A primer on herbarium-based phenological research

Brian Haggerty, Alisa Hove, and Susan Mazer

University of California, Santa Barbara

Funding for the development of these materials was provided by the US Geological Survey and the USA National Phenology Network

PHENOLOGICAL LITERACY

UNDERSTANDING THROUGH SCIENCE & STEWARDSHIP

Table of contents

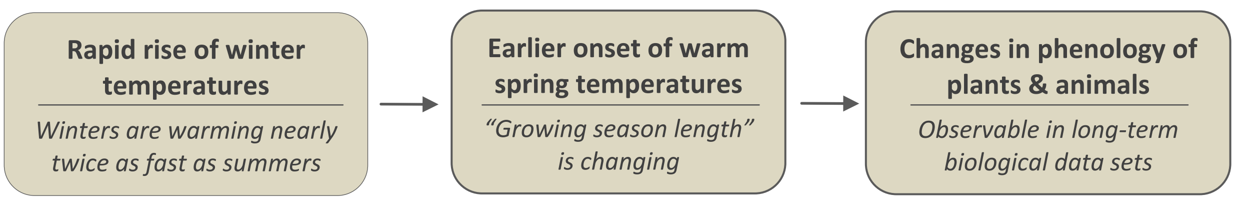
More phenology education materials and activities are available online, including a step-by-step guide to downloading and analyzing a real herbarium-based phenological data set using Microsoft Excel. Other available materials include annotated lectures for universities and the public; guides to establishing and using phenology gardens; seminar modules for undergraduate/graduate students; and standards-aligned K-12 lesson plans. To learn more and to download materials, visit the Education section of the California Phenology Project website (www.usanpn.org/cpp/education) or the USA National Phenology Network (www.usanpn.org/education).

All photos by Brian Haggerty

Background & Introduction

Our understanding of the global climate system has advanced remarkably in recent decades, along with our awareness that changes in the climate system have affected plant and animal activities and geographic distributions. Biologists have been able to detect some of the impacts of climate change on both wild and cultivated species by exploring data sets from long-term monitoring efforts. Repeated annual surveys of events such as wildflower blooms, animal migrations, insect emergence, and crop harvests – which have been monitored for many reasons – have informed our understanding of the biological effects of climate change.

In a landmark 2007 report, the Intergovernmental Panel on Climate Change (IPCC)[1] summarized 28,671 long-term data series from around the world documenting significant biological impacts of climate change. This report concluded, among other things, that "Phenology – the timing of seasonal activities of animals and plants – is perhaps the simplest process in which to track changes in the ecology of species in response to climate change." Although phenological responses to climate change (and the processes that drive them) differ among species and geographic regions, the general process is as follows:



Our understanding of phenological responses to climate change in North America, and our ability to predict them, are limited by the availability and scope of long-term data sets. For example, only 3.4% (or about 1,000) of the long-term data series that the IPCC summarized were generated in North America. While impressive progress has been made in the U.S. in recent years to establish a nationwide long-term biological monitoring program for academic, government, and citizen scientists (the USA National Phenology Network – www.usanpn.org), researchers have begun to seek additional sources of long-term historical data that may document phenological changes in plants and animals. These research projects help scientists to understand more completely species responses to climate change, and also help to illuminate the historical context in which phenological monitoring programs are being established. One of the most powerful and extensive sources of available historical data is herbarium records.

An **herbarium** is a research and education center much like a library or a museum, but where plants collected from wild or cultivated populations are curated and stored in a stable environment. After plants are collected, pressed, and dried, they are mounted on acid-free,

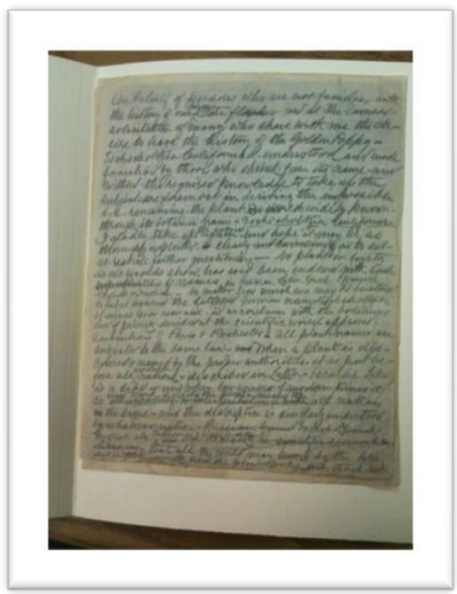
[1] Formed in 1988 by the World Meteorological Organization and the United Nations Environmental Programme, the IPCC is the international body of scientists assessing the available scientific, technical, and socio-economic information relevant to understanding climate change, its impacts, and options for mitigating its consequences on wild and managed systems.

museum-quality paper with a label that documents when and where the specimen was collected, who collected it, in what type of habitat it occurred, and any brief observations that the collector may have made about the species that co-occurred with the specimen. The specimen (or "sheet") is then assigned an accession number and stored with other specimens in a systematic manner in cabinets.

There are over 600 herbaria in the U.S., collectively containing over 260 million specimens. Most herbaria are located at universities and botanic gardens but they may also be found at National Parks and in private collections. Given the depth and breadth of collections, herbaria are perhaps our best repositories of historical information that can be used to study biological responses to environmental changes, including climate change.

In this guide we provide a brief background on herbaria (i.e., how they're organized and who uses them) before addressing how they have been used to study a wide range of topics, including phenology and climate change. Our goal is to inform readers about what to expect at and from an herbarium, how they can support the many activities of herbaria, and how to study phenology using herbarium specimens.

In a separate guide available on the Education section of the California Phenology Project website (www.usanpn.org/cpp/education), or from the USA National Phenology Network (www.usanpn.org/education), we provide step-by-step instructions on the basic processes of exploring long-term phenological data sets. Using a sample data set derived from California poppy (*Eschscholzia californica*) herbarium specimens collected between 1906-2009, you'll be guided step-by-step through the processes of organizing, summarizing, visualizing, and analyzing the data using Microsoft Excel.



You might be surprised what you'll come across at an herbarium. The photo on the left shows one herbarium sheet containing flowers of California poppy collected at one site, on one day, by one collector (think of the time it took to carefully press, dry, and glue each flower to the sheet!). The photo on the right is of a letter housed at the Jepson Herbarium at UC Berkeley in which Sara Plummer Lemmon – who is largely responsible for California poppy becoming California's state flower – described the many values of adopting this charismatic species as the state flower.

What is an herbarium, how is it used?

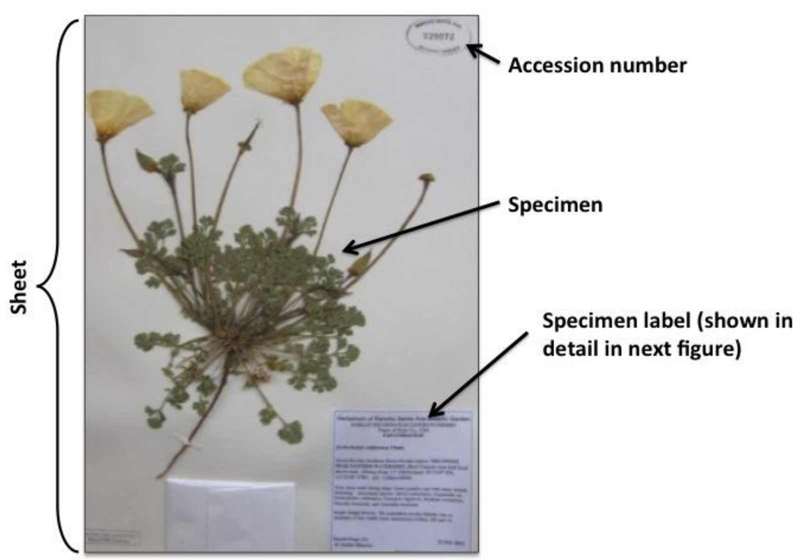
An herbarium is an institution that functions much like a hybrid between a museum and a library. While herbaria contain irreplaceable collections that are curated and stored in a systematic order, they also support a wide range of research and education projects by allowing access to the collections (with permission and training) and by providing loans to other herbaria. Additionally, given that herbaria are typically located at universities and botanic gardens where there's a broad interest in nature, educational outreach is the focus of many herbarium activities. There often is a community of botanists associated with the herbarium who organize collection trips and plant conservation efforts. In short, herbaria are hubs of botanical activity and education.

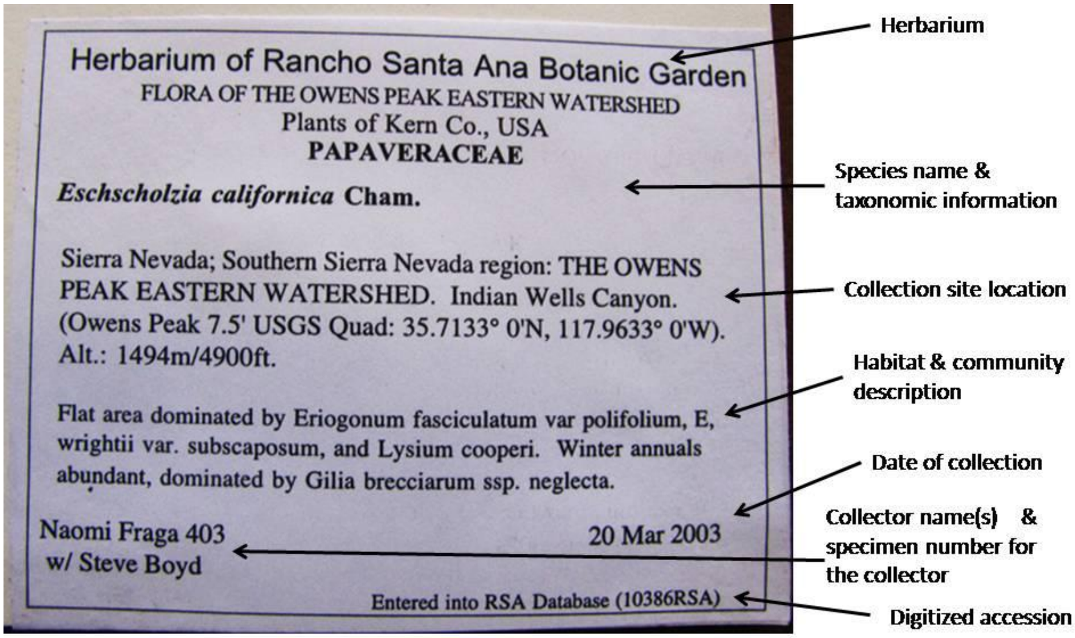
Herbarium collections

Herbarium specimens

An herbarium specimen, also referred to as a **sheet**, is the basic unit of organization and study in an herbarium. After plants are collected, pressed, and dried, they are mounted on museum-quality paper (a sheet). An information label also is mounted on the sheet, including information pertinent to the collection such as location, date, collector, and a description of the surrounding habitat. The sheet is then assigned an accession number and stored with others in a systematic manner in insect-resistant cabinets; specimens in the same species are held together in a large folder, and members of the same family and genus are also shelved together, usually in alphabetical order.

A sheet might have mounted on it an entire plant (including the roots, stems, leaves and reproductive structures), or just portions of a plant. In the latter case, it could be because the entire plant was too large to fit on the sheet, or because it was the only portion of a plant in which the collector was interested. Also found on the sheet might be a "fragment pack" – a small envelope into which broken stems, leaves, fruits, flowers, and/or are placed.





The format of herbarium sheet labels may differ among herbaria, collectors, and time periods. Most labels will include much of the information shown here, with some modern labels having GPS coordinates for the collection site as well (as shown here).

Organization of collections

Most herbaria follow the same basic convention in organizing their collections. In this taxonomic organizational scheme, specimens first are grouped in a folder by species. If several folders are filled for a species then each folder is given a number (subspecies and varieties also receive their own folder). Then, all the species in a genus are alphabetized; all the genera in a family are alphabetized, and the families are then alphabetized.

In some herbaria, all plant families are grouped together and then alphabetized regardless of the higher taxa (e.g., subfamily, order, class, or Division) to which they belong. In other herbaria, families are segregated among higher-order taxonomic groupings (i.e., monocot vs. eudicot, gymnosperm vs. angiosperm, lower plants vs. higher plants). Additionally, many herbaria contain specimens from all over the world. In these cases, all the folders of a given species are organized by geographic region – the first set of folders might represent plants collected in the contiguous U.S., with successive folders representing plants collected in North America and other continents. Each folder is then given a colored tag or symbol representing the geographic region of collection. There are variations among herbaria for these higher order organizational schemes, which is why it is essential to communicate with herbarium staff before accessing the collections.

Scenes from the herbarium at California Academy of Sciences (San Francisco, CA).

A) Specimens are spread out on a working desk in the collections room. Note the compactor system with rows of cabinets that can be compacted to increase storage area. B) An open cabinet with stacked folders. Here, the genus *Eschscholzia* is shown, with several folders per species arranged alphabetically. C) The history of botany in a region is often discernable from herbarium labels – here, a label created by two of California's famous botanists David Keck & Jens Clausen. D) A label from 1841 – notice how little collection information was included!

Types of collections

An herbarium may contain several types of botanical collections, reflecting the diverse research and collection interests of its contributors. Plants that have been collected for any reason in the past may appear on the shelves, long as they have been pressed, dried, mounted, labeled, and accessioned. There may also be various shelves and cabinets dedicated to plants that are in the process of being prepared for mounting and storage. Generally, an herbarium may have some or all of the following types of collections represented:

- **General collections –** Most specimens fall into this category. A specimen might be a single plant that a person collected for no other reason than that they wanted to preserve it. On the other hand, a specimen (or several specimens of the same or different species) might have been accessioned by a researcher to document that the plant was collected as part of a research project (called a voucher specimen). Some career botanists have contributed thousands of specimens from the regions that they have studied intensely.
- **Type specimens –** These are specimens that serve a critical role in defining species attributes. There are different kinds of "types" – for example, the *holotype* is the original specimen from which the species description was written, whereas the *neotype* is a replacement for a holotype that was destroyed.
- **Floristic surveys –** These are expeditions to specific geographic regions with the primary goal of documenting species occurrences and/or abundances. Generally, one to several specimens of each species in an area will be collected and accessioned. Floristic surveys have been used to write regional field guides, to document biological diversity, to locate species that may be used elsewhere for habitat restoration, to seek rare or endangered species, and to guide educational activities. Often, floristic surveys are conducted in regions that are threatened by human activities. For example, floristic surveys have been conducted throughout the U.S. preceding the construction of new housing developments destined to displace the indigenous plants. Strikingly, in the Three Gorges Dam region of China, a floristic survey was conducted before the dam was completed and the area flooded.
- **Rare plants –** Though it might seem odd to collect plants that are known to be rare, the collection can be performed so that only a small portion of a plant is collected, thereby documenting its occurrence without threatening its persistence.
- **Botanical structures –** While an accessioned specimen is likely to have several reproductive or vegetative structures visible, some botanists focus collections on a particular structure such as seeds, fruits, stems, or roots. For example, the U.S. National Seed Collection is housed at the USDA and comprises over 120,000 seed samples representing 390 plant families, 13,000 genera, and 27,000 species (www.ars.usda.gov/is/np/systematics/seeds.htm).
- **Liquid preservation –** These collections typically preserve plants in jars of alcohol or another preservative, and include aquatic plants or non-plant materials (some algae, fungi, or diatoms)
- **Non-plant materials –** Many herbaria contain additional collections of algae, lichens, fungi, or diatoms. Such collections often represent the specimens accumulated over a researcher's entire career, and are then donated by the investigator for research and education.

Storage of specimens varies among herbaria

Once properly curated and accessioned, herbarium specimens are stored in a stable environment that deters pests such as beetles, booklice, silverfish, and fungi. The ideal storage environment is one comprised of closed cabinets held at low humidity and low temperature (an example can be seen at University of California, Davis). This is prohibitively expensive for most herbaria, which may store specimens in closed cabinets that aren't climate controlled. In most herbaria, cabinet doors are lined with pest-deterring materials and chemicals.

The number of specimens in collections varies substantially across herbaria. Moderately sized herbaria might have 100,000 - 200,000 specimens and fill cabinets in a room approximately 20' x 40', whereas the largest herbaria in the U.S. have 3 million specimens and feel more like warehouses (the largest herbarium it the world – located in the National Museum of Natural History in Paris, France – has over 8 million specimens organized on several floors). Some herbaria have raised sufficient funds to install space-saving compacting cabinet systems, in which cabinets are placed on a linear wheel and pulley system that allows only two cabinets to be separated from each other (permitting their doors to open) at any one time.

Visiting an herbarium – what to expect

While the breadth, depth, and organization of collections vary considerably among herbaria, the process of visiting and using an herbarium is quite similar. In all cases, the first step is to contact someone at the herbarium in order to schedule a visit. It's possible that the species or collections in which you're interested have been loaned out to another herbarium, so it's worthwhile taking the time to check before arriving.

Upon your arrival, a member of the herbarium staff is likely to provide a brief orientation. This introduction gives them the opportunity to convey their local rules for use, and it gives you the orientation you'll need to access the collections efficiently. People managing herbaria are generally welcoming and excited to see people making use of their collections, and a brief conversation will set you off on the right foot. Some herbarium staff also have entertaining stories about their collections and the people who've used them!

Here are a few common-sense reminders about etiquette when visiting herbaria:

- Plan your visit and contact the herbarium in advance
- Always wash your hands before and after handling specimens
- Never have food or beverage out while working with specimens
- When pulling a folder out of a cabinet, place a "drop tag" or sheet of paper in its place so you replace the folder in the proper position. Remove only one folder at a time to ensure you are holding it securely.

- Keep cabinet doors closed at all times unless you are removing or replacing specimens.
- Treat each sheet like a sacred irreplaceable record (because in fact that's what it is!). Never flip through folders like you would a book; instead, carefully lift a sheet and place it nearby (with the specimen facing up), ensuring that it doesn't tilt or drag across the sheets below it.
- Don't touch the specimens unless you've confirmed that this is acceptable to the herbarium's curator or director. Also take care not to sneeze or cough on the sheets.
- If you find signs of insect damage or anything else that seems peculiar, alert herbarium staff. If there's a new pest outbreak, they'll want to treat it at the source (in the cabinet from which the specimen came) as well as on the sheet.

Activities at herbaria – much more than meets the eye!

Herbaria have been used by a wide array of people in a remarkable number of ways. A brief list of people using herbaria would include many types of scientists; educators; students; ethnobotanists; environmental consultants; veterinarians; forensic botanists (crime scene investigators); artists; and medical doctors. Some of the most common visitors to herbaria include:

- ✓ Plant taxonomists working to define evolutionary relationships among species;
- ✓ Plant ecologists trying to identify locations where particular species have been found in the past;
- ✓ Environmental consultants trying to determine whether particular species have been found in particular locations in the past (e.g., whether rare wetland species have been collected in the past from a proposed development site);
- ✓ Nature enthusiasts trying to learn the local flora in preparation for spending time outdoors or publishing a local field guide;
- ✓ Educators helping their students to learn about scientific collections and about plants in their area; and
- ✓ Researchers studying a wide range of topics including climate change.

Undergraduates in UC Santa Barbara's Botany Field & Lab course receive an introduction to the herbarium and the new compactor system.

One herbarium researcher maintains an ongoing list of herbarium uses; you can view Dr. Vicky Funk's complete list at this website: http://www.virtualherbarium.org/vh/100UsesASPT.html.

Several other important activities are taking place at many herbaria, in part due to a precipitous drop in funding levels in recent decades (this trend is true for natural history collections in general). To avoid the expense and risks associated with the long-distance loaning of herbarium specimens, many herbaria are creating internet-accessible databases of their collections by making digital scans of each sheet. The label information also is provided in these electronic databases, allowing researchers to search or to analyze large georeferenced data sets very quickly. Many herbaria also are joining together in consortia to enhance their collections management efforts as well as their collective probability of securing funding. For example, the Consortium of California Herbaria operates in ways that are similar to an interconnected library system, and it manages its georeferenced database for public access (http://ucjeps.berkeley.edu/consortium/).

In 2004 a group of researchers (Primack et al., 2004) published methods for using herbarium specimens to study phenology in the context of climate change. Since then, several other research teams have published papers testing those methods for various plants in different parts of the world. One general method seems to be emerging for data collection, although different approaches have been used for data analyses. We've taken care to distill this body of research into a brief summary. We encourage readers to obtain and read the original research papers.

Herbarium-based phenological research:
Establishing average annual phenophase windows for species & communities

Herbarium specimens can be used to determine the average time period during which phenophases are expected to occur. For example, herbarium records can help us to determine when allergy-inducing plants such as western ragweed (*Ambrosia psilostachya*) produce flowers in our region. This approach to “within-year” phenology is quite useful for many different research and educational goals:

- Researchers hoping to study a species’ current phenophases can use herbarium specimens to identify an approximate time window that can make their searching more efficient. In exploring the herbarium sheets, they also learn about the locations where each species has been found in the past.
- Botanists trying to create a local field guide
- Classrooms interested in learning more about their local flora
- Entomologists who are curious about the floral resources available to pollinators in a given area
- Place-based environmental education centers that want to promote upcoming events centered around plant phenology (*Come join us when the [insert species] are in full bloom!*).

The methods for this approach are quite simple; let’s use western ragweed flowering times as an example. To estimate the average time of year that western ragweed is flowering, we would sort through the available sheets and, for those specimens that have visible floral structures (specifically pollen-producing anthers), we would record the month and day of collection. Once all the specimens have been observed, we could then estimate an average date of flowering for the species.

If our interest was to estimate the *onset* of flowering of western ragweed, then we would become more precise about the flowering phenophase of each flowering specimen. Specifically, we would classify flowering specimens into categories representing plants that are:

(a) producing their first flowers (e.g., only the basal buds on their flowering stems are open),
(b) at or near peak flowering (e.g., ~50% of the flower buds distributed among their stems or flower clusters have opened), or
(c) producing their last flowers (e.g., the terminal flower buds on their branches are open).

After recording collection dates for specimens in each of these categories, we could then estimate the average onest, peak, and end of flowering. As a result, we would have a better estimate of when the ragweed allergy season begins, peaks, and finishes!

Additional research and education ideas:

- Following this simple approach, create a phenological calendar of local native plant species for one to several plant communities in your region. Combine this information with photos of each phenophase and species, and you've created a very useful field guide.
- Try completing this activity for other species in your region that are known allergens.
- Take a pollinator's perspective – make a list of all the hummingbird-pollinated wild plants in your region (or select another pollinator of your choice), and then determine their average flowering times. This can give you an idea of the resources available to hummingbirds throughout the year, and might guide efforts to install native plants in public or residential gardens in order to support hummingbirds.

Herbarium-based phenological research: exploring long-term changes in phenology

Long-term observational records indicate that plants are responding to warmer spring temperatures by changing the timing of their phenophases such as leaf-out, flowering, and fruiting. Species differ in their response rates however, and researchers have become interested in documenting which species are (or are not) responding rapidly and why (or why not). With over 260 million potential data points in the U.S. alone, herbaria harbor immense vaults of untapped information about how plants are responding to climate change and are perhaps the most powerful and extensive sources of historical data for studying this process.

Methods for using herbarium specimens to study historical changes in flowering times were first published by Dr. Richard Primack and his colleagues (Primack et al., 2004). Since then, several other research groups have published papers testing and further developing those herbarium-based methods for various plants in different parts of the world. One general method seems to be emerging for data collection, although different approaches have been taken for data analyses. We've taken care to distill this body of research into a brief summary. We encourage readers to obtain and read the original research papers.

The first published methods

Richard Primack and colleagues published the first methods for using herbarium specimens to assess phenological responses to climate change. They recorded flowering times of plants that had been collected since 1885 at Harvard University's Arnold Arboretum, and then compared the historical flowering times to those directly observed in 2003 on living trees in the same location.

The species they included in the study satisfied five criteria; 1) they were available for observation in both the arboretum and the herbarium; 2) they produced conspicuous, easily recognizable flowers; 3) they flowered for a relatively short duration; 4) they were wild species (not cultivars); and 5) they were represented by at least one herbarium specimen during "full flower". The date of full flower was determined for both living and preserved plants – "full flower" was defined as a plant (living or preserved) displaying at least 50% of its buds in full bloom. For living plants this was determined by tracking twice-weekly the flowering display for an individual in the arboretum; for preserved plants it was determined by first locating specimens that were collected at full flowering and then recording the date on which they were collected. The researchers then estimated the change in flowering time by subtracting the current year's flowering date from the historical flowering date. All estimated changes in flowering time were then included as the dependent variable in a regression analysis in which the independent (or predictor) variable was either the year of collection *or* the average temperature during the period of February-May preceding flowering.

In sum, Primack and colleagues defined the phenophases and species of interest, compared historical to current flowering times, and then examined the data set for long-term changes in flowering time as a function of either time (year) or recent temperature. What they discovered is that plants flowered on average 8 days earlier at the end of the century (1980-2002) than at the beginning of the century (1900-1920), and that the change in flowering time is associated with warming temperatures.

Further development of methods & analyses

Over the next few years other researchers published papers following similar methods but with different scopes of research and more rigorous analyses. For example, Lavoie & Lachance (2006) focused on just one abundant species (Coltsfoot, *Tussilago farfara*) and expanded the geographic area of herbarium collections to include the area of southern Quebec, Canada. After examining specimens from seven different herbaria in the region, and after including snowmelt date in their analyses to account for the environmental variation spanning the study area, they found that plants were flowering 15-31 days earlier at the end of the century compared to the beginning of the century.

Gallagher et al. (2009) conducted a rigorous analysis of phenological data derived from herbarium records, using extensive climate modeling to account for variation in temperatures across the study region across years. By matching the spatial locations of herbarium specimens to a gridded temperature data set for the region, they were able to identify several species whose flowering response was sensitive to temperature. Interestingly, they obtained their datasets electronically from three herbaria in Australia by querying their online databases (they also took several important steps to refine the data, which they describe in their paper).

Most recently, Robbirt et al. (2011) evaluated the capacity for herbarium specimens to serve as a proxy for field-collected data. They compared a long-term observational data set of the flowering phenology of an orchid (*Ophrys sphegodes*) with its herbarium-based phenological records and discovered similar trends in both data sets. They also included climatic data in their analyses, with which first they determined the window of time each year during which temperatures had the strongest association with the flowering records.

The progression and increasing sophistication of herbarium-based phenological research are apparent in its brief seven-year history. Extending upon and improving the methods demonstrated by Primack et al. (2004), several other researchers have evaluated: how to include climatic data more explicitly, how to select species, and whether herbarium data approximate field-collected data. The emerging theme is that not only are herbaria highly useful resources for climate change research, but that rigorous analyses can be conducted with climatic data to fully explore trends in the data sets. In the next section, we'll focus on the data collection step.

Methods in herbarium-based phenological research

The general approach to studying phenology with herbarium specimens is to compare the biological information determined from a specimen with the collection information provided on its label. Specifically, the phenological status of a specimen at the time of collection is compared with its collection date. This approach can be used with two major goals in mind:

- **Establish the average seasonal timeframe of a phenophase**

  Example: *When, on average, do allergy-inducing plants flower in my region?*

  Methods:

  1. Sort through the folders and sheets of one species in an herbarium. For each specimen, assign a phenophase category to describe how far it has progressed through flowering (e.g., pre-flowering, first flowering, peak flowering, last flowering, after flowering). For annual species, avoid specimens for which only a portion of the above-ground stems have been collected. For longer-lived species, you will have to assume that the stems preserved on the herbarium sheet represent the flowering phenophase of the entire plant from which it came.
  2. Record the month and day of collection, as well as any pertinent geographic information.

- **Explore long-term changes in plant phenology**

  Example: *Is flowering time of your focal species changing over recent decades?*

  Methods: Follow the same steps as above, but also record the year of collection for each specimen.

Both of these approaches to herbarium-based phenology research entail recording similar data from the herbarium sheets. A data sheet for both of these activities might appear as follows:

Collection information		Phenology	Collection date			Additional geographic information			
Accession	Collector	Phenophase	Month	Day	Year	Elevation	Latitude	Longitude	County

Once a row has been filled out for each specimen in a species' folder, data collection is complete! For an introduction on how to analyze data collected in this way, please see the accompanying exercise on the analysis of herbarium-based phenological data.

Additional considerations for herbarium-based phenological research and analysis

Once a phenological data set has been generated from herbarium specimens, or from any historical or current records, researchers should consider sharing and archiving the data set with the USA National Phenology Network. Publications resulting from these data sets are, of course, incredibly valuable for many reasons. When data sets are archived, future researchers may obtain and combine several herbarium-based phenological data sets in order to conduct more extensive analyses. Visit the USA-NPN website to learn more: www.usanpn.org.

Project planning and data collection phases

Depending on the goals of the research, it may be necessary to exclude some specimens from the data set. It may also be desirable to study particular phenophases or phenological events (e.g., the date of first flower), species, or geographic regions. Consider the following factors in the project planning and data collection phases:

Specimen quality – Examine each specimen carefully for missing structures. Establish criteria for including specimens in the data set. For example, you may want to include only specimens where every structure is visible and available for observation. If structures are missing, then it's impossible to know the plant's true phenophase. This is particularly important for research projects focused on reproductive structures because they tend to be very fragile when dried and often are missing on specimens.

Collection information – Many specimens, especially older specimens, only have a month and year recorded on the label. Other specimens may not have any geographic information. Some of these specimens with missing collection information might still be helpful to some projects, but missing collection information is likely to be cause for exclusion from the data set.

Geographic limitations – Often the geographic information reported on herbarium sheets is incomplete. This may complicate interpretation of results because, without accurate geographic information, it is impossible to discern whether a specimen is, say, flowering early because it was a warm year or because it was collected at a low elevation or latitude. Additionally, only since 1983 have GPS satellites and coordinates been accessible by the public, and only since May 1, 2000 have high-resolution coordinates been available to the public (any GPS coordinates available from specimens from 1983 to 2000 may be off by at least 100 meters). In the absence of reliable and accurate geographic information, it may be better to ignore a specimen than to spend the time recording its phenophase.

Reducing noise in the data – potential sources of error

There are several inherent issues associated with using historical collections for phenology research. Here we briefly describe some of the potential sources of error that are commonly addressed in peer-reviewed publications.

Collector bias – Collectors select plants to collect based on many criteria, none of which we will ever know. A collector might be fond of collecting first flowering plants, or peak flowering plants, or plants producing their first new leaves. Thus, the phenology of any one specimen may not accurately reflect the predominant phenological status of the plants with which it was growing. The resolution to this issue is to include specimens from as many collectors as possible, thereby reducing the potential for collector bias to play a role in the distribution of the data set.

Plant size & collector choice – Many plants are too large to be collected in their entirety and mounted on one herbarium sheet. This means that an available specimen represents a choice that the collector(s) made during the collection. Did they collect a branch representative of the entire plant, or did they collect a branch because it was one of the few that exhibited a particular phenophase? We will never know the answers to these questions, but it is important to consider them when interpreting results. One way to control for variation in phenological data due to the collection of partial specimens is to select for analysis only small-statured plants or species whose entire body (or at least the above-ground portions) fits on the sheet, although this limits the number of species available for research.

Flowering duration of species – The potential to detect phenological changes in a given population or species may depend on its flowering duration. Some authors have argued that short-flowering species that are in bloom for only a couple of weeks each year might be particularly suitable candidates for detecting changes in flowering time. Similarly, long-flowering species that produce flowers for several months each year may be more difficult to assess for long-term changes in phenology. Some authors have argued (using appropriate statistical tests) that flowering duration does not influence the probability of detecting changes in phenology. At this point, the issue seems unresolved and would be a fruitful area of research.

Choice of phenophase – It has been argued that the "peak" phenophase, such as peak flowering, is a more reliable phenophase compared to the "first" phenophase, such as first flowering. This is because first events in a population tend to be statistical outliers – in a population of 100 or 10,000 plants, the first few plants to flower are statistically very early compared to the population average. Combined with the effects of collector bias described above, some authors have argued that first events should be avoided in favor of peak events. Indeed most researchers have focused on peak phenophases, but the issue has not yet been addressed with statistical rigor.

Fluctuating collection efforts – Most data sets from herbarium specimens will have collection gaps – years or even decades where there are very few specimens available. There also has been a general decline in collection efforts over recent decades, likely attributable to the lack of funding available for basic collections and their curation and storage. Gaps in the data set can affect the statistical inference and conclusions we draw from the data set. Some researchers have addressed gaps in their data sets either by a) first demonstrating a change in temperature over the study period and then focusing analyses on the relationship between phenology and temperature regardless of year, or b) contrasting groups of data to each other (for example, comparing early-century and late-century flowering times in order to detect a statistically significant temporal change in mean flowering times).

A brief reading list for herbarium-based phenological research

In addition to the references made throughout this primer, we've also included other pertinent papers.

Bolmgren, K., and K. Lönnberg. 2005. Herbarium data reveal an association between fleshy fruit type and earlier flowering time. *International Journal of Plant Science.* 166(4): 663-670.

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.

Lavoie, C., and D. Lachance. 2006. A new herbarium-based method for reconstructing the phenology of plant species across large areas. *American Journal of Botany.* 93(4): 512-516.

Miller-Rushing, A.J., R.B. Primack, D. Primack, and S. Mukunda. 2006. Photographs and herbarium specimens as tools to document phenological changes in response to global warming. *American Journal of Botany.* 93(11): 1667-1674.

Miller-Rushing, A.J., R.B. Primack, P.H. Templer, S. Rathbone, and S. Mukunda. 2009. Long-term relationships among atmospheric CO_2, stomata, and intrinsic water use efficiency in individual trees. *American Journal of Botany.* 96(10): 1779-1786.

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 with the orchid *Ophrys sphegodes*. *Journal of Ecology.* 99: 235-241.

Primack, D., C. Imbres, R.B. Primack, A.J. Miller-Rushing, and P. Del Tredici. 2004. Herbarium specimens demonstrate earlier flowering in response to warming in Boston. *American Journal of Botany*. 91: 1260–1264.

Primack, R.B., and A.J. Miller-Rushing. 2009. The role of botanical gardens in climate change research. *New Phytologist.* 182: 303-313.