Supplemental materials: Understanding growing degree days to predict spring phenology under climate change

Authors:

Catherine J. Chamberlain, ORCID: 0000-0001-5495-3219 [1,2] & E. M. Wolkovich [1,2,3]

Author affiliations:

[1]Arnold Arboretum of Harvard University, 1300 Centre Street, Boston, Massachusetts, USA 02131;

[2]Organismic & Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts, USA 02138;

[3]Forest & Conservation Sciences, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4

[*]Corresponding author: 248.953.0189; cchamberlain@g.harvard.edu

Methods S1: Using simulations to test Bayesian models in Rstan

To test all models, we used data simulations but simulation data can only take us so far in model interpretation. We first needed to be certain our models were operating correctly. To test our model function, we built test data and manipulated the effects of the parameters and sigmas around these parameters to make sure the models were reporting this information accurately. Building test data is easy and we encourage readers to use the Shiny App and GitHub repository to use as a building block.

Methods S2: Data analysis and model equations

Using Bayesian hierarchical models, we estimated the effects of site (i.e., 'forest' sites modeled as '0' versus 'urban' sites modeled as a '1'), method (i.e., hobo logger climate data modeled as '0' and weather station climate data modeled as '1') and the interaction between site and method effects as predictors with species modeled hierarchically as grouping factors:

$$y_i = \alpha_{species[i]} + \beta_{site_{species[i]}} X_{site} + \beta_{method_{species[i]}} X_{method} + \beta_{sitexmethod_{species[i]}} X_{sitexmethod} + \epsilon_i \tag{1}$$

$$\epsilon_i \sim N(0, \sigma_y)$$

The α and each of the five β coefficients are modeled at the species level, as follows:

$$\alpha_{species} \sim N(\mu_\alpha, \sigma_\alpha)$$
$$\beta_{site_{species}} \sim N(\mu_{site}, \sigma_{site})$$
$$\beta_{method_{species}} \sim N(\mu_{method}, \sigma_{method})$$
$$\beta_{sitexmethod_{species}} \sim N(\mu_{sitexmethod}, \sigma_{sitexmethod})$$

where i represents each unique observation, *species* is the species, α represents the intercept, β terms represent slope estimates, and y is the number of growing degree days.
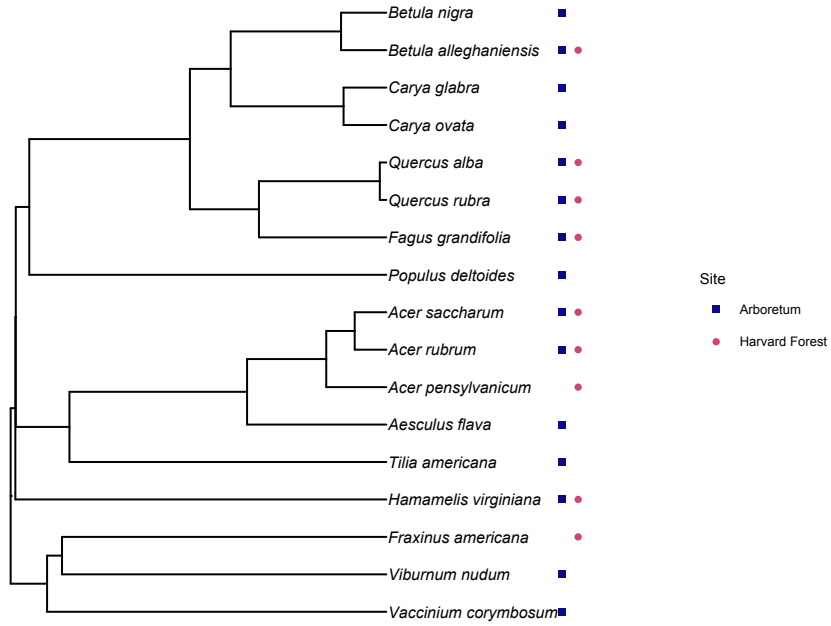
Supplemental tables and figures

Figure S1: Phylogeny indicating species across the two sites

Table S1: **Estimates from noisy weather station simulations**. We present posterior means, as well as 50% and 95% uncertainty intervals from models in which the predictors have been standardized so that they are comparable.

	mean	25%	75%	2.5%	97.5%
μ_α	307.45	303.99	310.89	297.13	318.19
μ_{site}	0.07	-0.50	0.66	-1.54	1.69
μ_{method}	4.93	2.35	7.60	-3.09	12.52
$\mu_{sitexmethod}$	-0.39	-1.16	0.36	-2.81	1.84
σ_{site}	0.68	0.26	0.98	0.02	1.93
σ_{method}	14.63	12.37	16.33	9.81	22.36
$\sigma_{sitexmethod}$	0.86	0.33	1.23	0.03	2.59
σ_α	21.79	18.43	24.37	14.47	33.19
σ_y	15.55	15.42	15.67	15.16	15.95
N_{sp}	15.00				

Table S2: **Estimates from noisy hobo logger simulations**. We present posterior means, as well as 50% and 95% uncertainty intervals from models in which the predictors have been standardized so that they are comparable.

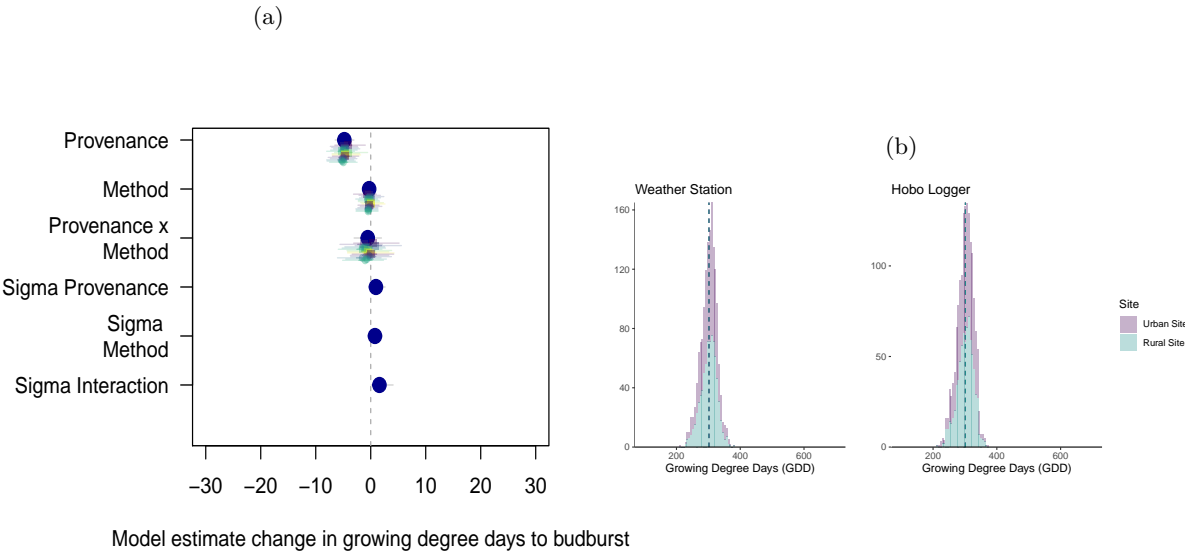|	mean	25%	75%	2.5%	97.5%
μ_α	316.41	311.01	321.54	299.27	333.71
μ_{site}	-1.20	-1.78	-0.62	-2.89	0.51
μ_{method}	-6.41	-9.15	-3.68	-14.64	1.86
$\mu_{sitexmethod}$	2.15	1.36	2.91	-0.10	4.47
σ_{site}	1.00	0.42	1.43	0.03	2.73
σ_{method}	15.14	12.73	17.02	10.11	23.17
$\sigma_{sitexmethod}$	1.20	0.48	1.70	0.04	3.31
σ_α	31.87	27.47	35.30	22.11	46.96
σ_y	15.05	14.91	15.19	14.67	15.46
N_{sp}	15.00				

(a)

(b)

Figure S2: Using simulated data to test provenance latitude, we show (a) the effects of provenance latitude and climate data method (weather station data as '1' or hobo logger data as '0') on simulated growing degree days (GDDs) until budburst. More positive values indicate more GDDs are required for budburst whereas more negative values suggest fewer GDDs are required. Dots and thin lines show means and 90% uncertainty intervals and thicker lines show 50% uncertainty intervals. See Table
reftab:prov for full model output. We also show (b) histograms of GDDs at the Arboretum and Harvard Forest using weather station data and hobo logger data.

Table S3: **Estimates from microclimate simulations**. We present posterior means, as well as 50% and 95% uncertainty intervals from models in which the predictors have been standardized so that they are comparable.

	mean	25%	75%	2.5%	97.5%
μ_α	322.94	318.50	327.45	310.33	335.58
μ_{site}	-0.32	-0.91	0.30	-2.11	1.42
μ_{method}	-6.78	-8.07	-5.50	-10.55	-3.07
$\mu_{sitexmethod}$	1.05	0.26	1.85	-1.33	3.38
σ_{site}	1.14	0.47	1.65	0.06	3.12
σ_{method}	6.82	5.63	7.72	4.26	10.92
$\sigma_{sitexmethod}$	1.23	0.46	1.77	0.04	3.60
σ_α	24.85	21.32	27.48	17.17	36.72
σ_y	16.73	16.59	16.87	16.32	17.14
N_{sp}	15.00				

Table S4: **Estimates from urban simulations**. We present posterior means, as well as 50% and 95% uncertainty intervals from models in which the predictors have been standardized so that they are comparable.

	mean	25%	75%	2.5%	97.5%
μ_α	307.37	303.97	310.87	296.13	317.88
μ_{site}	-19.23	-19.81	-18.64	-20.98	-17.50
μ_{method}	0.27	-0.31	0.90	-1.47	1.92
$\mu_{sitexmethod}$	-0.41	-1.21	0.36	-2.59	1.87
σ_{site}	1.15	0.47	1.64	0.05	3.16
σ_{method}	1.03	0.46	1.47	0.05	2.65
$\sigma_{sitexmethod}$	0.77	0.30	1.10	0.02	2.33
σ_α	21.55	18.47	23.96	14.96	31.33
σ_y	15.56	15.42	15.70	15.17	15.97
N_{sp}	15.00				

Table S5: **Estimates from real data**. We present posterior means, as well as 50% and 95% uncertainty intervals from models in which the predictors have been standardized so that they are comparable.

	mean	25%	75%	2.5%	97.5%
μ_α	424.96	413.29	437.11	387.75	458.74
μ_{site}	-31.66	-39.63	-23.74	-54.08	-9.28
μ_{method}	-0.50	-7.29	6.22	-19.96	20.45
$\mu_{sitexmethod}$	-40.35	-48.17	-32.56	-63.08	-17.80
σ_{site}	21.46	11.46	29.44	1.25	49.34
σ_{method}	17.09	7.80	24.21	0.88	43.70
$\sigma_{sitexmethod}$	25.62	16.26	33.80	2.66	52.91
σ_α	62.95	53.51	71.16	40.91	93.12
σ_y	71.54	69.73	73.28	66.63	76.75
N_{sp}	18.00				

Table S6: **Estimates from an urban effect and noisy weather station simulations.** We present posterior means, as well as 50% and 95% uncertainty intervals from models in which the predictors have been standardized so that they are comparable.

	mean	25%	75%	2.5%	97.5%
μ_α	306.04	302.37	309.82	294.54	316.85
μ_{site}	-17.13	-20.24	-14.11	-26.63	-7.89
μ_{method}	5.65	3.45	7.87	-1.17	12.40
$\mu_{sitexmethod}$	-13.47	-14.65	-12.27	-16.97	-10.01
σ_{site}	18.38	15.64	20.55	12.45	27.44
σ_{method}	12.16	10.23	13.76	7.95	18.51
$\sigma_{sitexmethod}$	3.69	2.06	5.07	0.31	8.34
σ_α	21.03	17.96	23.41	14.71	31.42
σ_y	20.36	20.18	20.54	19.86	20.85
N_{sp}	15.00				

Table S7: **Estimates from provenance latitude simulations.** We present posterior means, as well as 50% and 95% uncertainty intervals from models in which the predictors have been standardized so that they are comparable.

	mean	25%	75%	2.5%	97.5%
μ_α	307.61	304.40	310.79	296.64	318.14
μ_{site}	-4.80	-5.38	-4.22	-6.47	-3.05
μ_{method}	-0.29	-0.69	0.11	-1.50	0.90
$\mu_{sitexmethod}$	-0.54	-1.41	0.32	-3.03	1.96
σ_{site}	0.94	0.39	1.35	0.04	2.61
σ_{method}	0.77	0.33	1.09	0.03	2.13
$\sigma_{sitexmethod}$	1.58	0.72	2.23	0.08	4.08
σ_α	20.54	17.57	22.71	14.07	30.22
σ_y	15.57	15.43	15.71	15.19	15.96
N_{sp}	15.00				