

# How well do first flowering dates measure plant responses to climate change? The effects of population size and sampling frequency

Abraham J. Miller-Rushing<sup>1,2\*</sup>, David W. Inouye<sup>1,2</sup> and Richard B. Primack<sup>3</sup>

<sup>1</sup>Rocky Mountain Biological Laboratory, PO Box 519, Crested Butte, CO 81224, USA; <sup>2</sup>Department of Biology, University of Maryland, College Park, MD 20742, USA; and <sup>3</sup>Department of Biology, Boston University, 5 Cummington Street, Boston, MA 02215, USA

## Summary

1. First flowering dates are occurring earlier than they did in the past in many locations around the world. It is sometimes assumed, implicitly or explicitly, that the changes in first flowering dates describe the phenological behaviour of entire populations. However, first flowering dates represent one extreme of the flowering distribution and may be susceptible to undesirable confounding effects.
2. We used observations of flowering in Colorado and Massachusetts to test whether changes in population size and sampling frequency affect observations of first flowering dates.
3. We found that the effect of population size on first flowering dates depended on location. Changes in population size were strongly related to the dates on which first flowering was observed in Massachusetts but not in Colorado. The lack of a significant effect in Colorado may reflect the rapid onset of spring after snowmelt and fixed developmental schedules of the plants at this sub-alpine site, or the scale of the plots sampled during the study.
4. We also found that changes in sampling frequency can influence observed changes in first flowering dates and other aspects of the flowering distribution. Similar to the effect of declines in population size, lower sampling frequency caused later observations of first flowering. However, lower sampling frequency, if maintained consistently throughout a study, did not significantly affect estimates of changes in flowering dates over time or in response to climate.
5. *Synthesis.* Researchers should consider the effects of changes in population size and sampling frequency when interpreting changes in first flowering dates. In some cases, past results may need to be reinterpreted. When possible, researchers should observe the entire flowering distribution or consider tracking peak or mean flowering dates to avoid the confounding effects of population size and sampling frequency.

**Key-words:** climate change, Colorado, Concord, date of first flowering, global warming, Massachusetts, phenology, Rocky Mountain Biological Laboratory, sampling frequency

## Introduction

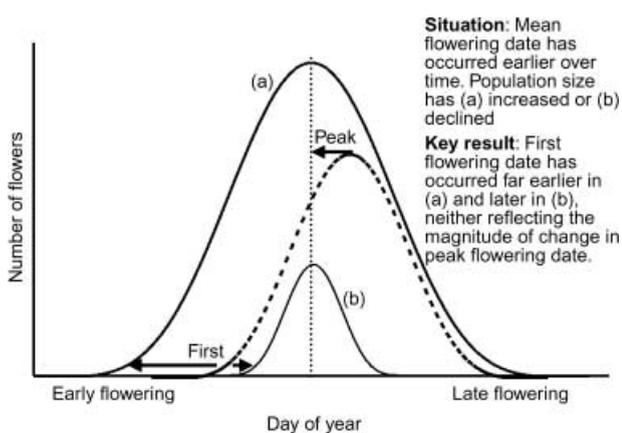
Climate change is altering flowering times world-wide. In most areas studied to date, plants are flowering earlier (Root *et al.* 2003; Parmesan 2007), although these patterns can be species specific (Fitter *et al.* 1995; Fitter & Fitter 2002; Miller-Rushing & Primack 2008) and depend on local conditions and changes in climate (Menzel *et al.* 2006; Zhang *et al.*

2007). In some cases, plants are not flowering earlier or may even be flowering later (Fitter *et al.* 1995; Inouye *et al.* 2003; Miller-Rushing & Primack 2008). Understanding these changes is critical, as plant reproductive phenology is central to many ecological relationships. Moreover, flowering phenology is linked to the timing of other plant life-history traits, such as germination, leaf out and fruiting (Primack 1987; Donohue 2002; Bolmgren & Lönnberg 2005). Thus, changes in flowering phenology can indicate changes in the timing of many events integral to ecological relationships and processes across a variety of seasons and scales. The ecological significance

\*Correspondence author. E-mail: abe@rmbll.org

of phenology has been increasingly recognized recently (e.g. Schwartz 2003; Visser & Both 2005; Post & Inouye 2008).

Many past and recent studies have used first flowering dates to describe changes in the flowering phenology of plant populations (e.g. Sparks & Carey 1995; Bradley *et al.* 1999; Fitter & Fitter 2002; Inouye *et al.* 2002; Inouye *et al.* 2003; Miller-Rushing & Primack 2008). Although researchers would generally prefer to measure changes in entire flowering distributions or mean or peak flowering dates (for reasons discussed below), it is often necessary to rely on first flowering dates because they may be the only data available. It is far easier for an observer to note the date that a species first flowers rather than monitoring the progression of flowering for an entire population, which could last for weeks or even months. However, first flowering dates occur at one extreme of the flowering distribution and those observations may be affected by population size and sampling effort (i.e. number of observers or hours of observation, or frequency of observations). If flowering dates were approximately normally distributed, we would expect to have a greater probability of observing a very early flower in a year when a population size is large or sampling effort is great than in a year with a small population size or a diminished sampling effort (Fig. 1). Changes in the distribution of flowering dates (e.g. from a normal to skewed distribution) or changes in spatial patterns of microclimate could additionally alter the observation of first flowering dates. We would expect that changes in population size or sampling effort would have less of an effect on observations of mean or peak flowering dates, or the date that a certain percent of the plants have flowered. Thus, changes in first flowering dates may reflect changes in population size or sampling effort in addition to or instead of the population's overall phenological response to climate change (Fig. 1).



**Fig. 1.** The theoretical effect of changes in population size on changes in first flowering dates. Dashed curve shows the distribution of flowering dates for a population in a year in the past. The solid curves show two scenarios in which the mean flowering date of the population occurs earlier now than it did in the past: the population size has either (a) increased or (b) declined. Arrows highlight changes in peak and first flowering dates over time. Increases or decreases in sampling effort could similarly affect observed flowering dates. Figure is not drawn to scale and is conceptual only.

It is possible that some previous studies have confounded the effects of climate change with changes in sampling effort and population size. For example, if researchers censused plants for first flowering twice a week for the first 20 years of a study, then sampled 7 days a week for the next 20 years, they might see a pattern of earlier flowering that was unrelated to climate change. Similarly, if a population declined over 20 years, the first flowers might appear later over time even if the peak flowering dates were occurring earlier (Fig. 1). The impact of sampling intensity and population size have been considered in studies of bird arrival times (Tryjanowski & Sparks 2001; Knudsen *et al.* 2007; Miller-Rushing *et al.* 2008), which often rely on records of first arrivals rather than mean arrivals or other measures of migration time (e.g. Bradley *et al.* 1999; Butler 2003; Sparks & Tryjanowski 2007), but to our knowledge, they have not been previously considered in the studies of plant phenology other than in the work of Aldo Leopold (Leopold & Jones 1947).

Here we use long-term phenological records from two different locations – Colorado and Massachusetts – to test empirically whether population sizes and sampling frequency affect observations of first flowering dates, as well as overall flowering distributions. We also test the ability of changes in first flowering dates to predict changes in peak flowering dates, addressing the question: Do first flowering dates serve as an adequate proxy for peak flowering dates?

## Methods

### ROCKY MOUNTAIN BIOLOGICAL LABORATORY, COLORADO

We examined records of flowering for eight species, a subset of those that were observed approximately every 2 days throughout 33 growing seasons (May–September). The species were: *Delphinium nuttallianum* Pritz. ex Walp. (low larkspur), *Erigeron speciosus* (Lindl.) DC. (aspen fleabane), *Eriogonum umbellatum* Torr. (sulphur-flower buckwheat), *Hydrophyllum capitatum* Douglas ex Benth. (ballhead waterleaf), *Lathyrus lanszwertii* var. *leucanthus* (Rydb.) Dorn (aspen peavine), *Potentilla hippiana* Lehm. (woolly cinquefoil), *Taraxacum officinale* F.H. Wigg. (common dandelion), and *Viola praemorsa* Douglas ex Lindl. (canary violet). Each species was observed in an average of 5–13 permanent 2 × 2 m plots each year from 1973 to 2007 (number of plots depended on the species: no observations were made in 1978 or 1990). These plots were located in a sub-alpine meadow habitat at the Rocky Mountain Biological Laboratory (RMBL) in Gothic, Colorado at an elevation of 2900 m (GPS coordinates and metadata for plots and census methods are available at <www.rmbll.org>). The plots contained plants growing naturally and were never manipulated. The number of flowers open for each species was counted every second day with a few exceptions where the interval may have been 3 or 4 days. We used the peak number of flowers observed on a single day as an estimate of the total population of flowers, distinct from the number of plants and the total number of flowers produced in a season. A few plants with a large number of flowers could potentially have been responsible for much of the peak number of flowers. However, the species we examined were relatively small and typically have a small number of flowers per individual. Most of the species can produce more than one flower per individual and many plants may not flower in any given year, as is true for most long-lived perennials such as these.

## CONCORD, MASSACHUSETTS

We examined records of first flowering in Concord, Massachusetts as observed by Alfred Hosmer (Hosmer, A. W. Alfred W. Hosmer Botanical Manuscripts, 1878–1903, William Munroe Special Collections, Concord Free Public Library) from 1888 to 1902 and by Miller-Rushing & Primack (2008) from 2004 to 2006. Hosmer and Miller-Rushing & Primack (2008) observed first flowering dates for species throughout Concord, without any permanent transects. Hosmer made observations about 4 days/week, while Miller-Rushing and Primack made observations about 2.5 days/week, with two separate sets of observers in different locations. Hosmer's phenological observations were made based on more days per week, but fewer total observer days per week than Primack and Miller-Rushing. For detailed descriptions of both data sets, see Miller-Rushing & Primack (2008). We calculated the change in first flowering dates for three groups of species: those with declining population sizes ( $n = 11$ ), increasing population sizes ( $n = 5$ ), and relatively unchanging population sizes ( $n = 34$ ). We considered a species to be declining in abundance if Hosmer (1888–1902) considered it common and Primack *et al.* (unpublished data) considered it now rare (i.e. we only found it in Concord at a single location). Species increasing in abundance were rare in the older period and are now common or frequent (i.e. occur at three or more locations in Concord). We selected the species with relatively unchanging population sizes from those used in Miller-Rushing & Primack (2008) that were relatively common in the past and are still common in Concord today. The group of species with relatively unchanging population sizes acted as our control, with minimal effects of changes in population size. It is important to note that due to the different types of data collected, population size for Concord refers to the number of locations where a species occurred in the town – more locations was interpreted as a larger population size – and for Colorado refers to the number of flowers observed in specific plots.

## EFFECTS OF POPULATION SIZES

*Colorado*

We used multiple linear regression to determine the relationship between first flowering date (response variable) and peak flowering date and the peak number of flowers (explanatory variables) for each species. We used this method to test whether the peak number of flowers explained any of the variation in first flowering dates beyond that explained by peak flowering date. In addition, we used an  $F$  test to compare the residual sum of squares among three models that varied in their level of restriction: (i) intercepts and slopes held constant across species, (ii) slopes held constant across species, but intercepts allowed to vary, and (iii) slopes and intercepts allowed to vary across species (as with individual regressions for each species).

We also ran simple linear regressions with peak flowering date as the response variable and first flowering date as the explanatory variable for each species. For each species, we then used a  $t$ -test to determine if the slope of the relationship was significantly different from one. If first flowering date were a one-to-one predictor of peak flowering date, as is sometimes assumed, we would expect the slope to equal one.

*Massachusetts*

We used single-factor ANOVA to compare changes in first flowering dates among the three categories of species – declining, increasing

and unchanging (control). If changes in population sizes affected first flowering dates, we expected that the flowering dates of the declining species would advance more slowly than the control group, whereas the flowering dates of the species increasing in population size would advance more rapidly than the control group (Fig. 1). Before performing the calculations, we adjusted the flowering dates to days after the vernal equinox to account for directional changes in the timing of the equinox over time (Sagarin 2001).

## EFFECTS OF SAMPLING FREQUENCY

We set out to answer three questions regarding the effects of sampling frequency on the observations of flowering dates: (i) Does sampling frequency alter the apparent distribution of flowering times? (ii) Does sampling frequency affect estimates of the change in flowering dates over time or the relationship between flowering dates and the timing of snowmelt? (iii) Does sampling frequency affect the ability to detect trends in flowering dates over time? For the empirical portion of this analysis we used only observations made at RMBL in Colorado. The original observations of flowering for eight species (see Rocky Mountain Biological Laboratory, Colorado) were made every second day. We then degraded this data set to create a set where observations were made every sixth day (i.e. days  $n$ ,  $n + 6$ ,  $n + 12$ , etc.), thus decreasing the sampling frequency. Several phenology studies have previously relied on observations made this often or less frequently (e.g. Bradley *et al.* 1999; Miller-Rushing *et al.* 2006).

To answer the first question, we compared the flowering distributions of the original, intensely sampled data with the degraded data. Specifically, we used paired  $t$ -tests to compare the dates of first, peak and last flowering as well as the peak number of flowers observed on any day.

To address the second question, we used panel analysis (Hsiao 2003) with both data sets together. We tested whether the frequency of sampling affected estimates of the change in first flowering date, peak flowering date, last flowering date, flowering duration and peak number of flowers over time. Panel analysis is a form of multiple regression that allowed us to consider all of the plant species in a single model. Panel analysis increased the power of our models to find statistically significant relationships, improved the efficiency of the models' estimates, and controlled for estimation biases. In each model the response variable was a characteristic of the flowering distribution (first, peak, or last flowering date, duration of flowering or peak number of flowers), and the explanatory variables were year and a year-sampling frequency interaction term. We performed an identical test to determine whether sampling frequency affected estimates of the response of first flowering dates to changes in date of snowmelt. In this location, snowmelt is the primary environmental cue for flowering phenology (Inouye & McGuire 1991; Inouye *et al.* 2002; Dunne *et al.* 2003; Inouye 2008). If the panel model indicated that the regression coefficients varied significantly among species (as determined by an  $F$ -test comparing models with varying levels of restriction), we performed regressions for each species individually.

For the third question, we used Monte Carlo techniques to estimate the ability of various sampling frequencies to detect changes in flowering dates in the future. We used the following equation to generate one thousand experimental data sets for spring temperature for the next 50 years:

$$T_Y = \alpha + \beta Y + \mu_Y$$

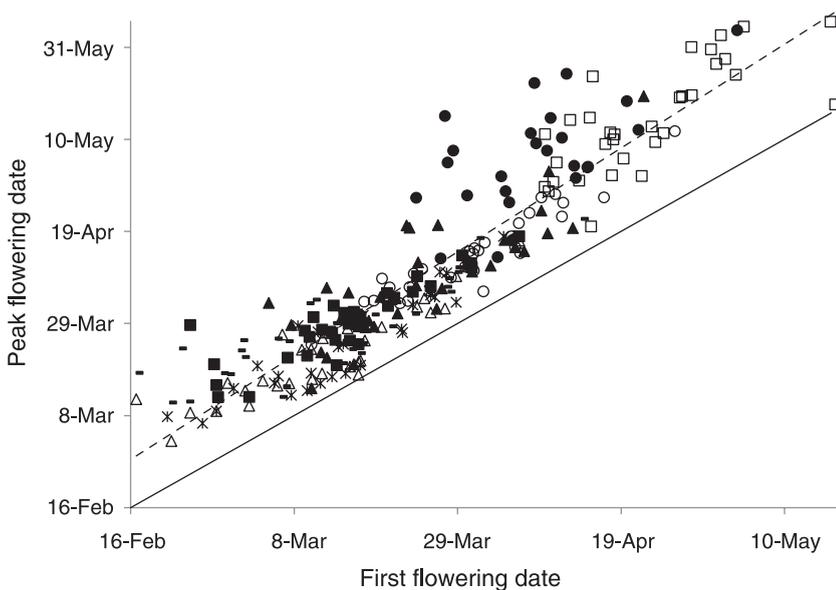
in which  $T$  was temperature in year  $Y$  of the experimental data set,  $\alpha$  was a constant,  $\beta$  represented the linearized annual rate of warming and  $\mu$  was an error term. We set  $\beta = 0.028$  °C/year, a mid-range

estimate of warming made by the Intergovernmental Panel on Climate Change (IPCC 2007). This warming scenario is in line with the downscaled predictions for annual and seasonal warming in the New England region over the next 100 years (New England Regional Assessment Group 2001; Hayhoe *et al.* 2007). The error term was drawn randomly from a normal distribution with a mean of zero and a SD of 1.2, which was the SD of January, April and May temperatures in Boston during 1831–2004. Temperatures in these months are significantly associated with flowering dates in many Massachusetts plants, whereas temperatures in other months are correlated with the flowering dates of relatively few species (Miller-Rushing & Primack 2008). This procedure assumed that mean January, April and May temperature variability will remain constant in the future. We used Massachusetts and not Colorado as a setting for this simulation because we have much longer temperature records for Massachusetts than we have snowmelt records for Colorado, which allowed us to estimate more precisely long-term inter-annual climate variation.

We then used these temperature simulations to test whether we could detect a change in flowering dates over time. For each experimental data set, we calculated flowering date as:

$$FD_Y = \alpha + \beta T_Y + \mu_Y$$

where FD was the flowering date in year  $Y$ ,  $\alpha$  was a constant, and  $\mu$  was an error term. We considered nine scenarios (three temperature effects on flowering  $\times$  three sampling frequencies). We set  $\beta$ , the linearized effect of temperature ( $T$ ) on flowering date to one of three magnitudes: 6 days earlier/ $^{\circ}\text{C}$  warming, 3 days earlier/ $^{\circ}\text{C}$ , or 1 day earlier/ $^{\circ}\text{C}$ . These values are all realistic flowering responses to temperature in Concord, Massachusetts (Miller-Rushing & Primack 2008). The error term contained information about the sampling frequency. We tested three sampling frequencies: every 2 days, every 7 days and every 14 days. Finally, we used ordinary least squares regression to test in each year whether we could detect a significant change ( $P < 0.05$ ) in flowering dates for each experimental data set. Because one anomalously warm year might create a significant trend that would disappear the next year, we recorded the point at which the trend had been statistically significant for five consecutive years. We tested for changes in flowering dates for a period of 50 years.



**Fig. 2.** The relationship between first flowering date and peak flowering date for eight montane wildflower species in Colorado. Each point represents one species in one year. *Delphinium nuttallianum*: solid squares; *Eriogonum speciosus*: open squares; *Eriogonum umbellatum*: open circles; *Hydrophyllum capitatum*: asterisks; *Lathyrus lanszwertii* var. *leucanthus*: solid triangles; *Potentilla hippiana*: solid circles; *Taraxacum officinale*: solid bars; *Viola praemorsa*: open triangles. Solid line is one-to-one line. Dashed line indicates the least squares best-fit line for all species combined.

## Results

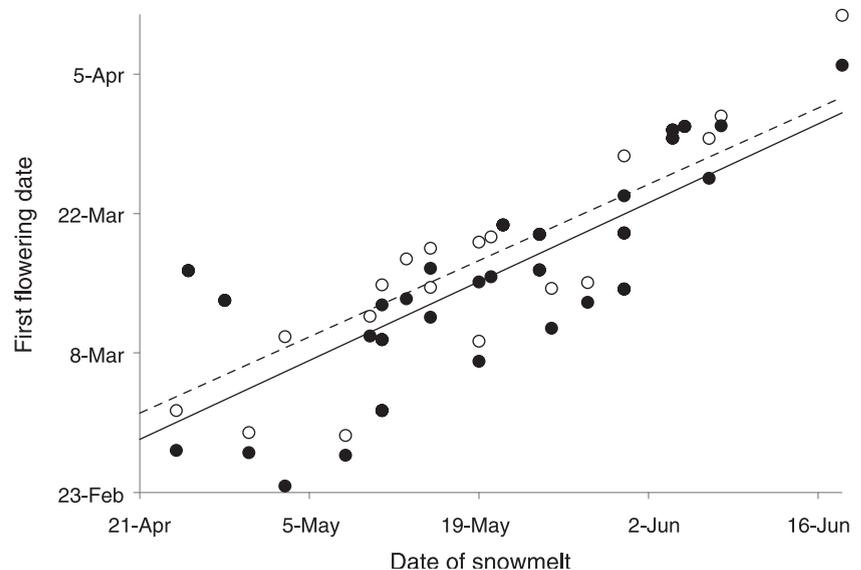
### EFFECTS OF CHANGING POPULATION SIZES

#### Colorado

An  $F$ -test indicated the relationships between first flowering date and peak number of flowers varied significantly among species ( $F$ -test  $P < 0.001$ ,  $H_0$ : slope of relationship did not differ among species). Thus, we evaluated the relationship with individual regressions for each species. Of the eight species we examined at RMBL, only one of them had a significant relationship between first flowering date and peak number of flowers, as determined by multiple regression with peak flowering date and peak number of flowers as explanatory variables. The first flowers of *L. lanszwertii* var. *leucanthus* opened  $0.56 \pm 0.23$  ( $\pm$  SE) days earlier for each additional 10 flowers ( $P = 0.023$ ; peak flowering date: slope = 0.697 days/day,  $P < 0.001$ ; adjusted  $R^2 = 0.73$ ). Similarly, an  $F$ -test indicated that the relationship between flowering duration and peak number of flowers differed among species ( $F$  test  $P < 0.010$ ). After evaluating each species individually, only *E. umbellatum* had a significant relationship between duration of flowering and the peak number of flowers ( $3.5 \pm 0.95$  longer duration/10 flowers,  $P = 0.001$ ). By chance we would have expected 5% of species to show a significant relationship in both instances, suggesting that the significant relationships may have occurred by chance (although the very low  $P$ -value for the relationship between flowering duration and peak number of flowers relationship for *E. umbellatum* suggests that this relationship is real).

We next tested the ability of first flowering date to predict peak flowering date. An  $F$ -test indicated that the slope of the relationship did not vary significantly among species ( $P = 0.274$ ). Peak flowering dates occurred  $0.85 \pm 0.04$  days earlier for each day earlier that the first flower was observed ( $P < 0.001$ , adjusted  $R^2 = 0.68$ ; Fig. 2). Intriguingly, the slope

**Fig. 3.** Estimates of the relationship between the timing of snowmelt and first flowering date for *Delphinium nuttallianum* using two sampling frequencies: observations made every 2 days (solid circles and solid line) and every 6 days (open circles and dashed line). Each point represents the first flowering date for this native wildflower at the Rocky Mountain Biological Laboratory in Colorado. Data are shown for the years 1975–2007, except 1978 and 1990, for which data are missing. The slopes of the two least squares best-fit lines are not statistically distinguishable (sampling frequency  $\times$  snowmelt interaction:  $P = 0.861$ ; change in flowering date:  $0.6 \pm 0.1$  days earlier/day earlier snowmelt,  $P < 0.001$ ). Estimates of first flowering dates using observations made every 6 days are consistently later than estimates using observations made every 2 days.



of the relationship was significantly less than one ( $t = 4.09$ ,  $P < 0.001$ ) and first flowering dates explained 68% of the variation in peak flowering dates, as determined by the adjusted  $R^2$ .

### Massachusetts

Directional changes in population size had a significant relationship with changes in first flowering times in Concord, Massachusetts, as determined by ANOVA ( $P = 0.010$ ). Species with declining population sizes flowered  $7.2 \pm 6.1$  days later in the period 2004–06 than they did in the period 1888–1902. In comparison, species with increasing population sizes flowered  $12.0 \pm 4.3$  days earlier over the same time period, while species with relatively unchanging population sizes flowered  $5.1 \pm 1.5$  days earlier.

### EFFECTS OF CHANGING SAMPLING FREQUENCY

As expected, a relatively low frequency of observations (making observations every 6 days) in Colorado significantly delayed the observation of first flowering dates (average delay =  $2.5 \pm 0.2$  days,  $t = -16.62$ , two-tailed  $P < 0.001$ ), advanced the observation of last flowering dates (average advance =  $3.5 \pm 0.5$  days,  $t = 7.56$ , two-tailed  $P < 0.001$ ), and shortened the observed duration of flowering (average shortening =  $6.0 \pm 0.5$  days,  $t = 11.79$ , two-tailed  $P < 0.001$ ) relative to making frequent observations (every 2 days), as determined by paired  $t$ -tests. Maintaining a low frequency of observations also significantly decreased the peak number of flowers observed (average decrease =  $43.8 \pm 4.6$  flowers,  $t = -9.57$ , two-tailed  $P < 0.001$ ), but did not affect the observed date of peak flowering (average advance =  $0.4 \pm 0.3$  days,  $t = 1.43$ , two-tailed  $P = 0.153$ ). Overall, making observations with low frequency (every 6 days rather than every 2 days) caused the distribution of flowering dates to shrink.

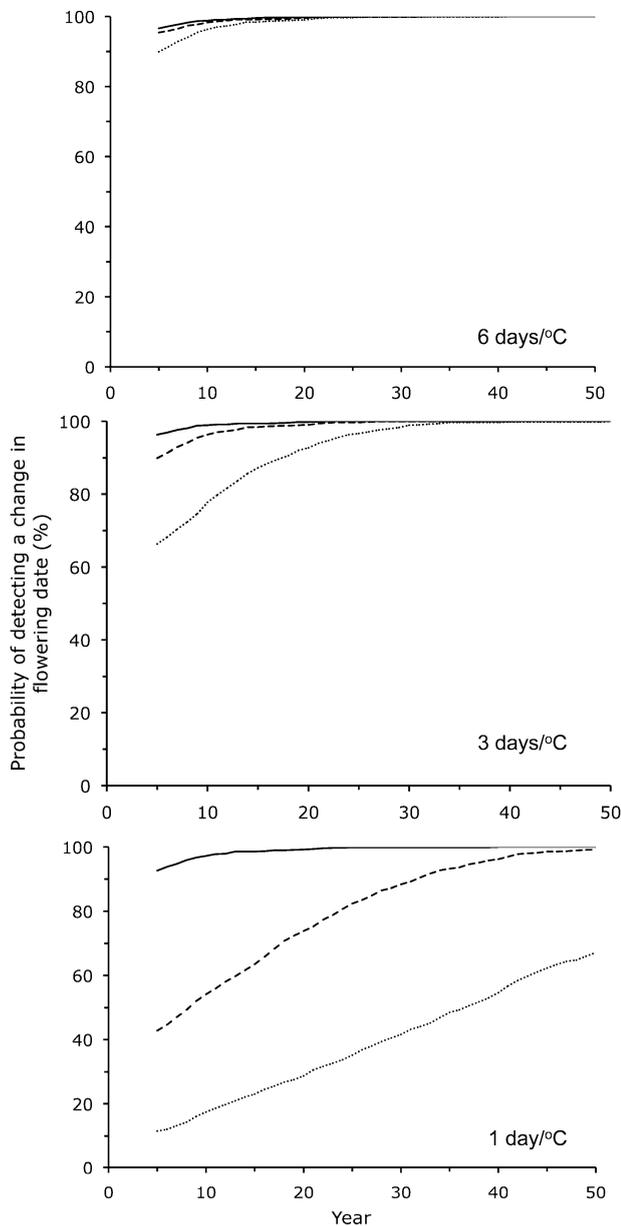
Sampling less frequently in Colorado, however, did not significantly affect estimates of changes in first, peak, or last

flowering dates, duration, or the peak number of flowers over time, as determined by the interaction between sampling frequency and year in random effects panel models ( $P > 0.82$  for each interaction term). Neither did sampling frequency significantly affect estimates of the relationships of flowering dates nor duration with the timing of snowmelt ( $P > 0.65$  for each interaction term in random effects panel models). The relationship between the peak number of flowers and the date of snowmelt varied among species ( $F$ -test  $P < 0.001$ ), but individual regression models for each species indicated that sampling frequency did not affect the estimates of the relationship for any of the species ( $P > 0.13$  for each species' interaction term) (Fig. 3).

Finally, sampling less frequently in Colorado substantially reduced the ability to detect a change in flowering dates for a species that flowered just one day earlier for each  $1^\circ\text{C}$  warming, but did not have much effect on the ability to detect a change in flowering date for a species that flowered 6 days earlier for each  $1^\circ\text{C}$  warming (Fig. 4). After 10 years of observation, there was a 96–99% chance of detecting a 6 day/ $^\circ\text{C}$  advance in flowering date and a 78–99% chance of detecting a 3 day/ $^\circ\text{C}$  advance, depending on sampling frequency. When the actual change in flowering date was just 1 day/ $^\circ\text{C}$ , however, an observer would have a 97% chance of detecting a significant trend over 10 years if sampling every 2 days, a 54% chance if sampling every 7 days, and an 18% chance if sampling every 14 days (Fig. 4).

### Discussion

We found that population sizes and sampling frequency may substantially affect observations of first flowering dates and estimates of changes in first flowering dates. Surprisingly, the presence of an effect depended on the location and method of the study. In Concord, Massachusetts, changes in population size appeared to alter observed changes in first flowering dates. While the first flowering dates for the control group are



**Fig. 4.** The effect of sampling frequency on the ability to detect changes in flowering dates. Each curve represents the percent of simulated experimental data sets in which the study detected a significant ( $P < 0.05$ ) change in flowering date for five consecutive years. Three sampling frequencies are shown in each panel: every 2 days (solid curve), every 7 days (dashed curve), and every 14 days (dotted curve). Each panel shows the results for a plant species with a flowering date that responds to temperature at a different rate: 6 days/°C, 3 days/°C, and 1 day/°C.

occurring about 4 days earlier than they did 100 years ago, the first flowers for species with increasing population sizes are opening 12 days earlier. The first flowering dates for species with declining population sizes are occurring seven days later than they did 100 years ago. It is also possible, however, that the populations of species that are not responding phenologically to climate change are declining, possibly due to mistimed ecological relationships (Willis *et al.* unpublished data). The

relationship may exist in both directions. Because in Massachusetts a species with an increasing population size occurs in a larger number of locations over time, the species also covers more environmental variation, including variation in temperature caused by shading, aspect, soils and other microsite features. Individuals growing at warmer sites will generally flower earlier than those at cooler sites. Our finding that species with declining population sizes have flowered later particularly suggests that population size is causing the later first flowering dates, because warming temperatures would generally be expected to lead to earlier or unchanging plant phenology in the Massachusetts climate, where winters are generally cold enough to meet chilling requirements (Schwartz 1998; Chuine 2000; Zhang *et al.* 2007). Additionally, many plants and animals in eastern Massachusetts are active earlier in the spring now than they have been in the past (Ledneva *et al.* 2004; Miller-Rushing *et al.* 2006; Miller-Rushing *et al.* 2008; Miller-Rushing & Primack 2008), so it is reasonable to expect that timing-based mismatches may be occurring for those species not active earlier in the spring (Stenseth & Mysterud 2002; Visser & Both 2005).

At RMBL in Colorado, however, changes in population size did not substantially affect first flowering dates. At that location, changes in first flowering dates provided fairly good estimates of changes in peak flowering dates (Fig. 2). It is important to note, though, that first flowering dates did not provide a one-to-one prediction of peak flowering dates; first flowering dates explained 68% of the variation in peak flowering dates and peak flowering dates occurred just  $0.85 \pm 0.04$  days earlier for each day earlier first flowering. We suspect that the lack of an effect of population size on first flowering dates at RMBL may have reflected the relatively small area of fixed space that was sampled (i.e. the same  $2 \times 2$ -m plots each year) and the rapid onset of the growing season after snowmelt in this sub-alpine environment (Inouye & McGuire 1991; Inouye *et al.* 2002; Dunne *et al.* 2003). When population sizes increased in the plots, they did not cover an appreciably greater range of microclimates, as occurred in Massachusetts. In addition, a skewed flowering distribution (e.g. many early-flowering individuals with a long tail of late-flowering individuals) (Thomson 1980) could have minimized the effect of population size on the distribution of flowering times. However, the flowering distributions of the species we observed in Colorado were not generally skewed (data not shown).

Making observations every 6 days, instead of every 2 days, caused the observed distribution of flowering times to shrink. First flowering was recorded later, last flowering occurred earlier, and the peak number of flowers observed declined, while the date of peak flowering did not change significantly. Importantly, sampling frequency did not significantly alter estimates of changes in flowering dates, duration of flowering, or peak number of flowers observed over time, nor did it affect estimates of the relationship between those variables and the timing of snowmelt. For example, a low sampling frequency resulted in later observations of first flowering, but did not affect estimates of how first flowering dates changed over time

or how they responded to the date of snowmelt (Fig. 3). However, as expected, a low sampling frequency could substantially reduce the chances of detecting a significant change in flowering dates over time by increasing the variability in the date that flowering is observed (Fig. 4). The strength of this effect is most pronounced for species with flowering dates that are not changing very rapidly.

These findings have important implications for researchers examining phenological change in plant populations. First, if the only data available are first flowering dates, researchers should account for changes in population size and sampling effort. In many cases population sizes or sampling effort might change directionally over time, and these changes can significantly alter changes in first flowering dates (Fig. 1), although they do not always. Increases in population size or sampling effort can lead to earlier first flowering dates, while declines in population size or sampling effort can delay first flowering dates. It is possible that monitoring relatively small, fixed plots or marked individuals may minimize the effect of changes in population sizes, as we observed in Colorado, however further research is needed to confirm that this finding is not simply due to the rapid onset of flowering in this area. Conceptually, population size or sampling effort could affect the observation of first flowering even when measuring the phenology of individual marked plants over time if the number of flowers produced or sampling effort varies significantly among years (Primack 1985).

Second, studies that differ only in sampling frequency will find different first flowering dates on average, but should find the same change in first flowering dates over time and the same flowering responses to snowmelt or temperature. For example, consider a case in which two researchers studied first flowering dates for the same species in the same location for 20 years but used different sampling frequencies – one sampled every 2 days, the other every 6 days. Our results show that sampling frequency alone would not cause the two studies to differ in their estimates of change in first flowering dates. Without other confounding factors, the trends in flowering dates would be plotted as parallel lines (Fig. 3). Other factors, such as changes in population size, nonlinear changes in climate, or nonlinear flowering responses to climate, might still confound comparisons between the two studies if they were carried out in different locations or over different time periods.

Third, sampling frequency can substantially affect the ability of a study to detect changes in flowering dates. This point may seem obvious, but it suggests that studies that fail to detect changes in flowering dates over short time periods or after using relatively infrequent sampling may simply lack the power to detect changes that are actually occurring. It requires fairly frequent sampling to detect changes in flowering dates given that the phenologies of most plants studied to date are changing relatively slowly (Parmesan 2007) and that there is high inter-annual variability in weather that cues the flowering dates for many plant species (Cleland *et al.* 2007). For species with very short flowering durations, frequency of sampling may be particularly important. This result also shows that

future studies of phenological change should carefully consider sampling frequency as a part of their study design.

Fourth, results could be difficult to interpret when two or more factors are affecting first flowering times. For example, if flowering dates are becoming earlier because of warming temperatures, but declining population sizes are causing first flowering to occur later, the two shifts could cancel each other. No overall change in first flowering would be observed. Or if flowering phenology and abundance did not change, but sampling intensity increased during the study, then researchers might erroneously conclude that climate change was affecting phenology.

In summary, population size and sampling frequency can affect observations of changes in first flowering dates. The effects are not always intuitive, nor are they always present. To avoid the confounding effects of population size and sampling effort, researchers should record the entire flowering distribution whenever possible, or consider observing mean or peak flowering dates to control for undesired confounding effects. Observing mean or peak flowering dates requires observing the entire flowering season, which involves greater effort than observing just first flowering, but it results in data less susceptible to the influences of confounding factors. If first flowering dates are the only data available, researchers must consider the effects of population size and sampling effort when interpreting their results.

## Acknowledgements

Authors thank Kjell Bolmgren, Jessica Forrest and Mark D. Schwartz for providing valuable comments on this manuscript. Funding and research assistance for this project were provided by the National Science Foundation (dissertation improvement grant, grants DEB 75-15422, DEB 78-07784, BSR 81-08387, DEB 94-08382, IBN-98-14509, DEB-0238331, and DEB-0413458), Sigma Xi, an NDEA Title IV predoctoral fellowship, research grants from the University of Maryland's General Research Board and Boston University, and assistance from Earthwatch and its Research Corps [DWI]. RMBL provided research facilities and access to study sites. Snowpack data were provided by Billy Barr.

## References

- Bolmgren, K. & Lönnberg, K. (2005) Herbarium data reveal an association between fleshy fruit type and earlier flowering time. *International Journal of Plant Sciences*, **166**, 663–670.
- Bradley, N.L., Leopold, A.C., Ross, J. & Huffaker, W. (1999) Phenological changes reflect climate change in Wisconsin. *Proceedings of the National Academy of Sciences USA*, **96**, 9701–9704.
- Butler, C.J. (2003) The disproportionate effect of global warming on the arrival dates of short-distance migratory birds in North America. *IBIS*, **145**, 484–495.
- Chaine, I. (2000) A unified model for the budburst of trees. *Journal of Theoretical Biology*, **207**, 337–347.
- Cleland, E.E., Chuine, I., Menzel, A., Mooney, H.A. & Schwartz, M.D. (2007) Shifting plant phenology in response to global change. *Trends in Ecology and Evolution*, **22**, 357–365.
- Donohue, K. (2002) Germination timing influences natural selection on life-history characters in *Arabidopsis thaliana*. *Ecology*, **83**, 1006–1016.
- Dunne, J.A., Harte, J. & Taylor, K.J. (2003) Subalpine meadow flowering phenology responses to climate change: integrating experimental and gradient methods. *Ecological Monographs*, **73**, 69–86.
- Fitter, A.H. & Fitter, R.S.R. (2002) Rapid changes in flowering time in British plants. *Science*, **296**, 1689–1691.
- Fitter, A.H., Fitter, R.S.R., Harris, I.T.B. & Williamson, M.H. (1995) Relationships between flowering date and temperature in the flora of a locality in central England. *Functional Ecology*, **9**, 55–60.

- Hayhoe, K., Wake, C.P., Huntington, T.G., Luo, L., Schwartz, M.D., Sheffield, J. *et al.* (2007) Past and future changes in climate and hydrological indicators in the US Northeast. *Climate Dynamics*, **28**, 381–407.
- Hsiao, C. (2003) *Analysis of Panel Data*, 2nd edn. Cambridge University Press, Cambridge.
- Inouye, D.W. (2008) Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. *Ecology*, **89**, 353–362.
- Inouye, D.W. & McGuire, A.D. (1991) Effects of snowpack on timing and abundance of flowering in *Delphinium nelsonii* (Ranunculaceae): implications for climate change. *American Journal of Botany*, **78**, 997–1001.
- Inouye, D.W., Morales, M.A. & Dodge, G.J. (2002) Variation in timing and abundance of flowering by *Delphinium barbeyi* Huth (Ranunculaceae): the roles of snowpack, frost, and La Niña, in the context of climate change. *Oecologia*, **130**, 543–550.
- Inouye, D.W., Saavedra, F. & Lee, W. (2003) Environmental influences on the phenology and abundance of flowering by *Androsace septentrionalis* L. (Primulaceae). *American Journal of Botany*, **90**, 905–910.
- IPCC (2007) Climate Change 2007: the Physical Science Basis. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (eds S. Solomon *et al.*), pp. 747–845. Cambridge University Press, New York.
- Knudsen, E., Lindén, A., Ergon, T., Jónzén, N., Vik, J.O., Knappe, J., Røer, J.E. & Stenseth, N.C. (2007) Characterizing bird migration phenology using data from standardized monitoring at bird observatories. *Climate Research*, **35**, 59–77.
- Ledneva, A., Miller-Rushing, A.J., Primack, R.B. & Imbres, C. (2004) Climate change as reflected in a naturalist's diary, Middleborough, Massachusetts. *Wilson Bulletin*, **116**, 224–231.
- Leopold, A. & Jones, S.E. (1947) A phenological record for Sauk and Dane Counties, Wisconsin, 1935–1945. *Ecological Monographs*, **17**, 81–122.
- Menzel, A., Sparks, T.H., Estrella, N., Koch, E., Aasa, A., Ahas, R. *et al.* (2006) European phenological response to climate change matches the warming pattern. *Global Change Biology*, **12**, 1969–1976.
- Miller-Rushing, A.J., Lloyd-Evans, T.L., Primack, R.B. & Satzinger, P. (2008) Bird migration times, climate change, and changing population sizes. *Global Change Biology*, **14**, 1959–1972.
- Miller-Rushing, A.J. & Primack, R.B. (2008) Global warming and flowering times in Thoreau's Concord: a community perspective. *Ecology*, **89**, 332–341.
- Miller-Rushing, A.J., Primack, R.B., Primack, D. & Mukunda, S. (2006) Photographs and herbarium specimens as tools to document phenological changes in response to global warming. *American Journal of Botany*, **93**, 1667–1674.
- New England Regional Assessment Group (2001) *New England Regional Assessment*. The University of New Hampshire, Institute for the Study of Earth, Oceans, and Space, Durham.
- Parmesan, C. (2007) Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Global Change Biology*, **13**, 1860–1872.
- Post, E. & Inouye, D.W. (2008) Phenology: response, driver, and integrator. *Ecology*, **89**, 319–320.
- Primack, R.B. (1985) Patterns of flowering phenology in communities, populations, individuals, and single flowers. *The Population Structure of Vegetation* (ed. J. White), pp. 571–593. Dr. W. Junk Publishers, Dordrecht.
- Primack, R.B. (1987) Relationships among flowers, fruits, and seeds. *Annual Review of Ecology and Systematics*, **18**, 409–430.
- Root, T.L., Price, J.T., Hall, K.R., Schneider, S.H., Rosenzweig, C. & Pounds J.A. (2003) Fingerprints of global warming on wild animals and plants. *Nature*, **421**, 57–60.
- Sagarin, R. (2001) False estimates of the advance of spring. *Nature*, **414**, 600.
- Schwartz, M.D. (1998) Green-wave phenology. *Nature*, **394**, 839–840.
- Schwartz, M.D., ed. (2003) *Phenology: An Integrative Environmental Science*. Kluwer Academic Publishers, Dordrecht.
- Sparks, T. & Tryjanowski, P. (2007) Patterns of spring arrival dates differ in two hirundines. *Climate Research*, **35**, 159–164.
- Sparks, T.H. & Carey, P.D. (1995) The responses of species to climate over two centuries: an analysis of the Marsham phenological record 1736–1947. *Journal of Ecology*, **83**, 321–329.
- Stenseth, N.C. & Mysterud, A. (2002) Climate, changing phenology, and other life history traits: nonlinearity and match-mismatch to the environment. *Proceedings of the National Academy of Sciences, USA*, **99**, 13379–13381.
- Thomson, J.D. (1980) Skewed flowering distributions and pollinator attraction. *Ecology*, **61**, 572–579.
- Tryjanowski, P. & Sparks, T.H. (2001) Is the detection of the first arrival date of migrating birds influenced by population size? A case study of the red-backed shrike *Lanius collurio*. *International Journal of Biometeorology*, **45**, 217–219.
- Visser, M.E. & Both, C. (2005) Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings of the Royal Society B*, **272**, 2561–2569.
- Zhang, X., Tarpley, D. & Sullivan, J.T. (2007) Diverse responses of vegetation phenology to a warming climate. *Geophysical Research Letters*, **34**, doi:10.1029/2007GL031447.

Received 15 March 2008; accepted 23 July 2008

Handling Editor: Richard Bradshaw