1

Time to branch out? Application of hierarchical survival models in plant phenology

3

4   Sarah C. Elmendorf [1,2,*], Theresa M. Crimmins [3,4], Katharine L. Gerst [3,4], and Jake F. Weltzin [3,5]

5

6   [1] Institute for Arctic and Alpine Research (INSTAAR), University of Colorado, Boulder, CO, USA

7   [2] Ecology and Evolutionary Biology Department, University of Colorado, Boulder, CO, USA

8   [3] National Coordinating Office, USA National Phenology Network, Tucson, AZ, USA

9   [4] School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, USA

10   [5] U.S. Geological Survey, Ft. Collins, CO, USA

11

12   *Corresponding author: Sarah Elmendorf (sarah.elmendorf@colorado.edu)

13

18

19

20   **DECLARATIONS OF INTEREST:** none

21

22

23

ABSTRACT

The sensitivity of phenology to environmental drivers can vary across geography and species. As such, models developed to predict phenology are typically site- or taxon-specific. Generation of site- and taxon-specific models is limited by the intensive in-situ phenological monitoring effort required to generate sufficient data to parameterize each model. Where in-situ phenological observations exist, the data are often subject to analytical issues due the limited duration of any individual monitoring program, spotty site- and species- level coverage, lack of standardized methodology, and infrequent or variable census intervals. Together, these characteristics constrain our ability to make phenological inferences outside of select sites and taxa where long-duration, intensive monitoring has occurred.

In this study, we leveraged two national, standardized phenology datasets to develop a multi-species and multi-site state-space survival model of the onset of deciduous tree and shrub spring (leaf out) and fall (leaf-color) events across temperate ecoregions of the United States. We used data from two national-scale phenological databases, a 9-year, broadly distributed dataset from the USA National Phenology Network and a 4-year dataset from the National Ecological Observatory Network, to quantify regional and interspecific variation in sensitivity to environmental drivers for both spring and fall leaf phenophases. Spring leaf out was generally promoted by longer days, spring growing degree day accumulation, overwinter chilling, and was suppressed by frost events, whereas fall leaf color was promoted by shorter days and cold accumulation. The sensitivity to most environmental drivers tended to be more variable among species than among the regions as defined here (EPA ecoregions of North America, excluding desert and tropical areas). The results of this study lay the groundwork for incorporating the growing collection of phenological observations into a generalized framework for predicting the transition states for any species, in any location.

50

51

52 **KEYWORDS:** citizen science, phenology, temperature, day length, USA-NPN, NEON

53

54 **HIGHLIGHTS:**

55 • Hierarchical survival models of plant phenology were developed

56 • Citizen science and national-scale monitoring phenology data were analyzed

57 • Chilling, spring warming, longer days, and less frost promote spring leaf out

58 • Cold accumulation and shorter days promote fall leaf coloration.

59

60 **1. INTRODUCTION**

61

62 The ability to predict the timing of life cycle events for individual species at given locations is

63 beneficial to many aspects of resource and species management (e.g. optimally timing of land

64 prescribed burns or chemical treatments; (Enquist et al., 2014)) and human health (e.g.

65 forecasting the onset or peak of airborne pollen in allergenic taxa (Siljamo et al., 2013).

66 Understanding the drivers to phenology across species also improves our ability to anticipate

67 future changes to ecosystem services as climate conditions change, due the impact of plant

68 phenology on carbon dynamics (Richardson et al., 2013), hydrology (Morisette et al., 2009), and

69 albedo (Williamson et al., 2016).

70

71 In recent years, many predictive models have been developed to elucidate the drivers of plant

72 phenology, ranging from empirical approaches such as regression (Hufft et al., 2018) to more

73 complex process-based models that attempt to explain underlying physiological responses

74 (Basler, 2016). Model construction is frequently data-limited; phenological datasets with both

75   broad taxonomic and geographic coverage that are consistently sampled over decades do not

76   exist. Accordingly, phenology modeling efforts typically fall into one of a few categories: those

77   that analyze one or more species at a single site or small region (Bertelsen's data from the

78   Santa Catalina mountains; the Marshall phenology record from central England, or Thoreau's

79   Concord records (Crimmins et al., 2010; Miller-Rushing and Primack, 2008; Roberts et al.

80   2015), those that combine phenological observations across different species, regions, or

81   sampling to describe general patterns using meta-analysis type approaches (e.g Cook et al.,

82   2012; Oberbauer et al., 2013), and those that use space-for-time substitution to parameterize

83   models for single species using a relatively short time-series across multiple sites (Liang and

84   Schwartz, 2014; Melaas et al., 2016; Taylor et al., 2019). Relatively simple phenology models

85   often show equivalent predictive performance with more complex models. However, with all

86   models, transferability of fitted parameters among sites is imperfect (Basler, 2016).

87

88   An additional challenge in modeling phenology is that observations often are conducted

89   infrequently because of the effort required in regular censusing, introducing uncertainty into the

90   timing of when an event occurred. Because of this, phenological models have typically been

91   developed separately for the few species or handful of sites with a rich phenological record,

92   which excludes the remaining "long-tail" of phenological observations: those collected over a

93   short duration; those that have few replicates for any species or location; and those that have

94   lower sampling frequency. For phenological studies, we are fortunate that many of these "long-

95   tail" datasets have been collated and publicly archived, making such observations readily

96   available for inclusion in analyses. What remains to be developed is an analytical approach that

97   leverages these data, which can broaden the diversity of species, sites, and time-frames

98   represented, and can enable national-scale inference.

99

100   Survival analysis approaches are a promising solution for capitalizing on the long tail of unused

101   phenological data. Survival models are a class of statistical models that are well-suited to

102   accommodating the typical shortcomings of phenology datasets and that have the potential to

103   outperform traditional regression models in predicting phenological events (Templ et al., 2017).

104   Survival analysis has a long history in engineering studies of failure time, as well as in the

105   medical literature in estimating time to death or disease recurrence where observations are

106   right-censored (e.g. Cox 1972). The analogy to plant monitoring data is clear: Like patients,

107   individual plants may be monitored at long or irregular intervals and in some cases the event of

108   interest occurs after the monitoring period is over. Only a handful of studies have used survival

109   approaches to model climate drivers of plant phenology, despite its potential (e.g. Allen et al.,

110   2015; Diez et al., 2014; Templ et al., 2017, Terres et al., 2013).

111

112   We additionally considered the best ways of including data from multiple species and multiple

113   sites. Interspecific variation in phenology dates can be substantial, commonly on the order of

114   several weeks (Lechowicz, 1984). Further, individual species can show different sensitivity to

115   environmental drivers across locations (Ma et al., 2018; Parker et al., 2017; Prevéy et al., 2017;

116   Way and Montgomery, 2015), but see MacKenzie et al. (2018) for a counterexample. However,

117   despite clear differences among species in the dates of phenological transitions, experimental

118   work has shown the phenology of a diversity of deciduous taxa responds similarly to a common

119   set of environmental cues (Flynn and Wolkovich, 2018) and many successful models have been

120   developed by aggregating phenological data using space-for-time substitution. The potential for

121   heterogeneity in sensitivity to environmental drivers among species and regions argues for

122   phenological forcing models that specifically allow for this variability. At the other extreme,

123   modeling each species and region entirely separately dramatically reduces sample sizes.

124   Therefore, to the degree which cues are common across species and regions, these common

125   characteristics should be leveraged to improve models for all species and regions.

126

127   Using hierarchical models, it is possible to incorporate variation among species and regions as

128   random effects (Clark 2007, Bolker 2008). In this framework, the estimated sensitivity of any

129   individual species or region is influenced by the sensitivities of other taxa and locations,

130   because inferences on species- and region-specific coefficients "borrow strength" from species

131   and regions. Using a multi-species, multi-region approach, the modeled sensitivity of any

132   individual species or region reflects both the sensitivity that would be estimated analyzing each

133   species or region separately and the average sensitivity over all species and regions (Gelman

134   and Hill, 2007). Thus, integrating survival models into a hierarchical framework further broadens

135   the ability to incorporate additional datasets into the analysis by accommodating observations

136   made on different species and regions.

137

138   New, geographically extensive and taxonomically diverse in-situ phenological datasets collected

139   over broad geographic regions broadens possibilities for developing and improving phenological

140   models. The USA National Phenology Network (USA-NPN) and the National Ecological

141   Observatory Network (NEON) are national-scale programs with a primary (USA-NPN) or

142   substantial (NEON) focus on collecting phenological data using concordant, standardized

143   protocols that have substantially increased the availability of in-situ phenological observations

144   across the United States over the past decade (Denny et al., 2014; Elmendorf et al., 2016). The

145   USA-NPN dataset has been used for numerous applications to date, including the identification

146   of climate drivers of phenological transitions for individual species (e.g. Piao et al. 2015, Mazer

147   et al. 2015, Park and Mazer 2018; Melaas et al., 2016; Taylor et al. 2019), understanding spatial

148   and temporal phenological trends (e.g. Pearse et al. 2017, Liang 2019), and evaluation of

149   remotely sensed data (e.g. Peng et al., 2017; Zhang et al., 2018). However, none to date have

150   leveraged either the USA-NPN nor the NEON datasets to create hierarchical multi-species

151   models across broad geographic regions.

152

153   In this study, we outline a novel analytical approach for analyzing this growing collection of

154   phenology observations across a broad array of ecosystems and species. The objective of our

155   work was to develop hierarchical multi-region, multi-species survival models of the onset of

156   spring (leaf out) and fall (leaf coloration) for over 100 plant species within 15 temperature

157   ecoregions regions of the U.S. The resulting models demonstrate common drivers of deciduous

158   spring and fall phenology across species in temperate regions of the U.S. and provide estimates

159   of sensitivity for each species and region.

160

161   **2. MATERIALS and METHODS**

162

163   **2.1 Model overview**

164   Survival analysis is a modeling approach robust to the challenges that phenology data pose.

165   First, phenology data are essentially time-to-event data, and therefore show different statistical

166   properties than other typical ecological parameters such as size, counts, or concentration.

167   Second, the drivers of phenological transitions vary temporally so that including "average"

168   weather (e.g., April mean daily temperature) as a phenological predictor misrepresents the

169   biology of the system. For example, recent weather events are likely to have an outsized

170   influence on the near-term rate of phenological progression, and weather after a phenophase

171   transition obviously cannot have any effect on the date of the (past) event. Third, in the absence

172   of continuous monitoring, the date on which phenological transitions occur is typically reported

173   with varying degrees of uncertainty, depending on the frequency of observation. Phenological

174   datasets – especially those collected opportunistically or infrequently by volunteers -- can suffer

175   from both relatively long periods with no observations as well as variability in temporal

176   frequency of observations. Survival analysis, which can accommodate temporally varying

7

177   forcing variables and any intercensus interval, is well-suited to data with these characteristics

178   (Templ. et al., 2017).

179

180   Here, we implement survival modeling using discrete time models where the time-step is a

181   single day, which is typically the smallest intercensus sampling interval for plant phenology.

182   Borrowing from the nest-survival literature (Royle and Dorazio, 2008), we categorize a single

183   plant's spring (or fall) phenology as a state change from one phase (e.g., without leaves) to

184   another (e.g., with leaves). For the purpose of model development, we classify the states as a

185   binary variable that starts in state '1' and has a modeled probability of transitioning to state '0' on

186   each day of year. Mathematically, we can consider this daily probability of changing state as a

187   daily 'hazard' rate, or the complement of the daily 'survival' rate φ_d, which is the probability of

188   staying in the same (pre-transition) phenophase state on any date.

189

190   Thus, on each day (*d*), for each individual (*i*),

191   *Eq (1):*

192     $P(State_{i,d} = 1 | State_{i,d-1} = 1) = \varphi_{id}$

193

194   Once transitioned, the probability of remaining in the terminal state is 100%, i.e.

195   *Eq (2):*

196     $P(State_{i,d} = 0 | State_{i,d-1} = 0) = 1.$

197

198   After transitioning to the (seasonally) terminal state ('0'), reversals are not possible in this

199   framework (i.e., this is a single-season model). The influence of factors that vary at the level of

200   an individual (e.g. geolocation, species, or traits) and/or time period within individual (e.g. daily

201 weather at an individual's location) on phenological timing can be modeled by adding these

202 factors as predictors of the daily survival rates.

203

204 We combined data across multiple taxa and regions into a single hierarchical, multi-species,

205 multi-region space-state model separately for spring and for fall phenological events. We used a

206 state-space modeling framework where the daily probability of transitioning from the initial state

207 (leafless for the spring model; lacking colored leaves for the fall model) to the terminal state

208 (with leaves for the spring model; with colored leaves for the fall model) is a function of

209 environmental drivers that occurred in the time prior to each event. The sensitivities of each

210 species and region to each environmental driver were modeled as random effects. Further

211 details on the statistical modeling approach are provided in Appendix A.

212

213 **2.2 Environmental drivers**

214 To assemble a suite of candidate environmental drivers of spring leaf out and fall leaf color to

215 incorporate into multispecies, multi-region models, we surveyed the literature to prioritize the

216 putative drivers (Appendix B). Based on that analysis, we selected growing degree days (GDD),

217 daylength, chilling, precipitation and frost as environmental forcing variables for the spring leaf

218 model. Environmental forcing variables for the fall leaf color model included cold degree days

219 (CDD), precipitation and daylength. Climate data used to derive these environmental forcing

220 variables were extracted from Daymet (Thornton et al., 2018).

221 Environmental drivers used as covariates were calculated using an accumulation window

222 approach, reflecting cumulative and dynamic effects on organismal biology. Accumulation

223 windows can be calculated over a moving window period (e.g. the sum of heat over an n-day

224 moving window) or with a fixed start date (e.g. the sum of heat since Jan 1). While the latter is

225 more common, the calendar date selected (usually Jan 1) is fairly arbitrary and a wide range of

9

226   parameterizations have been used successfully. After considering both the geographic spread

227   of our data and parameterizations used in previous studies, we included as environmental

228   drivers the sum of frost events, cumulative precipitation, and cold degree days (CDD, base 20)

229   over a moving window 14 days to one day prior to each modeled timepoint, the sum of chilling

230   hours using the Utah model (Richardson et al., 1974), over a moving window 135 days to 45

231   days prior to each modeled timepoint, and growing degree days (GDD, base 0) over a moving

232   window 45 days to one day prior to each modeled timepoint and (Diez et al., 2014; Gill et al.,

233   2015; Gunderson et al., 2012), and the prior day's daylength (Appendix C). These windows

234   generally created forcing variables for growing degree days, cold degree days, frost, and chill

235   accumulation that were changing unidirectionally over the spring and fall periods. All covariates

236   were scaled to mean 0, unit variance prior to their inclusion in the models to facilitate model

237   interpretation and convergence.

238

239   **2.3 Filtering and pre-processing of phenological data**

240   We used publicly available datasets from the USA-NPN (USA National Phenology Network, 2018)

241   and NEON (National Ecological Observatory Network, 2018) to predict the onset of leaf out in

242   spring and leaf coloration in fall for deciduous tree and shrub taxa over the temperate regions of

243   the US. The USA-NPN developed standardized monitoring methods for plants and animal taxa

244   that include the assessment of the phenological status (presence or absence) of carefully

245   described, recognizable phenophases (Denny et al. 2014). Citizen science monitoring launched

246   in 2009 and is conducted by tens of thousands of professional scientists and volunteers. NEON

247   is a national-scale ecological observatory which conducts standardized measurements of

248   multiple ecological parameters in 47 terrestrial and 34 aquatics sites across the US. The

249   program has just entered its operational phase, but plant phenology measurements at some

250   sites are available from as early as 2014. USA-NPN and NEON monitoring protocols use

251   identical phenophase definitions and are status-based, such that the raw data on each day of

252   year is an observation of whether and individual is in a particular phenophase (i.e., has leaves)

253   or not (has no leaves) (Denny et al. 2014). At each NEON terrestrial site, plant phenology is at

254   regular intervals by trained technicians. We accessed all available data using the USA-NPN and

255   NEON APIs (application programming interface) and retained observations for species denoted

256   as "Deciduous Trees and Shrubs" by the USA-NPN.

257

258   Response variables in the space-state model include, for each monitored individual, in each

259   year, the latest day of year in which an individual was observed to be in the initial state (i.e,

260   leafless for the spring leaf model; having green leaves for the fall color model) and (optionally)

261   the earliest day of year in which an individual was observed to be in the terminal state (with the

262   phenophase "leaves" for the spring leaf model, and having the phenophase "colored leaves" for

263   the fall color model). We converted the raw status observations into 'last observed in initial state

264   (last no)' values and 'first observed in terminal state (first yes)' values based on the annual time-

265   series (Appendix D). In some cases, absence of the phenophase was not recorded (i.e., no 'last

266   no' values could be extracted from the observations). For these individuals, we assumed the

267   individuals lacked leaves in the middle of winter (Day of Year [DOY]=1), and had no colored

268   leaves in the middle of summer (DOY = 180). Because the survival analysis approach we took

269   can accommodate any inter-census interval, we did not further filter observations to those with a

270   particular duration between the 'last observed no' and 'first observed yes' values, as is typically

271   necessary for working with phenology data that include some observations with large inter-

272   census intervals.

273

274   We assigned each observation based on its location to an EPA level II ecoregion (United States

275   Environmental Protection Agency, 2018, discarding observations that fell into non-temperate

276   (i.e., Mediterranean, tropical and desert-type) regions. In total, we used 16,942 observations

277   (individual*year combinations) representing 128 species in the spring model and 16,244

278   observations representing 112 species in the fall model (the full list of species included can be

279   found in supplemental tables S1 and S3). Both models covered records across 15 ecoregions of

280   the US (the full list of regions used can be found in supplemental tables S2 and S4) with data

281   from 2009-2017 (USA-NPN) or 2014-2017 (NEON).

282

283   **2.4 Model evaluation**

284

285   Because our data included uncertainty in the response variable (the true date of the

286   phenophase transition may have occurred at any point between the observed last no and

287   observed first yes), we could not rely on traditional methods such as RMSE to assess model fit.

288   Instead, we include a visual assessment of model fit, plotting the model-predicted median

289   phenophase transition DOY vs. observed date range ('last observed no' → 'first observed yes')

290   for each individual*year combination. We also calculated the ranked probability score (RPS)

291   (Epstein, 1969; Murphy, 1970) for a quantitative assessment of relative model fit, following as

292   similar example in Terres et al. (2013). Specifically, we calculated a relative ranked probability

293   score ($RRPS_i$) for each individual observation, defined as the average of the squared difference

294   of the probability of having transitioned by any day of year (P_{it}) and the actual state (0,1) over

295   all dates prior to and including the 'last observed no' and after and including the 'last observed

296   yes.'

297

298    $$RRPS_i = \frac{\left(\sum_{t=minStart}^{t=lastNo}\left[P_{i,t}\right]^2 + \sum_{t=firstYes_i}^{t=maxEnd-1}\left[1 - P_{i,t}\right]^2\right)}{maxEnd - minStart - (firstYes_i - lastNo_i)}$$

299

300

301

302   where P_{it} is the modeled probability of having transitioned (state=1) by any given day of year (t).

303   Because the state is an ordered variable and P_{it} is a monotonically increasing function, $RRPS_i$

304   scores are always higher for individuals with greater measurement uncertainty (i.e., more days

305   between the 'last observed no' and 'first observed yes'). We averaged all $RRPS$ measurements

306   over all observations included in each model for an overall assessment of model quality ($\overline{RRPS}$).

307    $\overline{RRPS}$ scores are only valid when compared to the same input dataset, so, for example, a

308   comparison of the $\overline{RRPS}$ for the spring vs. fall models is not meaningful because the underlying

309   measurement uncertainty of observations included in the two datasets differs. When applied to

310   different models of the same dataset, lower $\overline{RRPS}$ values indicate better model performance.

311

312   **3. RESULTS**

313

314   **3.1 Model assessment**

315

316   For the spring leaf out model, we were able to fit a single model including growing degree days

317   (GDD), daylength, chilling, precipitation and frost as environmental forcing variables as well as

318   an intercept (baseline daily probability of remaining in the initial [leafless] state for the spring leaf

319   model ($\hat{R}$ for all parameters <1.1). Inclusion of frost days in the fall model created convergence

320   problems. We also experienced a lack of convergence when we included both intercept and

321   daylength as predictors in the same fall model. These issues with convergence were likely due

322   to the large number of days prior to the phenological event that included 0 sum frost days. We

323   therefore fit the fall model without frost as well as alternative forms of the fall model, including

324   either an intercept or daylength but not both. Also, to parallel the fall model, we constructed

325   alternative forms of the spring model with and without an intercept. Spring models that included

326   an intercept showed substantially better fit ($\overline{RRPS}$ = 0.026) than those without ($\overline{RRPS}$= 0.112).

13

327   The fall model that included daylength had a slightly better fit ($\overline{RRPS}$=0.063) than the model that

328   included an intercept ($\overline{RRPS}$=0.071).

329

330   Visual assessment of both spring and fall models indicated better predictive performance of the

331   spring vs. fall model (Fig 1). In particular, the fall model erroneously predicted a suite of fall

332   coloration events to occur very early in the season (before DOY 200). Spring data showed

333   better agreement between predicted and observed dates.

334

335

336

337   **Predicted Day of Transition (Median + 95% CI)**
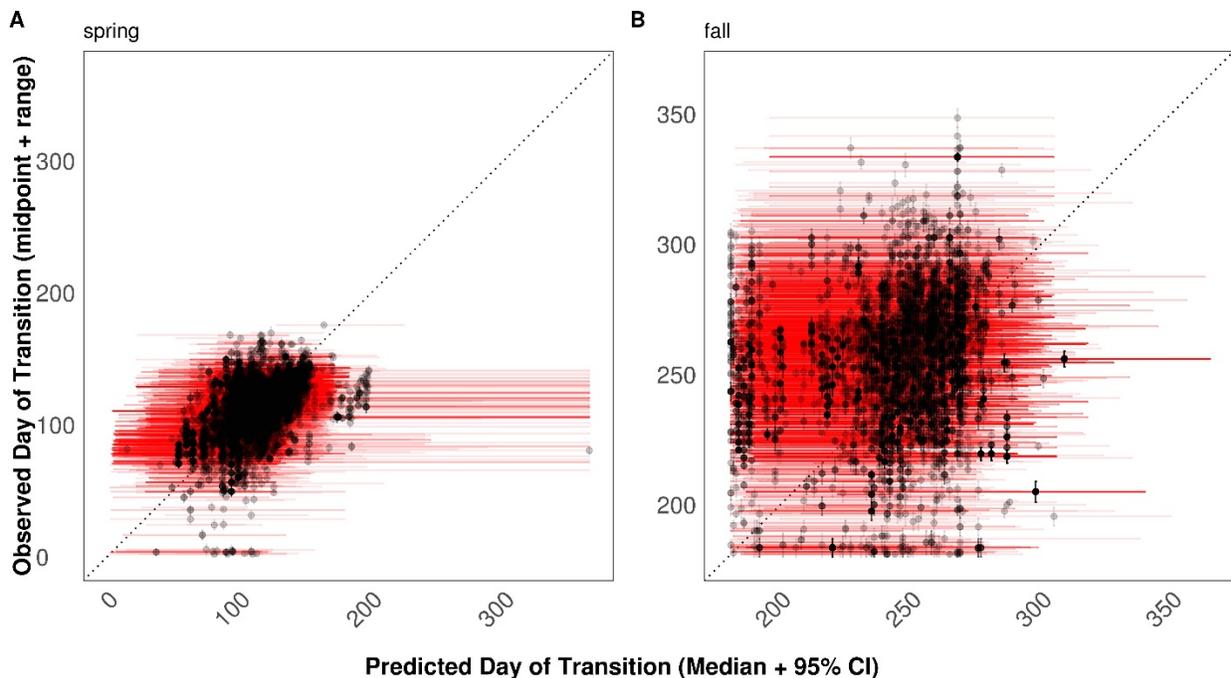
338

Figure 1. Predicted versus observed dates of spring leaf out for 128 species (A) and fall leaf coloration for 112 species (B). Figure shows median estimated date of leaf out or fall coloration (x axis, and horizontal 95% CI in red) against observed dates of transition (vertical black bars, lower = 'last observed no,' upper = 'first observed yes,' dot = midpoint). Perfect predictions would fall along the 1:1 line. For purposes of the figure, we included only observations with <10d intercensus intervals between the 'last observed no' and 'first observed yes,' though the model included all data.

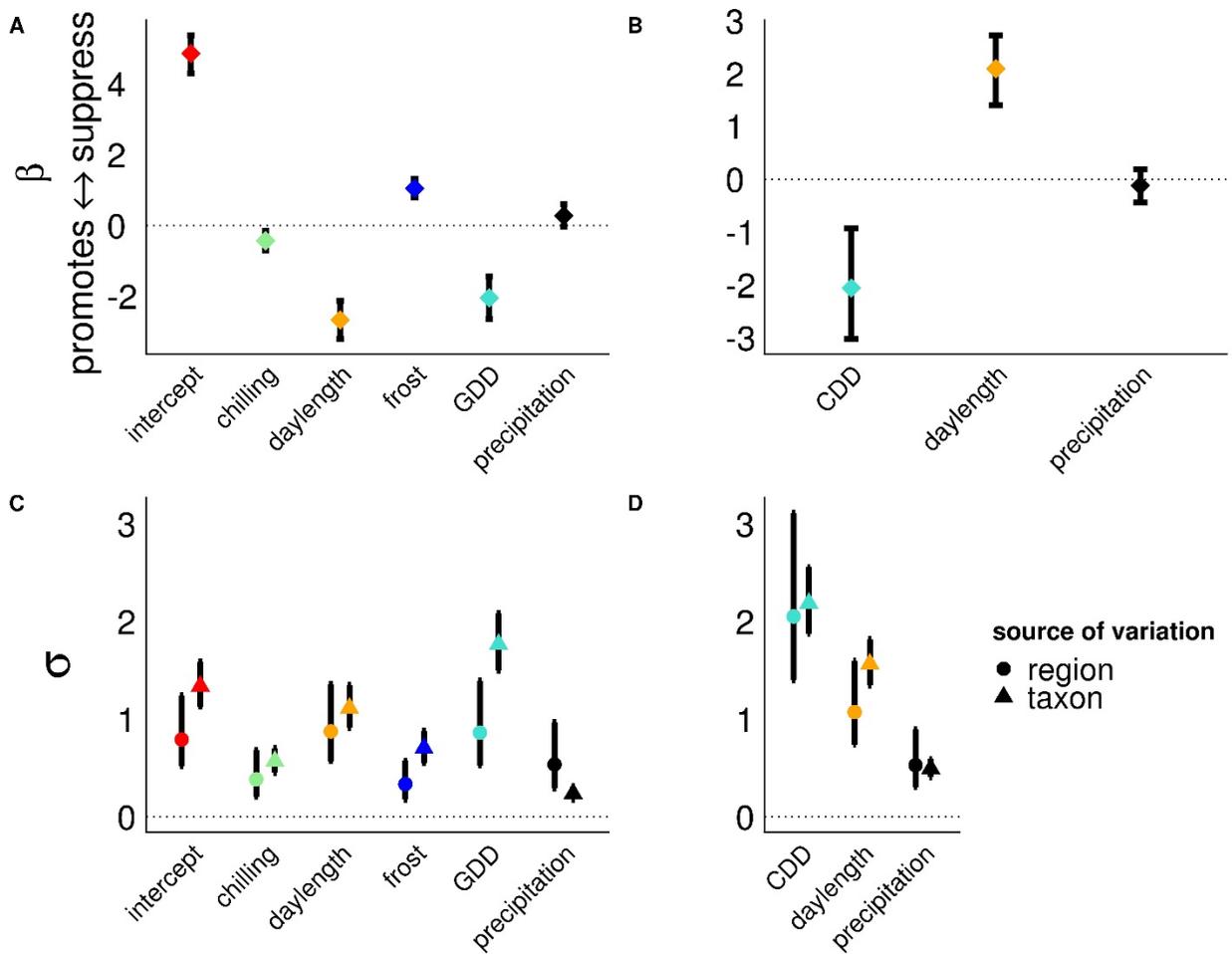342 **3.2 Sensitivity to Environmental Drivers**

343

344 **3.2.1 Spring leaf out**

345 On average, spring leaf out was promoted by longer days, higher growing degree day sums,

346 higher amounts of chilling, and was suppressed by recent frost events (Fig 2a). Precipitation

347 had a slight but nonsignificant suppressing effect on spring leaf out.

348

349

350

351

352

353

Figure 2. Mean sensitivities and variance in sensitivities among species and regions of deciduous trees and shrubs phenology to environmental drivers. Panels A (spring leaf out) and B (fall leaf coloration) show estimated mean sensitivity + 95% CI for each environmental driver, plus the baseline hazard rate (intercept). Positive coefficients indicate climate forcing variables that prolong the initial state (i.e., leafless state for spring; green-leaved state for fall), and therefore delay the onset of the phenological transition. Panels C and D show the variability in sensitivity of drivers for spring (C) and fall (D) models among geographic regions (squares) and taxa (triangles).

354   -

355

356   Models successfully generated estimates of sensitivity of each taxon and each region to the

357   environmental drivers. With the exception of precipitation, sensitivities of leaf out to

358   environmental drivers varied more among taxa than among regions (Fig 2c). To illustrate the

359   ability of the models to generate estimates for the large diversity of taxa included here,

360   sensitivities of spring leaf out to one environmental driver (GDD) is presented in Figure 3. The

361   full suite of parameter estimates for all environmental drivers for each species and region for

362   both the spring models are included in the Supplementary materials (Tables S.1-S.4).

Figure 3: Sensitivity of spring leaf out to GDD for each taxa (point estimate + 95% CI of each $\beta_{GDD,species}$) term. Negative parameter estimates reflect species where GDD accelerates spring leaf out.

364

365 **3.2.2 Fall leaf coloration**

366

367 On average, fall leaf coloration was promoted by short days and higher cold degree day sums.

368 Precipitation had a slight but nonsignificant suppressing effect on fall coloration (Fig 2B). For fall

369 coloration, we again saw higher variability in sensitivity among species than regions to

370 daylength, but more similar amounts of variability among species and regions for precipitation

371 and cold sensitivity (Fig 2D). The full suite of parameter estimates for all environmental drivers

372 for each species and region for both the spring models are included in Appendix F.

373

374 **4. DISCUSSION**

375

376 Recent growth and standardization of phenology observations across broad spatial scales over

377 multiple years has the potential to enhance our understanding of the drivers of temperature

378 plant phenology. Yet the patchiness of available data results in analytical challenges with

379 respect to spatial, temporal, and taxonomic coverage. We applied a novel analytical framework

380 to address these challenges and better understand the drivers of temperate leaf phenology and

381 quantify sources of variability. On average, spring leaf out was accelerated by longer days,

382 spring growing degree day accumulation, overwinter chilling, and was suppressed by frost

383 events, whereas fall leaf color was accelerated by shorter days and cold accumulation.

384 Comparisons across species and regions indicated climate sensitivity tended to be more

385 variable among species than among regions.

386

4.1 Advantages of a hierarchical survival modeling approach

388

389   We used hierarchical survival models to predict spring and fall phenological events across many

390   taxa and regions. Space-state phenology models have been developed using datasets with

391   frequent and regular census intervals (Allen et al., 2014; Diez et al., 2014; Xie et al., 2015).

392   These models can also be easily applied to datasets with variable and irregular observation

393   frequencies. Such flexibility makes state-space approaches particularly well-suited to the data

394   collected by volunteers, where infrequent and unevenly spaced and sampling intervals are the

395   norm. This approach overcomes a common obstacle to embarking on a phenological campaign,

396   namely, a requirement of high-frequency sampling regimes (Fitchett et al., 2015; Menzel, 2002).

397

398   In addition to the ability to make use of data with infrequent and/or variable intercensus-

399   intervals, our approach has several advantages over common methods to analyze phenological

400   data in terms of representing our understanding of the biological drivers of phenology. Modeling

401   the daily probability of an event occurring permits us to incorporate weather up to but not

402   beyond the date of a phenological transition as drivers of phenology, which is consistent with

403   our understanding of the biological basis of phenological cues. Using a moving window for

404   forcing accumulation also allows daily probabilities to vary non-monotonically over the growing

405   season; for example, a late spring frost event could suppress the leaf out probability after that

406   event occurs to levels below that before the frost event. Daily models require substantial data

407   volumes to fit; models for only a few species would achieve convergence if each species were

408   modeled separately. At the other extreme, combining datasets across taxa and regions (such as

409   is commonly done in meta-analyses) ignores potential important variability in the sensitivities to

410   environmental drivers among species and regions. Using a hierarchical framework allows us to

411   leverage a large volume of data collected over many sites and species, despite the fact that the

2

412 dataset as a whole contains relatively few long time series of a single species at a given site.

413 This approach allows us to gain holistic insight into general drivers of deciduous spring and fall

414 phenology across temperate regions of North America, quantify the degree of variability among

415 species and regions, as well as generate estimated phenological sensitivities for over 100

416 species and 15 ecoregions.

417

418 **4.2 Drivers of spring and fall phenology**

419

420 The dominant controls on deciduous tree phenology over most non-desert, temperate regions of

421 the world are generally considered winter chilling, day length and temperature (Korner and

422 Basler, 2010). Our analysis of over 100 deciduous taxa confirmed the importance of these

423 drivers, showing that on average spring leaf out was stimulated by longer days, growing degree

424 sums, and chilling. These results concur with those of Templ. et al. 2017, who found a strong

425 positive effect of GDD on flowering using survival modeling, but a weak to insignificant effect of

426 precipitation. The relationship between spring frost or precipitation and spring leaf out events

427 has been less well-studied than that of temperature and daylength. Our results indicated a

428 previously unappreciated role of spring frost in delaying leaf out and a slight trend (different from

429 0 for a few taxa) of spring precipitation to delay leaf out. The magnitude of the negative

430 (suppressive) effect of frost events was stronger than the positive (promoting) effect of

431 overwinter chilling. The physiological (and ensuing demographic effects) of frost damage are

432 well-documented (Inouye, 2000), so it is unsurprising that plants use recent frost events as a

433 proximal cue for the emergence of sensitive tissues. Our leaf out results parallel a recent

434 synthesis of flowering sensitivity to frost (Park and Mazer, 2018), with both studies highlighting

435 the importance of frost as a driver of the timing of spring phenological events.

436

3

437   The direction of the mean sensitivities of fall leaf coloration to climate drivers was partially

438   consistent with previous studies. Xie et al (2018) found rainfall in autumn cause earlier peak

439   color, while we found a non-significant trend in the same direction. Previous research on the

440   effects of temperature on the timing of fall onset have been conflicting, with evidence to support

441   both advances and delays in leaf coloration with warmer temperatures under natural conditions

442   (see Archetti et al., 2013 and references therein), whereas experimental manipulations tend to

443   suggest a more consistent delaying effect of summer and fall temperature on senescence (Fu et

444   al., 2018; Gunderson et al., 2012). The sign and strength of the relationship between

445   temperature change and onset of fall can vary depending on the time of year and fall

446   phenophase under consideration (Archetti et al., 2013)

447

448   We did not detect a strong role of precipitation in the onset of either spring or fall phenological

449   events in this study. In temperate regions such as our study area, plants likely experience

450   moisture stress or abundance through soil moisture, which is a function of soil properties and

451   land cover as well as incoming precipitation (Pan et al., 2003). For example, Terrestrial

452   biosphere models such as ORCHIDEE, CLM, LPJ-GUESS and Sheffield- DGVM commonly

453   assume soil moisture physiologically constrains leaf phenology (Abramoff and Finzi, 2015).

454   Improved satellite-based soil moisture products (e.g. Dorigo et al. 2017 or Reichle et al 2018),

455   when generated on fine-spatial scales for a sufficient number of years, might provide a

456   detectable soil moisture signal beyond the limited trends with precipitation found here.

457

458   Species showed consistently strong sensitivity to day length in both the spring (Tables S.1) and

459   fall models (Table S.2). These correlative results are consistent with experiments that show

460   photoperiod sensitivity to be fairly ubiquitous in spring events (Flynn and Wolkovich (2018). The

461   relationship of daylength to fall coloration has not been as thoroughly studied experimentally;

462 however, our results concur with the hypothesis that day length is likely a primary regulator of

463 fall phenology (Way and Montgomery, 2015).

464

465 **4.3 Variability in sensitivity**

466

467 Our results demonstrate that sensitivity to environmental drivers varies both by species and by

468 region. The inter-specific and inter-regional variation in sensitivity to environmental drivers found

469 in this study helps explain the heterogeneity of results of previous studies of phenological

470 sensitivity conducted on different species, which has limited transferability of phenological

471 models parameterized at one location to other locations (Basler 2016). Previous studies have

472 noted that interspecific variability can be as important as climate in predicting landscape level

473 phenological patterns (Xie et al., 2015). Our results support an important role for species

474 composition and location in understanding larger-scale phenological responses. Our process for

475 selecting regions and species for inclusion must be kept in mind when interpreting the

476 magnitude of the region and species variability. We constrained our geographic regions to

477 exclude desert and tropical ecoregions and included species only within a single life form. It is

478 likely that broadening the species pool or geographic locations included would increase the

479 amount of interspecific or inter-regional variability detected. Unsurprisingly, parameter

480 uncertainty declined with sample size, asymptoting at 250-1,000 individuals per species or per

481 region (Appendix E). This suggests that future data collection activities should focus on species

482 and regions with data densities below that threshold.

483

484 **4.4 Challenges and opportunities**

485

486 In this study, predictive capacity for the models was moderate, as evidenced by the fit of the

487 subset of observations with relatively small sampling error to the predicted phenological dates

488   (Fig 1). Consistent with the literature, our models performed better for spring than fall events

489   (Gallinat et al., 2015). This may reflect a higher degree of variability in the drivers of onset of

490   colored leaves such as disease or other stressors. Future applications using the analytical

491   framework outline here might consider alternative or novel methods for accumulating

492   phenological forcing variables. For example, modeling forcing variables using critical thresholds

493   rather than using a simple logit-linear relationship; accounting for the varying importance of

494   environmental drivers over different time periods, and/or modeling interactive effects between

495   drivers may yield improved results. An additional consideration is that our estimates of

496   sensitivity to environmental variation combine both sensitivity to interannual climate variability

497   and sensitivity to climate variability over space (due to different locations being sampled in

498   different years). While it is relatively common to combine shorter phenological time series in this

499   way (Liang and Schwartz, 2014; Melaas et al., 2016; Taylor et al., 2019), even within regions

500   there is some evidence that sensitivity to temperature over spatial gradients are stronger than

501   over time (Jocher et al., 2013). As more years of data accumulate from the core-resampled

502   (NEON) datasets here, it may be possible to address the generality of this pattern, and how to

503   best model both spatial and temporal sensitivity. Approaches that account for the relationship

504   between the timing of early-season and later-season phenological events may result in

505   increased performance (Keenan and Richardson, 2015). On the other hand, a similar study

506   (using probit-regression, which has a fairly similar sigmoid form relating linear terms of

507   probability of occurrence as logit-regression) yielded very accurate predictions of flowering date

508   (Allen et al., 2014). This suggests that some of the 'noise' inherent in many distributed

509   phenology datasets, and/or generalizing over such large regions and taxonomically diverse

510   species may have predictive limits.

511

512   A limitation of the framework used here is that models were slow to run. Re-running the models

513   over a large combination of forcing variables*periods*functional forms, as has been done for

514 some other efforts (Basler, 2016; Park and Mazer, 2018), is too computationally expensive at

515 present. A more tractable approach to including alternative forms of environmental drivers or

516 accumulation windows might be to use a smaller subset of species and regions for intermodel

517 comparison before refitting the most promising subset on the full dataset.

518

519 **5. CONCLUSIONS**

520

521 The development of phenological prediction tools requires advancing the field of phenology

522 beyond directional assessment, such as evaluating whether plants generally leaf out earlier with

523 warmer springs, to specific quantitative assessments, such as determining the number of days

524 earlier a species of plant growing in a particular location will leaf out with warming that occurs at

525 particular times of year (Dietze et al., 2018; Houlahan et al., 2017). We demonstrated a novel

526 approach to predicting the timing of phenological events such as leaf out utilizing survival

527 analysis. This hierarchical approach demonstrated strengths over traditional statistical

528 approaches, allowing for the incorporation of data across species and regions and leveraging

529 the large amount of data availability while also offering information on species and region level

530 variability. The results of this study offer a path forward for making widespread phenological

531 predictions by moving from the development of individual models tailored to a single site and

532 taxon to a generalized framework for predicting the transition states for any species, in any

533 location.

534

535 **6. ACKNOWLEDGEMENTS**

536

552

553

554

555

REFERENCES

556 **REFERENCES**

557 Abramoff, R.Z., and Finzi, A.C. (2015). Are above- and below-ground phenology in sync? New

558 Phytol. *205*, 1054–1061.

559 Allen, J.M., Terres, M.A., Katsuki, T., Iwamoto, K., Kobori, H., Higuchi, H., Primack, R.B.,

560 Wilson, A.M., Gelfand, A., and Silander, J.A. (2014). Modeling daily flowering probabilities:

561 expected impact of climate change on Japanese cherry phenology. Glob. Change Biol. *20*,

562 1251–1263.

563 Archetti, M., Richardson, A.D., O'Keefe, J., and Delpierre, N. (2013). Predicting Climate Change

564 Impacts on the Amount and Duration of Autumn Colors in a New England Forest. PLOS ONE *8*,

565 e57373.

566 Basler, D. (2016). Evaluating phenological models for the prediction of leaf-out dates in six

567 temperate tree species across central Europe. Agric. For. Meteorol. *217*, 10–21.

568 Bolker, B. M. (2008). Ecological Models and Data in R. Princeton University Press, Princeton,

569 New Jersey.

570

571 Clark, J. S. 2007. Models for ecological data: an introduction. Princeton University Press,

572 Princeton, New Jersey.

573 Cook, B.I., Wolkovich, E.M., and Parmesan, C. (2012). Divergent responses to spring and

574 winter warming drive community level flowering trends. Proc. Natl. Acad. Sci. *109*, 9000–9005.

575 Cox, D.R. (1972) Regression models and life-tables. J Royal Stat Soc Ser B *34*:187–220.
576
577

578 Crimmins, T.M., Crimmins, M.A., and Bertelsen, C.D. (2010). Complex responses to climate

579 drivers in onset of spring flowering across a semi-arid elevation gradient. J. Ecol. *98*, 1042–

580 1051.

581 Denny, E.G., Gerst, K.L., Miller-Rushing, A.J., Tierney, G.L., Crimmins, T.M., Enquist, C.A.F.,

582 Guertin, P., Rosemartin, A.H., Schwartz, M.D., Thomas, K.A., et al. (2014). Standardized

583 phenology monitoring methods to track plant and animal activity for science and resource

584 management applications. Int. J. Biometeorol. *58*, 591–601.

585 Dietze, M.C., Fox, A., Beck-Johnson, L.M., Betancourt, J.L., Hooten, M.B., Jarnevich, C.S.,

586 Keitt, T.H., Kenney, M.A., Laney, C.M., Larsen, L.G., et al. (2018). Iterative near-term ecological

587 forecasting: Needs, opportunities, and challenges. Proc. Natl. Acad. Sci. *115*, 1424–1432.

588 Diez, J.M., Ibáñez, I., Silander Jr, J.A., Primack, R., Higuchi, H., Kobori, H., Sen, A., and James,

589 T.Y. (2014). Beyond seasonal climate: statistical estimation of phenological responses to

590 weather. Ecol. Appl. *24*, 1793–1802.

591 Dorigo, W.A., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl,

592 M., Forkel, M., Gruber, A., et al. (2017) ESA CCI Soil Moisture for improved Earth system

593 understanding: State-of the art and future directions, Rem. Sens. of Envi. 203: 185-215.

594
595 Elmendorf, S.C., Jones, K.D., Cook, B.I., Diez, J.M., Enquist, C.A.F., Hufft, R.A., Jones, M.O.,

596 Mazer, S.J., Miller-Rushing, A.J., Moore, D.J.P., et al. (2016). The plant phenology monitoring

597 design for The National Ecological Observatory Network. Ecosphere *7*, n/a-n/a.

598 Epstein, E.S. (1969). A Scoring System for Probability Forecasts of Ranked Categories. J. Appl.

599 Meteorol. *8*, 985–987.

600   Fitchett, J.M., Grab, S.W., and Thompson, D.I. (2015). Plant phenology and climate change:

601   Progress in methodological approaches and application. Prog. Phys. Geogr. Earth Environ. *39*,

602   460–482.

603   Flynn, D.F.B., and Wolkovich, E.M. (2018). Temperature and photoperiod drive spring

604   phenology across all species in a temperate forest community. New Phytol. *219*, 1353–1362.

605   Fu, Y.H., Piao, S., Delpierre, N., Hao, F., Hänninen, H., Liu, Y., Sun, W., Janssens, I.A., and

606   Campioli, M. (2018). Larger temperature response of autumn leaf senescence than spring leaf-

607   out phenology. Glob. Change Biol. *24*, 2159–2168.

608   Gallinat, A.S., Primack, R.B., and Wagner, D.L. (2015). Autumn, the neglected season in

609   climate change research. Trends Ecol. Evol. *30*, 169–176.

610   Gelman, A., and Hill, J. (2007). Data Analysis Using Regression and Multilevel/Hierarchical

611   Models (Cambridge University Press).

612   Gill, A.L., Gallinat, A.S., Sanders-DeMott, R., Rigden, A.J., Short Gianotti, D.J., Mantooth, J.A.,

613   and Templer, P.H. (2015). Changes in autumn senescence in northern hemisphere deciduous

614   trees: a meta-analysis of autumn phenology studies. Ann. Bot. *116*, 875–888.

615   Gunderson, C.A., Edwards, N.T., Walker, A.V., O'Hara, K.H., Campion, C.M., and Hanson, P.J.

616   (2012). Forest phenology and a warmer climate - growing season extension in relation to

617   climatic provenance. Glob. Change Biol. *18*, 2008–2025.

618   Houlahan, J.E., McKinney, S.T., Anderson, T.M., and McGill, B.J. (2017). The priority of

619   prediction in ecological understanding. Oikos *126*, 1–7.

620   Hufft, R.A., DePrenger-Levin, M.E., Levy, R.A., and Islam, M.B. (2018). Using herbarium

621   specimens to select indicator species for climate change monitoring. Biodivers. Conserv. *27*,

622   1487–1501.

623   Inouye, D.W. (2000). The ecological and evolutionary significance of frost in the context of

624   climate change. Ecol. Lett. *3*, 457–463.

625   Jochner, S., Caffarra, A., and Menzel, A. (2013). Can spatial data substitute temporal data in

626   phenological modelling? A survey using birch flowering. Tree Physiol 33, 1256–1268.

627   Keenan, T.F., and Richardson, A.D. (2015). The timing of autumn senescence is affected by the

628   timing of spring phenology: implications for predictive models. Glob. Change Biol. *21*, 2634–

629   2641.

630   Korner, C., and Basler, D. (2010). Phenology under global warming. Science *327*, 1461–1462.

631   Lechowicz, M.J. (1984). Why do temperate deciduous trees leaf out at different times?

632   Adaptation and ecology of forest communities. Am. Nat. *124*, 821–842.

633   Liang, L. (2019). A spatially explicit modeling analysis of adaptive variation in temperate tree

634   phenology. Agric. For. Meteorol.*266–267*, 73-86.

635   Liang, L., and Schwartz, M.D. (2014). Testing a growth efficiency hypothesis with continental-

636   scale phenological variations of common and cloned plants. Int. J. Biometeorol. *58*, 1789–1797.

637   Ma, Q., Huang, J.-G., Hänninen, H., and Berninger, F. (2018). Reduced geographical variability

638   in spring phenology of temperate trees with recent warming. Agric. For. Meteorol. *256–257*,

639   526–533.

640 MacKenzie, C.M., Primack, R.B., and Miller-Rushing, A.J. (2018). Local environment, not local

641 adaptation, drives leaf-out phenology in common gardens along an elevational gradient in

642 Acadia National Park, Maine. Am. J. Bot. *105*, 986–995.

643 Mazer, S.J., Gerst, K.L., Matthews, E.R., and Evenden, A. (2015). Species-specific phenological

644 responses to winter temperature and precipitation in a water-limited ecosystem. Ecosphere 6,

645 art98.

646 McKinney, A.M., CaraDonna, P.J., Inouye, D.W., Barr, B., Bertelsen, C.D., and Waser, N.M.

647 (2012). Asynchronous changes in phenology of migrating Broad-tailed Hummingbirds and their

648 early-season nectar resources. Ecology *93*, 1987–1993.

649 Melaas, E.K., Friedl, M.A., and Richardson, A.D. (2016). Multiscale modeling of spring

650 phenology across Deciduous Forests in the Eastern United States. Glob. Change Biol. *22*, 792–

651 805.

652 Menzel, A. (2002). Phenology: Its Importance to the Global Change Community. Clim. Change

653 *54*, 379–385.

654 Miller-Rushing, A.J., and Primack, R.B. (2008). Global warming and flowering times in

655 Thoreau's Concord: A community perspective. Ecology *89*, 332–341.

656 Morisette, J.T., Richardson, A.D., Knapp, A.K., Fisher, J.I., Graham, E.A., Abatzoglou, J.,

657 Wilson, B.E., Breshears, D.D., Henebry, G.M., Hanes, J.M., et al. (2009). Tracking the rhythm of

658 the seasons in the face of global change: phenological research in the 21st century. Front. Ecol.

659 Environ. *7*, 253–260.

660 Murphy, A.H. (1970). The ranked probability score and the probability score: a comparison.

661 Mon. Weather Rev. *98*, 917–924.

662   [dataset] National Ecological Observatory Network. 2018. Data Product: NEON.DP1.10055.

663   Provisional data downloaded from http://data.neonscience.org on 2018-07-05. Battelle, Boulder,

664   CO, USA.

665

666   Oberbauer, S., Elmendorf, S., Troxler, T., Hollister, R., Rocha, A., Bret-Harte, M., Dawes, M.,

667   Fosaa, A., Henry, G., and Høye, T.T. (2013). Phenological response of tundra plants to

668   background climate variation tested using the International Tundra Experiment. Philos. Trans.

669   R. Soc. B Biol. Sci. *368*, 20120481.

670   Pan, F., Peters-Lidard, C.D., and Sale, M.J. (2003). An analytical method for predicting surface

671   soil moisture from rainfall observations. Water Resour. Res. *39*.

672   Park, I.W., and Mazer, S.J. (2018). Overlooked climate parameters best predict flowering onset:

673   Assessing phenological models using the elastic net. Glob. Change Biol. *24*, 5972–5984.

674   Parker, T.C., Tang, J., Clark, M.B., Moody, M.M., and Fetcher, N. (2017). Ecotypic differences

675   in the phenology of the tundra species Eriophorum vaginatum reflect sites of origin. Ecol. Evol.

676   *7*, 9775–9786.

677   Pearse, W.D., Davis, C.C., Inouye, D.W., Primack, R.B., and Davies, T.J. (2017). A statistical

678   estimator for determining the limits of contemporary and historic phenology. Nature Ecology &

679   Evolution *1*, 1876-1882.

680   Peng, D., Wu, C., Li, C., Zhang, X., Liu, Z., Ye, H., Luo, S., Liu, X., Hu, Y., and Fang, B. (2017).

681   Spring green-up phenology products derived from MODIS NDVI and EVI: Intercomparison,

682   interpretation and validation using National Phenology Network and AmeriFlux observations.

683   Ecological Indicators *77*, 323–336.

684

685
686 Piao, S., Tan, J., Chen, A., Fu, Y.H., Ciais, P., Liu, Q., Janssens, I.A., Vicca, S., Zeng, Z.,

687 Jeong, S.-J., et al. (2015). Leaf onset in the northern hemisphere triggered by daytime

688 temperature. Nature Communications 6, 6911.

689
690 Prevéy, J., Vellend, M., Rüger, N., Hollister, R.D., Bjorkman, A.D., Myers-Smith, I.H., Elmendorf,

691 S.C., Clark, K., Cooper, E.J., Elberling, B., et al. (2017). Greater temperature sensitivity of plant

692 phenology at colder sites: implications for convergence across northern latitudes. Glob. Change

693 Biol. *23*, 2660–2671.

694 Reichle, R., G. De Lannoy, R. D. Koster, W. T. Crow, J. S. Kimball, and Q. Liu. 2018. *SMAP L4*

695 *Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Analysis Update, Version*

696 *4*. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive

697 Center. doi: https://doi.org/10.5067/60HB8VIP2T8W.

698
699 Richardson, A.D., Keenan, T.F., Migliavacca, M., Ryu, Y., Sonnentag, O., and Toomey, M.

700 (2013). Climate change, phenology, and phenological control of vegetation feedbacks to the

701 climate system. Agric. For. Meteorol. *169*, 156–173.

702 Richardson, E.A., Seeley, S.D., and Walker, D.R. (1974). A model for estimating the completion

703 of rest for "Redhaven" and "Elberta" peach trees. HortScience *9*, 331-332.

704 Roberts, A.M.I., Tansey, C., Smithers, R.J., and Phillimore, A.B. (2015). Predicting a change in

705 the order of spring phenology in temperate forests. Glob Change Biol *21*, 2603–2611.

706 Royle, J.A., and Dorazio, R.M. (2008). Modeling Survival. In Hierarchical Modeling and

707 Inference in Ecology: The Analysis of Data from Populations, Metapopulations and

708 Communities, (Academic Press), pp. 347- 377.

709   Siljamo, P., Sofiev, M., Filatova, E., Grewling, Ł., Jäger, S., Khoreva, E., Linkosalo, T., Ortega

710   Jimenez, S., Ranta, H., Rantio-Lehtimäki, A., et al. (2013). A numerical model of birch pollen

711   emission and dispersion in the atmosphere. Model evaluation and sensitivity analysis. Int. J.

712   Biometeorol. *57*, 125–136.

713   Taylor, S.D., Meiners, J.M., Riemer, K., Orr, M.C., and White, E.P. (2019). Comparison of large-

714   scale citizen science data and long-term study data for phenology modeling. Ecology *100*,

715   e02568.

716   Templ, B., Fleck, S., and Templ, M. (2017). Change of plant phenophases explained by survival

717   modeling. Int. J. Biometeorol. *61*, 881–889.

718   Terres, M.A., Gelfand, A.E., Allen, J.M., and Silander, J.A. (2013). Analyzing first flowering

719   event data using survival models with space and time-varying covariates. Environmetrics *24*,

720   317–331.

721   [dataset] Thornton, P.E., Thornton, M.M., Mayer, B.W., Wei, Y., Devarakonda, R., Vose, R.S.,

722   and Cook, R.B. (2018). Daymet: Daily Surface Weather Data on a 1-km Grid for North America,

723   Version 3. Data set accessed 5 July 2018. ORNL DAAC Oak Ridge Tenn. USA.

724   https://doi.org/10.3334/ORNLDAAC/1328

725
726   [dataset] United States Environmental Protection Agency. 2018. Ecoregions of North America.

727   USA. Data set accessed 2018-07-05 at https://www.epa.gov/eco-research/ecoregions-north-

728   america

729
730   [dataset] USA National Phenology Network. 2018. Plant and Animal Phenology Data. Data type:

731   Status and Intensity (2009-2017). NPN, Tucson, Arizona, USA. Data set accessed 2018-07-05

732   at http://doi.org/10.5066/F78S4N1V

733

734   Way, D.A., and Montgomery, R.A. (2015). Photoperiod constraints on tree phenology,

735   performance and migration in a warming world. Plant Cell Environ. *38*, 1725–1736.

736   Williamson, S.N., Barrio, I.C., Hik, D.S., and Gamon, J.A. (2016). Phenology and species

737   determine growing-season albedo increase at the altitudinal limit of shrub growth in the sub-

738   Arctic. Glob. Change Biol. *22*, 3621–3631.

739   Xie, Y., Ahmed, K.F., Allen, J.M., Wilson, A.M., and Silander, J.A. (2015). Green-up of

740   deciduous forest communities of northeastern North America in response to climate variation

741   and climate change. Landsc. Ecol. *30*, 109–123.

742   Xie, Y., Civco, D.L., and Silander, J.A. (2018). Species-specific spring and autumn leaf

743   phenology captured by time-lapse digital cameras. Ecosphere *9*, e02089.

744   Zhang, X., Liu, L., Liu, Y., Jayavelu, S., Wang, J., Moon, M., Henebry, G.M., Friedl, M.A., and

745   Schaaf, C.B. (2018). Generation and evaluation of the VIIRS land surface phenology product.

746   Remote Sensing of Environment *216*, 212–229.

747

748

749

750

751

752

SUPPLEMENTAL TABLES

754

Table S.1: Species-specific parameter estimates for the spring leaf out model. Shown are the

sensitivity of plants in each taxon to each forcing variable (+ intercept). Included are point

estimates + 95% CI for each coefficient $\beta_{j,species}$). Random effects for species are u_j centered

(see appendix A, Statistical modeling details); negative parameter estimates reflect species

where higher values of the coefficient accelerates spring leaf out. See figure 3 in the main text

for an examplar graphical display of the same information for $\beta_{GDD,species}$.

761

Table S.2: Region-specific parameter estimates for the spring leaf out model. Shown are the

sensitivity of plants in each region to each forcing variable (+ intercept). Included are point

estimates + 95% CI for each coefficient $\beta_{j,region}$). Random effects for region are 0-centered (see

appendix A, Statistical modeling details). Negative parameter estimates therefore reflect regions

where higher values of the coefficient accelerates spring leaf out more *relative to other regions*.

767

Table S.3: Species-specific parameter estimates for the fall leaf-coloration model. Shown are

the sensitivity of plants in each taxon to each forcing variable. Included are point estimates +

95% CI for each coefficient $\beta_{j,species}$). Random effects for species are u_j centered (see appendix

A, Statistical modeling details); negative parameter estimates reflect species where higher

values of the coefficient accelerate fall leaf coloration leaf out.

773

774 Table S.4: Region-specific parameter estimates for the fall leaf-coloration model. Shown are the

775 sensitivity of plants in each region to each forcing variable. Included are point estimates + 95%

776 CI for each coefficient $\beta_{j,region}$). Random effects for region are 0-centered (see appendix A,

777 Statistical modeling details). Negative parameter estimates therefore reflect regions where

778 higher values of the coefficient accelerate fall leaf coloration more *relative to other regions*.

779

780

781