

## Tree phenology responses to winter chilling, spring warming, at north and south range limits

James S. Clark<sup>\*,1,2</sup>, Carl Salk<sup>1,3,4</sup>, Jerry Melillo<sup>5</sup> and Jacqueline Mohan<sup>6</sup>

<sup>1</sup>Nicholas School of the Environment, Duke University, Durham, North Carolina 27705, USA; <sup>2</sup>Department of Statistical Science, Duke University, Durham, North Carolina 27705, USA; <sup>3</sup>Institute of Behavioral Science, University of Colorado, Boulder, Colorado 80309, USA; <sup>4</sup>International Institute for Applied Systems Analysis, Laxenburg, A-2361, Austria; <sup>5</sup>The Ecosystems Center, Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA; and <sup>6</sup>Odum School of Ecology, University of Georgia, 517 BioSciences Bldg, Athens, Georgia 30602, USA

### Summary

1. Increases in primary production may occur if plants respond to climate warming with prolonged growing seasons, but not if local adaptation, cued by photoperiod, limits phenological advance. It has been hypothesized that trees with diffuse-porous xylem anatomy and early successional species may respond most to warming. Within species, northern populations may respond most due to the fact that growing seasons are relatively short. Species most sensitive to spring temperature may show little overall response to warming if reduced chilling in fall/winter offsets accelerated winter/spring development.

2. Because current thermal models consider only highly aggregated variables, for example degree-days or chilling units (temperature sums for a season or year), they may not accurately represent warming effects. We show that assumptions contained in current thermal (degree-day) models are unrealistic for climate change analysis. Critical threshold parameters are not identifiable, and they do not actually have much to do with thresholds for development. Traditional models further overlook the discrete nature of observations, observation error and the continuous response of phenological development to temperature variation. An alternative continuous development model (CDM) that addresses these problems is applied to a large experimental warming study near northern and southern boundaries of 15 species in the eastern deciduous forest of the USA, in North Carolina and Massachusetts.

3. Results provide a detailed time course of phenological development, including vernalization during winter and warming in spring, and challenge the basic assumptions of thermal models. Where traditional models find little evidence of a chilling effect (most are insignificant or have the wrong sign), the continuous development model finds evidence of chilling effects in most species.

4. Contrary to the hypothesis that northern populations respond most, we find southern populations are most responsive. Because northern populations already have a compressed period for spring development, they may lack flexibility to further advance development. A stronger response in the southern range could allow residents to resist northward migration of immigrants as climate warms. If potential invaders fail to exploit a prolonged growing season to the same degree as residents, then there is a resident advantage.

5. Hypothesized effects of warming for xylem anatomy and successional status are not supported by the 15 species in this study.

**Key-words:** degree-day models, forests, hidden Markov model, phenology, warming

### Introduction

Climate warming will increase productivity of forests if trees can exploit longer growing seasons (Goulden *et al.*

1996; Keeling, Chin & Whorf 1996; Chuine & Beaubien 2001; Nemani *et al.* 2003; Baldocchi *et al.* 2005; Churkina *et al.* 2005; Richardson *et al.* 2009). On the one hand, geographic variation (Badeck *et al.* 2004; Yang *et al.* 2012) and advances in budbreak with warming trends over time (Schwartz 1998; Menzel & Fabian 1999; Fitter & Fitter

\*Correspondence author. E-mail: jimclark@duke.edu

2002; Davi *et al.* 2006; Menzel *et al.* 2006; Van Vliet & Wielgolaski 2006; Cleland *et al.* 2007; Jeong *et al.* 2011; Wang *et al.* 2011; Yang *et al.* 2012) both suggest that growth will respond to warming unless offset by limited moisture. On the other hand, photoperiod cues can dampen phenological advance (Wareing 1956; Ashby *et al.* 1992; Mimura & Aitken 2007; Aldrete, Mexal & Burr 2008; Lopez *et al.* 2008; Körner & Basler 2010; Cooke, Eriksson & Junttila 2012). The interactions and conflicting evidence from different types of data (Parmesan 2007; Diez *et al.* 2012; Wolkovich *et al.* 2012; Yang *et al.* 2012) have frustrated efforts to predict responses to climate change. We implement a hierarchical model of phenological response to warming that accounts for the sources of uncertainty. Application of a new approach to a large experimental warming study shows the extent of the response for each species, northern vs southern seed sources within species, depending on temperature variation over the previous seasons. We evaluate species responses in the context of different ecological traits, such as xylem anatomy and successional status, and population differences resulting from seed sources from northern and southern portions of the range.

Traditional thermal models (including degree-day models) can misrepresent how warming affects phenology, in part because they collapse temperature time series into a cumulative sum or mean value for a year or season. The temperature threshold in these models defines onset of phenological development. Such a mechanism could protect vulnerable tissues from exposure to late frost. The cumulative sum of temperatures above the threshold, degree-days (DD), is taken to be a requirement that must be fulfilled before a discrete phenological event (e.g. budbreak) happens. Threshold warming and chilling parameters in thermal models (e.g. Chuine 2010; Polger & Primack 2011; Cook, Wolkovich & Parmesan 2012; Cooke, Eriksson & Junttila 2012) are hard to quantify on the basis of a single event like time of budbreak (Hunter & Lechowicz 1992; Hänninen 1995; Chuine, Cour & Rousseau 1998; Bailey & Harrington 2006; Allen *et al.* 2013). Interactions involving chilling followed by subsequent warming are even harder to model (Hänninen 1995; Hänninen, Slaney & Linder 2007). For example, Cook, Wolkovich & Parmesan (2012) hypothesized that warming has offsetting effects, delaying a chilling requirement to break dormancy while accelerating spring development. They hypothesize that if this interaction is general, the largest physiological effects of warming might be experienced by species that show the least phenological effect. However, the analysis did not find evidence of chilling requirements for 80% of species tested. A recent study that treats budbreak as a time-to-event process (as opposed to linear regression), also using degree-days as the predictor (Terres *et al.* 2013), could not converge on both threshold parameters, requiring instead a model-selection approach, that is with fixed parameter values. We begin by showing why aggregating temperature

variation into a single number (date of budbreak or flowering) precludes parameter estimation and, more generally, why such thresholds do not actually represent the onset of development.

A second limitation of current models concerns treatment of data. Phenological states are observed at intervals, but models interpret them as transitions, often as continuous variables (but see Terres *et al.* 2013; Allen *et al.* 2013 for binary treatment). Phenological observations at intervals of days to several weeks do not document when an individual reached a given state, only the interval between dates when it occurred. Thus, not only are observations *ordinal*, they are also *interval-censored*. Finally, we are unaware of models that allow for errors in the developmental process and in the observations.

Understanding how climate change will impact phenology could benefit from two innovations, (i) accommodation of the discrete, ordinal and censored nature of observed events and (ii) an underlying process that admits the continuous nature of phenology development. Modeling (unobserved) underlying development from discrete observations allows us to exploit the advantages of state-space representations within a hierarchical framework (Calder *et al.* 2003; Clark *et al.* 2011a). The model developed here, a continuous development model (CDM), allows us to directly quantify and test the effects of environment on spring phenology.

Firstly, we consider the role of timing, but unlike a previous analysis of spring development (Clark *et al.* 2014b), we focus on the relationship between spring warming vs chilling requirements. The importance of timing is expected to vary among species. Early successional species might risk early budbreak due to already unreliable conditions for colonizing species (Körner & Basler 2010), although effects of successional status are not obvious in the large study of Lopez *et al.* (2008). The capacity to advance budbreak could be limited in species having ring-porous xylem because leaf expansion follows vessel formation (Wang, Ives & Lechowicz 1992; Sass-Klaassen, Sabajo & den Ouden 2011). However, ring-porous *Quercus* and *Fraxinus* sometimes appear more responsive to warming than diffuse-porous species (Vitasse *et al.* 2009).

Secondly, we examine the role of latitude, from northern and southern ranges for eastern temperate tree species. Timing of warming and species attributes could interact in different ways at different latitudes (Cooke, Eriksson & Junttila 2012). Migration in response to climate change could be enhanced if northward-migrating species benefit from warming more than resident competitors (Ibáñez, Clark & Dietze 2009). Individuals in the northern part of their range might also suffer negative carbon balance in years with short growing seasons (Morin *et al.* 2009) and thus might risk early budburst when spring temperatures permit. Such close climate tracking puts the individual at increased risk of late frost, but that risk could be necessary when living near a physio-

logical limit. Finally, we evaluate the hypothesis that the largest effects of warming could be missed because they entail offsetting responses in fall and spring (Cook, Wolkovich & Parmesan 2012).

The application we describe is experimental. Observational studies have the advantage of comparatively low cost and, in some cases, long duration (Menzel *et al.* 2006; Polger & Primack 2011; Allen *et al.* 2013). Remote sensing can provide spatial perspective, but does not control for species composition; when year-to-year timing of budbreak varies with latitude and elevation, it may be unclear whether climate gradients or varying species abundance is the cause (Badeck *et al.* 2004; Hwang *et al.* 2011). More generally, observational data do not control for source population or for correlated climate variables (Diez *et al.* 2012; Clark *et al.* 2014b). They do not address the effects of warming outside the variation that occurred during the observation period. Experiments provide opportunity to control for population source and to isolate the effects of warming from other sources of variation. A recent report that experiments underestimate the effect of warming on phenology (Wolkovich *et al.* 2012) came from an analysis that applied different metrics to the two types of studies. We are unaware of evidence to indicate the experiments systematically underestimate warming effects (Clark *et al.* 2014a,b,c).

We use a continuous development model with experimental data to show why the thermal inefficiency assumption of degree-day models can lead to inconsistent results. A previous paper introduced this new approach, focused on why experiments and observational studies could be misinterpreted and showing that there is a period in late winter when many species are most sensitive to warming (Clark *et al.* 2014a,b,c). Here, we provide technical background on the model and use it to evaluate the relationship between warming effects and chilling requirements and how they may differ among species and geographically. To motivate this approach, we begin by showing why existing degree-day models have limited application to climate change. We find a relationship between responses to spring warming and winter chilling that differs from previous studies. Finally, we evaluate differences in development rates of northern vs southern species and northern vs southern seed sources of the same species, in xylem anatomy, and successional stage.

## Why thermal (degree-day) models for spring phenology are misinterpreted

### THE BASIC DEGREE-DAY MODEL

Thermal or degree-day models are used to quantify the advance in a phenological event, such as budbreak, depending on a temperature time series (Murray, Cannell & Smith 1989). Phenological development is assumed to proceed when temperatures exceed a threshold  $T'$ . In its simplest form the degree-day model is

$$DD_i = \sum_{t=d_0}^{d_i} (T_{i,t} - T') I(T_{i,t} > T') \quad \text{eqn 1}$$

where  $T_{i,t}$  is the mean temperature for individual  $i$  on day  $t$ , and  $I()$  is an indicator equal to 1 when its argument is true and 0 otherwise. Degree-day counting begins on day  $d_0$ , which can be an arbitrary date sometime after a chilling requirement is met (see below) or when the thermal threshold  $T'$  is first reached.  $DD_i$  is the degree-day or *thermal requirement* for budbreak, which occurs for individual  $i$  on day of the year (DOY)  $d_i$ . Two types of data are needed to evaluate  $DD_i$  for each of  $i = 1, \dots, n$  individuals. These are timing of budbreak  $\{d_i; i = 1, \dots, n\}$  and daily temperatures  $\{T_{i,t}\}$ . The sequence of temperatures does not influence  $DD_i$ , which depends only on the cumulative temperature above the threshold  $T'$ . More complex versions of eqn. 1 are used (examples include Richardson *et al.* 2006; Caffarra, Donnelly & Chuine 2011), but the fundamental assumptions we address below apply to all models of this basic form.

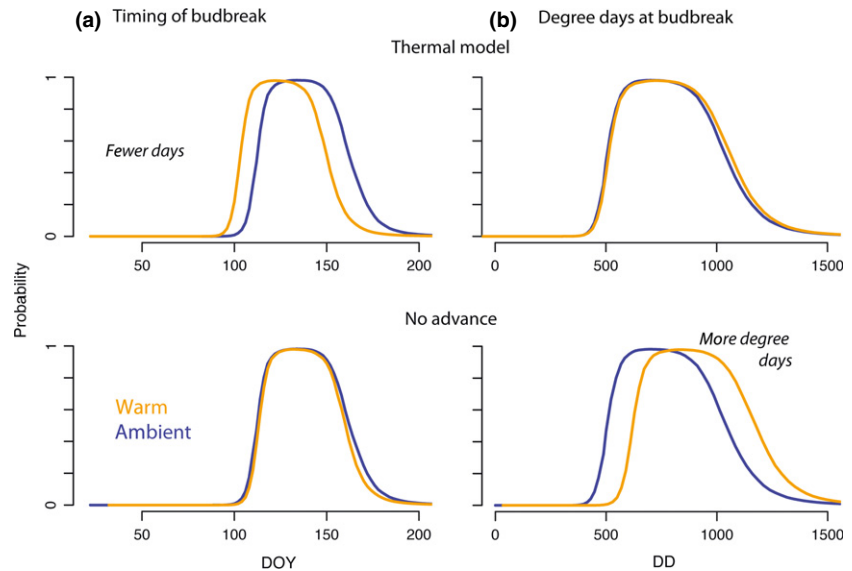
The additional variable most commonly included in thermal models is a chilling requirement

$$CU_i = \sum_{t=c_0}^{c_i} (T'_C - T_{i,t}) I(T_{i,t} < T'_C) \quad \text{eqn 2}$$

where number of days or cumulative daily temperatures below threshold  $T'_C$  are counted, beginning at a date in autumn, after onset of dormancy and continuing until an unobserved date  $c_i$ , when the chilling requirement  $CU_i$  is assumed to be satisfied. Some models accumulate warming DDs only after the chilling requirement is satisfied ( $d_0 \geq c_i$ ); others accumulate warming DDs and chilling units simultaneously ( $d_0 < c_i$ ) (Linkosalo, Häkkinen & Hänninen 2006). Because the date where such a transition might occur is unobserved, it is difficult to find a biological justification for either assumption (Cooke, Eriksson & Junttila 2012).

Degree-day models are increasingly used not only to explain geographic variation in onset of the growing season (Ibáñez *et al.* 2010; Yang *et al.* 2012; Terres *et al.* 2013) but also to quantify responses to climate variation (e.g. Menzel *et al.* 2006; Wolkovich *et al.* 2012; Polger & Primack 2011). If the degree-day model is accurate, then plots of budbreak dates against DD in different temperature regimes should be identical (Fig. 1b, top), despite differences in timing of budbreak (Fig. 1a, top). Conversely, phenology will not advance with warming if adaptive mechanisms against late frost damage cue to photoperiod oppose early budbreak (Fig. 1a, bottom).

If phenology responds to temperature in different ways at different phenological stages (Campoy, Ruiz & Egea 2011), then the responses could be more complex than the dichotomy represented by the hypotheses in Fig. 1. In other words, a specific value of  $DD_i$  does not discriminate between a warm winter followed by a cool spring, or vice versa. If phenological development varies in sensitivity



**Fig. 1.** Two predictions for the probability of being in a developmental state, such as budbreak, plotted against day of the year (a) and degree-days (b). Probability increases before the state is reached and declines thereafter. The thermal (degree-day) model (above) predicts advance with warming when plotted against day of year (a) and constant effect of degree-days (b). If warming does not advance phenology (lower panels), then warming results in budbreak at greater degree-days (b).

over time, we should not expect these two temperature sequences to result in budbreak at the same  $DD_i$  value. We test the hypothesis that warming accelerates phenology as predicted by the degree-day model, but first discuss why the degree-day model can miss important responses to climate change.

#### SOURCES OF CONFUSION

When temperature varies (in contrast to agronomic germination studies where temperature is held constant, e.g. Welbaum, Tissaoui & Bradford 1990; Bradford & Still 2004), the threshold  $T'$  is difficult to identify, because only the sequence of temperatures  $T_i$  is known in eqn. 1. Timing of onset  $d_0$  is not known. Including a cooling threshold temperature  $T'_C$  and unknown termination of cooling  $c_i$  worsens parameter identifiability problems (Terres *et al.* 2013; Clark *et al.* 2014a). Even so, estimates of chilling and warming thresholds are widely reported in the literature.

In fact, parameter identifiability can be viewed as only part of a larger problem, the fact that degree-day models do not reflect the biological assumptions that most biologists attribute to them. A motivation for the threshold temperature centres on conditions required to break dormancy, as low temperatures satisfy a chilling requirement and then rise above a threshold for development in winter/spring. Adaptive explanations focus on *delayed onset* (DO) of development, which could reduce risk of frost damage. In Appendix S1 (Supporting information), we show that degree-day models are in fact insensitive to events early in the year when temperatures are close to hypothesized thresholds. Instead, the dominant effect in the model is to change the effect of temperatures when they are above the threshold. A high threshold in the model reduces the effect of all temperatures above that threshold. On the other hand, a high threshold has little effect on the degree-days that are counted near the time when development is

assumed to begin, on day  $d_0$  in eqn. 1. A number of hypotheses have been suggested to explain why a threshold temperature might be adaptive for delaying onset of development (e.g. Campoy, Ruiz & Egea 2011, Cooke *et al.* 2012). However, we are unaware of biological justification for the assumption that is actually built into the DD model, that a high threshold for *onset* of development reduces the effect of temperatures on development *rate* after the threshold is passed.

#### Methods

In this section, we extend the approach of Clark *et al.* (2014a) that allows for responses to continuous weather variation, the continuous development model (CDM), followed by a description of experimental methods.

#### MODEL DEVELOPMENT

Consider continuous phenological development of an individual plant, initiated by fulfilment of a chilling requirement followed by warming temperatures in winter/spring, and that proceeds irregularly due to fluctuating daily temperatures, its current developmental state, and traits of the individual plant, such as genetic variation associated with seed source. While this development is continuous, it is observed only as recognizable discrete states at times when plants are observed. In fact, the state changes themselves are not discrete, nor is the precise timing of changes between them. Clark *et al.* (2014a) developed a hierarchical Bayes model to coherently integrate discrete observations of a continuous process.

Appendix S2 describes the three levels of our CDM, each having a different role. The first level consists of discrete, ordinal observations of phenological states,  $S_{iy,t}$ , with values 1–6 corresponding to no bud activity (1) through full leaf expansion (Norby, Hartz-Rubin & Verbrugge 2003). The observations are obtained on individual  $i$  in year  $y$  and day  $t$ . All variables have the same subscripts; so, we hereafter omit them. The second level consists of 'true' ordinal states  $s$ , which progress monotonically from 1 through 6 as a result of underlying continuous development  $h(t)$ , a continuous function of time  $t$  (in days). The separation of observation  $S$  from true state  $s$  allows for observation error, the fact



that an observer cannot always assign precisely the same discrete state to traits that vary continuously. The observed stages need not be monotonic, but the true states are monotonic. The continuous developmental state  $h(t)$  constitutes level three, which allows for the fact that development responds continuously to fluctuating temperatures and other factors in the environment and endogenous to the plant. Diagnostics are contained in Appendix S3.

To evaluate the effects of temperature on budbreak, we quantify its effects in combination with other variables in the model in Appendix S4. The quantity  $\gamma_k$  is a sensitivity coefficient describing the effect of a unit change in temperature on the rate of progress towards stage  $k = 1, \dots, 6$  (e.g. budbreak is stage  $k = 3$ ; Table 1). There is a value of  $\gamma_k$  for each individual and time. We use index  $\gamma$  to identify times of high sensitivity, that is when temperatures have a large effect on phenological development.

EXPERIMENTAL DESIGN AND DATA COLLECTION

Experimental warming was implemented near the southern and northern range limits of North America's eastern deciduous forest (Fig. 2). The southern site, Duke Forest (DF), North Carolina (36°0' N, 79°1' W; elevation 180 m), has a mean annual temperature of 14.5 °C and annual precipitation of 1208 mm, nearly all of which falls as rain. The northern site, Harvard Forest (HF), Massachusetts (42°5' N, 72°2' W; elevation 340 m), has a mean annual temperature of 7.5 °C with 1183 mm of precipitation. At Harvard Forest, snow cover is common from December to March.

The experiment used a factorial design to provide replication for all combinations of three temperatures (ambient, ambient +3 °C and ambient +5 °C assigned randomly) in a total of 18 open-top chambers per site. The 17-m<sup>2</sup> rectangular chambers were gridded into 14 columns and 30 rows at 15-cm spacing, resulting in 420 planting locations. Thus, each chamber supports two 4.6 m × 1.05 m heated areas. Chamber walls consist of transparent plastic greenhouse sheeting attached to wooden frames, with heights of 2.5 m. Six additional control plots had no chamber.

Soil and air were heated with independent systems to track ambient temperatures with the consistent +3 °C and +5 °C offsets. Soil was heated with electric resistance cable buried 10 cm deep at 20-cm spacing (Melillo *et al.* 2002). Temperature offsets were maintained by an automated tracking system. Ambient chambers received buried cables as a disturbance control, but they were not heated. Air was heated indirectly with propane. Heated water with non-toxic propylene glycol was pumped to a heat-exchange coil and airflow system in +3 and +5 chambers. Air circulating through the ambient chambers was not heated. The temperature offset was maintained by a defined heat delivery rate. Environmental variables monitored in each chamber included soil (10 cm deep) and air temperatures (30 cm above-ground). Control of soil temperature in the heated chambers was excellent ( $\pm$ ), whereas control of the air temperature was less precise, in part due to air scooping on

windy days. The combinations of chambers, sites, treatments and years provide substantial variation in temperatures.

Cohorts were established annually from seed, beginning in 2009, obtained from sites across eastern North America. For purposes of seed provenance analysis, they were classified as north or south of 40°N latitude (Table 2). Seeds from each source were planted in equal numbers at each site. Planting occurred at the times of seed dispersal for each species in the mineral soil horizon at grid locations. Each year existing cohorts were amended with new planting, replacing losses due to death and selective harvest of some individuals for analysis. Seedlings naturally recruited from the surrounding forest were also marked and monitored at both sites.

Weekly censuses enabled us to quantify germination, demographic and phenological responses to warming in each experimental year (2009–2012). The opening of buds and development of leaves in the spring were scored on a scale of 1 (no bud activity) to 6 (fully hardened and expanded leaves, Table 1) using the Norby Scale (Norby, Hartz-Rubin & Verbrugge 2003). The analysis omits the germination year, to avoid effects that could be related to the precise timing of planting and germination.

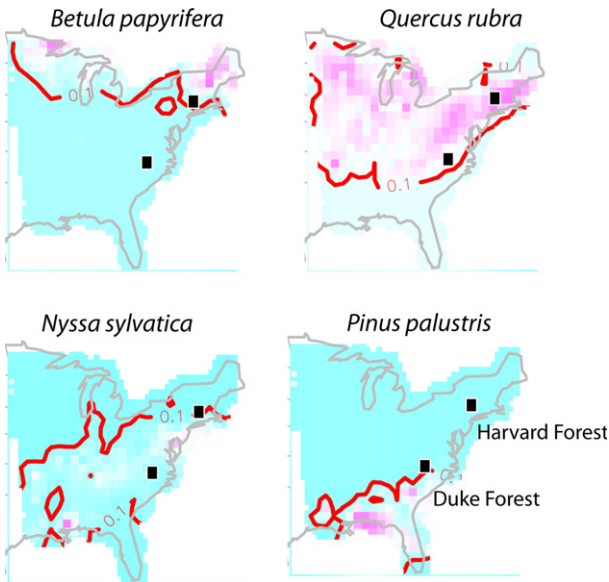
Species were selected to include those that are dominant at one or both sites and to span the range of xylem anatomy (ring porous, diffuse porous and conifer) and successional statuses (Table 2). No experimental manipulation of temperature can accommodate enough species to test for the many ways in which different functional traits control phenology. Our large study of > 4000 surviving seeds and seedlings includes species representative of the different functional classes in Table 2, but recognizes them as representative only.

ANALYSIS

The CDM was fitted jointly for all individuals of each species having sufficient sample size. Due to low germination rates, survival, or both, not all species could be analysed. Seed source was included in the model only for species having more than 10 surviving

**Table 1.** Six stages of budbreak used in this study, following Norby, Hartz-Rubin & Verbrugge 2003

Stage	Hardwoods	Conifers
1	No visible bud swelling	No bud expansion
2	Bud swelling	Stem elongating
3	Bud break has occurred	Needles have emerged from sheaths
4	Leaves unfolding	Needles partly elongated
5	Leaves open, not fully expanded	Needles mostly elongated but unhardened
6	Fully expanded	Needles hardened



**Fig. 2.** Study sites near northern and southern transitions for eastern deciduous forest, shown with basal area surfaces for representative species in the study, the smoothed 0.1 basal area contour in red (m<sup>2</sup> ha<sup>-1</sup>). Blue indicates absence. Data are from the USFS Forest Inventory and Analysis Program (Zhu *et al.* 2014).

individuals of both northern and southern seed sources (Table 2). The sample sizes in Table 2 refer to post-germination year individuals. A joint posterior distribution was simulated using Markov chain Monte Carlo (MCMC) (Appendices S3, S4). Development occurs on days above the threshold temperature. However, the threshold influences only when development can occur and not the magnitude of the effect – it has no contribution to thermal efficiency. Thus, unlike degree-day models, the precise value used did not affect results.

Diagnostics involve model checking through in-sample prediction. Model selection is based on the Deviance Information Criterion (DIC). We evaluate the model based on its ability to predict the entire seasonal development of phenology for each individual, starting from the initial state of each individual ( $s = 1$ ,  $h = 1$ ) in spring.

#### COMPARISON WITH TRADITIONAL DEGREE-DAY MODELS

To compare our results with a traditional degree-day model, we regressed the response for budbreak date against growing degree-days with chilling units of the preceding winter. The literature includes many different regression models – we followed the approach of Cook, Wolkovich & Parmesan (2012) as a recent example, using the DD sum for the first 3 months of the year. Unlike their study, we did not standardize DD in the regression (we did not divide DD by its standard deviation), which makes the response dependent on the range of values that happened to occur during a specific study and species. Because each species could experience a different range of temperatures, the scale for standardized coefficients is not comparable across species. Unstandardized coefficients have the common scale across species in days per °C. Our use of two geographically separated sites and temperature treatments ensured a large range of DD as inputs.

## Results

#### ENVIRONMENTAL VARIATION AND EXPERIMENTAL CONTROL

Field sites at Duke Forest and Harvard Forest maintained a 5–10 °C temperature difference, the difference being greatest in summer, both above- and below-ground (Fig. 3a,b). At the end of pretreatment year 2009, warmed chambers tracked ambient at +3 °C and +4 °C for air and +3 °C and +5 °C for soil. Air temperatures for Harvard return to ambient over winter when snow cover is present, and air is unheated (Fig. 3a). Winter temperatures varied substantially between years, with chilling greatest at Harvard Forest in 2010–11. We used the average of air and soil chilling units as inputs for the model.

#### INFERENCE

Inference yields estimates of discrete and continuous states,  $s$  and  $h$ , observation errors in discrete states, effects of predictors on development in parameter vector  $\beta$ , and parameters that connect discrete and continuous states (Appendix S2). For clarity, we begin with the dynamic context, summarizing relationships over time and between observations, followed by estimates and model diagnostics.

Figure 4 shows relationships between the latent developmental state  $h(t)$  for an individual tree compared with the observed discrete states  $S_k$ ,  $k = 1, \dots, 6$ . Development begins in state 1 (Fig. 4a). The continuous developmental

**Table 2.** Sample sizes and models having the lowest deviance information criterion for the CDM

Species	Xylem anatomy	Number of trees	Tree years	Number of populations	
				Southern	Northern
<i>T</i>					
<i>Magnolia grandifolia</i>	D	69	153	2	1
<i>T, C</i>					
<i>Betula papyrifera</i>	D	123	359	0	1
<i>Liquidambar styraciflua</i>	D	124	383	4	0
<i>Liriodendron tulipifera</i>	D	311	933	6	1
<i>Nyssa sylvatica</i>	D	148	339	5	3
<i>Pinus palustris</i>	C	40	105	3	0
<i>Pinus taeda</i>	C	283	838	3	0
<i>T, C, T × C</i>					
<i>Betula alleghaniensis</i>	D	344	950	0	3
<i>Pinus resinosa</i>	C	62	196	0	1
<i>Pinus strobus</i>	C	241	685	2	3
<i>T, C, S</i>					
<i>Quercus alba</i>	R	980	2659	5	2
<i>T, C, T × C, S, T × S</i>					
<i>Fraxinus americana</i>	R	161	321	4	3
<i>T, C, S, T × S</i>					
<i>Acer rubrum</i>	C	1649	4260	4	3
<i>Acer saccharum</i>	D	191	391	5	3
<i>Quercus rubra</i>	R	246	622	4	4

Xylem anatomy codes are D – diffuse porous, R – ring porous, C – conifer. Model codes are *T* – temperature, *C* – chilling units, *S* – seed source. Species are listed in order of model complexity selected by DIC.

state  $h$  is estimated with uncertainty (95% CIs are dashed lines in Figure Fig 4a), depending on observations  $S$  (dots).

Development is nonlinear in state and in time. The connection between the development state  $h$  and the discrete states  $s$  is shown in Fig. 4b and c. The scale in Fig. 4b expands and contracts in  $h$  space to accommodate the relationship between underlying development  $h$  and the discrete observations – for this individual stage 5 occupies a larger portion of the  $h$  scale than do stages 2–4. The differences between posterior probabilities (solid lines in Fig. 4b) and prior probabilities (dotted lines) are important, because they show that the prior has been updated by data. Development is nonlinear not only on the developmental scale (Fig. 4b), but also on the time scale (Fig. 4c).

Using DIC for model selection, we arrived at a preferred model for each species (Table 2). All models included an intercept and temperature. Selected models for all but *Magnolia* included chilling units. Seed source was selected for all species having multiple seed sources, with the exception of *Nyssa*. Interactions involving temperature, chilling and seed source were included in selected models for seven of the 15 species.

The model was evaluated by predicting the entire course of observations for all individuals starting from dormant state 1 through fully expanded state 6 (Fig. 5). This is a more rigorous standard than the one-step-ahead prediction often used to evaluate fitted time-series models. There is large variability, but most 68% (1 standard deviation of the mean) and all 95% predictive intervals spanned the 1:1 line. Nonetheless, models for some species tend to over-predict the first stage or under-predict the last stage.

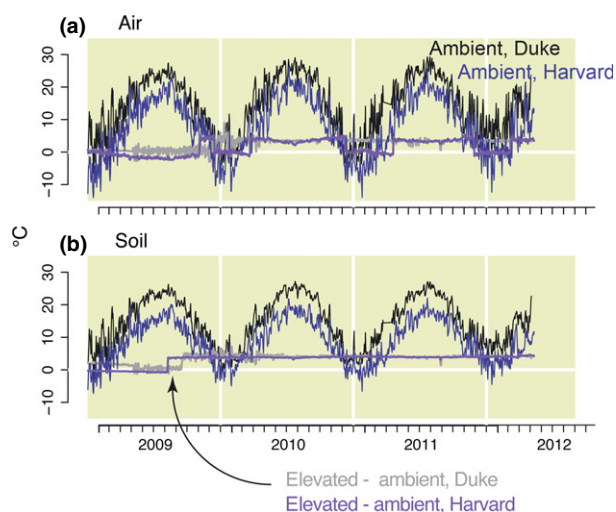
Parameter estimates were well resolved in the model. Effects of temperature (Fig. 6, 7a) and chilling units (Fig. 7a) differ from zero, with the exception of chilling

units for *Liquidambar*. Both spring temperatures and chilling units have positive effects on (accelerate) development. The signs of temperature effects are reversed in Figs 6 and 7a, because parameters in Fig. 7a are transformed to permit comparison with degree-day regression models (Appendices S1–S5). Positive chilling parameters in Fig. 7a are consistent with a chilling requirement – a positive relationship between winter chilling and spring development rate.

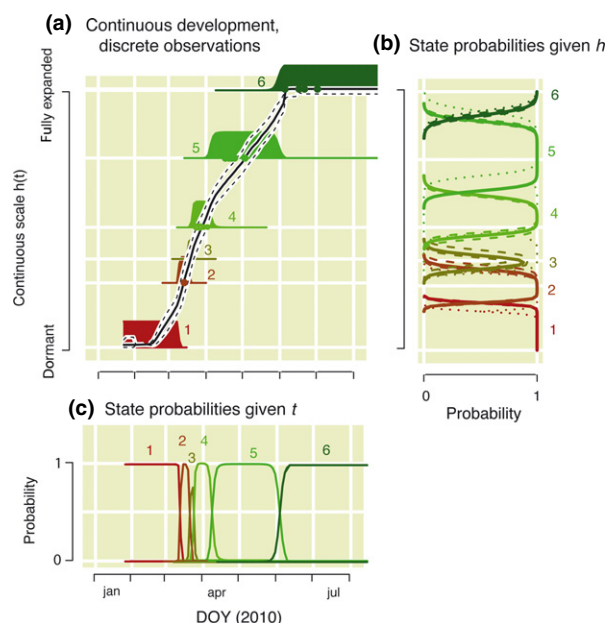
#### COMPARISON WITH DEGREE-DAY MODELS FOR WARMING AND CHILLING

To permit comparison of our CDM with traditional degree-day models, we transformed parameters in the CDM in terms of effects on days rather than development rate (the two have an inverse relationship – Appendix S5). The predictive interval includes parameter error, model error and individual variation.

The contrast between warming and chilling effects in the CDM vs the degree-day model is dramatic (Fig. 7). The CDM shows low correlation in predictive densities (parameter ellipses do not have a tendency to orient in a particular direction in Fig. 7a) as expected when there is an adequate distribution of data. Parameter estimates are well resolved, with clear warming and chilling effects, including differences between species. By contrast, the regression approach shows large correlation within and between parameter estimates (Fig. 7b).



**Fig. 3.** Daily air (a) and soil (b) temperature averaged over ambient and elevated chambers for Harvard and Duke Forest. Elevated temperatures are shown as departures from ambient.



**Fig. 4.** Relationship between continuous development  $h(t)$  and discrete states ( $k = 1, \dots, 6$ ) for a *Liriodendron* individual. (a) Observations (dots), inferred continuous state ( $h(t)$ , black lines, including 95% credible interval) and probabilities for each discrete state (shaded curves). (b) Posterior probabilities for discrete states evaluated in  $h(t)$  space (solid lines with 95% credible intervals), with prior means (dotted lines). (c) The same posterior probabilities shown in time  $t$  space. Both are highly nonlinear.

## HYPOTHESIZED EFFECTS OF GEOGRAPHY

Within a species, northern populations (unshaded densities in Fig. 6a) respond less to temperature than southern populations (shaded). In northern climates, populations break bud after their southern counterparts, but still achieve budbreak on fewer degree-days. In other words, plants originating from northern populations appear capable of completing development in a short interval within which fewer degree-days accumulate. Perhaps for this reason, seeds from northern populations tended to respond less to warming. The limited response of northern populations to warming is counter to the hypothesis that northern populations should be more responsive.

## HYPOTHESIZED SPECIES AND FUNCTIONAL TYPE DIFFERENCES

Phenology of *Nyssa sylvatica* most closely matched the prediction of DD models, that the advance in budbreak would be predicted by DD (Fig. 8). This is shown as the probability for state 3 (budbreak) plotted against DOY in the centre panels and against degree-days at right of Fig. 8. These are predictive mean probabilities for an individual exposed to the temperatures for an ambient and an elevated chamber, shown at left in blue and orange, respectively. Warming advances budbreak date (centre) to a degree predicted by the DD, indicated by the overlapping curves at right. *Pinus strobus* is typical of the remaining species, which show a large advance in timing (lower centre) that is not well predicted by DD (lower right).

Taken across all species, the ranking of response to warming or chilling was not related to xylem anatomy (Table 2). The ring-porous *Quercus alba*, *Q. rubra* and *Fraxinus* tended to be among the more responsive to warming, particularly southern populations of *Quercus*

(Fig. 6a), but intermediate in response to chilling (Fig. 6b). The conifers *Pinus palustris*, *P. resinosa*, *P. strobus* and *P. taeda* were intermediate in response to both.

## Discussion

One of the largest effects of warming on net primary production is expected to come from prolonged growing seasons, but warming will be uneven throughout the year, and counteracting effects of winter chilling vs spring warming frustrate prediction. Traditional degree-day models that are increasingly used to evaluate the impacts of climate change can misrepresent impacts of this variation. Degree-day models aggregate fluctuating temperatures in ways that hide or distort responses. They are essentially uninfluenced by onset of warming in spring, but assume instead a temperature threshold acting as a chronic penalty on development. The continuous development model (CDM) recasts development as a continuous response to fluctuating temperature informed by observations of discrete responses at discrete time intervals. In doing so, we circumvent the problem of explaining a discrete event like budbreak on the basis of climate aggregated over months. Clear species and population differences emerge from a model that incorporates realistic assumptions concerning observations and underlying process.

## WHY PHENOLOGY MODELS NEED TO INCORPORATE UNCERTAINTY

The effects of species, seed source, site, resources and weather variation over entire seasons combine to determine single events like flowering or budbreak. At a time, when global change scientists increasingly appreciate the heterogeneity of climate change, evidence for climate effects on phenology is increasingly influenced by

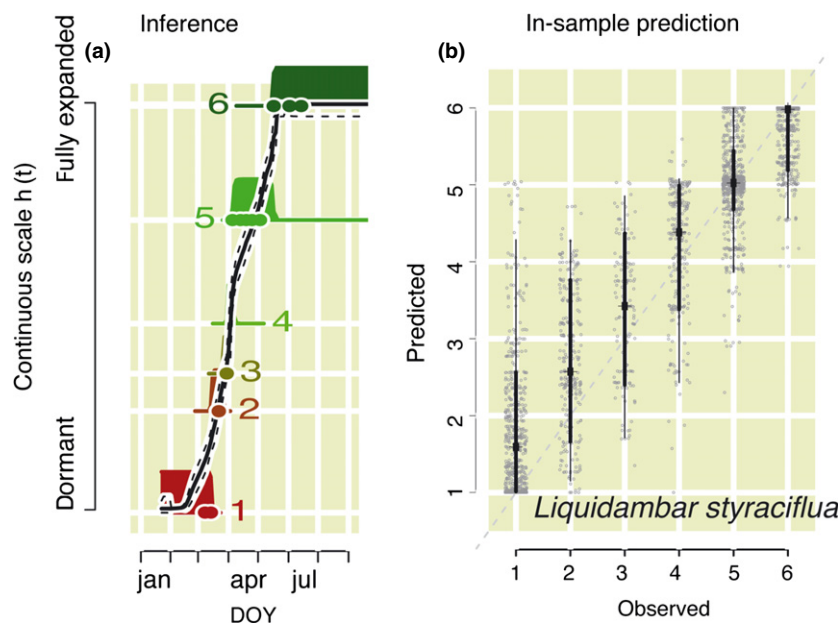


Fig. 5. Development of an individual (a) compared with predictive capacity for all *Liquidambar* trees (b), summarized as observations (large dots in (a) for the individual, small dots in (b) for all individuals), predictive means (cross)  $\pm 1$  standard deviation (thick bars in b) and 95% of predictions (thin bars in b).



meta-analyses of published results, where aggregation often precludes process-level modelling and uncertainty. Like survival analysis, phenology is especially prone to aggregation problems, because information is limited to time-to-event. Such analyses can be useful, provided the limitations are understood.

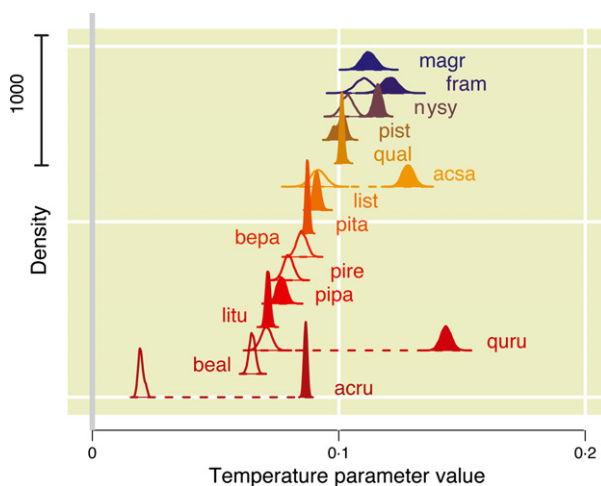
The discrete ordinal nature of data is likewise important. Increasing the number of states that are reported in data sets can benefit inference, for example six (this study), rather than two (see also Norby, Hartz-Rubin & Verb- rugge 2003). However, the more states, the more important it becomes to allow for observation error, because it becomes increasingly difficult to consistently assign classes when there are more of them. As with any such classifications, the six developmental states we recognize are not linearly related to temperature or to the underlying

continuous developmental rate (Fig. 4). There is no reason to expect a simple relationship in light of the fact that observed states are defined based on capacity to confidently identify them. The CDM finds the relationship between discrete observations in time and state that explains the changing development of large numbers of individuals. It benefits from large numbers of observations for temperature and response.

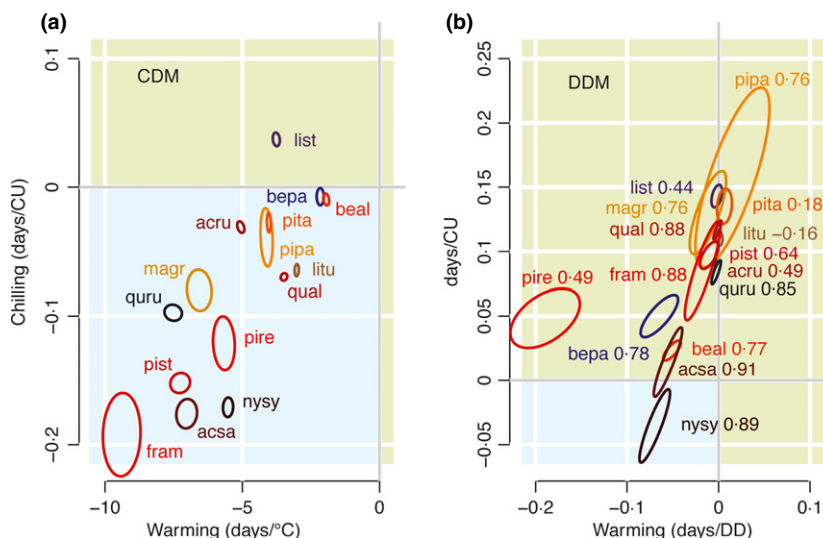
#### WARMING EFFECTS ON PHENOLOGY

The large differences between species in response to warming and chilling identified here (Figs 6 and 7) are not predicted by functional type differences hypothesized in the literature. Of course, even a large experimental study like this cannot include enough species of each functional type combination to fully evaluate effects of xylem anatomy and successional status. Our results should be taken as a first assessment where the full effects of continuously varying temperature are considered. However, it would not be correct to interpret lack of correlation with warming response to mean that traits like xylem anatomy do not matter. Xylem anatomy may indeed limit responses of ring-porous species, but they compensate in other ways. More experiments are needed to fully evaluate these effects. The importance of this study is not to question that functional relationships exist (they must); rather it questions the existence of simple correlations between traits and phenology.

Contrary to the hypothesis that northern populations should respond most, results show that southern populations show the largest responses (Fig. 6). Chuine (2010) and Morin *et al.* (2010) suggest that northern range limits could be determined by inability to complete fruit maturation rather than killing frost, while southern limits result from the inability to break dormancy due to a lack of chilling temperatures. In the northern part of the range, populations might be prone to taking greater risk during

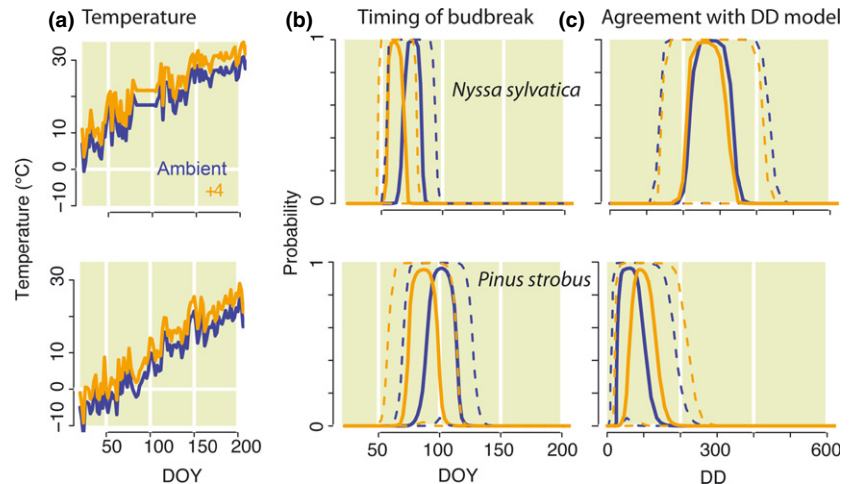


**Fig. 6.** Posterior densities for temperature parameters. Positive values accelerate development. Unshaded densities are for northern seed sources. Shaded are for southern seed sources or where seed source was not significant. The vertical density scale is shown at left.



**Fig. 7.** The relationship between the warming and chilling effects estimated from our continuous model (a) compared with a traditional regression against degree-days (b). Ellipses in (a) are 95% predictive intervals marginalized over all other parameters in the model (see text). Ellipses in (b) are 95% confidence intervals for the regression of budbreak date against DD and chilling units. Included in (b) are parameter correlations. Blue background indicates the case where chilling units and warming both advance budbreak, as expected from physiological studies.

**Fig. 8.** Temperatures from a specific ambient and elevate chamber (a) with predictive distributions (mean solid lines and dashed 95%) for probability of being in state 3 (budburst) from the continuous development model. *Nyssa* (above) is the only species closely predicted by the degree-day model – warming advances budbreak (b) and is unchanged with respect to degree-days (c). (compare with Fig. 1b) *Pinus strobus* shows an advance in days (b), but substantially less than predicted by the degree-day model (c). In other words, warming results in a larger degree-day requirement – it is not consistent with the DD model.



warm springs. Our results showing the opposite trend may be explained by the fact that northern populations already develop later in calendar days, but earlier in degree-days. If populations in the north must develop rapidly, they may be unable to accelerate development further.

#### OFFSETTING EFFECTS OF SPRING AND FALL WARMING

Cook, Wolkovich & Parmesan (2012) reported that vernalization sensitivity, which could result in a tendency for warming in fall/winter to delay spring phenology, was most common in the species that showed the least response to warming. The acceleration effects of a warm spring could be offset by the delay caused by reduced chilling in winter. If so, the species most sensitive to warming might be those showing no effect at all. In that study, a traditional degree-day calculation was applied to three-month windows of temperature data from fall through spring and compared with flowering times using regression. The analysis concentrates on two outcomes, a minority having a vernalization requirement, termed 'divergent', and a majority showing no such requirement. A vernalization response was only assigned to species where estimates had the correct sign, less than 20% of those tested. Thus, the non-responder category included the majority of species.

By applying both models to the same data, we demonstrate how the traditional degree-day model can be misleading. The expected responses are contained in the lower left (blue) quadrants in Fig. 7a,b, that is those for which both chilling units and spring warming advance phenology. Given that chilling requirements are known to operate in most species (e.g. Campoy, Ruiz & Egea 2011; Cooke *et al.* 2012), capacity to identify them is a basic test of phenology models. As in Cook, Wolkovich & Parmesan (2012) analysis, the degree-day model applied to our data fails this test by assigning to most species the wrong sign (Fig. 7b), suggesting the opposite of a chilling requirement. And most spring warming effects are not different from zero (horizontal axis in Fig. 7b). Both contradict known physiology.

But there is additional evidence in Fig. 7 that aids interpretation. The strong correlations in parameter estimates (many exceed 0.5 in Fig. 7b and evident in the 95% ellipses) and among species in Fig. 7b come from the correlations in winter and spring temperatures and, thus, in CU and DD. This results from the fact that warm winters tend to occur with warm springs. This is a standard feature of regression – parameter correlations are determined by the distribution of data. This correlation is unavoidable in the degree-day model, because temperature is aggregated over months. The loss of information and distortion that comes from aggregating an entire season of variation into a single number (Clark *et al.* 2011b, 2014a) is evident in the large number of non-significant warming effects in Fig. 7b (more than half of the warming effects straddle zero). Furthermore, the chilling effects mostly have the wrong sign – chilling is predicted to delay, not advance budbreak. In the CDM, correlations between fall and spring temperatures break down, because development responds to daily variation. Phenology is accelerated by both fall chilling and spring warming (Fig. 7a). Across broad taxonomic coverage (broadleaf deciduous, broadleaf and needleleaf evergreen, ring- and diffuse-porous, shade tolerant and intolerant) parameters are well resolved and have low correlation. All show negative warming coefficients, and only *Liquidambar* shows phenological delays due to chilling.

The distribution of data in regression explains the shapes of ellipses in the degree-day model, and the aggregation of temperatures explains how the order of species and signs of effects become distorted. Comparison of Fig. 7a and 7b further shows that the ordering of responses by the two methods differs substantially. If the degree-day method is detecting a positive correlation in responses to warming and chilling it appears not to be the correct one, and it is mostly predicting that fall chilling delays budbreak, contrary to a dormancy-breaking effect. Still, the CDM suggests a positive relationship between warming and chilling across species. This trend must be treated with caution. Although the CDM does not suffer

from the aggregation of temperatures for spring warming, it still aggregates chilling temperatures over a defined period (before 1 January). This was done because we cannot identify from data when chilling stops and warming begins. Even in the CDM, we cannot rule out the possibility that correlation structure in data affects estimates. From our analysis, it is clear there is a chilling requirement that differs between species, but the precise relationship to warming effects requires experimental decoupling of temperatures at different periods of time.

#### HOW TO ANTICIPATE UNEVEN WARMING

How important is the finding that traditional degree-day models can be misleading? We expect the largest problems to occur where the pattern of warming is uneven. If, as in some agricultural experiments, temperatures do not vary, and we simply increase temperatures the same amount at all times, then poor predictive performance could be attributed solely to variation in developmental sensitivity over time. It is possible that aggregating temperatures into a single cumulative value may not introduce large errors in such cases. However, if warming is uneven, there is an interaction between developmental state and seasonality of warming. The interaction could produce surprises at least as extreme as documented in the comparison in Fig. 7.

#### Conclusions

By accommodating the effects of temperature variation, we show why traditional degree-day models provide limited insight. Development may not be closely linked to mean or cumulative temperature, because it responds to fluctuations in ways that depend on the developmental state. There is variation in response within and among years that does not survive the data aggregation from daily fluctuations to seasonal sums. Given that future warming will vary by season, it is important to move beyond the assumption that all temperatures have the same impact, regardless of developmental state.

Species attributes such as xylem anatomy and successional status can influence phenological response without resulting in trait correlations. Such correlations might be 'significant' in a sample of hundreds rather than 15 species. Our finding that they are not apparent in our modest sample of species indicates that knowledge of traits like xylem anatomy will likely not be a good predictor of phenology response over the range of variation in this study.

The fact that southern populations respond most to temperature could be viewed as evidence that residents will resist northward migration of immigrants as climate warms. If potential invaders fail to exploit a prolonged growing season to the same degree as residents, then there is a resident advantage. This possibility, combined with the observation that tree migration has lagged well behind climate change (Zhu *et al.* 2012), suggests further study of the role of phenology in tree migration.

#### Acknowledgements

The project was funded by the Department of Energy. For comments on the manuscript, we thank Matthew Kwit, Bradley Tomasek and three anonymous reviewers. For field assistance, we thank Lindsay Scott, Becky Roper, Maria Terres and Vicki Woltz. The authors claim no conflicts and no interests.

#### Data accessibility

Data deposited in the Dryad repository: <http://doi.org/10.5061/dryad.38r08>.

#### References

- Aldrete, A., Mexal, J.G. & Burr, K.E. (2008) Seedling cold hardiness, bud set, and bud break in nine provenances of *Pinus greggii* Engelm. *Forest Ecology and Management*, **255**, 3672–3676.
- Allen, J.M., Terres, M.A., Katsuki, T., Iwamoto, K., Kobori, H., Higuchi, H. *et al.* (2013) Japanese cherry flowering phenology in a survival analysis framework. *Global Change Biology*, **20**, 1251–1263.
- Ashby, W.C., Bresnan, D.F., Roth, P.L., Preece, J.E. & Huetteman, C.A. (1992) Nursery establishment, phenology and growth of silver maple related to provenance. *Biomass and Bioenergy*, **3**, 1–7.
- Badeck, F.-W., Bondeau, A., Böttcher, K., Doktor, D., Lucht, W., Schaber, J. *et al.* (2004) Responses of spring phenology to climate change. *New Phytologist*, **162**, 295–309.
- Bailey, J.D. & Harrington, C.A. (2006) Temperature regulation of bud-burst phenology within and among years in a young Douglas-fir (*Pseudotsuga menziesii*) plantation in western Washington, USA. *Tree Physiology*, **26**, 421–430.
- Baldocchi, D.D., Black, T.A., Curtis, P.S., Falge, E., Fuentes, J.D., Granier, A. *et al.* (2005) Predicting the onset of net carbon uptake by deciduous forests with soil temperature and climate data: a synthesis of FLUXNET data. *International Journal of Biometeorology*, **49**, 377–387.
- Bradford, K.J. & Still, D.W. (2004) Applications of hydrotime analysis in seed testing. *Seed Technology*, **26**, 75–85.
- Caffarra, A., Donnelly, A. & Chuine, I. (2011) Modelling the timing of *Betula pubescens* budburst. II. Integrating complex effects of photoperiod into process-based models. *Climate Research*, **46**, 159–170.
- Calder, K., Lavine, M., Mueller, P. & Clark, J.S. (2003) Incorporating multiple sources of stochasticity in population dynamic models. *Ecology*, **84**, 1395–1402.
- Campoy, J.A., Ruiz, D. & Egea, J. (2011) Dormancy in temperate fruit trees in a global warming context: a review. *Scientia Horticulturae*, **130**, 357–372.
- Chuine, I. (2010) Why does phenology drive species distribution? *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 3149–3160.
- Chuine, I. & Beaubien, E.G. (2001) Phenology is a major determinant of tree species range. *Ecology Letters*, **4**, 500–510.
- Chuine, I., Cour, P. & Rousseau, D.D. (1998) Fitting models predicting dates of flowering of temperate-zone trees using simulated annealing. *Plant Cell and Environment*, **21**, 455–466.
- Churkina, G., Schimel, D., Braswell, B.H. & Xiao, X. (2005) Spatial analysis of growing season length control over net ecosystem exchange. *Global Change Biology*, **11**, 1777–1787.
- Clark, J.S., Agarwal, P., Bell, D.M., Flikkema, P., Gelfand, A.E., Nguyen, X. *et al.* (2011a) Inferential ecosystem models, from network data to prediction. *Ecological Applications*, **21**, 1523–1536.
- Clark, J.S., Bell, D.M., Hersh, M.H., Kwit, M., Moran, E., Salk, C. *et al.* (2011b) Individual-scale variation, species-scale differences: inference needed to understand diversity. *Ecology Letters*, **14**, 1273–1287.
- Clark, J.S., Salk, C., Melillo, J. & Mohan, J. (2014a) Data from: tree phenology responses to winter chilling, spring warming, at north and south range limits. *Dryad Digital Repository*, <http://doi.org/10.5061/dryad.38r08>.
- Clark, J.S., Melillo, J., Mohan, J. & Salk, C. (2014b) The seasonal timing of warming that controls onset of the growing season. *Global Change Biology*, **20**, 1136–1145.
- Clark, J.S., Gelfand, A.E., Woodall, C.W. & Zhu, K. (2014c) More than the sum of the parts: forest climate vulnerability from joint species distribution models. *Ecological Applications*, **24**, 990–999.



- Cleland, E., Chuine, I., Menzel, A., Mooney, H. & Schwartz, M. (2007) Shifting plant phenology in response to global change. *Trends in Ecology and Evolution*, **22**, 357–365.
- Cook, B.I., Wolkovich, E.M. & Parmesan, C. (2012) Divergent responses to spring and winter warming drive community level flowering trends. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 9000–9005.
- Cooke, J.E.K., Eriksson, M.E. & Junttila, O. (2012) The dynamic nature of bud dormancy in trees: environmental control and molecular mechanisms. *Plant, Cell and Environment*, **35**, 1707–1728.
- Davi, H., Dufrène, E., Francois, C., le Maire, G., Loustau, D., Bosc, A. *et al.* (2006) Sensitivity of water and carbon fluxes to climate changes from 1960 to 2100 in European forest ecosystems. *Agricultural and Forest Meteorology*, **141**, 35–56.
- Diez, J.M., Ibáñez, I., Miller-Rushing, A.J., Mazer, S.J., Crimmins, T.M., Crimmins, M.A. *et al.* (2012) Forecasting phenology: from species variability to community patterns. *Ecology Letters*, **15**, 545–553.
- Fitter, A.H. & Fitter, R.S.R. (2002) Rapid changes in flowering time in British plants. *Science*, **296**, 1689–1691.
- Goulden, M.L., Munger, J.W., Fan, S.-M., Daube, B.C. & Wofsky, S.C. (1996) Exchange of carbon dioxide by a deciduous forest: response to interannual climate variability. *Science*, **271**, 1576–1578.
- Hänninen, H. (1995) Effects of climatic change on trees from cool and temperate regions: an ecophysiological approach to modelling of bud burst phenology. *Canadian Journal of Botany*, **73**, 183–199.
- Hänninen, H., Slaney, M. & Linder, S. (2007) Dormancy release of Norway spruce under climatic warming: testing ecophysiological models of bud burst with a whole-tree chamber experiment. *Tree Physiology*, **27**, 291–300.
- Hunter, A.F. & Lechowicz, M.J. (1992) Predicting the timing of budburst in temperate trees. *Journal of Applied Ecology*, **29**, 597–604.
- Hwang, T., Song, C., Vose, J.M. & Band, L.E. (2011) Topography-mediated controls on local vegetation phenology estimated from MODIS vegetation index. *Landscape Ecology*, **26**, 541–556.
- Ibáñez, I., Clark, J.S. & Dietze, M. (2009) Estimating performance of potential migrant species. *Global Change Biology*, **15**, 1173–1188.
- Ibáñez, I., Primack, R.B., Miller-Rushing, A.J., Ellwood, E., Higuchi, H., Lee, S.D. *et al.* (2010) Forecasting phenology under global warming. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **365**, 3247–3260.
- Jeong, S.-J., Ho, C.-H., Gim, H.-J. & Brown, M.E. (2011) Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008. *Global Change Biology*, **17**, 2385–2399.
- Keeling, C.D., Chin, J.F.S. & Whorf, T.P. (1996) Increased activity of northern vegetation inferred from atmospheric CO<sub>2</sub> measurements. *Nature*, **382**, 146–149.
- Körner, K. & Basler, D. (2010) Phenology under global warming. *Science*, **327**, 1461–1462.
- Linkosalo, T., Häkkinen, R. & Hänninen, H. (2006) Models of the spring phenology of boreal and temperate trees: is there something missing? *Tree Physiology*, **26**, 1165–1172.
- Lopez, O.R., Farris-Lopez, K., Montgomery, R.A. & Givnish, T.J. (2008) Leaf phenology in relation to canopy closure in southern Appalachian trees. *American Journal of Botany*, **95**, 1395–1407.
- Melillo, J.M., Steudler, P.A., Aber, J.D., Newkirk, K., Lux, H., Bowles, F.P. *et al.* (2002) Soil warming and carbon-cycle feedbacks to the climate system. *Science*, **298**, 2173–2176.
- Menzel, A. & Fabian, P. (1999) Growing season extended in Europe. *Nature*, **397**, 659.
- 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.
- Mimura, M. & Aitken, S.N. (2007) Adaptive gradients and isolation-by-distance with postglacial migration in *Picea sitchensis*. *Heredity*, **99**, 224–232.
- Morin, X., Lechowicz, M.J., Augspurger, C., O'Keefe, J., Viner, D. & Chuine, I. (2009) Leaf phenology in 22 North American tree species during the 21st century. *Global Change Biology*, **15**, 961–975.
- Morin, X., Roy, J., Sonie, L. & Chuine, I. (2010) Changes in leaf phenology of three European oak species in response to experimental climate change. *New Phytologist*, **186**, 900–910.
- Murray, M.B., Cannell, M.G.R. & Smith, R.I. (1989) Date of budburst of fifteen tree species in Britain following climatic warming. *Journal of Applied Ecology*, **26**, 693–700.
- Nemani, R.R., Keeling, C.D., Hashimoto, H., Jolly, W.M., Piper, S.C., Tucker, C.J. *et al.* (2003) Climate-Driven Increases in Global Terrestrial Net Primary Production from 1982 to 1999. *Science*, **300**, 1560–1563.
- Norby, R.J., Hartz-Rubin, J.S. & Verbrugge, M.J. (2003) Phenological responses in maple to experimental atmospheric warming and CO<sub>2</sub> enrichment. *Global Change Biology*, **9**, 1792–1801.
- Parmesan, C. (2007) Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Global Change Biology*, **13**, 1860–1872.
- Polger, C.A. & Primack, R.B. (2011) Leaf-out phenology of temperate woody plants: from trees to ecosystems. *New Phytologist*, **191**, 926–941.
- Richardson, A.D., Bailey, A.S., Denny, E.G., Martin, C.W. & O'Keefe, J. (2006) Phenology of a northern hardwood forest canopy. *Global Change Biology*, **12**, 1174–1178.
- Richardson, A.D., Hollinger, D.Y., Dail, D.B., Lee, J.T., Munger, J.W. & O'Keefe, J. (2009) Influence of spring phenology on seasonal and annual carbon balance in two contrasting New England forests. *Tree Physiology*, **29**, 321–331.
- Sass-Klaassen, U., Sabajo, C.R. & den Ouden, J. (2011) Vessel formation in relation to leaf phenology in pedunculate oak and European ash. *Dendrochronologia*, **29**, 171–175.
- Schwartz, M.D. (1998) Green-wave phenology. *Nature*, **394**, 839–840.
- Terres, M.A., Gelfand, A.E., Allen, J.M. & Silander, J.A. (2013) Analyzing first flowering event data using survival models with space and time-varying covariates. *Environmetrics*, **24**, 317–331.
- Van Vliet, A.J.H., Wielgolaski, F.E., Zach, S. & Zust, A. (2006) European phenological response to climate change matches the warming pattern. *Global Change Biology*, **12**, 1969–1976.
- Vitasse, Y., Delzon, S., Dufrène, E., Pontailler, J.-Y., Louvet, J.-M., Kremer, A. *et al.* (2009) Leaf phenology sensitivity to temperature in European trees: do within-species populations exhibit similar responses? *Agricultural and Forest Meteorology*, **149**, 735–744.
- Wang, J., Ives, N.E. & Lechowicz, M.J. (1992) The relation of foliar phenology to xylem embolism in trees. *Functional Ecology*, **6**, 469–475.
- Wang, X., Piao, S., Ciais, P., Lic, J., Friedlingstein, P., Koven, C. *et al.* (2011) Spring temperature change and its implication in the change of vegetation growth in North America from 1982 to 2006. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 1240–1245.
- Wareing, P.F. (1956) Photoperiodism in woody plants. *Annual Review of Plant Physiology*, **7**, 191–214.
- Welbaum, G.E., Tissaoui, T. & Bradford, K.J. (1990) Water relations of seed development and germination in muskmelon (*Cucumis melo* L.): III. Sensitivity of germination to water potential and abscisic acid during development. *Plant Physiology*, **92**, 1029–1037.
- Wolkovich, E.M., Cook, B.I., Allen, J.M., Crimmins, T.M., Betancourt, J.L., Travers, S.E. *et al.* (2012) Warming experiments underpredict plant phenological responses to climate change. *Nature*, **485**, 494–497.
- Yang, X., Mustard, J.F., Tang, J. & Hong, X. (2012) Regional-scale phenology modelling based on meteorological records and remote sensing observations. *Journal of Geophysical Research*, **117**, G03029.
- Zhu, K., Woodall, C.W. & Clark, J.S. (2012) Failure to migrate: lack of tree range expansion in response to climate change. *Global Change Biology*, **18**, 1042–1052.
- Zhu, K., Woodall, C.W., Ghosh, S., Gelfand, A.E. & Clark, J.S. (2014) Dual impacts of climate change: forest migration and turnover through life history. *Global Change Biology*, **20**, 251–264.

Received 28 April 2014; accepted 10 June 2014  
Handling Editor: Niels Anten

## Supporting Information

Additional Supporting information may be found in the online version of this article:

- Appendix S1.** Thermal effects in the degree day model.
- Appendix S2.** The continuous development model of phenology.
- Appendix S3.** Diagnostics of the fitted model.
- Appendix S4.** From the development scale to discrete observations.
- Appendix S5.** From the development scale to event time.