Herbarium specimens can reveal impacts of climate change on plant phenology; a review of methods and applications

Casey A Jones $^{Corresp.,\ 1}$, Curtis C Daehler 1

¹ Department of Botany, University of Hawaii at Manoa, Honolulu, HI, United States

Corresponding Author: Casey A Jones Email address: jonesc22@hawaii.edu

Studies in plant phenology have provided some of the best evidence for large-scale responses to recent climate change. Over the last decade, more than thirty studies have used herbarium specimens to analyze changes in flowering phenology over time, although studies from tropical environments are thus far generally lacking. In this review, we summarize the approaches and applications used to date. Reproductive plant phenology has primarily been analyzed using two summary statistics, the mean flowering day of year and first-flowering day of year, but mean flowering day has proven to be a more robust statistic. Two types of regression models have been applied to test for associations between flowering, temperature and time: flowering day regressed on year and flowering day regressed on temperature. Most studies analyzed the effect of temperature by averaging temperatures from three months prior to the date of flowering. On average, published studies have used 55 herbarium specimens per species to characterize changes in phenology over time, but in many cases fewer specimens were used. Geospatial grid data are increasingly being used for determining average temperatures at herbarium specimen collection locations, allowing testing for finer scale correspondence between phenology and climate. Multiple studies have shown that inferences from herbarium specimen data are comparable to findings from systematically collected field observations. Understanding phenological responses to climate change is a crucial step towards recognizing implications for higher trophic levels and large-scale ecosystem processes. As herbaria are increasingly being digitized worldwide, more data are becoming available for future studies. As temperatures continue to rise globally, herbarium specimens are expected to become an increasingly important resource for analyzing plant responses to climate change.

1	Herbarium specimens can reveal impacts of climate change on plant phenology; a review of
2	methods and applications
3	
4	Casey A. Jones ¹ , Curtis C. Daehler ¹
5	
6	1 Department of Botany, University of Hawaii at Manoa, Honolulu, Hawaii
7	
8	Corresponding Author:
9	Casey Jones ¹
10	3190 Maile Way, Honolulu, HI 96822, U.S.A
11	Email address: jonesc22@hawaii.edu
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	

47 Abstract

Studies in plant phenology have provided some of the best evidence for large-scale responses to 48 recent climate change. Over the last decade, more than thirty studies have used herbarium 49 specimens to analyze changes in flowering phenology over time, although studies from tropical 50 environments are thus far generally lacking. In this review, we summarize the approaches and 51 52 applications used to date. Reproductive plant phenology has primarily been analyzed using two summary statistics, the mean flowering day of year and first-flowering day of year, but mean 53 flowering day has proven to be a more robust statistic. Two types of regression models have 54 55 been applied to test for associations between flowering, temperature and time: flowering day regressed on year and flowering day regressed on temperature. Most studies analyzed the effect 56 of temperature by averaging temperatures from three months prior to the date of flowering. On 57 average, published studies have used 55 herbarium specimens per species to characterize 58 changes in phenology over time, but in many cases fewer specimens were used. Geospatial grid 59 data are increasingly being used for determining average temperatures at herbarium specimen 60 collection locations, allowing testing for finer scale correspondence between phenology and 61 climate. Multiple studies have shown that inferences from herbarium specimen data are 62 63 comparable to findings from systematically collected field observations. Understanding phenological responses to climate change is a crucial step towards recognizing implications for 64 higher trophic levels and large-scale ecosystem processes. As herbaria are increasingly being 65 66 digitized worldwide, more data are becoming available for future studies. As temperatures continue to rise globally, herbarium specimens are expected to become an increasingly important 67 68 resource for analyzing plant responses to climate change.

70 Introduction

Carl Linnaeus pioneered the study of phenology when he outlined methods for investigating 71 associations between flowering and climate in the 1700s (Linnaeus 1751; von Linné, Freer, S., 72 trans, 2003; Puppi 2007). Around 1850, Charles Morren introduced the term "phenology" to 73 describe his observational studies of yearly flowering (Morren 1853 ;Demarée & Rutishauser 74 75 2009). Early field studies of plant phenology have been thoroughly reviewed by van Schaik et al. (1993), Fenner (1998) and Forrest et al. (2010). Long-term observations in field studies have 76 provided a valuable resource for analyzing phenological responses to recent climate change 77 78 (Walther et al. 2002; Parmesan & Yohe 2003). A growing need for historical data that allows for the exploration of ecological implications of climate change prompted researchers to look to 79 herbarium specimens. A few phenology studies such as Borchert et al. (1996) and Rivera & 80 Borchert (2001) used herbarium specimens to study flowering periodicity, but not in the context 81 of climate change. The first study to use herbarium specimens to understand phenological 82 83 responses to climate change was published in 2004 by Primack et al. (2004). Primack et al. (2004) analyzed 372 specimen records (1885-2002) and found flowering had advanced 84 approximately eight days over the last century. Between 2004 and 2017, more than 30 studies 85 86 were published using herbarium specimens to examine changes in phenology in response to climate change. 87

The most common approach found in studies using herbarium specimens follows the procedure set by Primack et al. (2004). This can be summarized as collecting Julian dates from herbarium specimens, collecting long-term temperature data from an independent source, and then using regression analyses to analyze correlations between Julian dates, temperatures and time (Primack et al. 2004; Miller-Rushing et al. 2006; Gallagher et al. 2009; Robbirt et al. 2011; Gaira et al.

2011; Molnár et al. 2012; Panchen et al. 2012; Park 2012; Primack and Miller-rushing 2012; Li
et al. 2013; Calinger et al. 2013; Hart et al. 2014; Rawal et al. 2015; Park & Schwartz 2015).
Primack et al. (2004) recorded the date of collection from each herbarium specimen and then
extracted Julian dates from the collection dates. A Julian date is a value between 1 and 365
corresponding to the day of year when the specimen was collected. Linear regression models are
also the most widely used statistical models in field studies investigating flowering phenology
(Zhao et al. 2013).

An early criticism of using herbarium specimens was that plant parts preserved as herbarium 100 specimens might not have been collected during their peak flowering season, potentially biasing 101 interpretations (Lamoureux 1972). Daru et al. (2017) also found spatial, temporal, trait, 102 phylogenetic, and collector biases among herbarium specimen samples. Daru et al. (2017) 103 concluded that while some of these biases can be accounted for using statistical approaches, 104 future herbarium collections should focus on filling large gaps in the data. Other studies have 105 106 found that large sample sizes afforded by herbarium specimens, and the use of mean flowering times (mean of Julian dates), could yield valid inferences, even if specimens were not collected 107 at the time of peak flowering (Primack 2004; Bertin 2015). Collector bias and plant size choice 108 109 have also been overcome by statistical analyses when mean flowering times were used as the variable of interest, rather than the date of first-flowering (Robbirt et al. 2011; Davis et al. 2015). 110 Most of the studies we reviewed used two types of linear regression models to show evidence of 111 112 associations between phenology and climate change (Table 1). These studies regressed flowering day on temperature (82%) and flowering day on year (64%) (Table 1). These studies 113 have primarily been conducted with specimens from herbaria in temperate latitudes such as the 114 115 Eastern Himalayas (Gaira et al. 2011; Li et al. 2013; Gaira et al. 2014; Hart et al. 2014),

Southern Australia (Gallagher et al., 2009; 2012; Rawal et al., 2015), Northern Europe (Robbirt 116 et al. 2011; Diskin et al. 2012; Molnár et al. 2012), and North America (Primack et al. 2004; 117 Lavoie and Lachance 2006; Miller-Rushing et al. 2006; Primack 2009; Neil et al. 2010; Panchen 118 et al. 2012; Park 2012; Primack and Miller-rushing 2012; Searcy 2012; Calinger et al. 2013; Park 119 2014; Park & Schwartz 2015; Bertin 2015; Davis et al. 2015). Although studies by Borchert 120 121 (1996) and Zalamea et al. (2011) analyzed flowering periodicity in tropical plants using herbarium specimens, we found no study to date that has used herbarium specimens to analyze 122 effects of recent climate change in a tropical region. In this review, we examined how studies 123 chose sample sizes, flowering specimens, temperature averages and geographical scale in their 124 analyses. We also examined how these studies validated the use of herbarium specimens and we 125 provide suggestions for methods to be used in future studies. 126

127

128 Survey Methodology

Between 2015 and 2017, we compiled and reviewed studies that used herbarium specimens to 129 assess climate change and flowering phenology. We searched Web of Science (1900 – present), 130 JSTOR (1665 – present) and Google Scholar for studies containing the terms herbarium, 131 132 specimen, phenology, and climate change. The methods of each study were reviewed for; sample size, determining flowering status of specimens, approach to determining temperatures, 133 134 geographic variation, and any validations of the use of herbarium specimens (e.g. comparisons to 135 field observations). Studies and methods were then categorized and a synthesis of each category is discussed; sample sizes and regression methods were also summarized (Table 1). 136

137

138 Specimen sample sizes

Sample size, or the number of specimens used per species, varied across studies (Table 1). The 139 minimum number of specimens used per species was occasionally as low as two or three records 140 (Searcy 2012). Some studies using herbarium data have set a minimum number of herbarium 141 specimens per species or a minimum time range for collections in order to more accurately 142 estimate phenologies and change over time. Calinger et al. (2013) and Gallagher et al. (2009) set 143 144 a minimum of 10 specimens in order to meet statistical assumptions of different models. Molnár et al. (2012), eliminated a species from analyses because collections only yielded dates across an 145 eight year time span. Park & Schwartz (2015) eliminated species with records that spanned less 146 than three years. Neil et al. (2009) organized species into functional groups (spring ephemerals, 147 spring shrubs, fall ephemerals, winter-spring ephemerals, and winter-spring shrubs) in order to 148 overcome the problem of low sample sizes for each species but found that responses of 149 individual species varied greatly within functional groups. 150

Several studies found sample size had a greater influence on first-flowering estimates than on 151 mean flowering estimates. Miller-Rushing & Primack (2008) used field data and found that 152 small sample sizes led to biased estimations of first-flowering dates, but mean flowering day was 153 not biased by sample sizes. Moussus et al. (2010) investigated sample sizes by simulating 10 154 known phenological estimators, such as mean flowering day and first-flowering date. After 155 comparing known phenological shifts from simulated sample data with shift estimations from 156 models using the same data, Moussus et al. (2010) concluded that first-flowering dates were 157 158 inaccurate because they showed much greater differences in comparisons than mean flowering day. Low sample sizes prompted Bertin (2015) to provide a detailed analysis of how sample size 159 affected mean, median, range, early flowering and late flowering summary statistics. In random 160 161 simulations comparing sample sizes, mean flowering day values deviated less than five days for

Manuscript to be reviewed

species with as few as four samples (Bertin 2015). Bertin (2015) concluded that the mean was a 162 more robust measure of phenology than other estimators of early flowering. Bertin (2015) also 163 showed that by increasing the sample size to 20, mean flowering times deviated only one to two 164 days. A recent study by Pearse et al. (2017) used a Weibull distribution to estimate the start of 165 the process of flowering rather than using only first-flowering observations. Pearse et al. (2017) 166 167 showed that by controlling for differences in sampling, first-flowering, peak-flowering (median) and cessation of flowering show similar changes over time in response to climate change. The 168 model used by Pearse et al. (2017) was also shown to be consistent with changes in mean-169 170 flowering from a separate sample using an early time period. Larger sample sizes may be required if phenology varies across a species' geographic range. In 171 order to analyze species distributions using herbarium specimens, van Proosdij et al. (2016) 172 found that the minimum number of herbarium specimens sampled should be between 14 and 25 173 depending on the geographical range of the species. The van Proosdij et al. (2016) study used 174 simulated species to assess the minimum herbarium samples required for acceptable model 175 performance in both virtual and real study areas. Some species with narrow geographical ranges 176 could be modeled with as few as 14 herbarium records while wide ranging species could be 177 satisfactorily modeled with a minimum of 25 records (van Proosdij et al. 2016). Based on these 178 studies, we recommend caution when interpreting results from samples sizes with fewer than 30 179 records (Miller-Rushing et al. 2008; Moussus et al. 2010; Bertin 2015). The average sample size 180 181 across studies in this review was about 55 records per species (Table 1.). We also recommend using the mean flowering day of year rather than averages of first flowering dates (Calinger et al. 182 183 2013; Gallagher et al. 2009; Pearse et al. 2017).

185 Determining flowering status of specimens

Some studies have simply recorded the presence or absence of flowers from herbarium 186 specimens as an indicator of flowering, but other studies have used more detailed criteria to 187 assess flowering status on specimens. Haggerty et al. (2012) provided a primer to assist 188 researchers with collecting data from herbarium specimens. Haggerty et al. (2012) suggested 189 190 researchers assign a phenophase for each specimen, such as pre-flowering, first-flowering or peak flowering. Haggerty et al. (2012) also noted that researchers must assume the stem on the 191 herbarium sheet represents the flowering phenophase for the entire plant. Past studies, such as 192 Diskin et al. (2011), have used a scoring system from 1 to 5 to categorize phenophase stages 193 raging from "no flowers" to "end of fruiting" on each specimen. Diskin et al. (2011) categorized 194 flowering as 50% of buds open on the specimen. Calinger et al. (2013) also categorized 195 flowering as 50% of flower buds in anthesis to ensure that the samples were in peak flowering. 196 For a species with an inflorescence, Davis et al. (2015) only counted specimens as flowering if 197 greater than 75% of flowers were open. Standardization of phenological terms remains a core 198 challenge of mining phenological data (Willis et al. 2017). Initiatives such as the Plant 199 Phenology Ontology (PPO) working group are currently structuring phenological terms for more 200 201 uniform application across studies (Willis et al. 2017).

Studies in temperate regions have used varying methods to determine flowering status for
species with long flowering durations. For example, Molnár et al. (2012) and Bertin (2015)
excluded species that flowered outside of the peak flowering season of the region, defined as the
period from late-spring to early-summer. Molnár et al. (2012) removed one species because its
peak flowering date was in September and focused on 40 other taxa that had flowering peaks
from in spring and early-summer. The excluded species was a strong outlier and it was

suggested that autumn climate events may affect species differently than spring climate events 208 (Molnár et al., 2012). Park (2012) also removed outlier records when flowering records fell 209 outside the peak regional flowering season. Flowering records before Julian day 45 and after 210 Julian day 310 were removed from analyses to reduce biases caused by winter flowering species. 211 Additionally, Park (2012) removed records that were 150 days after the median flowering date 212 213 for each species to reduce errors caused by any second flowerings that can happen in autumn months. Several other studies removed taxa with long flowering durations to reduce variance 214 among species. Bertin (2015) excluded native weedy species with flowering durations from 215 spring to fall. Gallagher et al. (2009) only used species with a flowering duration of less than 216 three months. Panchen et al. (2012) chose to use only species with clear beginning and ending 217 points to investigate long and short flowering duration. Panchen et al. (2012) found that plants 218 with shorter flowering durations required smaller sample sizes to produce significant results 219 when regressing flowering day on year. 220

Other studies such as Calinger et al. (2013) and Lavoie & Lachance (2006) disregarded the effect 221 of flowering duration and noted the results of Primack et al. (2004), which reported no bias 222 associated with long or short flowering durations when mean estimations are analyzed. Plants in 223 tropical regions often have long flowering durations (van Schaik 1993; Fenner 1998), but as long 224 as flowering is not continuous throughout the year, methods applied to temperate regions should 225 also yield valuable insight into effects of climate change on phenology in the tropics. While 226 227 studies using herbarium specimens to analyze long-term changes have been limited to temperate regions, future studies could use circular statistics to analyze long-term phenological changes in 228 229 tropical regions (Fisher 1993; Morellato et al. 2010). Circular statistics have been used to 230 analyze flowering phenology in several tropical field studies, but these studies lacked long-term

Manuscript to be reviewed

- climate change analyses (Novotny & Basset 1998; Morellato et al. 2000; Cruz et al. 2006;
- Rogerio & Araujo 2010; Tesfaye et al. 2011; Nadia et al. 2012; Nazareno & dos Reis 2012;
- 233 Staggemeier et al. 2012; Carvalho & SartorWe2015; Kebede & Isotalo 2016).
- 234

235 Averaging temperatures

236 The foundational study by Primack et al. (2004) examined temperature averages from three calendar months prior to the specimen flowering date, with the assumption that flowering date is 237 a function of temperatures experienced in past months. Field investigations such as Fitter et al. 238 (1995) have shown temperature averages from different sets of months preceding flowering 239 affected flowering phenology in different ways. More recently, Calinger et al. (2013) chose to 240 regress the month of flowering with temperature averages from each of the eleven months prior 241 to flowering. They found that temperature averages from three months prior to the date of 242 flowering showed the strongest correlations with flowering (Calinger et al. 2013). Robbirt et al. 243 244 (2011) investigated three sets of temperature averages over three month intervals and also found that three months prior to flowering had the most predictive power. Similarly, Rawal et al. 245 (2015) regressed flowering on temperature averages for each species from 1, 3, 6, 9, and 12 246 247 months prior to flowering, because responses can vary by species. Rawal et al. (2015) also found that mean temperatures three months prior had the greatest influence on flowering time for 248 249 all species. 250 Other studies have used average temperatures from spring months because spring temperatures

- 251 generally have the most predictive power for flowering date (Miller-Rushing & Primack 2008;
- 252 Primack et al. 2009; Robbirt et al. 2011; Calinger et al. 2013; Park 2014; Park & Schwartz
- 253 2015). Bertin (2015) found an interesting trend that supported the effect of spring temperatures:

the earlier a species' mean flowering time occurred in the spring, the more the species' mean 254 dates had shifted toward an earlier day of year over time. Robbirt et al. (2011) also found the 255 highest correlations of flowering day with spring temperature averages across March, April and 256 May. Calinger et al. (2013) found significant changes in flowering in response to average spring 257 temperatures (February-May) but not in response to summer temperatures (June to September). 258 259 Gaira et al. (2011) found the highest correlations between flowering and temperatures in earlier months from December-February in a Himalayan perennial. As an alternative to using mean 260 monthly temperatures, Diskin et al. (2011) investigated the averages of temperature anomalies, 261 or deviations from the overall long-term mean, for 2, 3, and 6 month periods from January to 262 June and found averages from six months prior to flowering had the strongest correlations. Park 263 (2014) used temperature averages across three month periods from early spring to late summer 264 and found a similar trend. Temperature averages were organized into early, mid, and late 265 seasonal classes within the months of February to October. Park (2014) found warming 266 267 temperatures had affected species in the early spring class more than other classes. Park & Schwartz (2015) also used early, mid and late seasonal classes for spring and summer and found 268 that mid-season phenology events should be modeled differently than early or late season events. 269 270 Hart et al. (2014) used annual temperatures and temperatures from each season (spring, summer, fall, and winter) and found significant correlations for annual and fall temperature averages, but 271 272 with opposite effects. Hart et al. (2014) discussed that warmer fall temperatures may delay the 273 chilling requirement for *Rhododendron* species, resulting in a delay in flowering while warmer annual temperatures will lead to advances in flowering overall. Other studies found annual 274 275 temperature means were as useful as spring temperatures. Davis et al. (2015) found similar 276 results between spring and annual temperature averages and used annual averages in analyses.

Gallagher et al. (2009) also used annual temperature means for analyses and explained thatseasonal means were correlated with annual means.

We recommend investigating the effect of temperature by analyzing averages from multiple sets 279 of months prior to flowering for each species rather than using only one fixed spring interval or 280 only annual temperatures (Diskin et al. 2011; Robbirt et al. 2011; Calinger et al. 2013). Caution 281 282 should be taken when analyzing temperature averages from the same months prior to flowering for all species when flowering month varies by species. For example, when analyzing the effect 283 of temperature averages from three months prior for all species, Calinger et al. (2013) found that 284 285 for many species, flowering was correlated with temperatures three months earlier, yet for species with an earlier mean flowering day in April, January temperatures (three months prior to 286 flowering) did not predict flowering date; instead, temperature averages from the months of 287 February, March and April were better predictors for those species. 288

289

290 Geographic variation

Among species that have broad geographic ranges, differences in climate in different parts of the 291 species' range can complicate attempts to correlate a species' flowering day with temperature. 292 293 Several methods have been used to account for climate variability across a species' range. An early study by Lavoie & Lachance (2006) investigated the effects of climate variation on the 294 phenology of Coltsfoot (Tussilago farfara L.) across a range of about 10,000 km2 in Quebec, 295 296 Canada. Temperature data from 88 meteorological stations were averaged together across this range. To account for early snow cover melt in the southern part of this range, flowering dates 297 from individuals in southern locations were normalized with individuals in northern locations by 298 299 subtracting extra periods of snow cover from individuals in the north. The adjusted dates

indicated flowering occurred 33 days earlier over the last century while original (unadjusted) 300 dates indicated flowering occurred 19 days earlier over the last century. 301 302 While the study by Lavoie & Lachance (2006) adjusted actual dates for analyses, more recent studies mostly account for climate variation using georeferenced climate data at various scales. 303 Calinger et al. (2013) accounted for climate variation across Ohio by using temperature averages 304 305 from ten US Climate Divisions across the state, each about 8,000 km2. A total of 344 Climatic Divisions were established across the contiguous United States in 1895 in order to monitor 306 climate records more accurately. These divisions have now accumulated about 100 years of 307 climate records (Guttman & Quayle 1996). A later study by Park (2014) used average 308 temperatures across the U.S. county where each specimen was collected. 309 Other studies accounted for climate variation across longitude, latitude, or elevation. Robbirt et 310 al. (2011) analyzed the geographical effect of longitude and found that flowering occurred 4.86 311 days earlier per degree of longitude in a westward direction across the southern coastal counties 312 313 of England (Robbirt et al. 2011). A later study by Bertin (2015) used Hopkins' bioclimatic law to normalize dates on specimens. Hopkins' (1918) generally stated that for every increase in a 314 degree of latitude, or increase of 121.92 m elevation, the life history events of plants and animals 315 316 were delayed by four days. Bertin (2015) found consistencies with Hopkins' bioclimatic law using latitude and elevation and chose to normalize flowering dates by adding expected 317 318 phenological deviations from both latitude and elevation. Gaira et al. (2011) also analyzed 319 climate variation using elevation when temperature data were not available, assuming a 6.5°C 320 change in temperature per 1000 m change in elevation in the Himalayan region. 321 Other studies used temperature averages across large regions. Li et al. (2013) used temperature 322 data that was averaged from 36 meteorological stations across the Tibet Autonomous Region.

Molnár et al. (2012) used temperature averages from 10 meteorological stations across Hungary 323 and stated that the data were statistically indistinguishable across stations (~93,030 km²). Park & 324 Schwartz (2015) averaged temperatures from 13 stations across South Carolina, USA (~82, 931 325 km²). A later study by Robbirt et al. (2014) used temperature averages from an area between 326 Bristol, Preston, and London, across the United Kingdom (~17, 000 km²). Robbirt et al. (2014) 327 328 used geographical divisions called Watsonian vice-counties specifically delineated for the purposes of collecting scientific data, much like the US Climate Divisions. Robbirt et al. (2014) 329 found temperature averages were sufficient because climate variation across the Watsonian vice-330 counties used in their study did not significantly differ. 331 In order to more accurately estimate temperature averages across a region, recent studies have 332 used Geographical Information Systems (GIS) to project finer-scale climate layers across a 333 region and extract temperature data from specific Global Positioning System (GPS) points. 334 Gallagher et al. (2009) referenced GPS locations for each specimen and extracted the 335 336 temperature averages at specimen GPS points from a gridded map of temperature averages across Australia (~5 km2 resolution). Hereford et al. (2017) also extracted climate data from 176 337 collection locations in order to analyze species distributions and phenology. Rawal et al. (2015) 338 339 used the nearest data point from gridded climate averages across Victoria, Australia. Edward & Still (2008) analyzed the climate envelopes of grasses by assigning GPS points to herbarium 340 341 specimen locations in order to extract temperature averages from gridded climate maps $(250m^2)$ 342 resolution). Kosanic et al. (2018) manually geo-referenced locations using herbarium specimen localities and provided a methodology for assigning GPS coordinates when analyzing species 343 344 distributions and phenology. Standardizing methods for geo-referencing localities of herbarium 345 records without GPS coordinates could allow for more specimen data and larger sample sizes.

Bloom et al. (2017) developed a comprehensive protocol for standardizing spatial accuracy of 346 geo-referenced specimen localities for species distributions. 347 Future studies of phenology could benefit from such geo-referencing methods because several 348 phenology studies only included data from specimens with GPS coordinates. Studies using GPS 349 data are able to account for climate variation with higher resolution, although accuracy still 350 351 depends on the underlying empirical data and modeling approach used to generate GIS climate layers. 352 We recommend using the most spatially precise temperature data available, such as climate 353 divisions (Calinger et al. 2013; Robbirt et al. 2015) rather than state or region averages (Li et al. 354 2013; Park & Schwartz 2015). Using GPS specimen data to identify local climate conditions 355 from GIS climate layers (Gallagher et al. 2009; Edward & Still 2008) is also now generally more 356 precise and convenient in comparison to making generic and coarse-scale corrections for 357 latitude, longitude or elevation (Gaira et al. 2011; Robbirt et al. 2011; Bertin 2015). If 358 temperature averages from larger areas are used, we recommend testing for climate variability 359 across smaller divisions before using averages across the larger area (Lavoie & Lachance 2006; 360 Molnár et al. 2012; Robbirt et al. 2015). 361

362

363 Validation: herbarium specimens versus field observations

Field data are often combined with herbarium specimen data in analyses, allowing for

365 comparison and sometimes allowing for validation of conclusions based on herbarium data

366 (Primack et al., 2004; Miller-Rushing et al. 2006; Bertin 2015). Primack et al. (2004) used

367 herbarium specimens for historic data and field observations for current data and combined the

two in analyses (Primack et al., 2004). Studies by Miller-Rushing et al. (2006) and Bertin (2015)

also compared herbarium specimen data with field observations. Miller-Rushing et al. (2006)
found that phenology inferences from herbarium specimens alone differed from the combined
data by only about one day.

An early study by Borchert, (1996) found that herbarium specimen data produced slightly longer 372 flowering durations than field data, but noted that durations were mostly similar overall. 373 374 Borchert, (1996) and Rivera & Borchert (2001) found phenology data from field sites largely overlapped that of herbarium specimens with only slight differences. The negligible differences 375 between herbarium specimen data and field data in these studies helped justify the use of 376 377 herbarium specimen data to analyze phenology in more recent studies. Nevertheless, several more recent studies specifically compared phenology estimates from field data to those made 378 from herbarium specimens. 379

Bolmgren & Lonnberg (2005) compared herbarium specimen data directly to field data and 380 found the two data sets were overall highly correlated with only minor differences. For example, 381 herbarium specimens showed a slightly earlier mean flowering for spring-flowering plants than 382 field data, but the difference was not significant (Bolmgren & Lonnberg, 2005). Later studies by 383 Robbirt et al. (2011) and Davis et al. (2015) also primarily focused on testing the validity of 384 using herbarium specimen data. Robbirt et al. (2011) used a principal axis regression analysis to 385 compare herbarium derived peak-flowering dates with field derived peak-flowering dates and 386 found a high degree of correlation. Robbirt et al. (2011) discussed how the high degree of 387 388 correlation between herbarium and field data also supports the notion that geographically different records will not significantly alter the robustness of either data set. A study by Davis et 389 al. (2015) used a paired t-test to compare mean first-flowering day between herbarium specimens 390

and field data and found no statistical difference. Davis et al. (2015) concluded that both 391 specimen and field data could be combined and used as a whole. 392 In order to increase sample sizes, Molnár et al. (2012) added about 2000 field observations to 393 about 5000 herbarium records, resulting in 70% herbarium records for the study. Similarly, 394 Panchen (2012) added about 2000 field records to about 1500 herbarium records, for a total of 395 396 43% herbarium records for the study. Searcy (2012) combined herbarium specimen and field data and then split the combined data into two time periods (1863–1935 and 1994–2008). 397 Herbarium specimen data may provide some advantages over field data. Bolmgren & Lonnberg 398 (2005) and Primack et al. (2004) noted that using herbarium specimens conserves time and 399 resources, especially when species are located in difficult to access geographical areas such as 400 mountain peaks or islands. Herbarium specimens are also collected over a greater period of time 401 from a larger geographical area while field data are often from specific localities over a shorter 402 time period (Primack et al. 2004; Bolmgren & Lonnberg 2005; Bertin 2015; Davis et al. 2015). 403 404 Herbarium specimens also provide long-term records that are widely accessible for multiple studies. Despite criticisms, herbarium specimen data have been shown to produce similar 405 enough results to field data that herbarium specimen data are now widely accepted in 406 407 phenological studies.

408

409 Conclusions

The use of herbarium specimens for the investigation of flowering phenology has grown
considerably during the past decade. As efforts to produce digital copies of specimens and label
information have amassed large datasets, new approaches for analyzing responses to climate
change are rapidly becoming available. Although small sample sizes have often been used in

early studies of phenology, various factors, such as wide geographic range, may require larger 414 sample sizes. Based on recent validations, estimations of mean-flowering should be used rather 415 than first-flowering because estimates of first-flowering are more sensitive to sampling. 416 Statistically modeling the start of the flowering process appears to be another promising 417 approach to investigating how climate change has affected the beginning of a flowering cycle 418 419 (Pearse et al. 2017). The use of GPS data appears to be the way forward for the advancement of methods in the study of phenology. GPS point data allow for correspondence with higher 420 resolution temperature data in climatically diverse geographical regions. Studies using 421 422 herbarium specimen data will continue to help us understand the impact of recent climate change on plant reproductive phenology. Other aspects of plant phenology that can be analyzed using 423 herbarium specimens, such as fruit ripening and spring leaf emergence, have important 424 implications for higher trophic levels, which may include rare animals dependent on plant 425 resources (Everill et al. 2014; Zohner & Renner 2014; Mendoza et al. 2017). Studies using 426 herbarium specimens have become an asset for long-term climate change vulnerability 427 assessment. These studies have begun to analyze the effects of climate change on community 428 composition (Miller-Rushing & Primack 2008, Park 2014), species distribution (Hereford et al. 429 430 2017; Kosanic et al. 2018), coevolved plant pollinator relationships (Molnár et al. 2012; Robbirt et al. 2014), functional groups (Miller-Rushing & Primack 2008; Panchen et al 2012; Calinger et 431 432 al. 2013; Munson and Long 2017), and phylogenetic relationships (Bolmgren & Lonnberg 2005; 433 Molnár et al. 2012; Primack and Miller-rushing 2012). Future studies investigating phylogenetic signals and plasticity are needed to further improve our understanding of adaptation and 434 435 resilience to climate change. As temperatures continue to rise globally, herbarium specimens 436 will continue to be crucial resources for analyzing phenological responses to climate change.

Manuscript to be reviewed

437								
438	References							
439	Bertin, R.I. 2015. Climate change and flowering phenology in Worcester County, Massachusetts.							
440	International Journal of Plant Sciences 176: 107–119.							
441								
442	Bloom, T.D.S., A. Flower, E.G. DeChaine. 2018. Why georeferencing matters: Introducing a							
443	practical protocol to prepare species occurrence records for spatial analysis. Ecology and							
444	Evolution 8:765–777.							
445								
446	Bolmgren, K., and K. Lonnberg. 2005. Herbarium data reveal an association between fleshy fruit							
447	type and earlier flowering time. International Journal of Plant Sciences 166: 663-670.							
448								
449	Borchert, R. 1996. Phenology and flowering periodicity of neotropical dry forest species:							
450	evidence from herbarium collections. Journal of Tropical Ecology 12: 65-80.							
451								
452	Calinger, K.M., S. Queenborough, and P.S. Curtis. 2013. Herbarium specimens reveal the							
453	footprint of climate change on flowering trends across North-Central North America. Ecology							
454	Letters 16: 1037–1044.							
455								
456	Daru, B.H., M. van der Bank, and T. J. Davies. 2017. Unravelling the evolutionary origins of							
457	biogeographic assemblages. Diversity and Distributions 66: 1014-1016							
458								

459	Davis, C.C., C.G. Willis, B. Connolly, C. Kelly, and A.M. Ellison. 2015. Herbarium records are							
460	reliable sources of phenological change driven by climate and provide novel insights into							
461	species' phenological cueing mechanisms. American Journal of Botany 102: 1599-1609.							
462								
463	Demarée, G.R., and Rutishauser, T., 2009. Origins of the word "phenology". Eos Transactions							
464	American Geophysical Union 90, 291							
465								
466	Diskin, E., H. Proctor, M. Jebb, T. Sparks, and A. Donnelly. 2012. The phenology of Rubus							
467	fruticosus in Ireland: herbarium specimens provide evidence for the response of phenophases to							
468	temperature, with implications for climate warming. International Journal of Biometeorology 56:							
469	1103–1111.							
470								
471	Edwards, E.J., and C.J. Still. 2008. Climate, phylogeny and the ecological distribution of C4							
472	grasses. Ecology Letters 11:266–76.							
473								
474	Everill, P.H., R.B. Primack, E.R. Ellwood, E.K. Melaas. 2014. Determining past leaf-out times							
475	of New England's deciduous forests from herbarium specimens, American Journal of Botany.							
476	101:1293-1300							
477								

Fenner, M. 1998. The phenology of growth and reproduction in plants. Perspectives in Plant
Ecology, Evolution and Systematics 1: 78–91.

481	Fitter, A. H., Fitter, R. S. R., Harris, I. T. B., Williamson, M.H. 1995. Relationships between
482	first-flowering date and temperature in the flora of a locality in central England. Functional
483	Ecology 9: 55–60.

484

Fisher, N. I. 1993. Statistical analysis of circular data. Cambridge University Press, Cambridge

487 Forrest, J., and A.J. Miller-rushing. 2010. Toward a synthetic understanding of the role of

488 phenology in ecology and evolution. Philosophical Transactions of the Royal Society B 365:

489 3101-3112.

1738.

490

Gaira, K.S., R.S. Rawal, B. Rawat, and D. Bhatt. 2014. Impact of climate change on the
flowering of Rhododendron arboreum in central Himalaya, India. Current Science 106: 1735–

494

493

Gaira, K.S., U. Dhar, and O.K. Belwal. 2011. Potential of herbarium records to sequence
phenological pattern: a case study of Aconitum heterophyllum in the Himalaya. Biodiversity and
Conservation 20: 2201–2210.

498

Gallagher, R. V., L. Hughes, and M.R. Leishman. 2009. Phenological trends among Australian
alpine species: using herbarium records to identify climate-change indicators. Australian Journal
of Botany 57: 1–9.

503 Guttman, N. B., and R.G. Quayle, 1995. A historical perspective of U.S. Climate Divisions.

504 Bulletin of the American Meteorological Society 77: 293–303.

505

- 506 Haggerty B.P, A.A. Hove, and S. J. Mazer 2012. Primer on herbarium-based phenological
- research. Guide for college and public audiences for understanding the use of preserved plants in

508 climate change research. California Phenology Project.

509

- 510 Hart, R., J. Salick, S. Ranjitkar, and J. Xu. 2014. Herbarium specimens show contrasting
- 511 phenological responses to Himalayan climate. Proceedings of the National Academy of Sciences

of the United States of America 111: 10615–9.

513

514 Hereford, J., J. Schmitt, and D.D. Ackerly. 2017. The seasonal climate niche predicts phenology

and distribution of an ephemeral annual plant, *Mollugo verticillata*. Journal of Ecology 105,
1323–1334.

517

Hopkins, A.D. 1918. Periodical events and natural law as guides to agricultural research and
practice. U. S. Department Of Agriculture Weather Bureau, Monthly Weather Review 643: 5–

520 42.

521

Kosanic, A., K. Anderson, S. Harrison, T. Turkington, and J. Bennie. 2018. Changes in the
geographical distribution of plant species and climatic variables on the West Cornwall peninsula

524 (South West UK). PLoS one. 13: e0191021. https://doi.org/10.1371/journal.pone.0191021

526	Lamoureux, C.H. 1973. Phenology and reproduction of Hawaiian plants, a preliminary report.
527	Island Ecosystems, US International Biological Program, Technical Report No 24.
528	
529	Lavoie, C., and D. Lachance. 2006. A new herbarium-based method for reconstructing the
530	phenology of plant species across large areas. American Journal of Botany 93: 512-516.
531	
532	Li, Z., N. Wu, X. Gao, Y. Wu, and K.P. Oli. 2013. Species-level phenological responses to
533	"global warming" as evidenced by herbarium collections in the Tibetan Autonomous Region.
534	Biodiversity and Conservation 22: 141–152.
535	
536	Linnaeus C. 1751, Philosophia botanica, Stockholm
537	
538	Mendoza, I., C.A. Peres, and L.P.C. Morellato. 2016. Continental-scale patterns and climatic
539	drivers of fruiting phenology: a quantitative neotropical review. Global and Planetary Change.
540	148: 227–241
541	
542	Miller-Rushing, A.J., and R.B. Primack. 2008. Global warming and flowering times in Thoreau's
543	Concord: a community perspective. Ecology 89: 332–341.
544	
545	Miller-Rushing, A.J., R.B. Primack, D. Primack, and S. Mukunda. 2006. Photographs and
546	herbarium specimens as tools to document phenological changes in response to global warming.
547	American Journal of Botany 93: 1667–1674.
548	

549	Molnár V, A., J. Tökölyi, Z. Végvári, G. Sramkó, J. Sulyok, and Z. Barta. 2012. Pollination							
550	mode predicts phenological response to climate change in terrestrial orchids: a case study from							
551	Central Europe. Journal of Ecology 100: 1141–1152.							
552								
553	Morellato, L. P. C. Alberti, L. F., and Hudson, I. L. 2010. Applications of circular statistics in							
554	plant phenology; a case studies approach. In Phenological Research; Methods for Environmental							
555	and Climate Change Analysis (ed. I. L. Hudson and M. R. Keatley), pp. 339-59.							
556								
557	Morren C. 1853. Souvenirs phénologiques de l'hiver 1852–1853. Bull de l'Académie royale des.							
558	Sciences, des Lettres et des Beaux-Artes de Belgique 20:160-186.							
559								
560	Moussus, J., R. Julliard, and F. Jiguet. 2010. Featuring 10 phenological estimators using							
561	simulated data. Methods in Ecology and Evolution 1: 140–150.							
562								
563	Munson, S.M., and A.L. Long. 2017. Climate drives shifts in grass reproductive phenology							
564	across the western USA. New Phytologist 213: 1945–1955.							
565								
566	Neil, K.L., L. Landrum, and J. Wu. 2010. Effects of urbanization on flowering phenology in the							
567	metropolitan phoenix region of USA: findings from herbarium records. Journal of Arid							
568	Environments 74: 440–444							
569								

570	Panchen, Z.A.	, R.B. Primac	k, T. Aniśko	, and R.E. Lyor	ns. 2012.	Herbarium sp	becimens,
		/		/			

- 571 photographs, and field observations show Philadelphia area plants are responding to climate
- change. American Journal of Botany 99: 751–756.
- 573
- 574 Papadopulos, A.S.T., and W.J. Baker. 2011. Speciation with gene flow on Lord Howe Island.

575 Proceeding of the National Academy of Sciences 108: 1–6.

576

- 577 Park, I.W. 2012. Digital herbarium archives as a spatially extensive, taxonomically discriminate
- 578 phenological record; a comparison to MODIS satellite imagery. International Journal of

579 Biometeorology 56: 1179–1182.

580

581 Park, I.W. 2014. Impacts of differing community composition on flowering phenology

throughout warm temperate, cool temperate and xeric environments. Global Ecology and

583 Biogeography 23: 789–801.

584

585 Park, I.W., and M.D. Schwartz. 2015. Long-term herbarium records reveal temperature-

dependent changes in flowering phenology in the Southeastern USA. International Journal of

587 Biometeorology 347–355.

588

Parmesan, C., C. Parmesan, G. Yohe, and G. Yohe. 2003. A globally coherent fingerprint of
climate change impacts across natural systems. Nature 421: 37–42.

- 592 Pearse, W.D., C.C. Davis, D.W. Inouye, R.B. Primack & T. J. Davies. 2017. A statistical
- estimator for determining the limits of contemporary and historic phenology. Nature Ecology &Evolution 1, 1876–1882
- 595
- 596 Primack, D., C. Imbres, R.B. Primack, A.J. Miller-Rushing, and P. Del Tredici. 2004. Herbarium
- 597 specimens demonstrate earlier flowering times in response to warming in Boston. American
- 598 Journal of Botany 91: 1260–1264.
- 599
- 600 Primack, R.B., and A.J. Miller-Rushing. 2009. The role of botanical gardens in climate change
- research. New Phytologist 182: 303–313.
- 602
- Primack, R.B., and A.J. Miller-rushing. 2012. Uncovering, collecting, and analyzing records toinvestigate the ecological impacts of climate change: a template from Thoreau's Concord.
- 605 BioScience 62: 170–181.

606

- Puppi, G. 2007. Origin and Development of Phenology as a Science. Italian Journal of
 Agrometeorology 12: 24–29.
- 609
- 610 Rawal, D.S., S. Kasel, M.R. Keatley, and C.R. Nitschke. 2015. Herbarium records identify
- sensitivity of flowering phenology of eucalypts to climate: implications for species response to
- 612 climate change. Austral Ecology 40: 117–125.

614	Rivera, G., and R. Borchert. 2001. Induction of flowering in tropical trees by a 30-min reduction
615	in photoperiod: evidence from field observations and herbarium specimens. Tree Physiology 21:
616	201–212.

617

618 Robbirt, K.M., A.J. Davy, M.J. Hutchings, and D.L. Roberts. 2011. Validation of biological

collections as a source of phenological data for use in climate change studies: a case study withthe orchid Ophrys sphegodes. Journal of Ecology 99: 235–241.

621

622 Robbirt, K.M., D.L. Roberts, M.J. Hutchings, A.J. Davy, R.B. Gardens, and S. Tw. 2014. Report

623 potential disruption of pollination in a sexually deceptive orchid by climatic change. Current

624 Biology 24: 2845–2849.

625

- 626 Searcy, K.B. 2012. Changes in the flora of the Mount Holyoke Range, Hampshire Co.,
- 627 Massachusetts over the past 150 years (1860–2010). Rhodora 114: 113–132.
- van Proosdij, A.S.J., M.S.M. Sosef, J.J. Wieringa, and N. Raes. 2016. Minimum required number
- of specimen records to develop accurate species distribution models. Ecography 39: 542–552.
- 631 van Schaik, C.P., J.W. Terborgh, and S.J. Wright. 1993. The phenology of tropical forests -
- adaptive significance and consequences for primary consumers. Annual Review of Ecology and
- 633 Systematics 24: 353–377.

- 635 von Linné, C. 2003. Linnäus' Philosophica Botanica. New York: Oxford University Press
- 636 (Translation by S. Freer).

Manuscript to be reviewed

637

Walther, G.R., E. Post, P. Convey, A. Menzel, C. Parmesan, T.J.C. Beebee, and J.M. Fromentin.
2002. Ecological responses to recent climate change. Nature 416: 389–395.

640

- 641 Willis, C.G., E.R. Ellwood, R.B. Primack, C.C. Davis, K.D. Pearson, A.S. Gallinato, J.M. Yost,
- 642 G. Nelson, S.J. Mazer, N.L. Rossington, T.H. Sparks, and P.S. Soltis. 2017. Old plants, new
- tricks: phenological research using herbarium specimens. Trends in Ecology & Evolution 32:531–546.

645

646 Zalamea, A.P., F. Munoz, P.R. Stevenson, C.E.T. Paine, C. Sarmiento, D. Sabatier, and P.

Heuret. 2016. Continental-scale patterns of Cecropia reproductive phenology: evidence from
herbarium specimens. Proceedings of the Royal Society 278: 2437–2445.

649

- 50 Zhao, M.F., C.H. Peng, W.H. Xiang, X.W. Deng, D.L. Tian, X.L. Zhou, and G.R. Yu. 2013.
- 651 Plant phenological modeling and its application in global climate change research: overview and

future challenges. Environmental Reviews 21: 1–14.

- Zohner, C.M. and S.S. Renner. 2014. Common garden comparison of the leaf-out phenology of
- woody species from different native climates, combined with herbarium records, forecasts long-
- 656 term change. Ecology Letters. 17, 1016–1025 (2014)

Table 1(on next page)

Methods of studies

The column "Flw Day ~ Temp" represents studies that conducted a type of regression analysis with flowering day (Flw Day) as the dependent variable and temperature average (temp) or year as the independent variable. The " $\Delta \bar{x}$ " symbol represents studies that analyzed a difference in the mean flowering day between historic and current time period groups rather than using a type of regression analysis **Reference** to the column "flw ~ temp" represents studies that conducted a type of regression analysis with flowering day (flw) as the dependent variable and temperature average (temp) or year as the independent variables. The " Δx " symbol represents studies that analyzed a difference in the mean flowering day between historic and current time period groups rather than using a type of regression analysis.

		specimen					
species	ecies specimens per species		authors	authors year geographic regi		$(flw \sim temp)$	$(flw \sim year)$
1	117	117	Gaira et al.	2011	Eastern Himalayas		Х
1	N/A	N/A	Gaira et al.	2014	Eastern Himalayas	Х	Х
1	192	192	Robbirt et al.	2011	Northern Europe	Х	
5	158	32	Rawal et al.	2015	Southern Australia	Х	Х
5	540	108	Diskin et al.	2012	Northern Europe	Х	Х
20	371	19	Gallagher et al.	2009	Southern Australia	Х	Х
20	1108	55	Davis et al.	2015	North America	Х	Х
28	1587	57	Panchen et al.	2012	North America	Х	Х
36	460	13	Hart et al.	2014	Eastern Himalayas	Х	
>37	372	10	Primack et al.	2004	North America	Х	Х
39	216	6	Lavoie & Lachange	2006	North America		Х
39	5424	139	Molnár et al.	2012	Northern Europe		Х
41	909	22	Li et al.	2013	Eastern Himalayas	Х	Х
42	142	3	Miller-Rushing et al. Primack &	2006	North America	Х	Х
43	N/A	N/A	Miller-Rushing	2012	North America		Х
87	N/A	N/A	Neil et al.	2010	North America		Х
141	5053	36	Calinger et al.	2013	North America	Х	
186	30,000	161	Bertin	2015	North America		$\Delta \bar{x}$
370	1125	3	Searcy	2012	North America		$\Delta \bar{x}$
1185	5949	5	Park	2012	North America		Х
>1700	19,328	11	Park	2014	North America	х	
24,105	823,033	34	Park & Schwartz	2015	North America	Х	Х