Abstract  Leaf out time is a widely used indicator of climate change and represents a critical transition point of annual seasonality in most temperate ecosystems. We compared three sources of data to determine the effect of spring temperature on tree leaf out: field observations, remotely sensed satellite data, and experimental warming. All three methods recorded earlier leaf out with warmer spring temperatures. However, leaf out timing was more than twice as sensitive to temperature in the field study (advancing at a rate of 6.1 days/°C), as under experimental warming (2.1 days/°C), with remote sensing intermediate (3.7 days/°C). Researchers need to be aware of the currently unexplained differences among methodologies when using phenological data to parameterize or benchmark models that represent ecosystem processes. The mechanisms behind these discrepancies must be better understood if we are to confidently predict responses of leaf out timing to future climates.

Keywords  Phenology · Remote sensing · Climate change · Experimental warming · Boston-Area Climate Experiment · Budburst

Introduction

Plant phenology has emerged as a key indicator of the biological effects of climate change (Menzel and Fabian 1999; Sparks et al. 2009). Spring plant phenology is considered to be of particular importance because the high sensitivity of plants to winter and spring temperatures allows plant phenology to respond quickly to warming, making it a powerful indicator of the effects of climate change. The importance of springtime phenology is underscored by its influence on ecosystem processes such as carbon sequestration, hydrological cycles, and trophic interactions (Menzel et al. 2006; Parmesan 2006; Piao et al. 2008). There is a long history of phenology monitoring among scientists, horticulturalists, and government agencies (Bradley et al. 1999; Ibanez et al. 2010; Menzel 2000). While traditional observations of tree leaf out dates in the field are still commonly made, additional techniques for monitoring leaf out and estimating the effects of a warming climate, such as experimental warming and remote sensing, are now being widely used (Cleland et al. 2007; Polgar and Primack 2011).

Over the past decade, hundreds of warming experiments have been established to investigate responses of plants and ecosystems, including phenological responses, to future climate change scenarios. In addition to temperature, researchers sometimes manipulate precipitation, CO₂ concentration, species composition, and other variables (Andresen et al. 2010; Beier et al. 2004; Morin et al. 2010). While there are many advantages to using warming experiments to study phenology, such as the ability to subject study species to predicted future conditions or other specific temperature regimes, there are
disadvantages as well. Most prominently, the expense of setting up and maintaining such a facility is high, and limits the area and number of species that can be treated. There are also questions of whether experimental warming affects plants in a manner comparable to plants in the wild. A recently published meta-analysis reports that experimental warming studies significantly underestimate the effects of warming on timing of spring plant phenology (Wolkovich et al. 2012). Nonetheless, results from these experimental studies are informing the next generation of climate change models and are being used to predict future responses to climate change of temperate forests and other ecosystems (Hanninen et al. 2007); therefore, it is imperative to know whether and when the results from disparate methodologies are comparable.

Whereas Wolkovich et al. (2012) compared phenological responses recorded using different methodologies by reviewing studies from around the world, we approached this question using experiments and observations in the same area, and using the same study species. In addition, we included remote sensing data in our analyses, as remote sensing is an increasingly common method for leaf out monitoring (Polgar and Primack 2011). Remote sensing has been used successfully to monitor leafing phenology in a number of recent studies (Shuai et al. 2013), but there are a number of difficulties associated with the practice, including coarse resolution making it difficult to monitor small geographic areas (Fisher and Mustard 2007; Fisher et al. 2006; Pouliot et al. 2011).

Using on-the-ground monitoring, experimental warming, and satellite data, we investigated the effects of climate change on the timing of leaf out of woody plants in the Concord, Massachusetts area in the United States. In this study we sought to determine whether the measured sensitivity of leaf out timing to temperature would be comparable using three distinct methods.

**Materials and methods**

**Field study**

In March, April, and May of the years 2009–2012 we visited Concord two or three times a week, conducting observations of first leaf date (FLD). Each spring we recorded the date on which we first saw each individual study species, *Acer rubrum* (red maple), *Betula lenta* (black birch), and *Quercus rubra* (red oak), in leaf. Red maple and black birch were seen in leaf for the first time in the Conantum area of Concord, while red oak was observed in leaf first at Walden Pond (Fig. 1). We used leaf out criteria similar to that of Project Budburst—a national citizen science plant phenology project. A species was considered to have leafed out when at least three branches on one individual plant had at least one fully unfolded leaf (Project Budburst; http://neoninc.org/budburst). We were not looking at particular individual plants, rather for the date on which any individual in Concord had met the criteria. Because individuals of the same species leaf out at similar times, the first leaf out date is representative of many individuals of the species. Dry bulb air temperatures recorded at the Blue Hill Meteorological Observatory in Milton, MA, a site 33 km southeast of Concord, provided temperature records closely correlated with temperatures in Concord (Miller-Rushing and Primack 2008).

**Experimental warming study**

We used the Boston-Area Climate Experiment (BACE) facility as the research site for our experimental warming study (Hoepner and Dukes 2012; Suseela et al. 2012). Located in Waltham, Massachusetts, about 14 km from Concord, BACE exposed an old-field plant community with tree seedlings to a factorial combination of precipitation and temperature manipulations. Warming treatments were nested within precipitation zones in each of three blocks. Each zone included one plot from each of four warming treatments: unwarmed, low, medium, and high. The plots used in this study all received ambient precipitation. Ceramic infrared heaters were mounted 1 m above the ground on all four corners of each experimental plot, and pointed at the centers of the plots at a 45° downward angle. Infrared radiometers (IRR-PN; Apogee Instruments, Logan, UT) measured the canopy surface temperature in the center of each ambient and high warming plot every 10 s, and LabView software (National Instruments, Austin, TX) controlled power output to heaters to achieve a target warming of 4 °C in the high warming plots relative to the controls, when conditions permitted. Intermediate levels of warming were achieved by using lower wattage heaters (200 and 600 W vs 1,000 W in the high warming treatments), with power to all four plots in a zone controlled simultaneously on a single circuit. Soil temperatures at 2 cm depth were monitored every 30 min with linear temperature sensors, and ratios of soil warming in intermediate treatments to high warming treatments were used to estimate canopy warming in the intermediate treatments. Plots were warmed throughout the entire year. In March and April of 2009, canopy temperatures in the low, medium, and high warming plots were an average of 0.88 °C, 2.06 °C, and 3.38 °C warmer than in the control plots, where the mean temperature was 7.03 °C; in 2010, these differences were 0.75 °C, 2.60 °C, and 3.38 °C respectively, and the ambient mean temperature was 9.31 °C.

Four seedlings (15–30 cm size) of *Acer rubrum*, *Betula lenta*, and *Quercus rubra* were planted into each of the three replicate plots per treatment in May 2008. A second set of four seedlings of each species was added in May 2009. Because *Betula lenta* seedlings were unavailable at the time of planting, *Betula nigra* (river birch) was planted in the second year instead. Spring phenological surveys of all trees were conducted every
2–3 days from 11 April (in 2009) and 15 March (2010) until all leaves had fully expanded (end of June). First leaf out was defined as the date when the first leaf of a seedling of a given species in a given plot had unfolded to the point that the petiole or leaf base had become visible. Because the seedlings were much smaller, with limited branches, than the trees in the field study, the definition of leaf out had to be changed from the Project BudBurst protocol used in the field study (where at least three unfolded leaves on different branches were required to be unfolded for an individual to be in leaf).

Remote sensing study

We made use of moderate-resolution imaging spectroradiometer (MODIS) remotely sensed data to capture the phenological activity of four plots in the Concord area (Fig. 1) and establish the date of spring green-up associated with onset of the vegetation growing period. The four plots, Brister’s Hill, Conantum, Middlesex School, and Punkatasset, named for the areas in Concord in which they occur, are in the same vicinity as the field-sampled data and the localities are dominated primarily by deciduous forest and understory vegetation, although there are some evergreen trees in the plots as well. The plots are based on the MODIS 500 m pixel centered over the forest of interest. Each plot was sampled on the ground at 50 points per site using transects to determine the dominant canopy species. Each plot was divided into five parallel transects; an investigator walked along each transect, stopping at ten evenly spaced intervals and recording the closest canopy tree to the stop. *Acer rubrum*, *Betula lenta*, and *Quercus rubra* accounted for approximately 63% of the canopy cover of these four plots.

Using the daily direct broadcast version of the MODIS albedo and bidirectional reflectance distribution function (BRDF) product (Lucht et al. 2000; Schaaf et al. 2002, 2008, 2011; Shuai et al. 2013; Wang et al. 2012), we generated daily nadir BRDF-adjusted surface reflectances (NBAR) for January–November for the years 2003–2010 and January–May for 2011. A best fit daily BRDF model is determined from all high quality, cloud-free, atmospherically corrected spectral surface reflectances available over a 16-day period and then coupled with the next subsequent single day to establish a daily retrieval. High quality BRDF retrievals are
achieved only if there are sufficient directional observations to adequately sample the surface anisotropy. This daily model (gridded at a 500 m spatial resolution) is then used to obtain a nadir surface reflectance that has been corrected for view-angle effects (an NBAR).

Using the daily NBAR data, we then derived daily NBAR-EVI (enhanced vegetation index) (Huete et al. 2002) data for each year. By applying the so-called XYZ logistic function (Zhang et al. 2003, 2004, 2006) to the temporal NBAR-EVI curve we were able to determine the inflection point or period of maximum curvature related to the spring green-up. The use of the NBAR data and the XYZ logistic function is similar to the method used for the global phenology of the operational MODIS Land Dynamics product (MCD12Q2) (Ganguly et al. 2010; Zhang et al. 2003, 2006). Because of persistent cloud cover over the Concord region during the spring of 2009, it was difficult to achieve sufficient high quality retrievals to generate a highly accurate NBAR-EVI temporal curve and therefore 2009 was dropped from further analysis.

Analysis

We were able to compare the rate of leaf out advance with warming temperature among the three methods because the results could be expressed in terms changes in the dates of leaf out for each 1 °C increase in temperature. For each study, we used simple regression analysis to determine the relationship between FLD and mean March and April temperature in days/°C (Fig. 2). More sophisticated models, such as those using growing degree-days or chilling days, are more difficult to compare across models. The regressions were calculated based on the mean FLD of three species across 4 years in the field study; the mean FLD of two species in four warming treatments for 2 years, and the mean FLD of two species in four treatments for 1 year each at BACE; and green up dates for four plots for 8 years, minus one missing value obtained from MODIS. We used multiple regression analysis to test for differences among the three slopes.

Results

Mean March and April temperatures, as measured at the Blue Hills Meteorological Observatory, varied widely among years, with 2009 and 2011 being approximately 2.5 °C colder than 2010 and 2012 (Table 1). In the field study, leaves emerged earlier in years with warmer temperatures in March and April (Fig. 2; $r^2=0.75$; $P<0.001$), advancing at a rate of 6.1 days/1 °C. At BACE, Acer rubrum, Betula lenta (only in 2009), Betula nigra (only in 2010), and Quercus rubrum together responded to experimental warming at a rate of 2.1 days/°C (Fig. 2; $r^2=0.47$; $P<0.01$). Using MODIS, greenup was calculated to advance at a rate of 3.7 days/°C (Fig. 2; $r^2=0.70$; $P=0.01$). Individual sites ranged in response between 2.1 and 5.7 days/°C.

In all three methods, plants leafed out significantly earlier with higher spring temperatures ($P<0.01$ for all methods). In a one-way ANOVA with a Tukey’s honestly significantly difference test of individual species slopes, the response to temperature calculated using field observations was significantly greater ($P=0.019$) than that produced by experimental warming. The remote sensing results yielded an intermediate response that did not significantly differ from the other two (Fig. 2).

The mean green up date of the MODIS plots in 2010 was 15 April, while in the field study the mean FLD of the study species was 9 April, and at BACE it was 2 May. In 2011 the mean green-up date of the MODIS plots was 28 April, while in the field study the mean FLD of the study species was 27 April. BACE data are not available for 2011 and MODIS data are not available for 2009 because of heavy cloud cover during much of that spring.

Tree seedlings in the warming experiment leafed out later than trees in the field; the difference was greater than 3 weeks under similar March/April temperatures of ~8 °C, but persisted even in the warmest experimental plots, which were ~5 °C warmer than the warmest year observed in the field.

Discussion

Terrestrial biosphere models are critically limited by what we still do not understand about phenology (Richardson et al. 2012). However, many approaches can help to build our understanding, including analyses of long-term phenology observations, climate change experiments and remote sensing. Each of these methodologies is useful, but discrepancies among their results are important to understand.

Each of the three methodologies we compared recorded an advance of leaf out with warmer spring temperature. This is unsurprising and agrees with the general consensus among experts that temperature is the main driver of spring leaf out of woody plants in temperate forests (Polgar and Primack 2011). Our results of leafing advancement of 3.7 days/°C from the remote sensing studies are the same as the response seen in a large study of leaf unfolding across Europe (Estrella et al. 2009). The response rates from all three of our studies fall between those predicted by Kramer of 2–3 days/°C and the advances of 6.7 days/°C seen in a study of four species across Europe (Chmielewski and Rotzer 2001; Kramer et al. 2000).

Our results support the findings of Wolkovich et al. (2012) that experimental warming significantly underestimates the sensitivity of leaf out to increasing spring temperature compared to observations in the field. In their meta-analysis they found that experimental warming advanced spring phenology by 1.9–3.3 days/°C while observational studies reported
sensitivities between 2.5 and 5 days/°C. Our results from BACE of an advancement rate of 2.1 days/°C, and from the field study of 6.1 days/°C are in line with those findings. Depending on the mechanism behind these differences, our results may bolster the conclusion that experimental warming sites underpredict the sensitivity of plant phenology to warming by showing that these results hold true within a single ecosystem as well as across many studies worldwide. However, the striking differences in results between observational monitoring and experimental warming in this study could have arisen through several mechanisms, some of which are unrelated to experimental technique, suggesting the need for a better physiological understanding of the link between temperature and phenology.

There are several possible explanations for the differences between observational and experimental studies. It is possible that warming in experimental plots during the fall and winter had subsequent consequences for springtime temperature sensitivity or leaf out timing (e.g., slowing leaf out by delaying the time at which winter chilling requirements are met); alternatively, average temperatures measured by infrared sensors in experimental plots may have differed from those sensed by trees due to within-plot structural heterogeneity leading to a somewhat uneven distribution of heating (Kimball et al. 2008), or to diurnal differences in temperature sensitivity. Warming often reached target temperatures at night, but often did not reach targets during the day, presumably because plant transpiration was greater and higher wind speeds led to greater convective heat losses from heaters and vegetation (Kimball et al. 2008).

Monthly average temperatures at the experimental plots were also higher than temperatures at Blue Hills Observatory; these temperature differences could have changed the slopes if species responses to temperature are non-linear. The discrepancy in responses might also be due to differences in the age of the plants (seedlings vs trees) and the resulting necessity to use differing definitions of green up, differences in soil conditions (forest vs field), degree of exposure to herbivory (likely higher at BACE), method of temperature measurement (leaf surface temperatures vs air temperatures), or sampling methods (e.g., many more trees were surveyed in the field than in the experiment). The temperatures used in the field study and remote sensing were also taken further away and may not reflect the temperature around individual trees as closely as they do at the BACE. Some of these differences, such as the use of seedlings rather than more mature trees in the BACE, are some of the compromises that must be made in setting up most climate change experiments.

In this study we also examined the use of remote sensing as another way to quantify the effect of temperature changes on the timing of leaf out phenology. Results from our remote sensing...
Remote sensing of vegetation phenology using satellite data has become more widely used in recent years and has been producing increasingly accurate results over the past decade (Kross et al. 2011). Studying phenology at this scale is beneficial because an entire forest or region of the country can be monitored by one person or research group and because trends across different types of geographical and topographical gradients can be compared more easily (Fisher and Mustard 2007). The tradeoff for the ability to examine a large area is that the resolution is coarser than that of traditional phenological monitoring. While remote sensing data can be downloaded and analyzed without the investigator needing to go into the field, there is always a risk that the data have undetected problems (Zhang et al. 2003). Features such as bodies of water, paved surfaces, evergreen plants, and grassy fields can all complicate analysis of data. The detection of understory greening by satellites can also lead to discrepancies between satellite-derived greening estimates and on-the-ground estimates of canopy, or dominant, tree species (Doktor et al. 2009). Finally, a string of cloudy days can cause an entire growing season of data to be unusable, as was the case in 2009 during this study.

Researchers studying or modeling responses of trees and other plants to climate should be aware of the differences in methodology and results when comparing and combining these three methods, and be cautious in predicting future leaf out based on either field observations or warming experiments until discrepancies such as these are better understood. As leaf out phenology continues to be used as an indicator for climate change, it is important to reconcile important differences among various monitoring methods. We found no significant difference in the sensitivity of leaf out to spring temperature between results obtained through traditional field monitoring and satellite remote sensing. This is promising in that satellite monitoring is likely to become a more widely used technology for phenology monitoring because it has the ability to cover a much broader area than on-the-ground monitoring. On the other hand, the sensitivity of leaf out to temperature under experimental warming was significantly lower than in the field study. Future studies should attempt to identify what causes these differences, so that results from warming facilities can be used more confidently to project future changes in phenological events.

Acknowledgments Funding came from Boston University; Purdue University; University of Massachusetts, Boston; National Aeronautics and Space Administration (NASA) MODIS NNX11AD58G; the National Science Foundation (DEB-0546670); and the United States Department of Energy’s Office of Science (Biological and Environmental Research), through the Northeastern Regional Center of the National Institute for Climate Change Research. C. Goranson and H. Emery provided technical support at BACE. A. Miller-Rushing, I. Ibanez, E. Wolkovich, and P. Templer provided comments. We would like to thank several anonymous reviewers for their helpful comments.

References

Kross A, Fernandes R, Seaquist J, Beaubien E (2011) The effect of the temporal resolution of NDVI data on season onset dates and trends...


