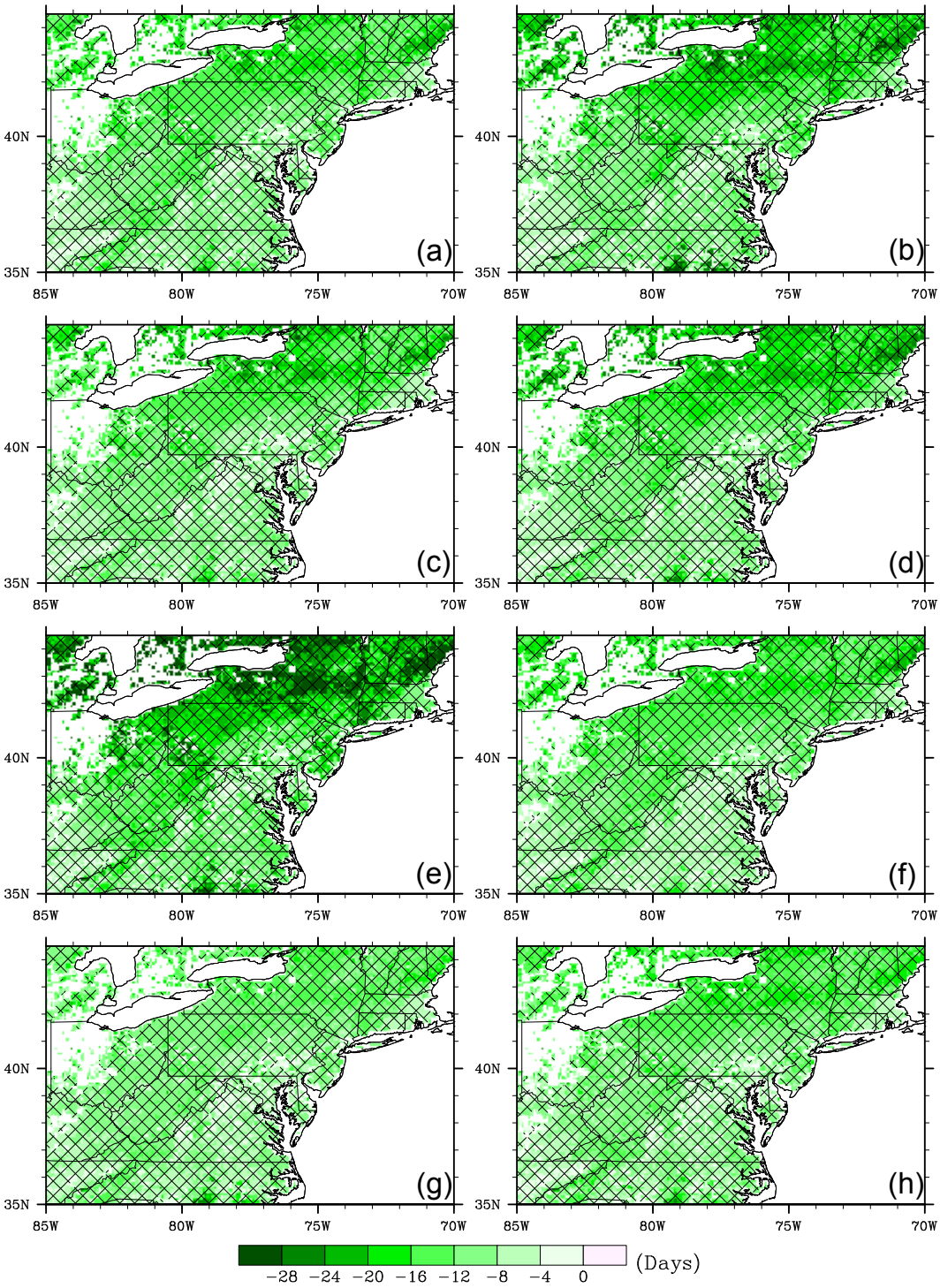


Figure 4.10 Total days of change in leaf onset over 2006-2100 simulated using SW under RCP8.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA. Hatching indicates trends are significant at $P<0.05$.

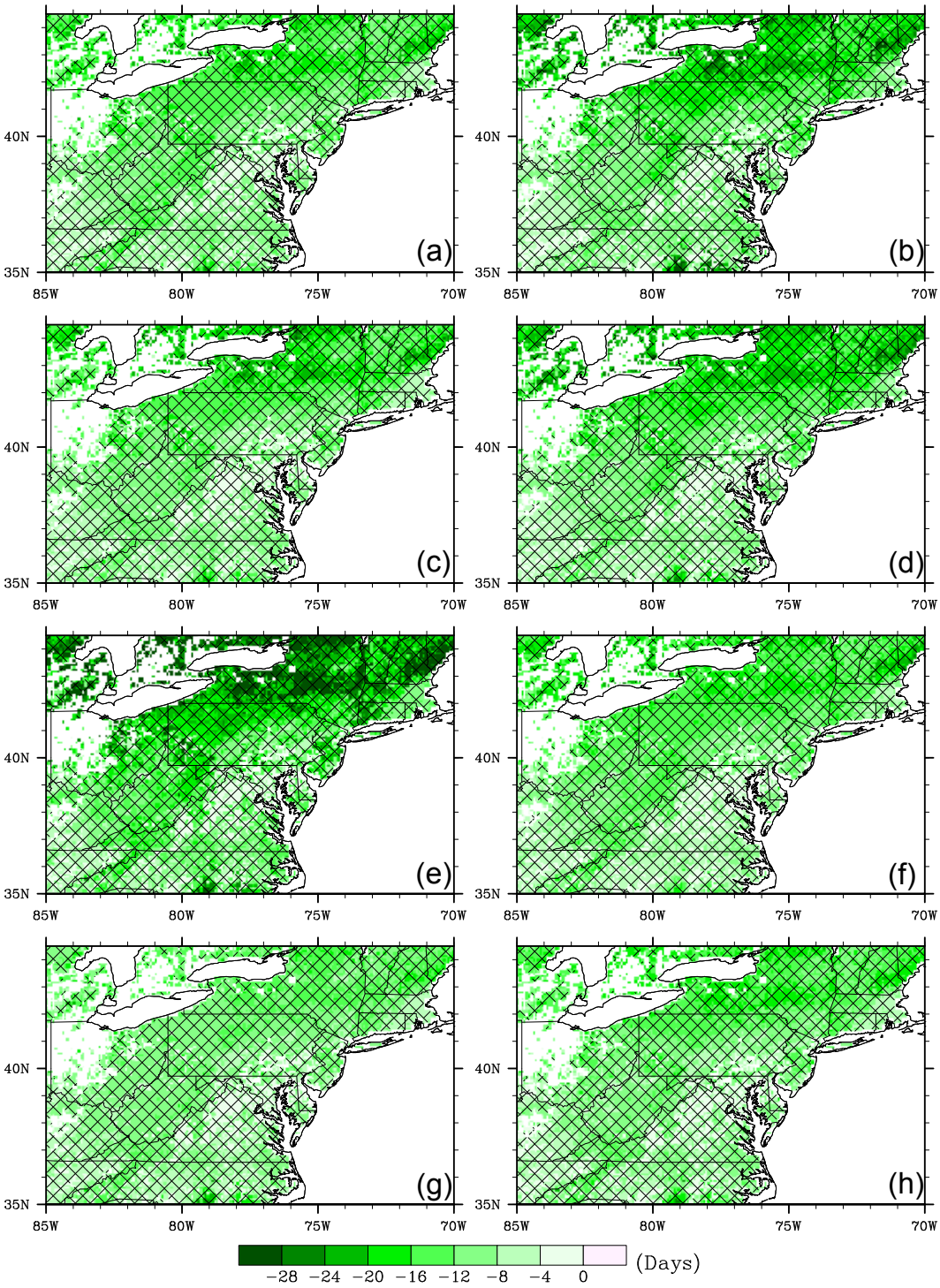


Figure 4.11 Total days of change in leaf onset over 2006-2100 simulated using SW2 under RCP8.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA. Hatching indicates trends are significant at $P<0.05$.

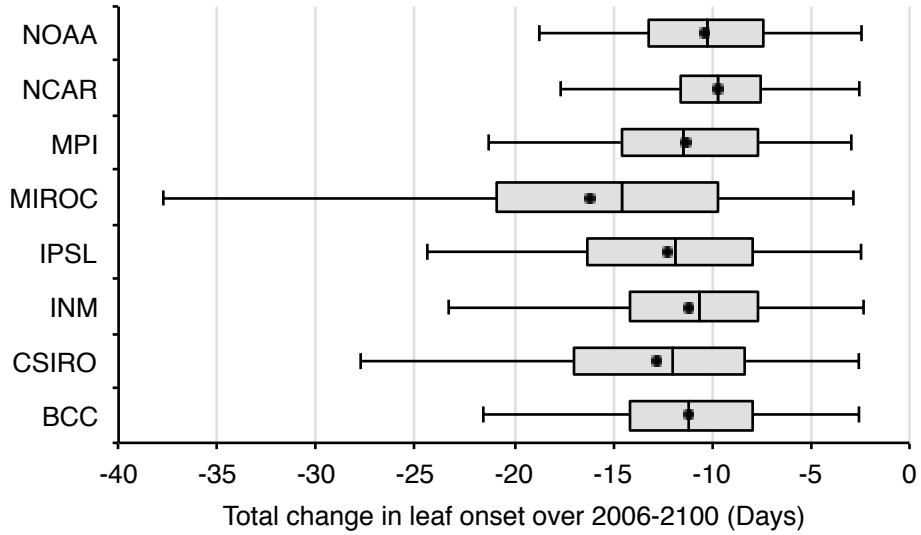


Figure 4.12 Box-and-whisker plot of the changes in leaf onset simulated using SW under RCP8.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

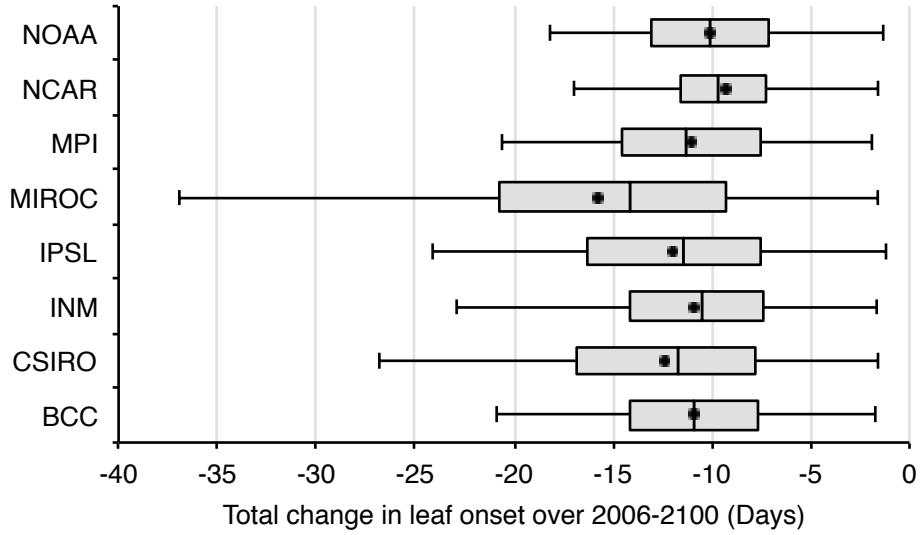


Figure 4.13 Box-and-whisker plot of the changes in leaf onset simulated using SW2 under RCP8.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

Table 4.4 Regional statistics of RMSD and AIC_c for leaf onset models

Model	Regional Average RMSD	Standard Deviation of RMSD	Regional Average AIC_c	Standard Deviation of AIC_c
SW (t_0 optimized)	3.14	1.92	28.58	6.91
SW-b (t_0 fixed)	5.19	2.22	37.30	7.10
SW2 (t_0 optimized)	3.07	1.88	37.54	6.98
SW2-b (t_0 fixed)	5.04	2.22	40.93	7.35

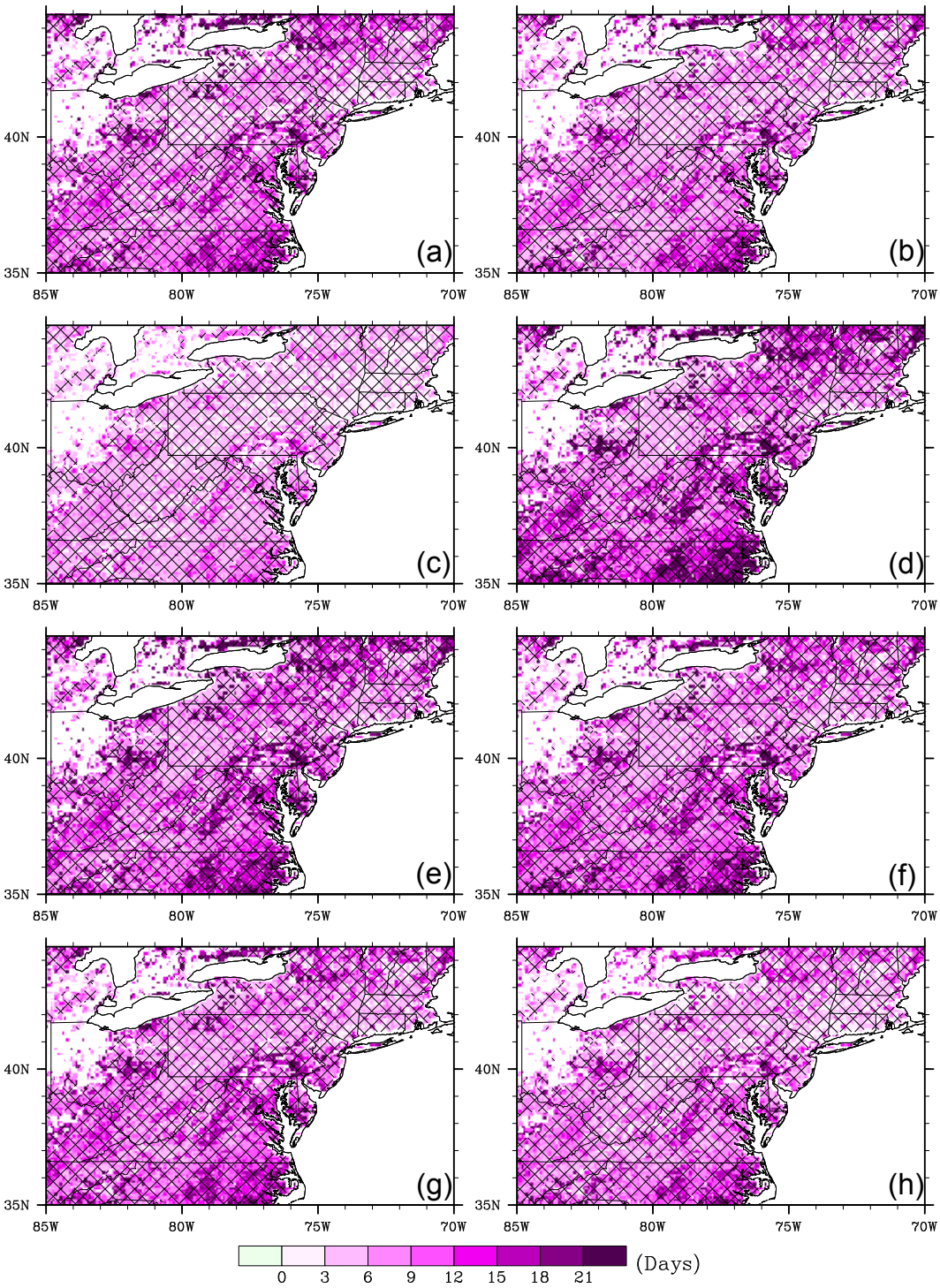


Figure 4.14 Total days of change in leaf offset over 2006-2100 simulated using DLP under RCP8.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA. Hatching indicates trends are significant at $P<0.05$.

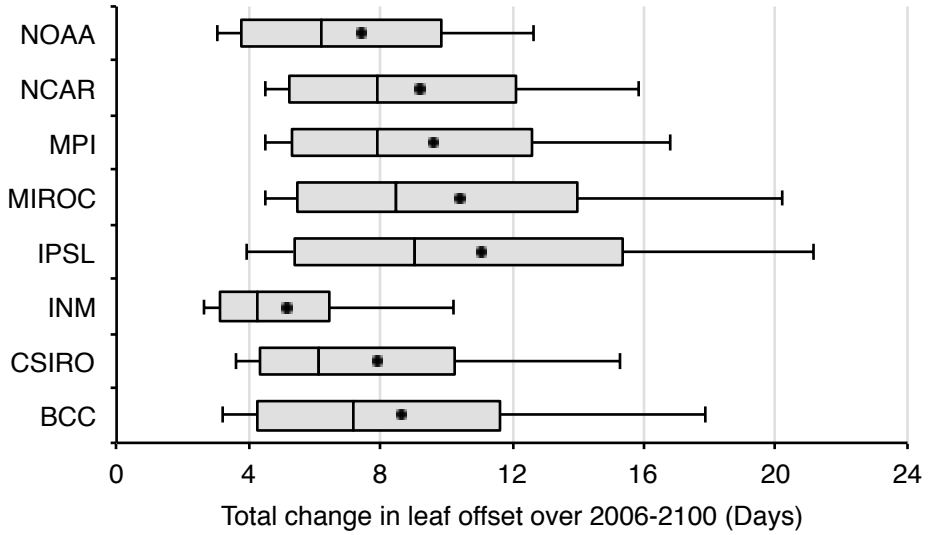


Figure 4.15 Box-and-whisker plot of the changes in leaf offset simulated using DLP under RCP8.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

regional average difference ranging between -0.22 and -0.09 days under RCP4.5, and between -0.44 and -0.27 days under RCP8.5 (Fig. 4.18 and 4.19).

When the parameter t_0 was fixed at January 1st, which eliminate the implicit photoperiod limitation, the minimized RMSD between modeled and remotely sensed leaf onset increased by about 2 days on average (Table 4.4). The reduction in the number of parameters that were optimized could not compensate the increase in RMSD, resulting in a higher regional average AIC_c score for both models (Table 4.4). The consistent higher RMSD and AIC_c score suggest that models with t_0 fixed at January 1st (i.e., SWb and SW2b) are less likely to be the true model than those with t_0 optimized (i.e., SW and SW2). Similar to the difference between the changes in leaf onset projected by SW and SW2 (Fig. 4.16-4.19), the difference between SWb and SW2b is also fairly small across the region under both scenarios (Fig. A4.12 - A4.15).

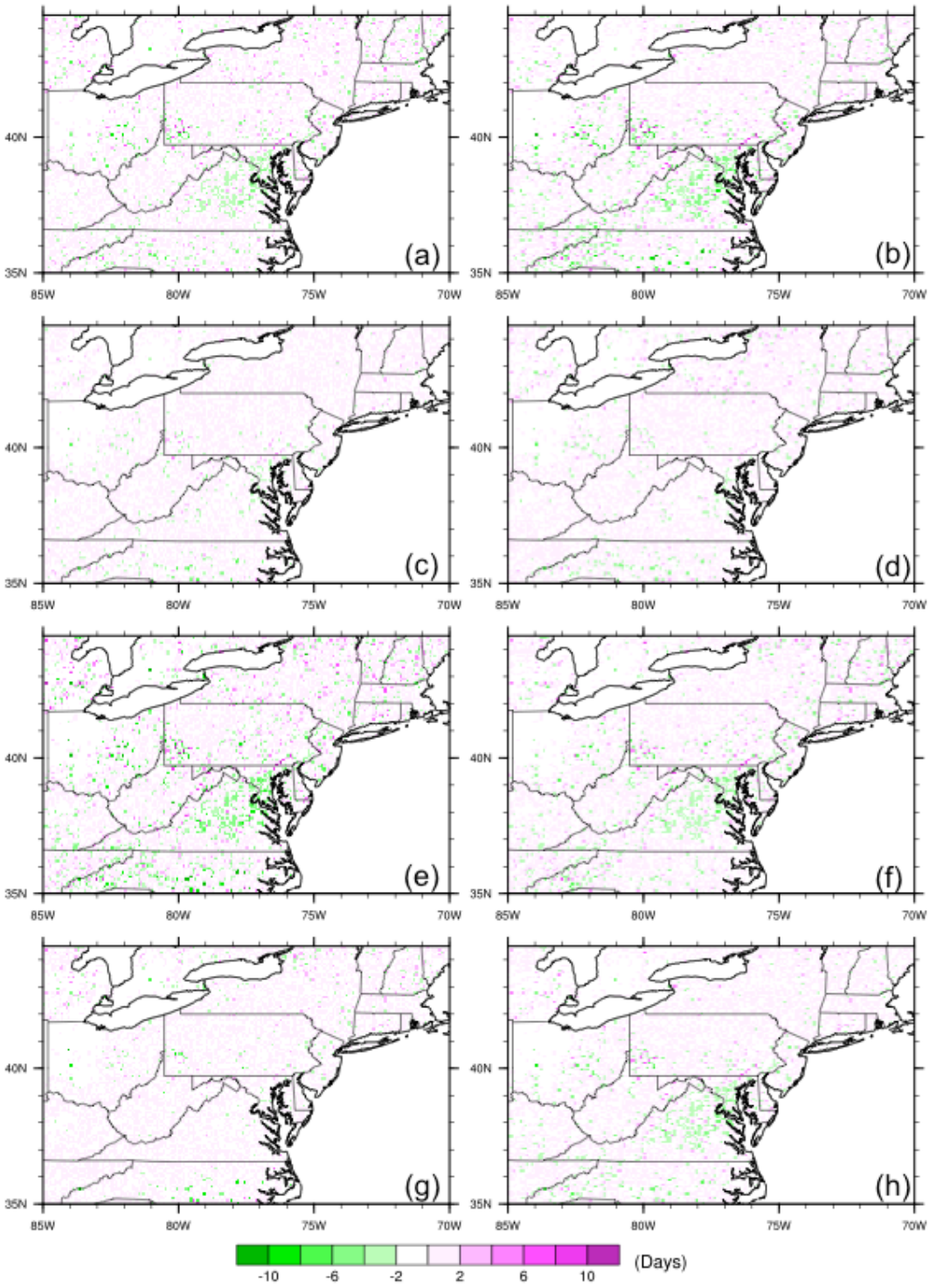


Figure 4.16 Difference (SW-SW2) in the change in leaf onset over 2006-2100 under RCP4.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA.

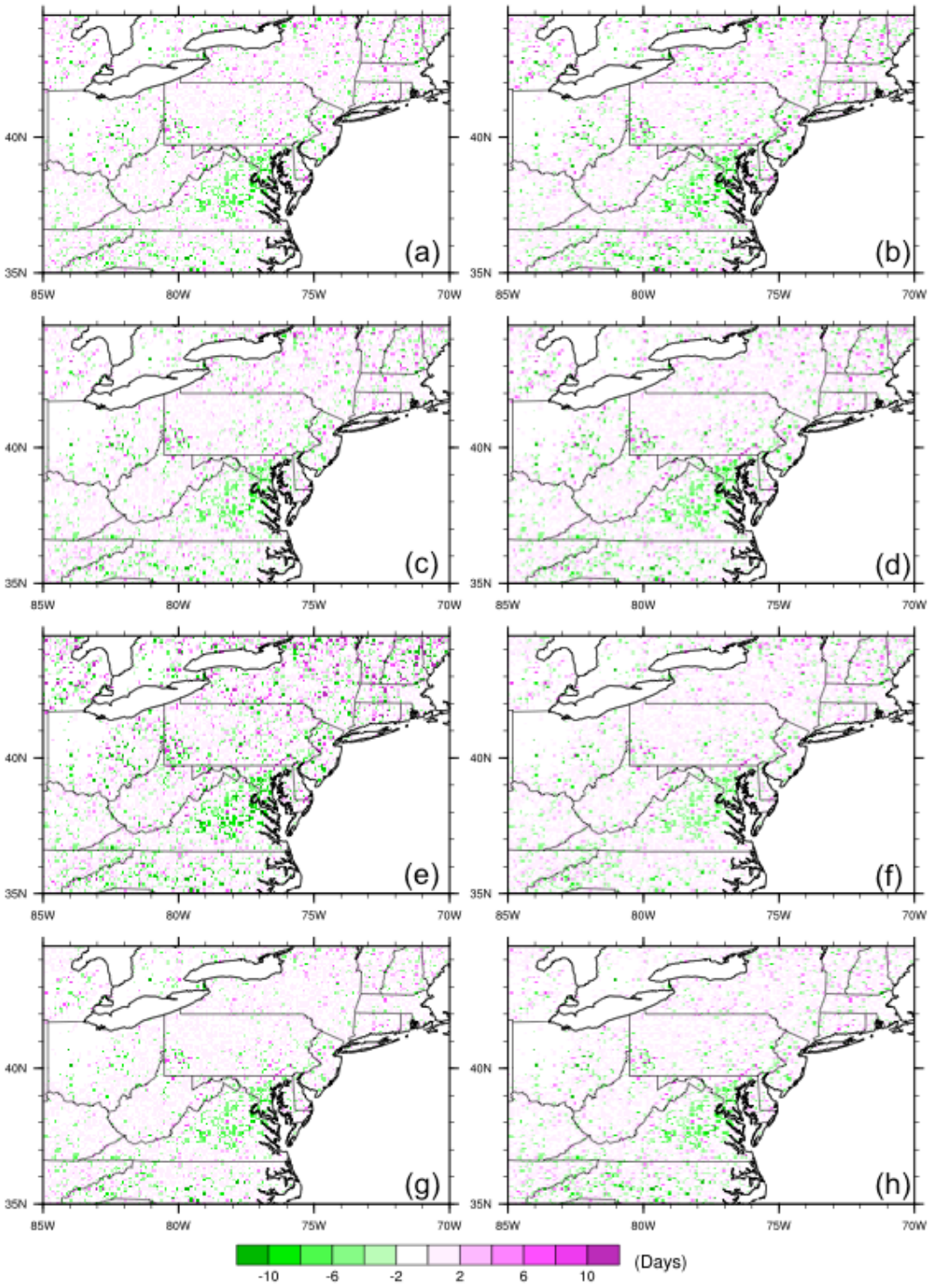


Figure 4.17 Difference (SW-SW2) in the change in leaf onset over 2006-2100 under RCP8.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA.

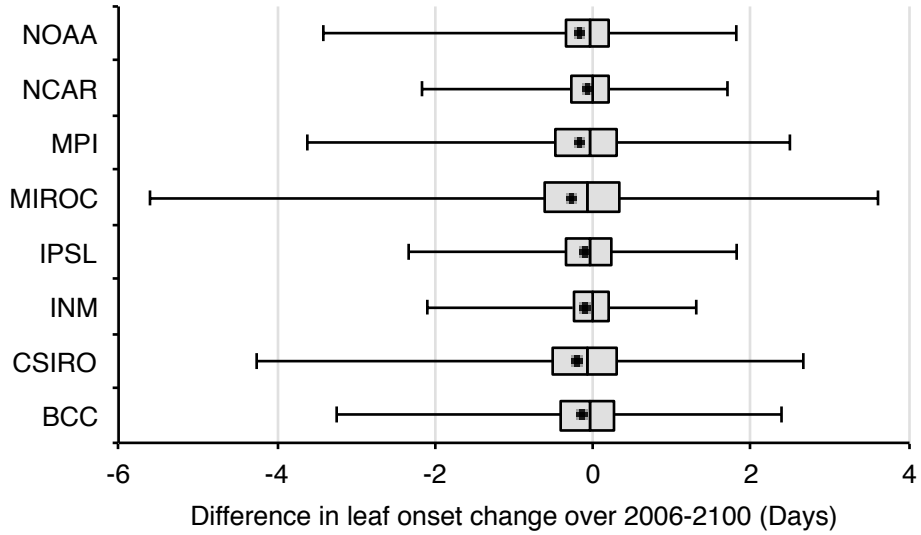


Figure 4.18 Box-and-whisker plot of the difference (SW-SW2) in leaf onset change under RCP4.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

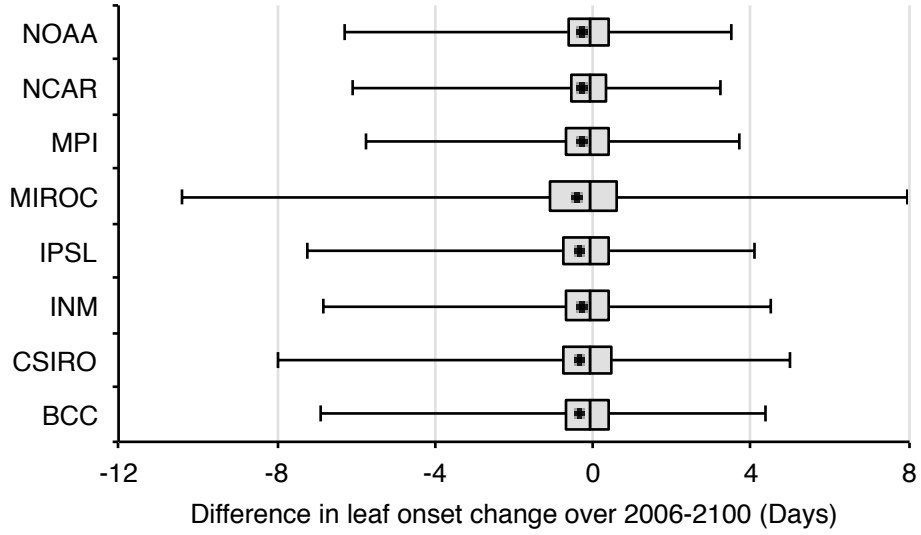


Figure 4.19 Box-and-whisker plot of the difference (SW-SW2) in leaf onset change under RCP8.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

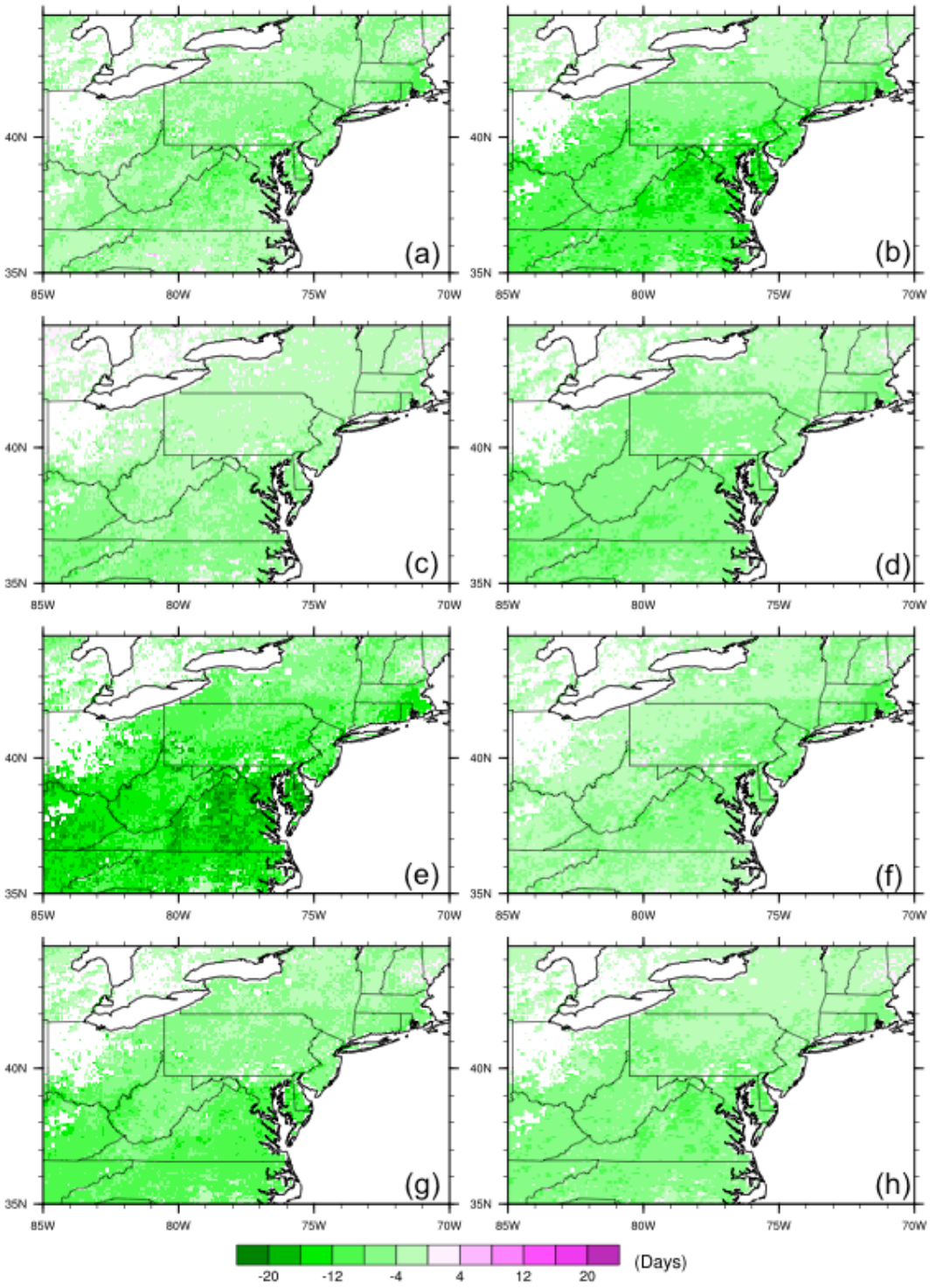


Figure 4.20 Difference (SWb-SW) in the change in leaf onset over 2006-2100 under RCP4.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA.

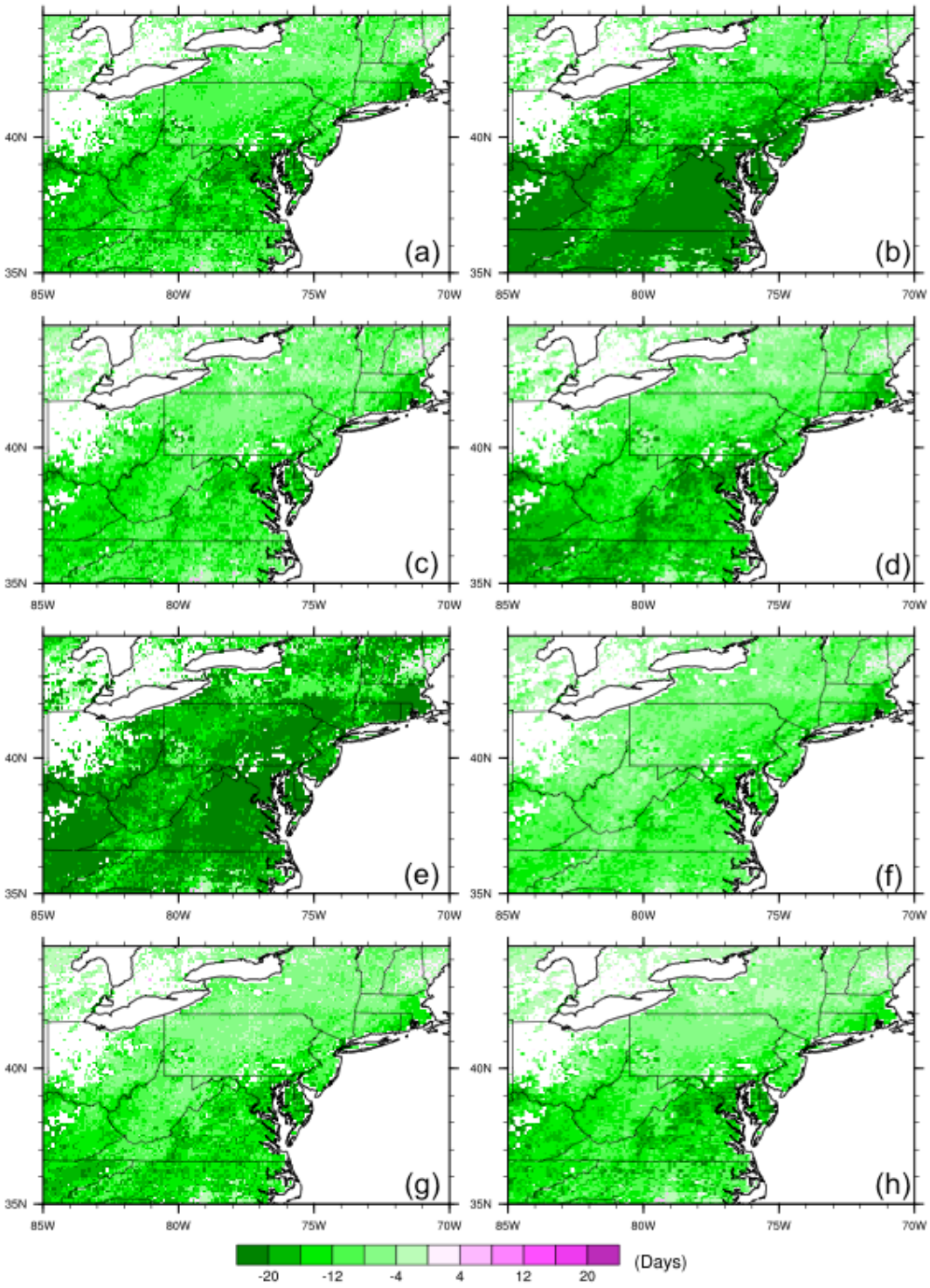


Figure 4.21 Difference (SWb-SW) in the change in leaf onset over 2006-2100 under RCP8.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA.

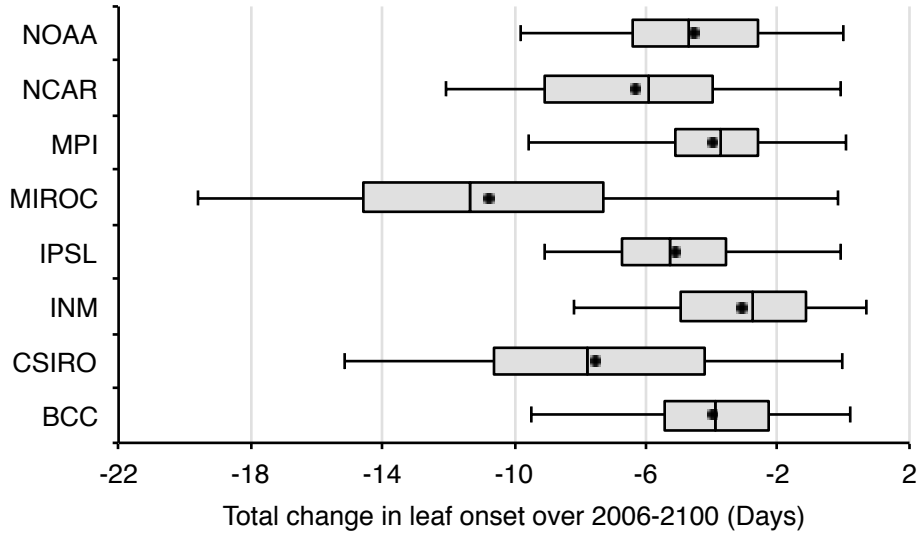


Figure 4.22 Box-and-whisker plot of the difference (SWb-SW) in leaf onset change under RCP4.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

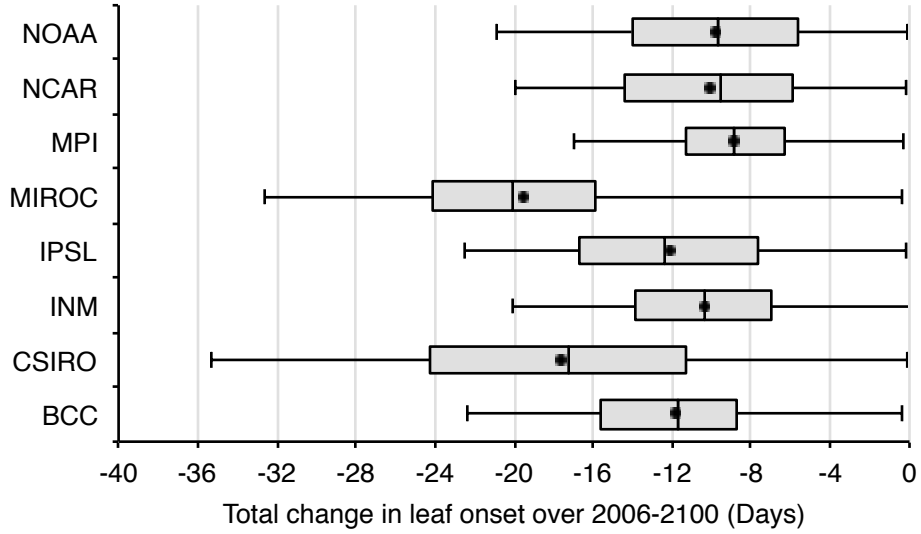


Figure 4.23 Box-and-whisker plot of the difference (SWb-SW) in leaf onset change under RCP8.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

However, the elimination of the implicit photoperiod limitation (i.e., t_0 fixed at January 1st) consistently resulted in relatively large difference in the projected change in leaf onset under both scenarios, no matter whether the explicit photoperiod term was

included. Figure 4.20 and 4.21 show the difference between the change in leaf onset projected by SWb and SW under RCP4.5 and RCP8.5, respectively. Most of the study region shows negative difference except for a few scattered grid cells, suggesting that the leaf onset tends to be further advanced without photoperiod limitation. (Similar for the difference between SW2b and SW2, Fig. A4.16 and A4.17). At the same location, the magnitude of further advancement in leaf onset is correlated with the change in temperature. For example, under the same scenario, larger advancement is usually found in the simulation with GCMs that show higher rate of warming (e.g., CSIRO, MIROC); while for simulations with same GCM, the further advancement is larger under RCP8.5 than under RCP4.5. On regional average, without the implicit photoperiod limitation, leaf onset may move earlier by another 3.2~11.0 days under RCP4.5 (Fig. 4.22), and 8.9~19.5 days under RCP8.5 (Fig. 4.23) (Also see Fig. A4.18 and A4.19 for the difference between SW2b and SW2). These further advancements are close to the total change in leaf onset projected by the models with the implicit photoperiod limitation. In other words, the earlier trend in leaf onset would double without photoperiod limitation.

4.4 Discussion

Temperature is considered to be the most important factor that drives the spring phenology of temperate and boreal plants (Linkosalo et al., 2006; Polgar and Primack, 2011), while recent studies claim that photoperiod also plays an important role in regulating spring phenology (Körner and Basler, 2010). Photoperiod limitation is therefore built into the widely used ‘Spring Warming’ model (Hunter and Lechowicz, 1992) either implicitly through optimizing the start date of GDD accumulation

(Migliavacca et al., 2012), or explicitly through introducing a photoperiod term as weight to adjust the forcing function (i.e., SW2 in this chapter) (Blümel and Chmielewski, 2012). By optimizing these two models over a relatively large region rather than at a single site, this study shows that including the explicit photoperiod term slightly improves the overall model performance in terms of lowering the RMSD (Table 4.4) as stated by Blümel and Chmielewski (2012). However, this might be merely a result of over-fitting rather than an effective mechanism according to the AIC theory (Migliavacca et al., 2012). The inter-comparison in this study suggests that at regional scale the explicit photoperiod term does not have a significant impact on the projected change in leaf onset in the 21st century, regardless of whether the implicit photoperiod limitation is included. The difference in the projected change in leaf onset is smaller when t_0 was fixed at January 1st (i.e., SWb vs. SW2b, Fig. A4.12 and A4.13) than that when t_0 was set to a later date to optimize model (i.e., SW vs. SW2, Fig. 4.16 and 4.17). Besides, those relatively large differences between SW and SW2 (Fig. 4.16 and 4.17) are usually found where the parameter t_0 of SW and SW2 show large difference (Fig. A4.20), suggesting that those differences in projected leaf onset should be partially attributed to the difference in t_0. The difference found in Blümel and Chmielewski (2012) might also be explained by the same reason.

When the implicit photoperiod limitation is considered, the temperature sensitivities of leaf onset in the 21st century derived from SW (-2.7 days per °C under RCP4.5, and -2.4 days per °C under RCP8.5) are similar to those found at Harvard Forest using similar models with photoperiod limitation (Migliavacca et al., 2012). The sensitivity under RCP4.5 is slightly lower than the historical sensitivity observed in a similar region

(Keenan et al., 2014), which support the evaluation that SW model has relatively good performance in representing reality. Considering the sensitivity is also lower under RCP8.5 than under RCP4.5, the lower-than-history sensitivity may be due to the nonlinearity of 'Spring Warming' model (Leaf onset becomes less sensitive as temperature gets higher). Compared with the observed historical sensitivity (Keenan et al., 2014), the average sensitivity of leaf offset derived from DLP is a little lower under RCP4.5 (1.4 days per °C), while is nearly the same under RCP8.5 (1.8 days per °C), suggesting that DLP model has a relatively good performance in reproducing the historical sensitivity of leaf offset.

This study shows that excluding the implicit photoperiod limitation substantially increases the model errors in the study region (Table 4.4), which might be explained by the fact that the dominant tree species in the study region (e.g., American beech, red oak) rely on both photoperiod and temperature to break dormancy (Caffarra and Donnelly, 2011). However, photoperiod limitation is usually not considered in the leaf onset models embedded in terrestrial biosphere models, which can potentially introduce large errors in projected leaf onset in the 21st century. The average sensitivities of leaf onset would become -5.5 days per °C under RCP4.5 and -4.9 days per °C under RCP8.5 when the implicit photoperiod limitation is omitted. That is, when photoperiod does act as a limiting factor and is overlooked, the change in leaf onset in the 21st century will be substantially overestimated. On the contrary, as not all species respond to photoperiod cues (Ghelardini et al., 2010; Körner and Basler, 2010), mistakenly including photoperiod limitation (e.g., from parameter optimization) can lead to large

underestimate in the change in leaf onset. The errors will propagate to the simulated exchange of carbon, water and energy in terrestrial biosphere models (Keenan et al., 2014; Richardson et al., 2012), and in turn misrepresent the feedbacks to the atmosphere when coupled to climate models (Levis and Bonan, 2004; Richardson et al., 2013).

Although mathematical measures such as RMSD and AIC_c score are useful for evaluating the model performance in the history (Migliavacca et al., 2012) and the observed sensitivity of phenology to temperature can be used to guide the development of phenological models (Keenan et al., 2013), the forecast of future phenology may still be subject to large uncertainties without understanding how plants really respond to environmental factors (Korner and Basler, 2010). As this study focused on quantifying the potential impact of photoperiod on leaf onset, the models with chilling requirements are not considered due to their relatively poor performance (Chapter 3). The reason that they showed relatively poor performance might be that in the past the chilling requirements were usually fulfilled before the photoperiod threshold (i.e., start date of GDD accumulation) was met. However, the warming in winter temperature in the future may delay the fulfillment of chilling, and in turn delay leaf onset once the fulfillment of chilling occurs later than the photoperiod threshold date (Murray et al., 1989). Therefore, chilling requirements may also need to be included in leaf onset model for phenological projection.

A single model was used over the entire region in this study, which might have ruled out better models at some locations. As the role of chilling requirements (Vitasse et al., 2009) and photoperiod (Korner and Basler, 2010; Ghelardini et al., 2010) in

determining leaf onset in woody plants differ among species, different model structures may need to be used at different locations in future studies.

4.5 Conclusion

Phenological models were used with projected daily temperature from CMIP5 to simulate leaf onset and offset in the 21st century. Different model structures and parameterizations were compared to quantify the potential impact of photoperiod on the sensitivity of leaf onset to spring temperature. The results suggest that, on regional scale, the explicit photoperiod term introduced into 'Spring Warming' model (Blümel and Chmielewski, 2012) does not significantly affect the projected change in leaf onset in the 21st century. Under the scenario with reduced greenhouse gas emission, RCP4.5, leaf onset would be advanced by 3.8~8.7 days regionally by the end of 21st century, with a sensitivity to spring temperature of -2.7 days per °C; while leaf offset would be delayed by 1.4~4.4 days with a sensitivity to autumn temperature of 1.4 days per °C. Under the scenario with high green house gas emission, RCP8.5, leaf onset would be advanced by 9.6~16.1 days with a smaller sensitivity of -2.4 days per °C; while leaf offset would be delayed by 5.1~11.1 days with a higher sensitivity of 1.8 days per °C. Model evaluation suggests that leaf onset is more likely to be limited by photoperiod in the study region. Omitting the implicit photoperiod limitation (i.e., the start date of GDD accumulation) in Spring Warming model would double the sensitivity of leaf onset and lead to large errors, which could propagate to the simulation of ecosystem processes and feedbacks to the climate system. The findings in this study highlight the need to develop or select phenological models based on the physiological traits of dominant species. However, for

most species, it is still unclear what combination of factors and the genes are involved in determining leaf onset (Howe et al., 2003) and offset (Delpierre et al., 2009; Vitasse et al, 2011). More experimental studies such as those involving environmental manipulation (Norby et al., 2003; Hanninen et al., 2007; Nakamura et al., 2010) are needed to provide additional information.

Chapter 5 General Conclusion

Vegetation phenology plays an important role in regulating photosynthesis and other terrestrial ecosystem processes through controlling the timing of leaf activities. Vegetation phenology is driven by climate variables such as temperature, precipitation and photoperiod. Vegetation phenology also exerts various feedbacks to the climate system through affecting biophysical properties (e.g., albedo) and biogeochemical processes (e.g., carbon cycle). Dynamic ecosystem model is a useful tool that is often applied to study ecosystem responses to climate variation, and coupled to climate models to study ecological feedbacks. However, the most important two phenophases of temperate and boreal forests, leaf onset and offset, are are usually poorly represented in dynamic ecosystem models. The misrepresentation can in turn lead to errors in the simulation of ecosystem processes in dynamic ecosystem models, as well as the feedbacks to the atmosphere in coupled earth system models. The goal of this dissertation is to improve phenological models through incorporating phenological information retrieved from satellite imagery, examine the phenological trends in both history and future with improved phenological models, and quantify the impact of phenology on carbon cycle using a dynamic ecosystem model.

This dissertation focuses on the phenology of temperate deciduous trees in northeastern U. S. forests. Research conducted in this dissertation is summarized as follows: First, ground phenological observations along with LAI and carbon flux measurement made at the Harvard Forest were used to evaluate six vegetation-index-based methods for retrieving leaf onset and offset. The propagation of biases in remotely

sensed phenology to simulated phenology, GPP and NEP were also examined using Agro-IBIS dynamic ecosystem model. Then, one of the evaluated methods was modified to retrieve historical leaf onset and offset over the northeastern U. S. forests, which were further used to parameterize multiple phenological models. Those models with relatively good performance were selected to conduct retrospective modeling of leaf onset and offset. Agro-IBIS experimental simulations were run to quantify the impact of phenological trends on GPP and NEP. Finally, the changes in phenology in the 21st century were quantified by using projected climate data to drive phenological models. Particularly, the potential impact of photoperiod on leaf onset was examined.

The evaluation of remotely sensed phenological metrics suggests that there are large discrepancies between the remotely sensed phenological metrics and ground phenology reference chosen according to the definition in Agro-IBIS, with two leaf onset metrics and one leaf offset metric showing relatively good performance. Those discrepancies can be attributed to the definition of phenological metrics, the parameters used to calculate the metrics, as well as the input data. When remotely sensed phenological metrics with bias were used to parameterize phenological models, the bias tends to be maintained in the simulated phenology, which will propagate to the simulated GPP and NEP. However, by adjusting parameters used, it is possible to improve remotely sensed phenological metrics to fit the specific need of research.

The evaluation of phenological models suggests that, in the study region, Spring Warming models (with and without the explicit photoperiod term) outperform simpler temperature threshold model as well as more complex models with chilling requirement

in terms of minimizing the RMSD between simulated and remotely sensed leaf onset; while the Delpierre model shows the best performance in simulating leaf offset. Retrospective modeling shows relatively small earlier trends in leaf onset and later trends in leaf offset. The trends in both leaf onset and offset are more pronounced in more recent decades in response to higher rate of warming. Agro-IBIS simulations suggest that, while the earlier trends in leaf onset have caused small increasing trends in annual GPP and NEP, the trends in GPP and NEP caused by the change in leaf offset are negligible. However, simulated GPP and NEP actually have strong correlation with phenology. The relationship between productivities and phenology varies across the region, and change in leaf onset generally shows larger effect on productivities than the same change in leaf offset.

The projection of phenology in the 21st century shows further advancement in leaf onset and delay in leaf offset. Changes in leaf onset simulated using Spring Warming model with and without the explicit photoperiod term are not significantly different. The sensitivity of both leaf onset and offset are similar to the observation. While leaf onset shows higher sensitivity under scenario with lower greenhouse gas emission, sensitivity of leaf offset is higher under high emission scenario. Extra parameterization of Spring Warming model suggests the implicit photoperiod threshold plays an important role in limiting the simulated change in leaf onset. Excluding the photoperiod limitation not only lower the performance of Spring Warming model in simulating historical leaf onset, but also double the sensitivity of leaf onset to spring temperature, which highlights the potential error in phenological projection and climate projections.

Based on the findings in this dissertation, future research can be conducted in following directions. First, to use phenological information derived from satellite imagery as a proxy of ground phenological observation, remotely sensed phenology needs to be further evaluated over wider spatial coverage and across different dominant species. Near-surface remote sensing (e.g., PhenoCam network) can potentially serve as standardized ground observation to address the scale issue of comparing remote sensing pixels with individual trees. Second, as different species respond to different combinations of climate variables, or to the same variables in different ways, different structures of phenological models need to be selected at different locations according to the physiology of dominant species. This requires experimental studies to further understand the combination of factors that drives phenology. Third, besides carbon cycle, phenological change has impact on other ecosystem properties and processes, which need to be examined quantitatively. For example, prolonged vegetation growing season may increase net radiation by lowering albedo, which has an effect opposite to potential increase in carbon assimilation. Advanced leaf onset may also increase evapotranspiration and have a cooling effect on local climate. In order to understand the net effect of these phenological feedbacks, simulation using coupled climate model is still needed.

Bibliography

Ahl, D.E. et al., 2006. Monitoring spring canopy phenology of a deciduous broadleaf forest using MODIS. Remote Sensing of Environment, 104: 88-95.

Anderson-Teixeira, K.J. et al., 2012. Climate-regulation services of natural and agricultural ecoregions of the Americas. Nature Climate Change, 2: 177-181.

Arora, V.K. and Boer, G.J., 2005. A parameterization of leaf phenology for the terrestrial ecosystem component of climate models. Global Change Biology, 11: 39-59.

Baldocchi, D., 2008. TURNER REVIEW No. 15. 'Breathing' of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. Australian Journal of Botany, 56: 1-26.

Balzter, H. et al., 2007. Coupling of Vegetation Growing Season Anomalies and Fire Activity with Hemispheric and Regional-Scale Climate Patterns in Central and East Siberia. Journal of Climate, 20: 3713-3729.

Birky, A.K., 2001. NDVI and a simple model of deciduous forest seasonal dynamics. Ecological Modelling, 143: 43-58.

Blümel, K. and Chmielewski, F.-M., 2012. Shortcomings of classical phenological forcing models and a way to overcome them. Agricultural and Forest Meteorology, 164: 10-19.

Botta, A., Viovy, N., Ciais, P., Friedlingstein, P. and Monfray, P., 2000. A global prognostic scheme of leaf onset using satellite data. Global Change Biology, 6: 7090-7725.

Burnham, K.P. and Anderson, D.R., 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer.

Caffarra, A. and Donnelly, A., 2011. The ecological significance of phenology in four different tree species: effects of light and temperature on bud burst. International journal of biometeorology, 55(5): 711-721.

Campbell, G.S. and Norman, J.M., 1998. An introduction to Environmental Biophysics.

Cannell, M.G.R. and Smith, R.I., 1983. Thermal time, chill days and prediction of budburst in picea sitchensis. Journal of Applied Ecology, 20: 951-963.

Chen, J. et al., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. Remote Sensing of Environment, 91: 332-344.

Chiang, J.-M. and Brown, K.J., 2007. Improving the budburst phenology subroutine in the forest carbon model PnET. ecological modelling, 205(3): 515-526.

Chmielewski, F.-M. and Rotzer, T., 2001. Response of tree phenology to climate change across Europe. Agricultural and Forest Meteorology, 108: 101-112.

Chuine, I., 2000. A unified model for budburst of trees. Journal of Theoretical Biology, 207: 337-437.

Chuine, I., Kramer, K. and Hänninen, H., 2003. Plant development models, Phenology: an integrative environmental science. Springer, pp. 217-235.

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

Churkina, G., Schimel, D., Braswell, B.H. and Xiao, X., 2005. Spatial analysis of growing season length control over net ecosystem exchange. Global Change Biology, 11: 1777-1787.

Cleland, E.E., Chuine, I., Menzel, A., Mooney, H.A. and Schwartz, M.D., 2007. Shifting plant phenology in response to global change. Trends in Ecology and Evolution, 22: 357-365.

de Beurs, K.M. and Henebry, G.M., 2005. Land surface phenology and temperature variation in the International Geosphere-Biosphere Program high-latitude transects. Global Change Biology, 11: 779-790.

de Beurs, K.M. and Henebry, G.M., 2008. Northern Annular Mode Effects on the Land Surface Phenologies of Northern Eurasia. Journal of Climate, 21: 4257-4279.

Delpierre, N. et al., 2009. Modelling interannual and spatial variability of leaf senescence for three deciduous tree species in France. Agricultural and Forest Meteorology, 149: 848-938.

Dragoni, D. and Rahman, A.F., 2012. Trends in fall phenology across the deciduous forests of the Eastern USA. Agricultural and Forest Meteorology, 157: 96-105.

Dragoni, D. et al., 2011. Evidence of increased net ecosystem productivity associated with a longer vegetated season in a deciduous forest in south-central Indiana, USA. Global Change Biology, 17: 886-897.

Dunn, A.L., Barford, C.C., Wofsy, S.C., Goulden, M.L. and Daube, B.C., 2007. A long-term record of carbon exchange in a boreal black spruce forest: means, responses to interannual variability, and decadal trends. Global Change Biology, 13: 577-590.

Estrella, N. and Menzel, A., 2006. Responses of leaf colouring in four deciduous tree species to climate and weather in Germany. Climate Research, 32: 253-267.

Fisher, J.I., Mustard, J.F. and Vadeboncoeur, M.A., 2006. Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. Remote Sensing of Environment, 100: 265-279.

Fisher, J.I., Richardson, A.D. and Mustard, J.F., 2007. Phenology model from surface meteorology does not capture satellite-based greenup estimations. Global Change Biology, 13.

Fitzjarrald, D.R., Acevedo, O.C. and Moore, K.E., 2001. Climatic consequences of leaf presence in the eastern United States. Journal of Climate, 14(4): 598-614.

Foley, J.A. et al., 1996. An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. Global Biogeochemical Cycles, 10: 603-628.

Friedl, M.A. et al., 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment, 83(1): 287-302.

Friedl, M.A. et al., 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sensing of Environment, 114(1): 168-182.

Friend, A., Stevens, A., Knox, R. and Cannell, M., 1997. A process-based, terrestrial biosphere model of ecosystem dynamics (Hybrid v3. 0). Ecological Modelling, 95(2): 249-287.

Friend, A.D. and White, A., 2000. Evaluation and analysis of a dynamic terrestrial ecosystem model under preindustrial conditions at the global scale. Global Biogeochemical Cycles, 14(4): 1173-1190.

Ghelardini, L., Santini, A., Black-Samuelsson, S., Myking, T. and Falusi, M., 2010. Bud dormancy release in elm (Ulmus spp.) clones—a case study of photoperiod and temperature responses. Tree physiology, 30(2): 264-274.

Hidalgo, H.G., Dettinger, M.D. and Cayan, D.R., 2008. Downscaling with constructed analogues: Daily precipitation and temperature fields over the United States. California Energy Commission PIER Final Project Report CEC-500-2007-123.

Hopkins, A.D., 1938. Bioclimatics: a science of life and climate relations. US Dept. of Agriculture.

Hu, J., Moore, D.J.P., Burns, S.P. and Monson, R.K., 2010. Longer growing seasons lead to less carbon sequestration by a subalpine forest. Global Change Biology, 16: 771-783.

Huete, A. et al., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment, 83: 195-213.

Hunter, A.F. and Lechowicz, M.J., 1992. Predicting the timing of budburst in temperate trees. Journal of Applied Ecology, 29: 597-604.

Kaduk, J. and Helmann, M., 1996. A prognostic phenology scheme for global terrestrial carbon cycle models. Climate Research, 6(1): 1-19.

Keenan, T. et al., 2012. Terrestrial biosphere model performance for inter‐annual variability of land‐atmosphere CO2 exchange. Global Change Biology, 18(6): 1971-1987.

Keenan, T.F. et al., 2014. Net carbon uptake has increased through warming-induced changes in temperate forest phenology. Nature Climate Change.

Kim, Y. and Wang, G., 2005. Modeling seasonal vegetation variation and its validation against Moderate Resolution Imaging Spectroradiometer (MODIS) observations over North America. Journal of Geophysical Research: Atmospheres (1984–2012), 110(D4).

Körner, C. and Basler, D., 2010. Phenology under global warming. Science, 327: 1461-1462.

Kramer, K., 1994. Selecting a model to predict the onset of growth of Fagus sylvatica. Journal of Applied Ecology, 31: 172-181.

Kucharik, C.J., 2003. Evaluation of a Process-Based Agro-Ecosystem Model (Agro-IBIS) across the U.S. Corn Belt: Simulations of the Interannual Variability in Maize Yield. Earth Interactions, 7: 1-33.

Kucharik, C.J. et al., 2006. A multiyear evaluation of a Dynamic Global Vegetation Model at three AmeriFlux forest sites: Vegetation structure, phenology, soil temperature, and CO2 and H2O vapor exchange. Ecological Modelling, 196: 1-31.

Kucharik, C.J. et al., 2000. Testing the performance of a Dynamic Global Ecosystem Model: Water balance, carbon balance, and vegetation structure. Global Biogeochemical Cycles, 14: 795-825.

Landsberg, J., 1974. Apple fruit bud development and growth; analysis and an empirical model. Annals of Botany, 38(5): 1013-1023.

Levis, S. and Bonan, G.B., 2004. Simulating Springtime Temperature Patterns in the Community Atmosphere Model Coupled to the Community Land Model Using Prognostic Leaf Area. Journal of Climate, 17: 4531-4540.

Liang, L. and Schwartz, M.D., 2009. Landscape phenology: an integrative approach to seasonal vegetation dynamics. Landscape Ecology, 24: 465-472.

Lieth, H., 1974. Purposes of a phenology book. In: H. Lieth (Editor), Phenology and Seasonality Modeling. Springer-Verlag, New York, pp. 3-19.

Linderholm, H.W., 2006. Growing season changes in the last century. Agricultural and Forest Meteorology, 137: 1-14.

Linkosalo, T., Häkkinen, R. and Hänninen, H., 2006. Models of the spring phenology of boreal and temperate trees: is there something missing? Tree Physiology, 26(9): 1165-1172.

Lüdeke, M., Janecek, A. and Kohlmaier, G.H., 1991. Modelling the seasonal CO2 uptake by land vegetation using the global vegetation index. Tellus, 43B: 188-196.

Maurer, E., Hidalgo, H., Das, T., Dettinger, M. and Cayan, D., 2010. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. Hydrology and Earth System Sciences, 14(6): 1125-1138.

Menzel, A. and Fabian, P., 1999. Growing season extended in Europe. Nature, 397: 659.

Menzel, A. et al., 2006. European phenological response to climate change matches the warming pattern. Global Change Biology, 12: 1969-1976.

Migliavacca, M. et al., 2012. On the uncertainty of phenological responses to climate change, and implications for a terrestrial biosphere model. Biogeosciences, 9: 2063-2083.

Moffat, A.M. et al., 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agricultural and Forest Meteorology, 147: 209-232.

Morin, X. et al., 2009. Leaf phenology in 22 North American tree species during the 21st century. Global Change Biology, 15(4): 961-975.

Moss, R.H. et al., 2010. The next generation of scenarios for climate change research and assessment. Nature, 463(7282): 747-756.

Motew, M.M. and Kucharik, C.J., 2013. Climate-induced changes in biome distribution, NPP, and hydrology in the Upper Midwest U.S.: A case study for potential vegetation. Journal of Geophysical Research: Biogeosciences, 118: 248-264.

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

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

O'Keefe, J., Harvard Forest Data Archive: HF003.

Parmesan, C. and Yohe, G., 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature, 421: 37-42.

Partanen, J., Koski, V. and Hänninen, H., 1998. Effects of photoperiod and temperature on the timing of bud burst in Norway spruce (Picea abies). Tree Physiology, 18(12): 811-816.

Peñuelas, J. and Filella, I., 2001. Phenology - Responses to a warming world. Science, 294: 793-795.

Peñuelas, J., Rutishauser, T. and Filella, I., 2009. Phenology Feedbacks on Climate Change. Science, 324: 887-888.

Piao, S., Friedlingstein, P., Ciais, P., Viovy, N. and 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: GB3018.

Polgar, C.A. and Primack, R.B., 2011. Leaf‐out phenology of temperate woody plants: from trees to ecosystems. New Phytologist, 191(4): 926-941.

Reclamation, B.o., 2013. Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with preceding Information, and Summary of User Needs, U.S. Department of the Interior, Bureau of Reclamation, Technical Services Center, Denver, Colorado.

Reed, B.C. et al., 1994. Measuring Phenological Variability from Satellite Imagery. Journal of Vegetation Science, 15: 703-714.

Riahi, K. et al., 2011. RCP 8.5—A scenario of comparatively high greenhouse gas emissions. Climatic Change, 109(1-2): 33-57.

Richardson, A.D. et al., 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.

Richardson, A.D., Bailey, A.S., Denny, E.G., Martin, C.W. and O'Keefe, J., 2006. Phenology of a northern hardwood forest canopy. Global Change Biology, 12: 1174-1188.

Richardson, A.D. et al., 2010. Influence of spring and autumn phenological transitions on forest ecosystem productivity. Philosophical Transcations of the Royal Society B, 365: 3227-3246.

Richardson, A.D., Braswell, B.H., Hollinger, D.Y., Jenkins, J.P. and Ollinger, S.V., 2009. Near-surface remote sensing of spatial and temporal variation in canopy phenology. Ecological Applications, 19: 1417-1428.

Richardson, A.D. et al., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. Agricultural and Forest Meteorology, 169: 156-173.

Richardson, A.D. and O'Keefe, J., 2009. Phenological differences between understory and overstory: Acase study using the long-term Harvard Forest Records. In: A. Noormets (Editor), Phenology of Ecosystem Processes: Applications in Global Change Research. Springer, New York, pp. 87-117.

Root, T.L. et al., 2003. Fingerprints of global warming on wild animals and plants. Nature, 421: 57-60.

Rosenzweig, C. et al., 2007. Assessment of observed changes and responses in natural and managed systems. In: M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson (Editors), Climate Change 2007: Impacts, Adaptation and Vulnerability: Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change Cambridge University Press, Cambridge, UK, pp. 79-131.

Sarvas, R., 1974. Investigations on the annual cycle of development of forest trees: II. Autumn dormancy and winter dormancy. Communicationes Instituti Forestalis Fenniae, 84: 1-101.

Schwartz, M.D., Ahas, R. and Aasa, A., 2006. Onset of spring starting earlier across the Northern Hemisphere. Global Change Biology, 12.

Schwartz, M.D. and Chen, X., 2002. Examining the onset of spring in China. Climate Research, 21: 157-164.

Schwartz, M.D., Reed, B.C. and White, M.A., 2002. Assessing satellite-derived start-of-season measures in the conterminous USA. International Journal of Climatology, 22: 1793-1805.

Schwartz, M.D. and Reiter, B.E., 2000. Changes in North American spring. International Journal of Climatology, 20: 929-932.

Sitch, S. et al., 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. Global Change Biology, 9(2): 161-185.

Sonnentag, O. et al., 2012. Digital repeat photography for phenological research in forest ecosystems. Agricultural and Forest Meteorology, 152: 159-177.

Stockli, R. and Vidale, P.L., 2004. European plant phenology and climate as seen in a 20-year AVHRR land-surface parameter dataset. International Journal of Remote Sensing, 25: 3303-3330.

Tateishi, R. and Ebata, M., 2004. Analysis of phenological change patterns using 1982–2000 Advanced Very High Resolution Radiometer (AVHRR) data. International Journal of Remote Sensing, 25: 2287-2300.

Twine, T.E. and Kucharik, C.J., 2008. Evaluating a terrestrial ecosystem model with satellite information of greenness. Journal of Geophysical Research - Biogeosciences, 113: G03027.

Twine, T.E. and Kucharik, C.J., 2009. Climate impacts on net primary productivity trends in natural and managed ecosystems of the central and eastern United States. Agricultural and Forest Meteorology, 149: 2143-2161.

Urbanski, S. et al., 2007. Factors controlling CO 2 exchange on timescales from hourly to decadal at Harvard Forest. Journal of Geophysical Research, 112: G02020.

Verseghy, D., McFarlane, N. and Lazare, M., 1993. CLASS—A Canadian land surface scheme for GCMs, II. Vegetation model and coupled runs. International Journal of Climatology, 13(4): 347-370.

Vitasse, Y. et al., 2011. Assessing the effects of climate change on the phenology of European temperate trees. Agricultural and Forest Meteorology, 151(7): 969-980.

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

White, M.A. and Nemani, R., 2003. Canopy duration has little influence on annual carbon storage in the deciduous broad leaf forest. Global Change Biology, 9: 967-972.

White, M.A., Running, S.W. and Thornton, P.E., 1999. The impact of growing-season length variability on carbon assimilation and evapotranspiration over 88 years in

the eastern US deciduous forest. International Journal of Biometeorology, 42: 139-145.

White, M.A., Thornton, P.E. and Running, S.W., 1997. A continental phenology model for monitoring vegetation responses to interannual climatic variability. Global Biogeochemical Cycles, 11: 217-234.

White, M.A., Thornton, P.E., Running, S.W. and Nemani, R.R., 2000. Parameterization and sensitivity analysis of the BIOME-BGC terrestrial ecosystem model: net primary production controls. Earth interactions, 4(3): 1-85.

Xu, H., Twine, T.E. and Yang, X., 2014. Evaluating Remotely Sensed Phenological Metrics in a Dynamic Ecosystem Model. Remote Sensing, 6(6): 4660-4686.

Yang, X., Mustard, J.F., Tang, J. and Xu, H., 2012. Regional-scale phenology modeling based on meteorological records and remote sensing observations. Journal of Geophysical Research, 117: G03029.

Zhang, X. et al., 2003. Monitoring vegetation phenology using MODIS. Remote Sensing of Environment, 84: 471-475.

Zhang, X., Tan, B. and Yu, Y., 2014. Interannual variations and trends in global land surface phenology derived from enhanced vegetation index during 1982-2010. International journal of biometeorology: 1-18.

Zhang, X., Tarpley, D. and Sullivan, J.T., 2007. Diverse responses of vegetation phenology to a warming climate. Geophysical Research Letters, 34: L19405.

Zhou, L. et al., 2001. Variation in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. Journal of Geophysical Research, 106: 20069-20083.

Zhu, W. et al., 2012. Extension of the growing season due to delayed autumn over mid and high latitudes in North America during 1982-2006. Global Ecology and Biogeography, 21: 260-271.

Appendix

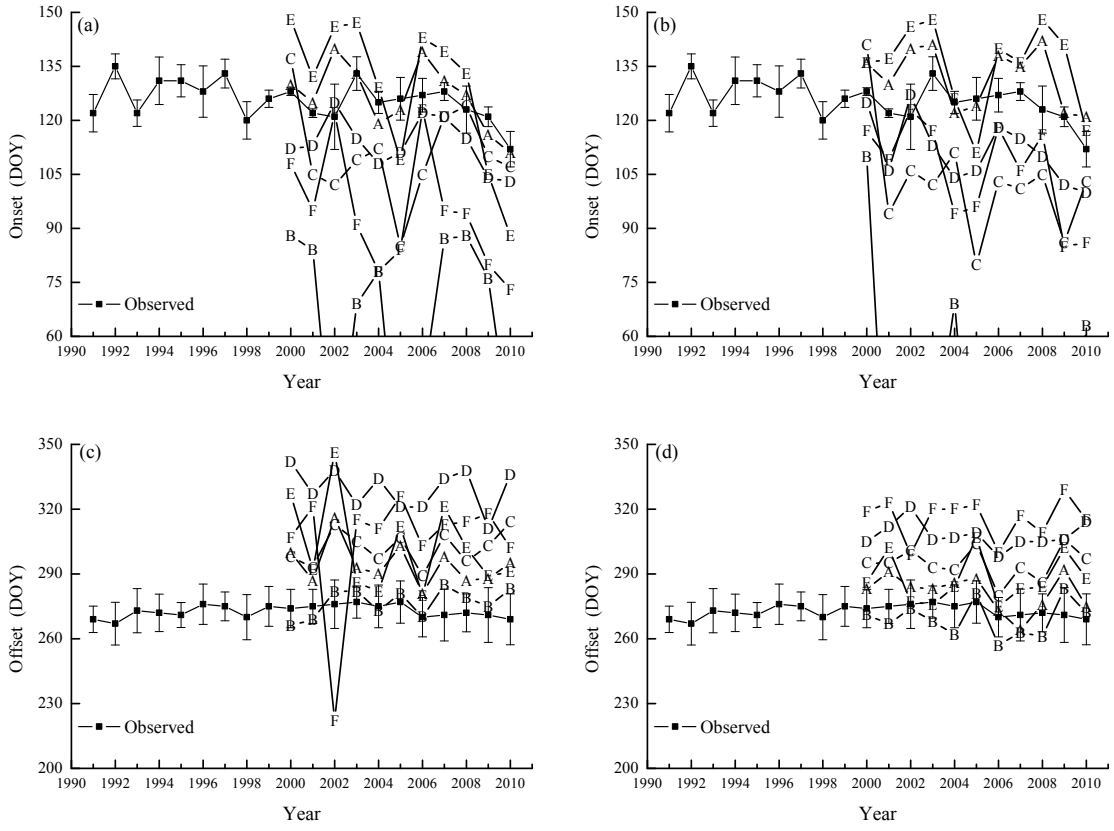


Figure A2.1 Ground observed phenology with remotely sensed onset from NDVI (a), onset from EVI (b), offset from NDVI (c), and offset from EVI (d). Satellite data used to retrieve phenology is MODIS MOD13A1 product. Error bars indicate the standard deviation of observation. Methods used to retrieve phenology are as follows: (A) MIDPOINT, (B) LOGISTIC1, (C) LOGISTIC2, (D) MOVING, (E) DERIVATIVE, and (F) CAMELBACK.

Table A2.1 Performance of remotely sensed onset using MOD13A1

Leaf	NDVI		EVI	
Onset	RMSD	Correlation	RMSD	Correlation
LOGISTIC1	67.3	0.30	102.7	0.27
LOGISTIC2	18.8	0.32	25.9	0.06
MIDPOINT	7.5	0.50	10.4	0.45
MOVING	12.4	0.37	15.0	0.50
DERIVATIVE	16.3	0.66	15.5	0.06
CAMELBACK	33.0	0.34	21.3	0.39

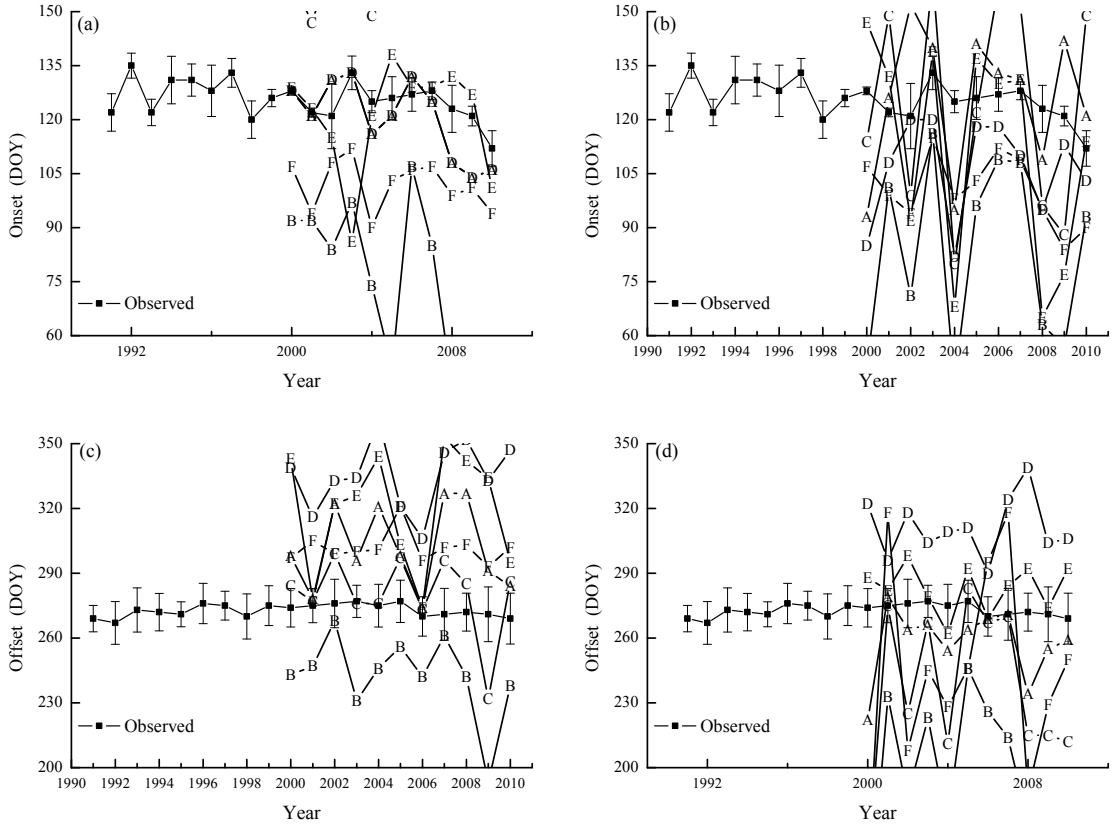


Figure A2.2 Ground observed phenology with remotely sensed onset from NDVI (a), onset from EVI (b), offset from NDVI (c), and offset from EVI (d). Satellite data used to retrieve phenology is MODIS MCD43A4 product. Error bars indicate the standard deviation of observation. Methods used to retrieve phenology are as follows: (A) MIDPOINT, (B) LOGISTIC1, (C) LOGISTIC2, (D) MOVING, (E) DERIVATIVE, and (F) CAMELBACK.

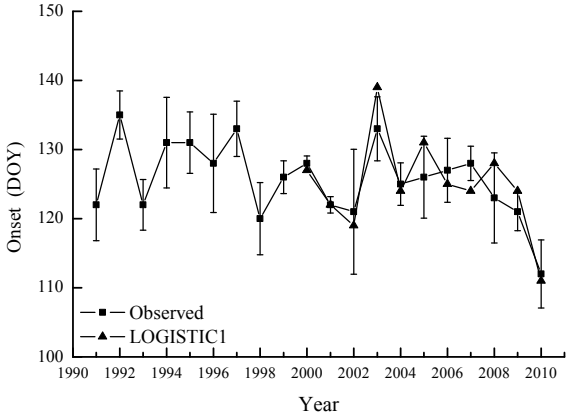


Figure A2.3 Ground observed phenology with remotely sensed onset using LOGISTIC1 with EVI (MOD09A1). The amplitude used to determine the onset is 25%. RMSD=3.3 days; ρ=0.74.

Table A2.2 Performance of remotely sensed offset using MOD13A1

Leaf	NDVI		EVI	
Onset	RMSD	Correlation	RMSD	Correlation
LOGISTIC1	7.6	-0.03	8.8	0.25
LOGISTIC2	29.8	0.06	22.4	0.20
MIDPOINT	22.6	0.41	10.6	0.52
MOVING	56.9	-0.03	35.1	0.37
DERIVATIVE	35.2	0.24	17.8	0.06
CAMELBACK	41.8	0.41	43.3	0.30

Table A2.3 Performance of remotely sensed onset using MCD43A4

Leaf	NDVI		EVI	
Onset	RMSD	Correlation	RMSD	Correlation
LOGISTIC1	54.6	0.68	47.6	0.39
LOGISTIC2	33.7	0.29	28.4	0.49
MIDPOINT	8.5	0.68	19.4	-0.17
MOVING	14.9	0.59	22.3	0.11
DERIVATIVE	15.5	0.24	30.2	0.62
CAMELBACK	22.9	0.56	24.3	0.93

Table A2.4 Performance of remotely sensed offset using MCD43A4

Leaf	NDVI		EVI	
Onset	RMSD	Correlation	RMSD	Correlation
LOGISTIC1	35.7	0.37	84.4	0.42
LOGISTIC2	18.2	0.24	53.2	0.33
MIDPOINT	33.6	0.26	22.1	0.07
MOVING	63.9	-0.15	40.1	0.09
DERIVATIVE	53.1	0.05	14.3	0.15
CAMELBACK	29.0	0.30	58.9	-0.25

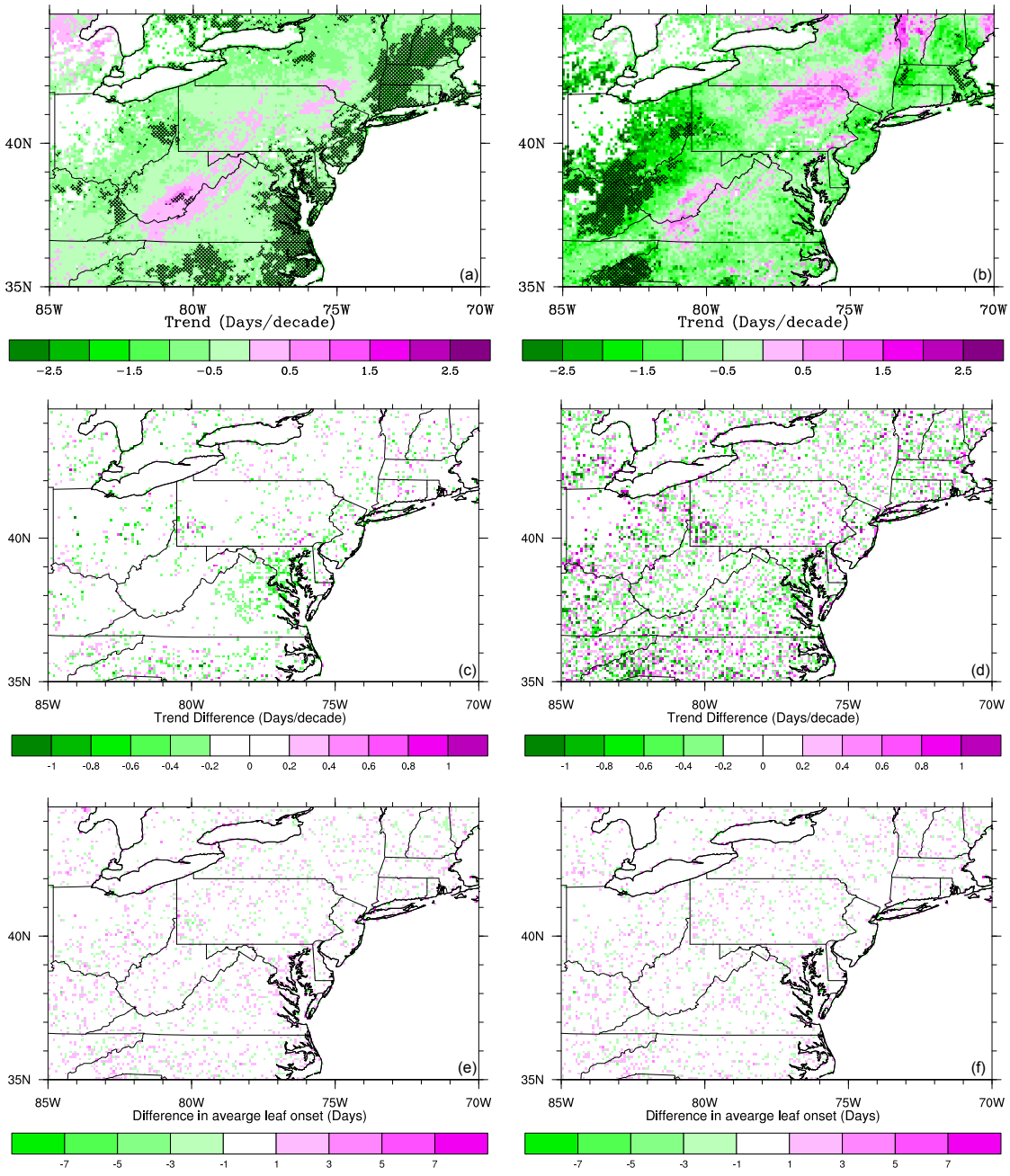


Figure A3.1 Historical trends in phenology simulated using SW and the difference from SW2 (a) Linear trends in leaf onset simulated using SW (1958-2007) (b) Linear trends in leaf onset simulated using SW (1983-2007) (c) Difference (SW-SW2) in trend in leaf onset (1958-2007) (d) Difference (SW-SW2) in trend in leaf onset (1983-2007) (e) Difference (SW-SW2) in long-term average leaf onset (1958-2007) (f) Difference (SW-SW2) in long-term average leaf onset (1983-2007)

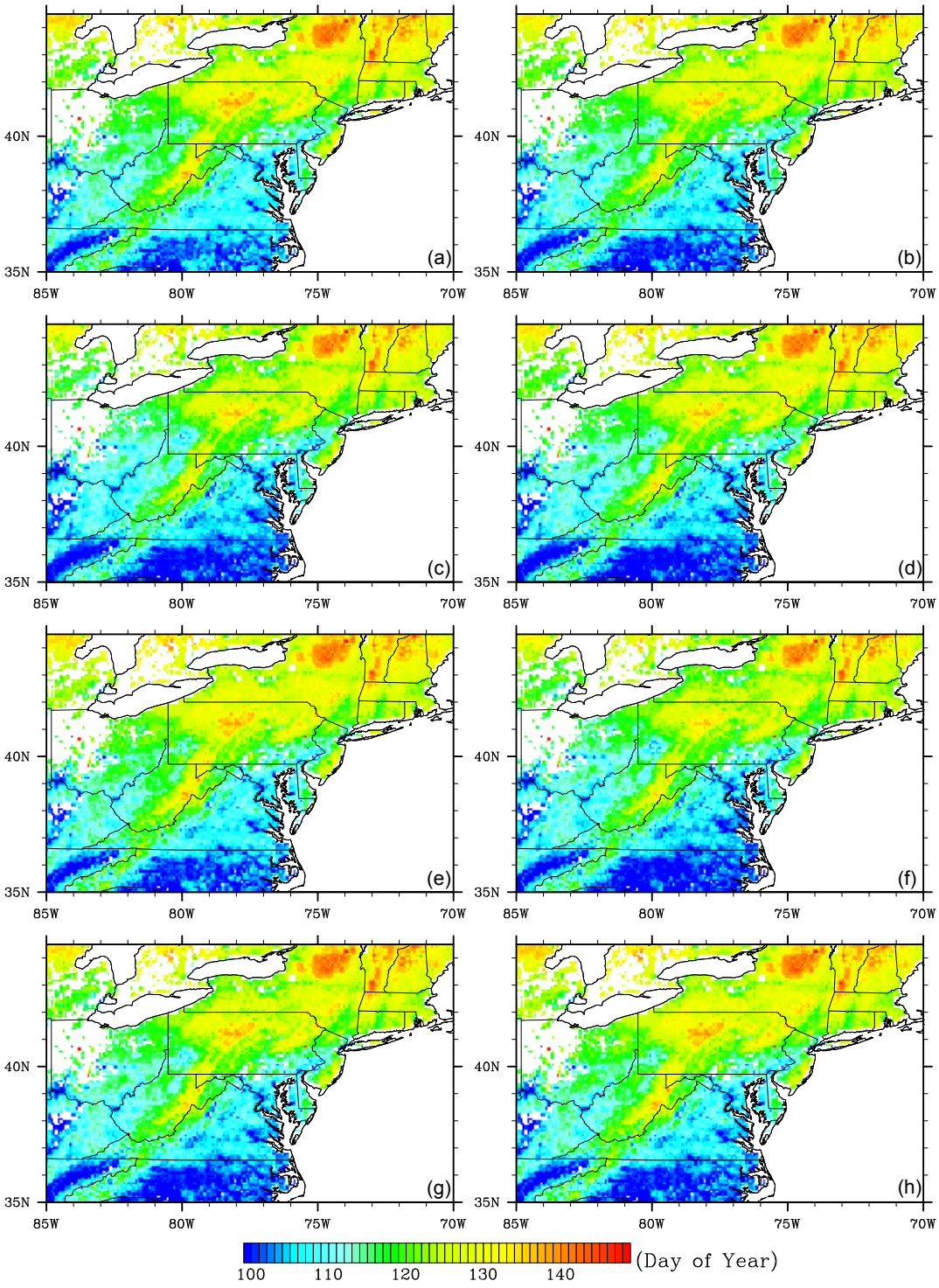


Figure A4.1 Average leaf onset date over 1986-2005 simulated using SW model with output of CMIP5 'historical' experiment from (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA

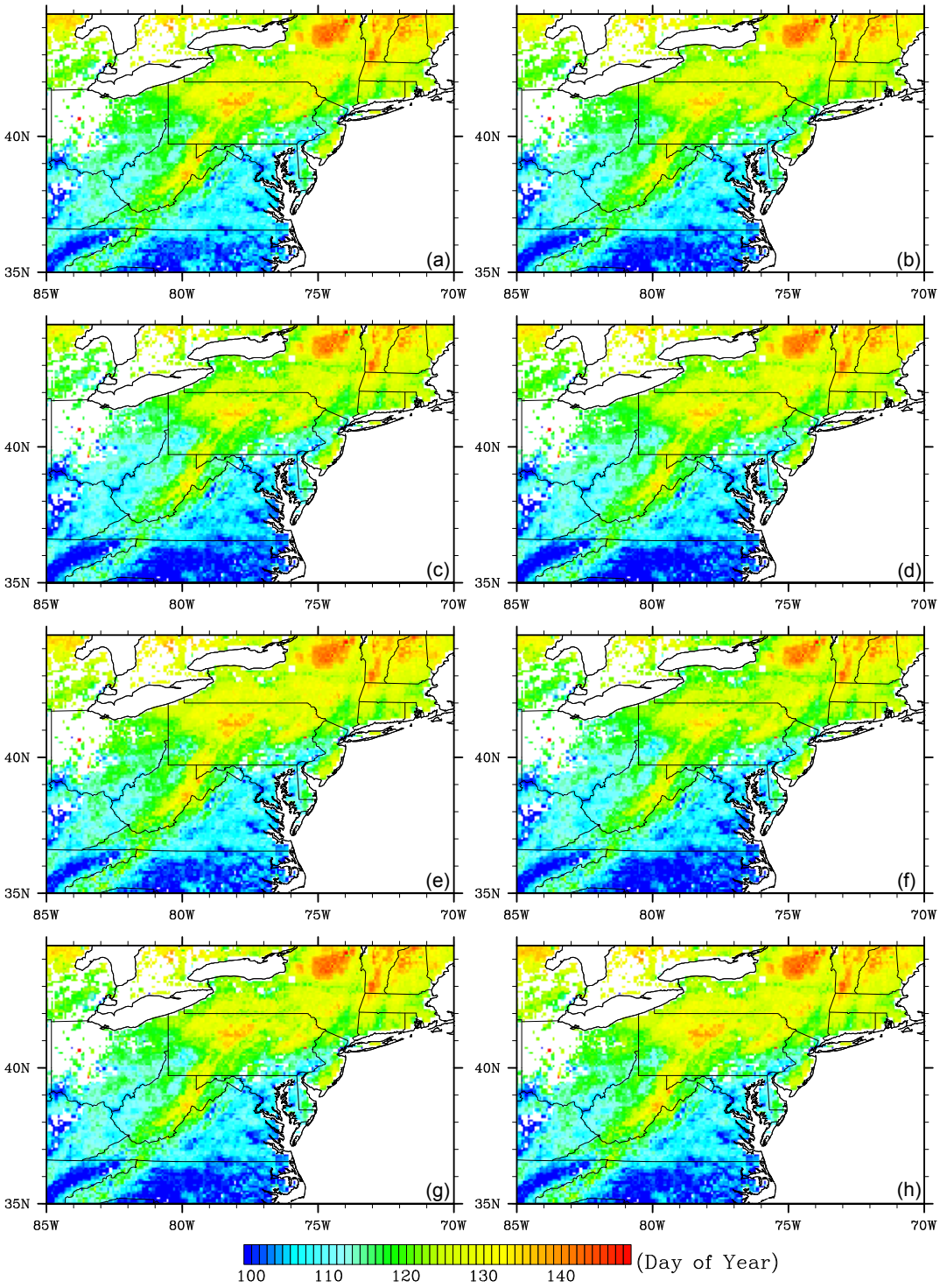


Figure A4.2 Average leaf onset date over 1986-2005 simulated using SW2 model with output of CMIP5 'historical' experiment from (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA

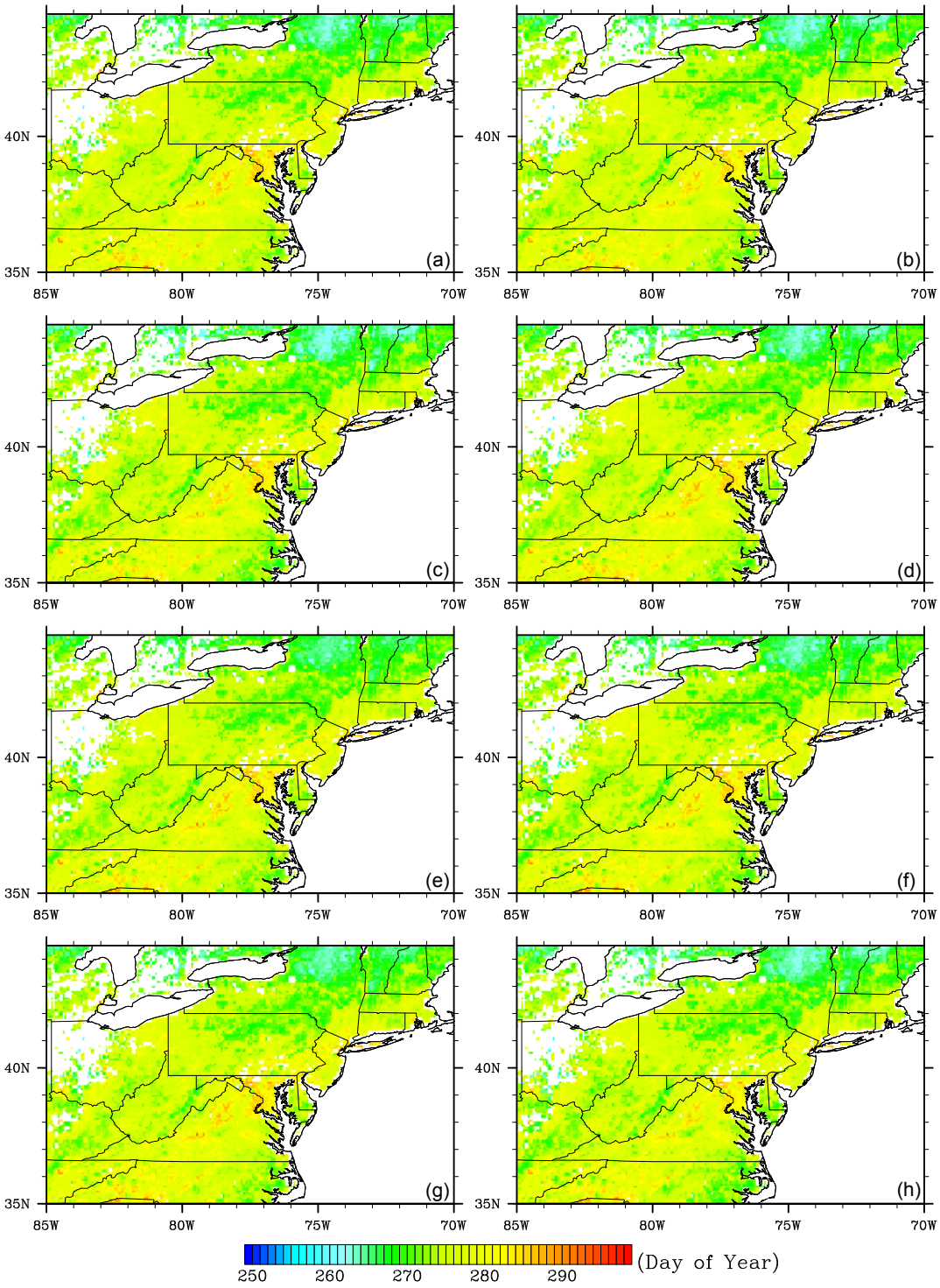


Figure A4.3 Average leaf offset date over 1986-2005 simulated using DLP model with output of CMIP5 'historical' experiment from (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA

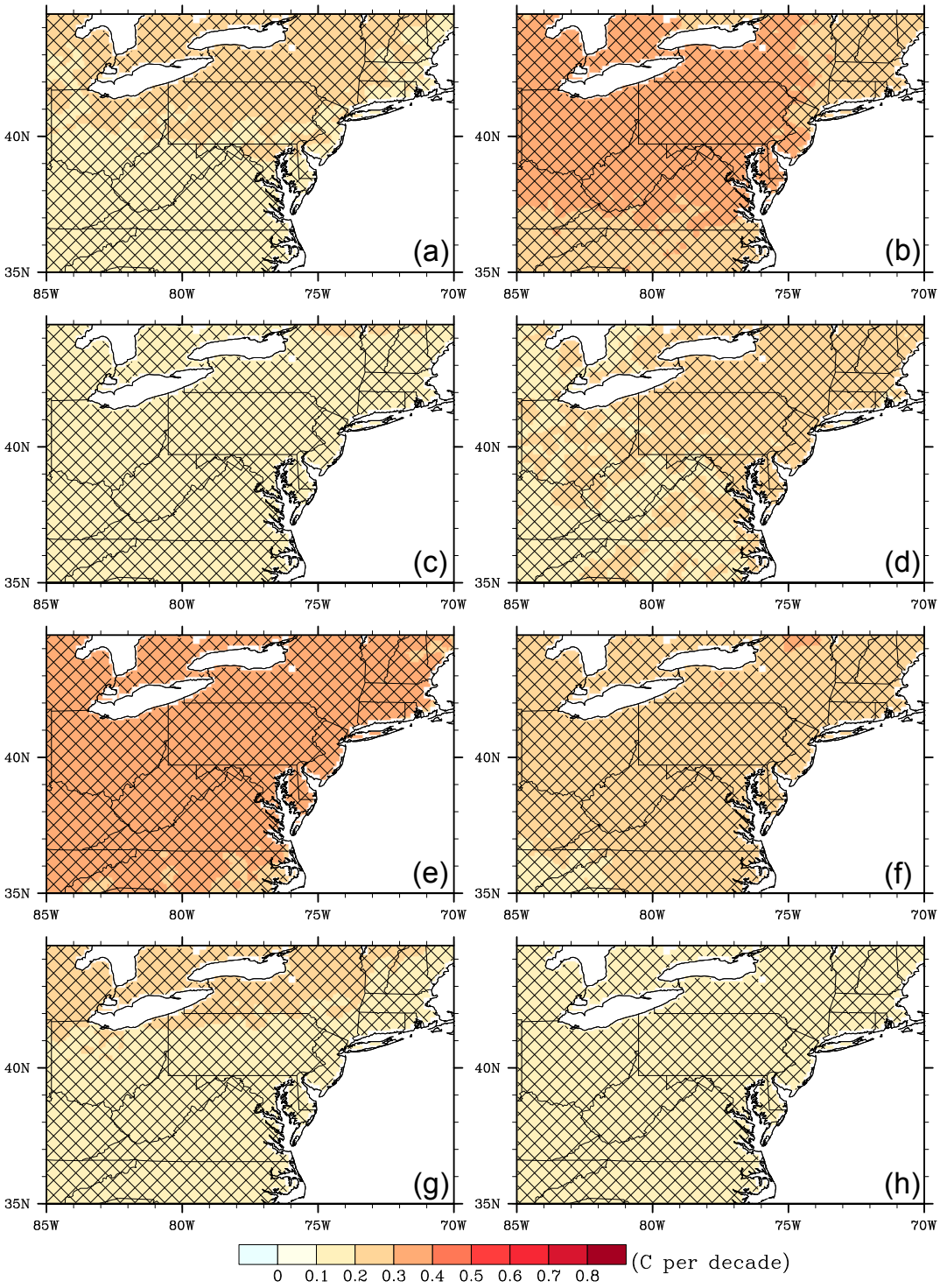


Figure A4.4 Trends in spring (MAM) temperature under RCP4.5 projected by following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA. Hatching indicates trends are statistically significant at P<0.01

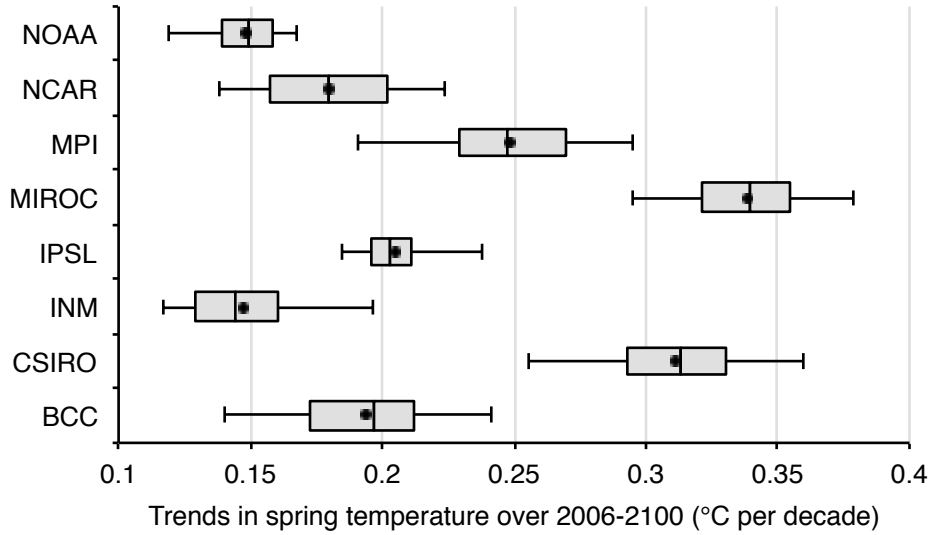


Figure A4.5 Box-and-whisker plot of the trends in spring (MAM) temperature under RCP4.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

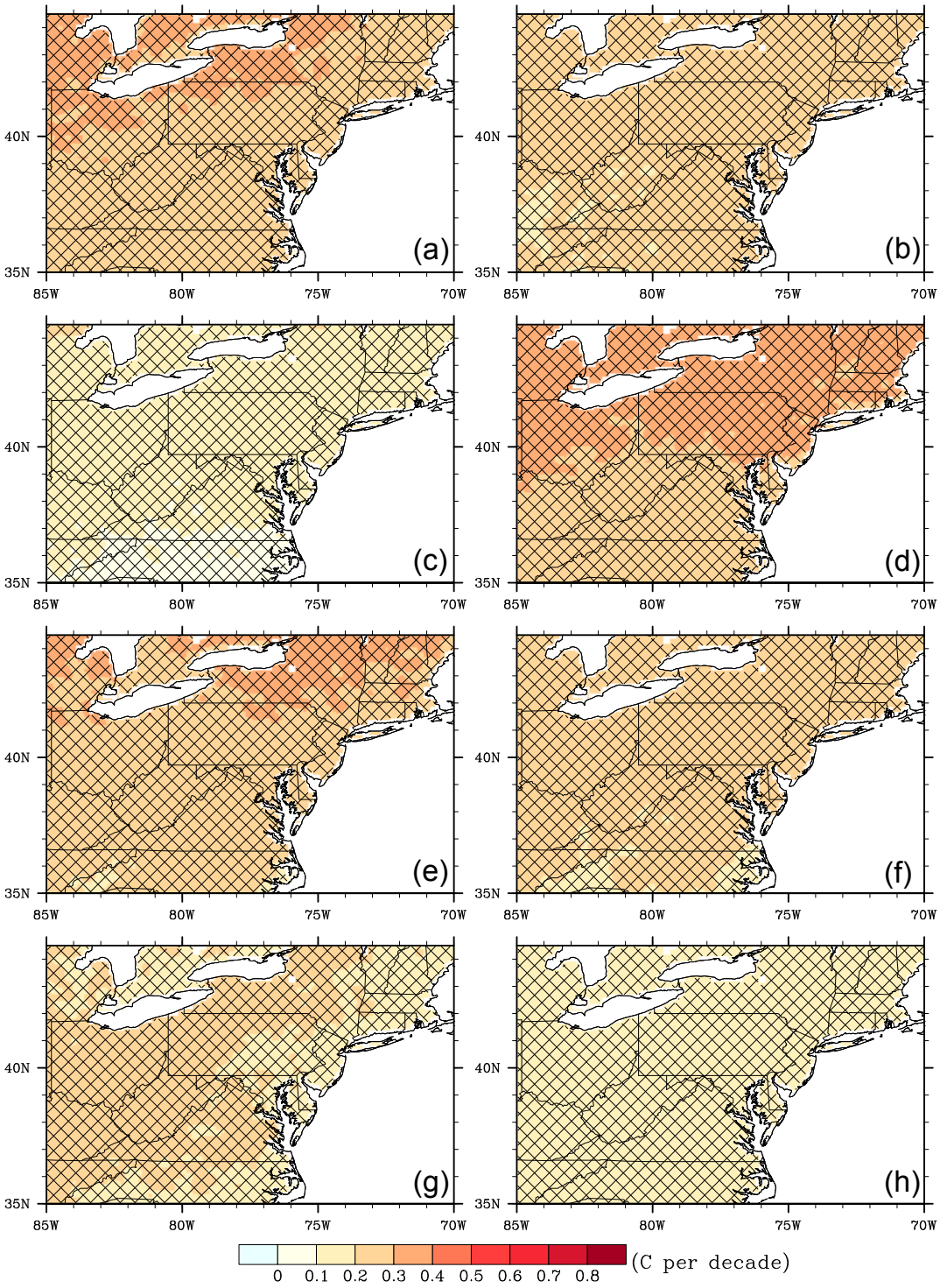


Figure A4.6 Trends in autumn (SON) temperature under RCP4.5 projected by following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA. Hatching indicates trends are statistically significant at P<0.01

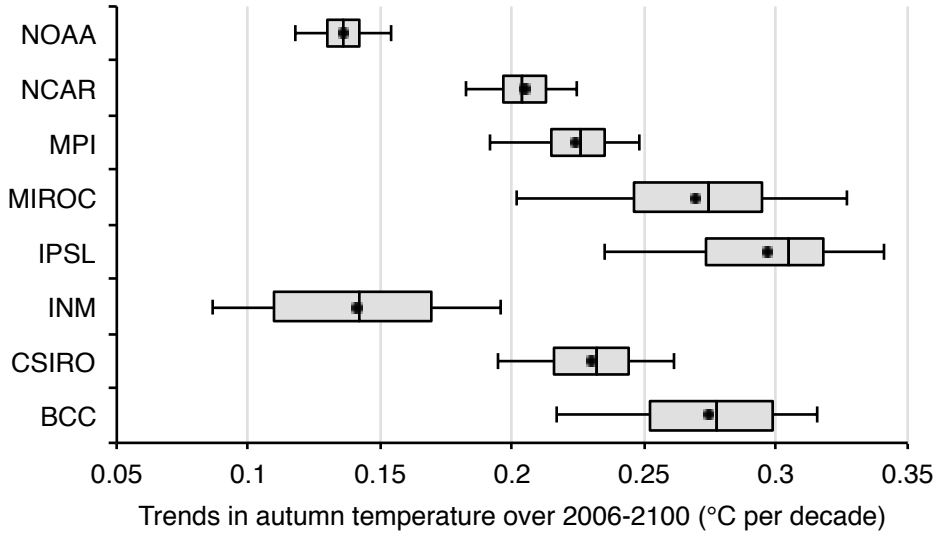


Figure A4.7 Box-and-whisker plot of the trends in autumn (SON) temperature under RCP4.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

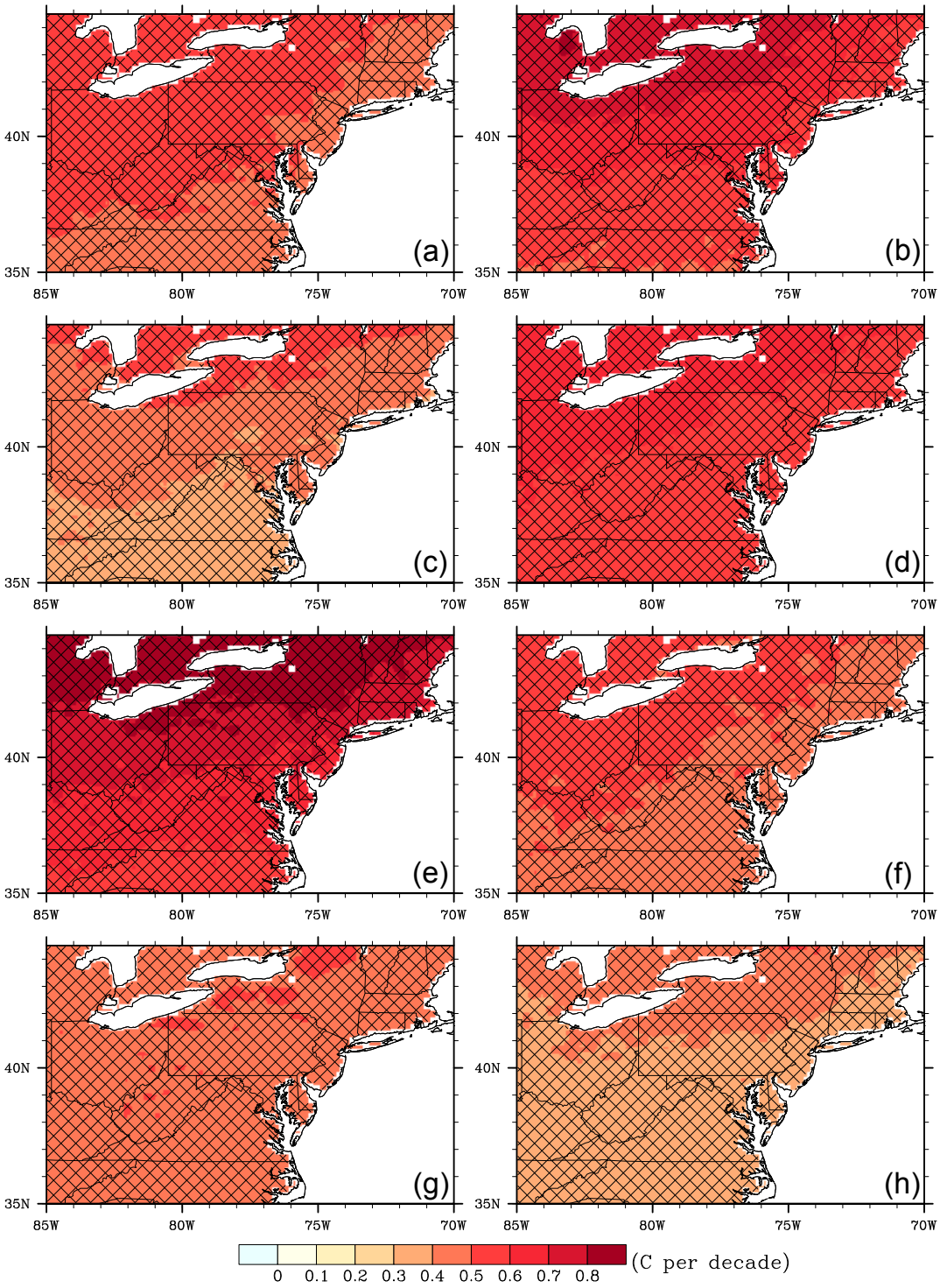


Figure A4.8 Trends in spring (MAM) temperature under RCP8.5 projected by following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA. Hatching indicates trends are statistically significant at P<0.01

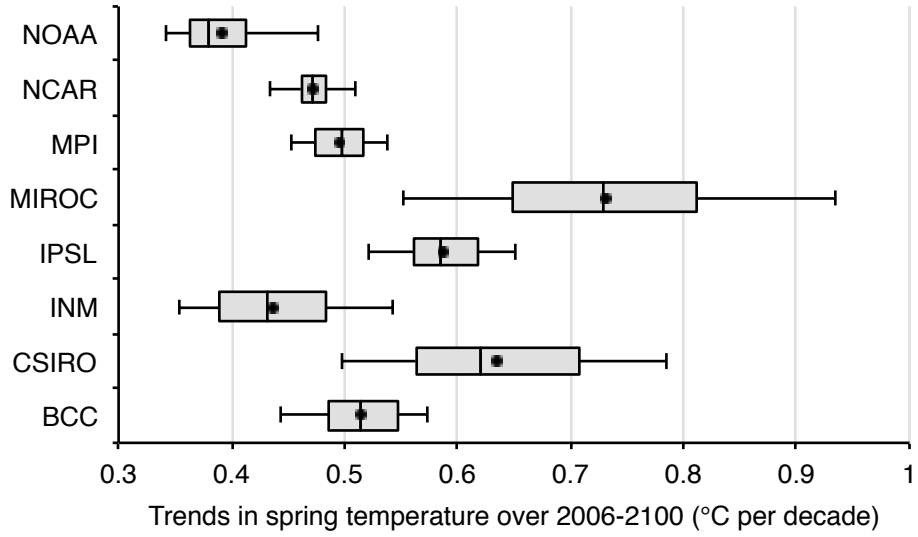


Figure A4.9 Box-and-whisker plot of the trends in spring (MAM) temperature under RCP8.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

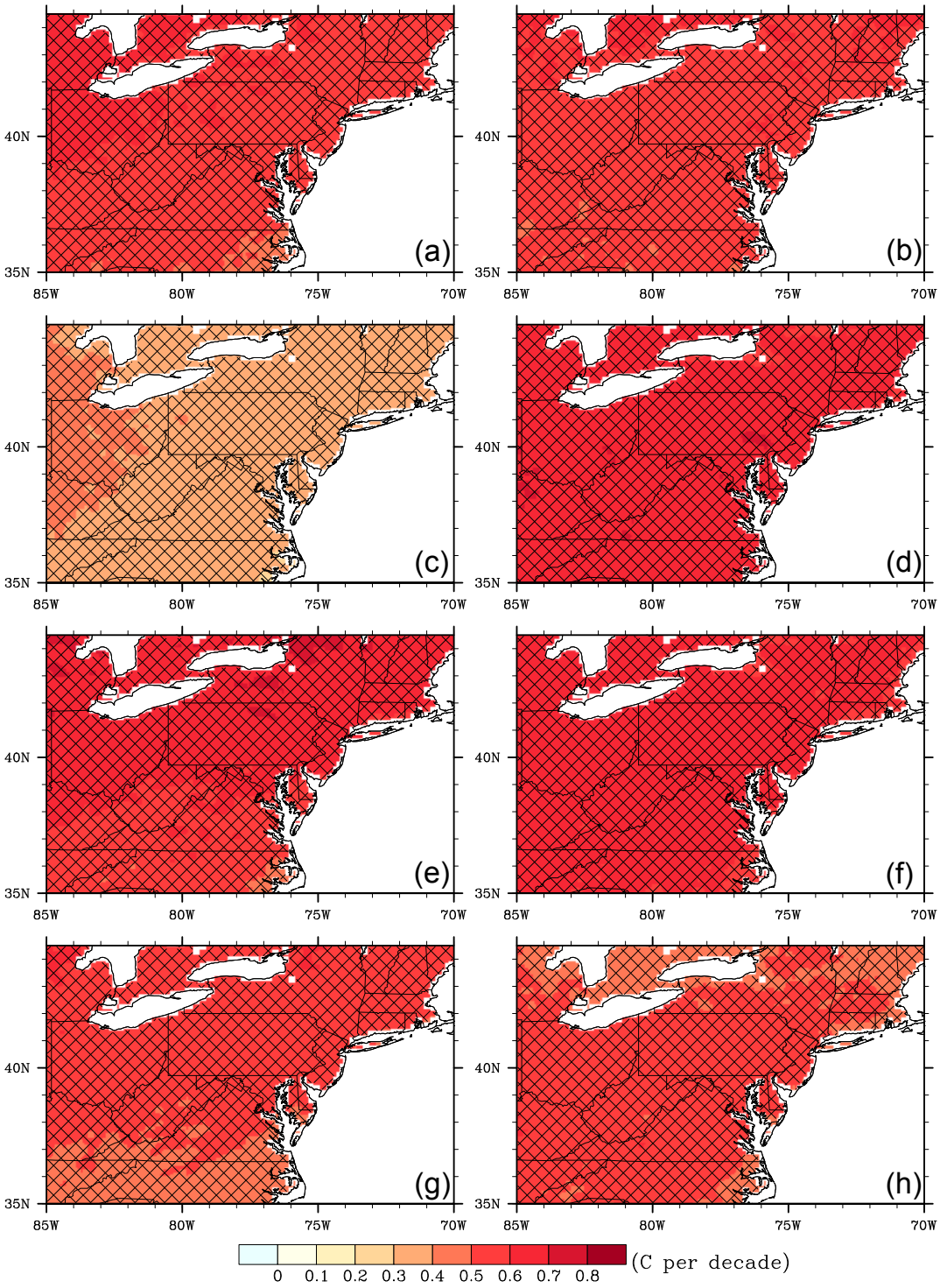


Figure A4.10 Trends in autumn (SON) temperature under RCP8.5 projected by following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA. Hatching indicates trends are statistically significant at P<0.01

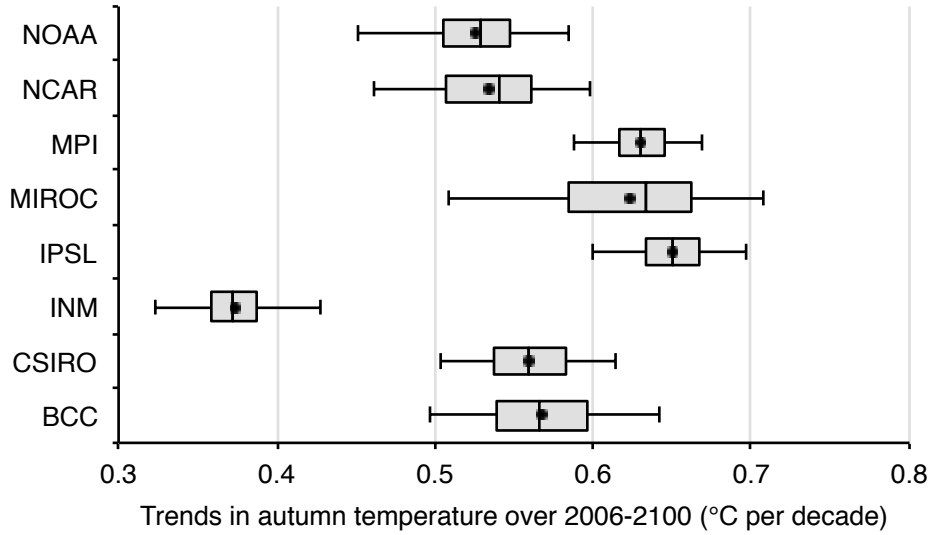


Figure A4.11 Box-and-whisker plot of the trends in autumn (SON) temperature under RCP8.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

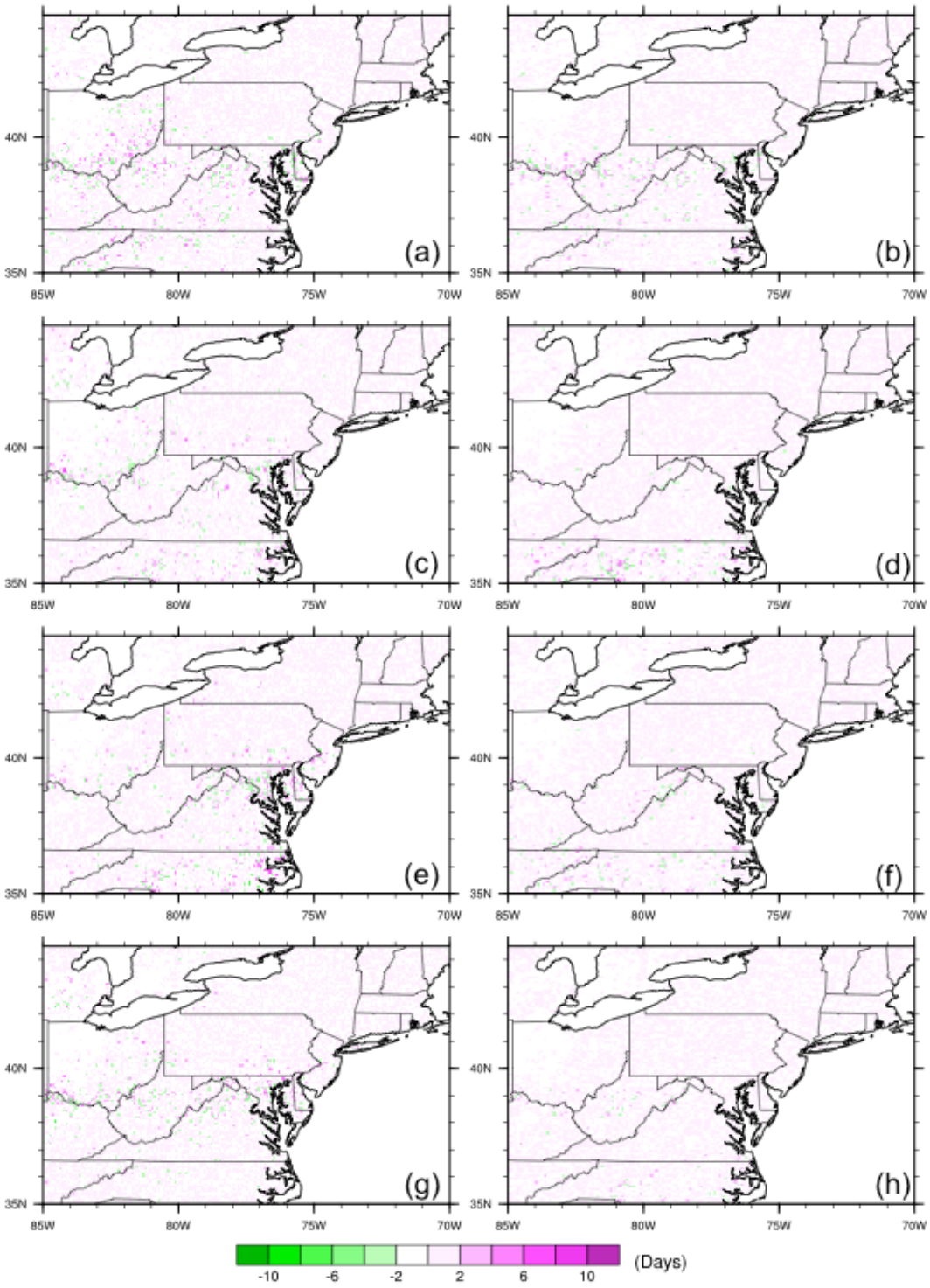


Figure A4.12 Difference (SWb-SW2b) in the change in leaf onset over 2006-2100 under RCP4.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA.

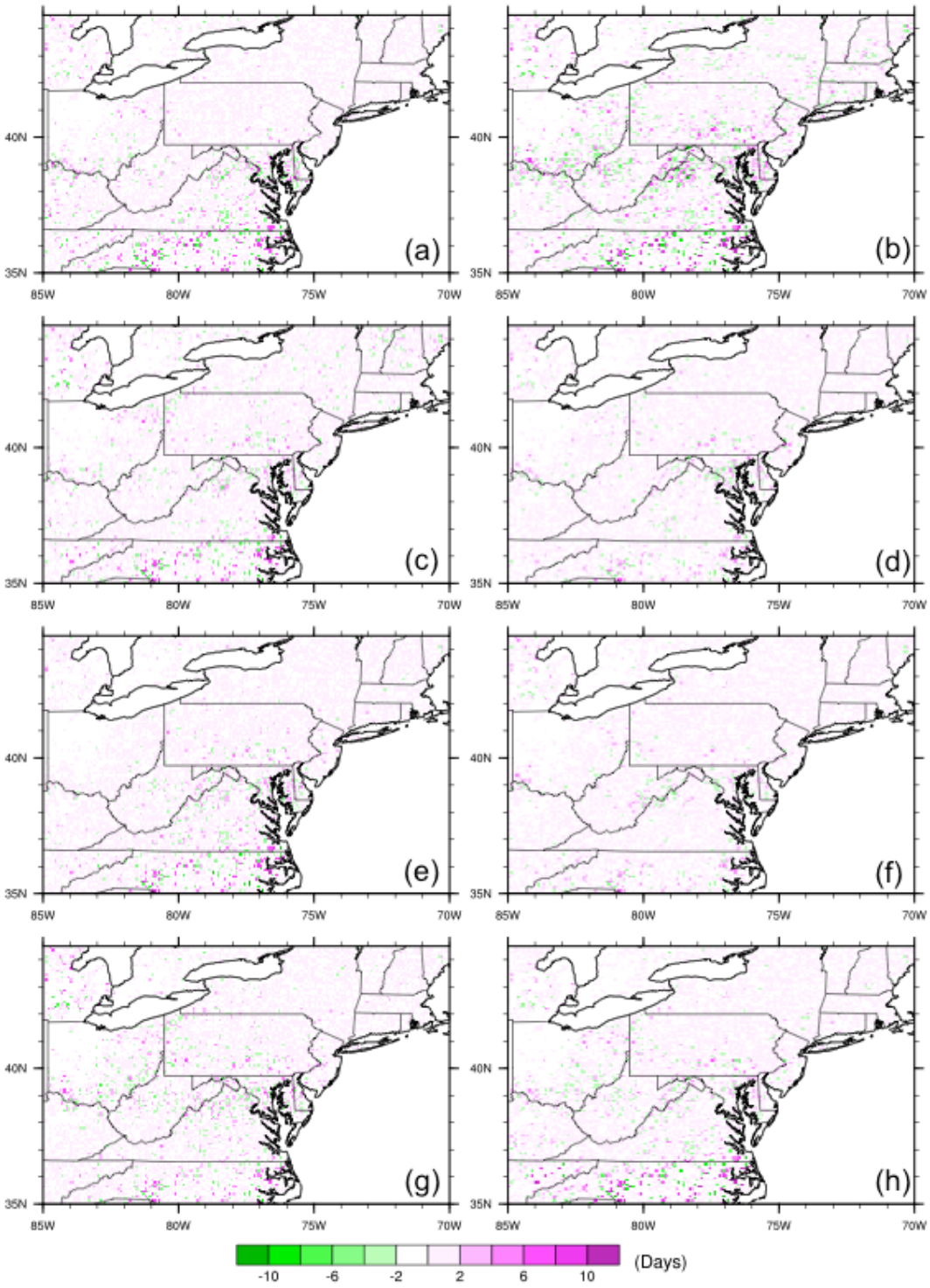


Figure A4.13 Difference (SWb-SW2b) in the change in leaf onset over 2006-2100 under RCP8.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA.

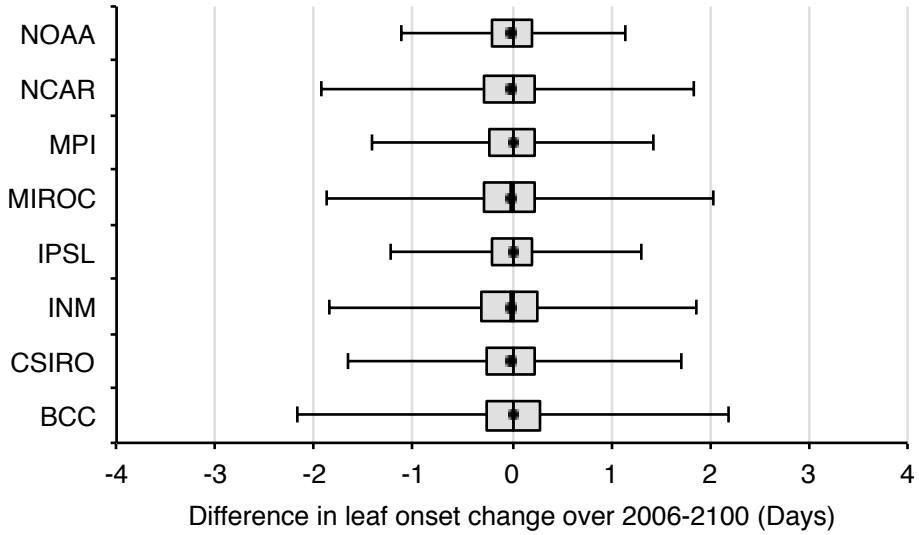


Figure A4.14 Box-and-whisker plot of the difference (SWb-SW2b) in leaf onset change under RCP4.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

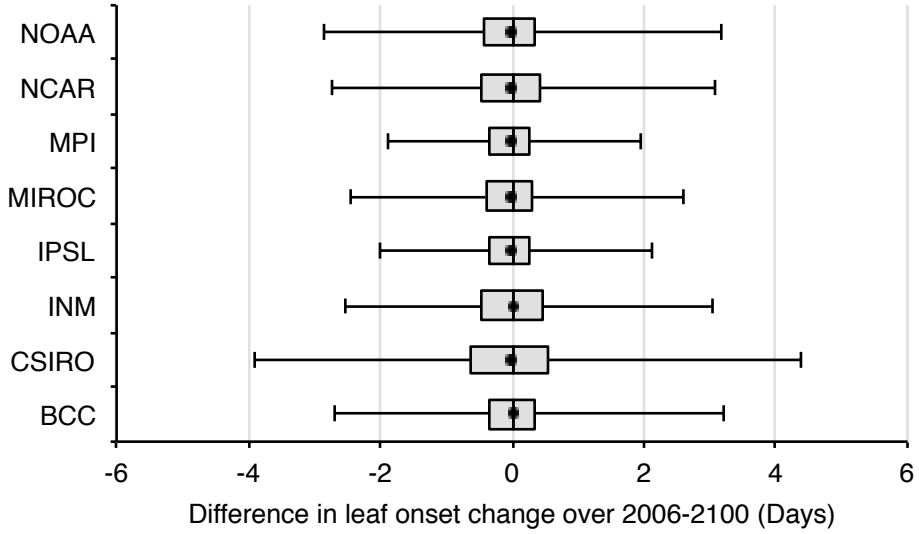


Figure A4.15 Box-and-whisker plot of the difference (SWb-SW2b) in leaf onset change under RCP8.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

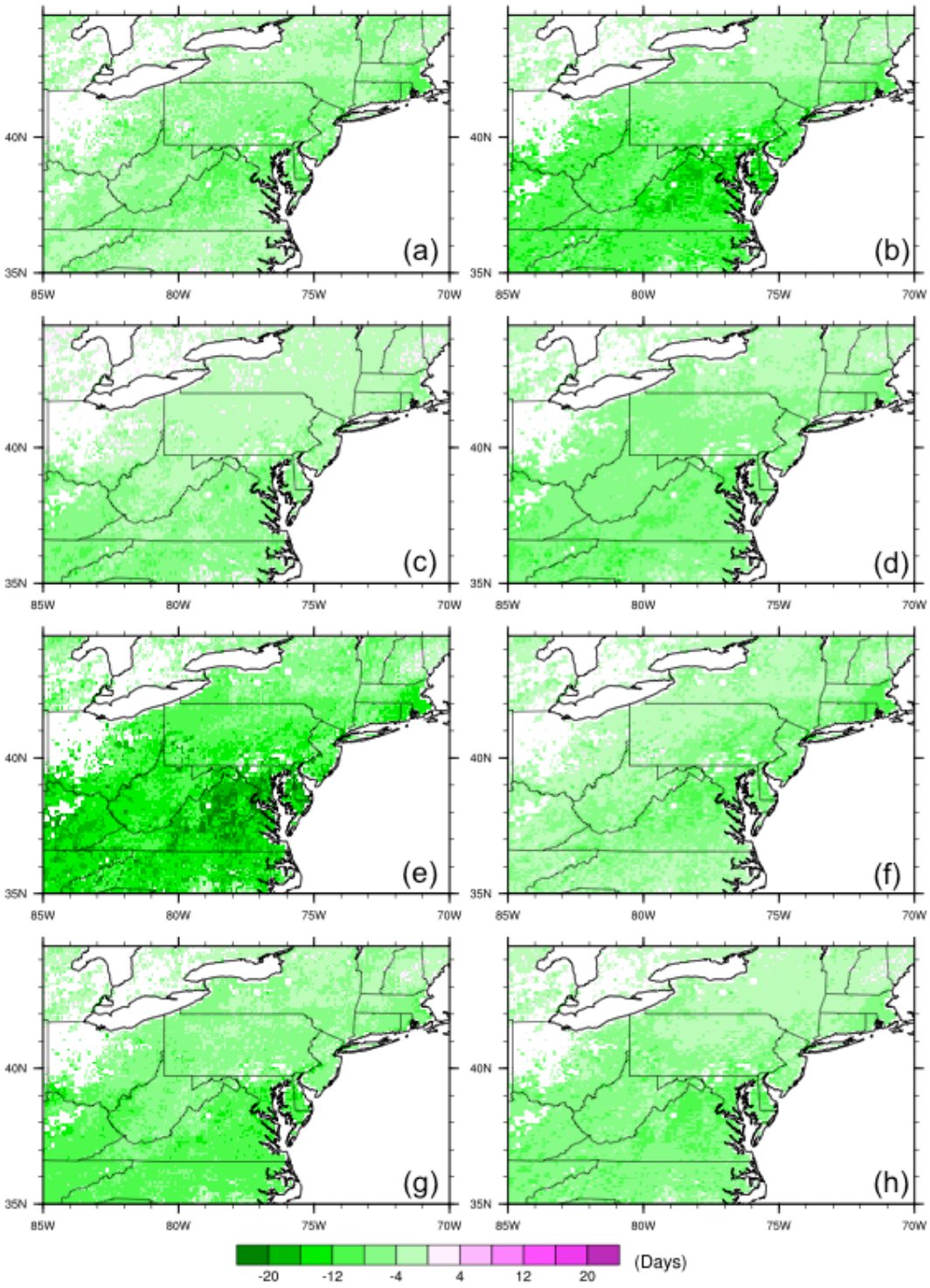


Figure A4.16 Difference (SW2b-SW2) in the change in leaf onset over 2006-2100 under RCP4.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA.

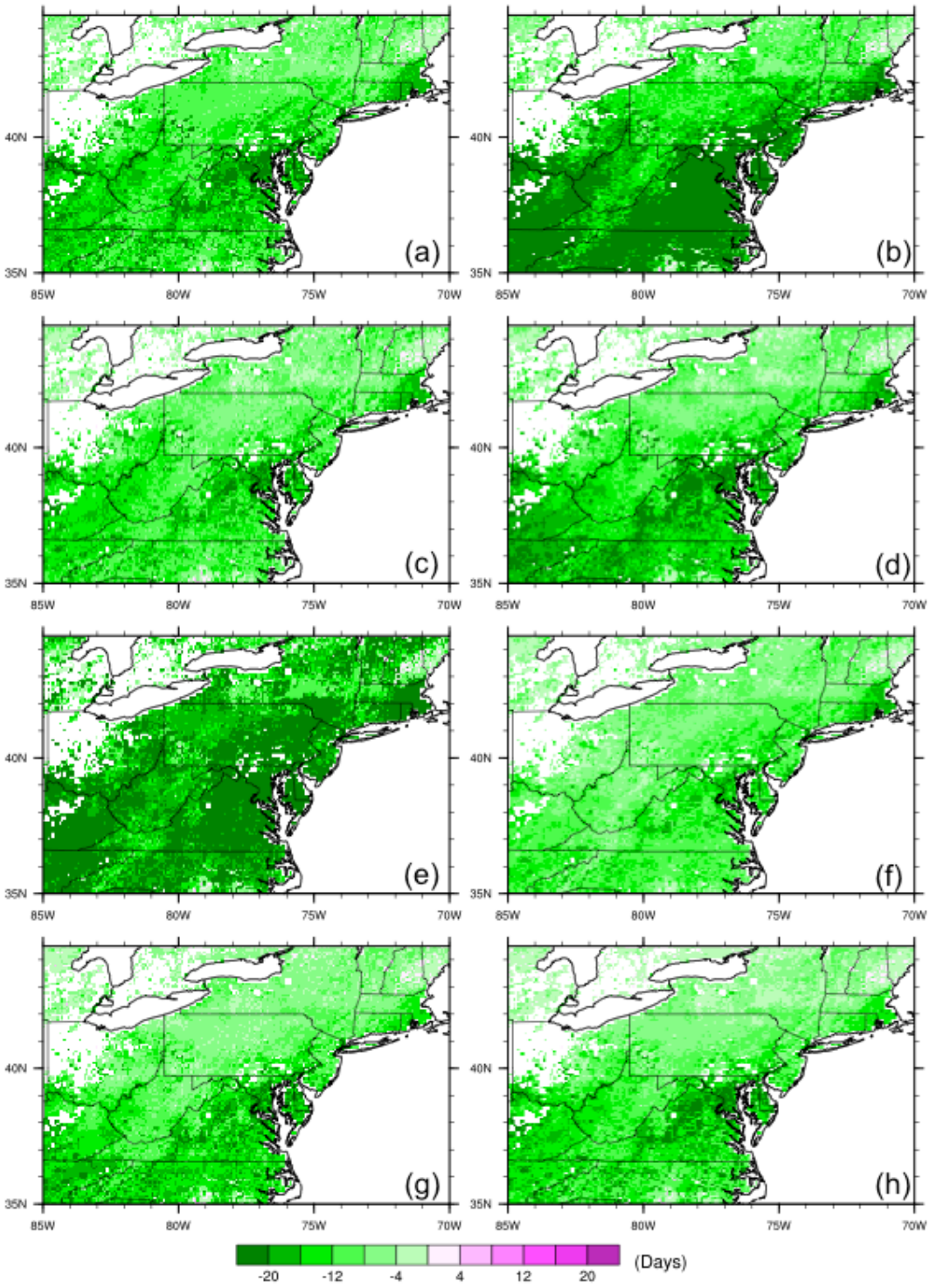


Figure A4.17 Difference (SW2b-SW2) in the change in leaf onset over 2006-2100 under RCP8.5 Driving data were the output from following GCMs: (a) BCC, (b) CSIRO, (c) INM, (d) IPSL, (e) MIROC, (f) MPI, (g) NCAR and (h) NOAA.

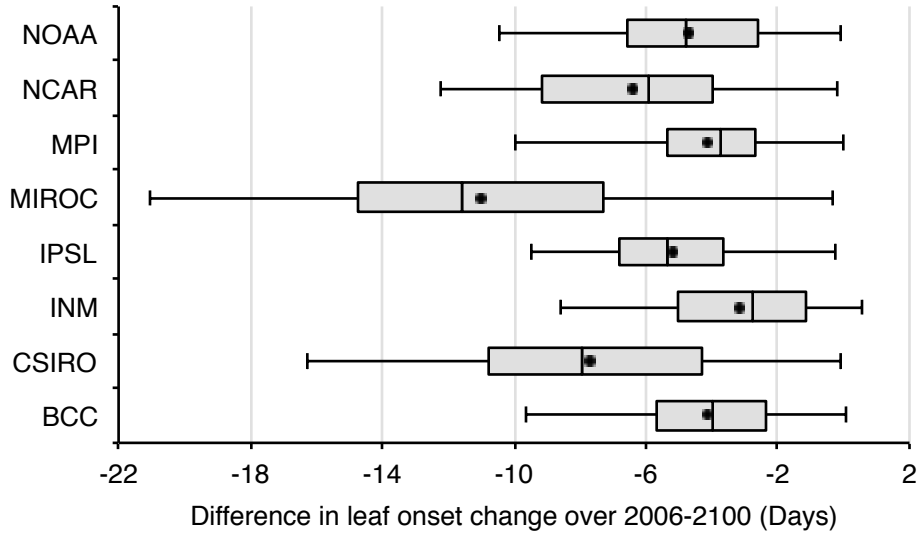


Figure A4.18 Box-and-whisker plot of the difference (SW2b-SW2) in leaf onset change under RCP4.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

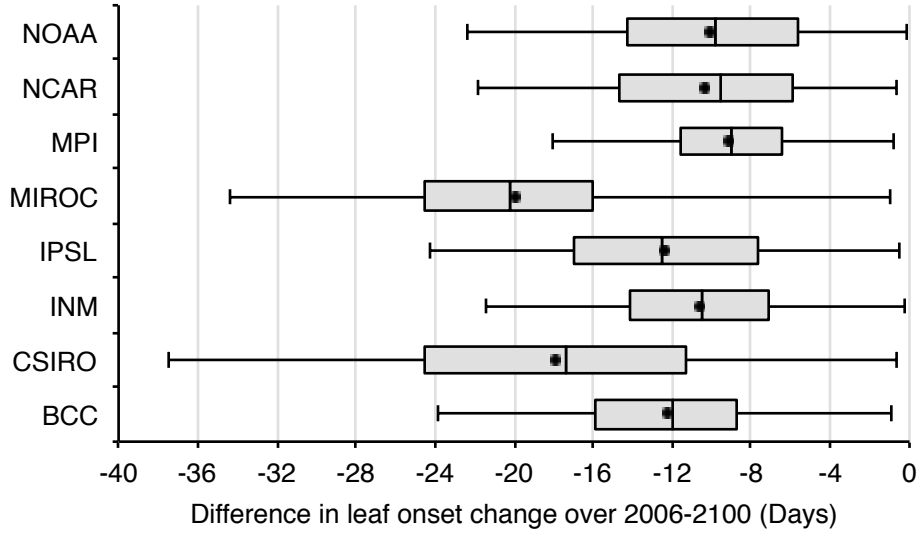


Figure A4.19 Box-and-whisker plot of the difference (SW2b-SW2) in leaf onset change under RCP8.5 The box is delimited by the 25th and 75th percentile, with the median indicated as line in between. The whiskers correspond to the 2nd and 98th percentile. The dots indicate the means.

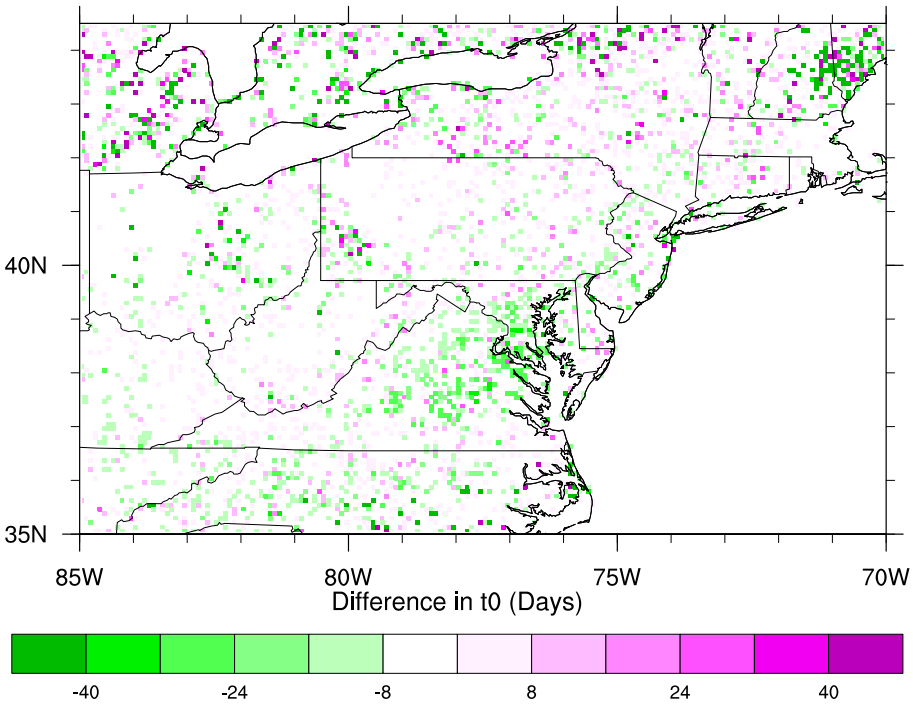


Figure A4.20 Difference in the parameter t_0 between SW and SW2