



USA National Phenology Network Observational Data Documentation

By Alyssa Rosemartin, Ellen G. Denny, Katharine L. Gerst, R. Lee Marsh, Erin E. Posthumus, Theresa M. Crimmins, and Jake F. Weltzin

Open-File Report 2018-1060

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
RYAN K. ZINKE, Secretary

U.S. Geological Survey
William H. Werkheiser, Deputy Director
exercising the authority of the Director

U.S. Geological Survey, Reston, Virginia: 2018

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit <https://www.usgs.gov/> or call 1-888-ASK-USGS (1-888-275-8747).

For an overview of USGS information products, including maps, imagery, and publications, visit <https://store.usgs.gov/>.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this information product, for the most part, is in the public domain, it also may contain copyrighted materials as noted in the text. Permission to reproduce copyrighted items must be secured from the copyright owner.

Suggested citation:

Rosemartin, A., Denny, E.G., Gerst, K.L., Marsh, R.L., Posthumus, E.E., Crimmins, T.M., and Weltzin, J.F., 2018, USA National Phenology Network observational data documentation: U.S. Geological Survey Open-File Report 2018-1060, 24 p., <https://doi.org/10.3133/ofr20181060>.

ISSN 2331-1258 (online)

Acknowledgments

Author contributions: Ellen G. Denny, Katharine L. Gerst, Alyssa Rosemartin, R. Lee Marsh, Erin E. Posthumus, Theresa M. Crimmins, and Jake F. Weltzin designed, developed and documented the data products described here; Alyssa Rosemartin wrote the first draft of the manuscript with major contributions, revisions, and final review provided by all authors.

These data products would not be possible without the dedication of the more than 12,000 contributors to *Nature's Notebook*, whose work is enabling science on new scales. We are grateful to our USA-NPN colleagues LoriAnne Barnett, Sharon Oliver, Sara Schaffer, and Jeff Switzer; former colleagues Carolyn Enquist, Jherime Kellermann, Abraham Miller-Rushing, and Kathryn Thomas; and USA-NPN co-founders Julio Betancourt and Mark Schwartz, for their significant contributions to the development and success of *Nature's Notebook*. We offer special thanks to Patricia Guertin for the depth of botanical knowledge and care that she brought to the development of the protocols and species profiles that underpin *Nature's Notebook*.

We appreciate the efforts of Katherine D. Jones and Ian Breckheimer who provided detailed reviews of the observational data products and documentation. We are also grateful to the members of the USA-NPN Research Coordination Network and Data Products Working Group, who fostered the vision for these products. Prepared in collaboration with the University of Arizona.

Contents

Acknowledgements	iii
Abstract	1
Introduction.....	1
Data Collection Protocols	3
Data Discovery, Access, Quality and Documentation.....	5
Data Discovery and Access.....	5
Ancillary Data	5
Data Quality.....	6
Quality Assurance and Quality Control	6
Results of Quality Control Investigations	6
Data Documentation.....	7
Data Types	7
Status and Intensity Data.....	9
Status and Intensity Data Usage Notes.....	10
Individual Phenometrics.....	10
Individual Phenometrics Usage Notes.....	11
Site Phenometrics.....	13
Site Phenometrics Usage Notes.....	13
Magnitude Phenometrics	14
Magnitude Phenometrics Usage Notes	16
Applications	17
References Cited.....	18
Appendix 1: Datafield Definition Tables.....	21
Appendix 2: Quality Assurance and Quality Control Measures.....	22

Figures

Figure 1. Example Observation Report. In this example observation using the USA-NPN protocols, an observer recorded data on a male phainopepla and a mesquite tree, in Tucson, Arizona in 2015.....	4
Figure 2. Data Type Comparison for Plants. This conceptual diagram compares Status and Intensity data, Individual Phenometrics, Site Phenometrics, and Magnitude Phenometrics for plants.....	8
Figure 3. Data Type Comparison for Animals. This conceptual diagram compares Status and Intensity data, Individual Phenometrics, Site Phenometrics, and Magnitude Phenometrics for animals	9
Figure 4. Status Data Phenology Calendars. These calendar displays from the USA-NPN visualization tool show status data for American beech, summarized across sites in the northeastern United States.....	10

Figure 5. Individual Phenometric Calendar. This calendar visualization enables observers to explore differences in individual plant phenology over time..... 11

Figure 6. Leaf Out by Latitude. In this scatterplot display from the USA-NPN Visualization Tool, onset date based on Site Phenometrics, for “Leaves” in American beech is plotted against latitude across the eastern United States, for the years 2010–2017..... 13

Figure 7. Phenological activity curves. This display from the USA-NPN Visualization tool is based on Magnitude Phenometrics and shows the degree to which a phenophase is expressed over time across individuals and (or) sites. 15

Tables

Table 1. Magnitude Phenometrics. These eight metrics enable the quantification of phenological activity across one or more individuals and sites, for both plants and animals..... 16

Abbreviations

API	Application Programming Interface
DOY	Day of Year
DMP	Data management plan
DOI	Digital Object Identifier
DPC	Data Product Catalog
FGDC	Federal Geospatial Data Committee
NCO	National Coordinating Office
NPDb	National Phenology Database
POP	Phenology Observation Portal
QA	Quality Assurance
QC	Quality Control
USA-NPN	USA National Phenology Network
USGS	U.S. Geological Survey

USA National Phenology Network Observational Data Documentation

By Alyssa Rosemartin,^{1,2} Ellen G. Denny,^{1,2} Katharine L. Gerst,^{1,2} R. Lee Marsh,^{1,2} Erin E. Posthumus,^{1,2} Theresa M. Crimmins,^{1,2} and Jake F. Weltzin^{2,3}

Abstract

The goals of the USA National Phenology Network (USA-NPN, www.usanpn.org) are to advance science, inform decisions, and communicate and connect with the public regarding phenology and species' responses to environmental variation and climate change. The USA-NPN seeks to advance the science of phenology and facilitate ecosystem stewardship by providing phenological information freely and openly. To accomplish these goals, the USA-NPN National Coordinating Office (NCO) delivers observational data on plant and animal phenology in several formats, including minimally processed status and intensity datasets and derived phenometrics for individual plants, sites, and regions. This document describes the suite of observational data products delivered by the USA National Phenology Network, covering the period 2009–present for the United States and accessible via the Phenology Observation Portal (<http://dx.doi.org/10.5066/F78S4N1V>) and via an Application Programming Interface. The data described here have been used in diverse research and management applications, including over 30 publications in fields such as remote sensing, plant evolution, and resource management.

Introduction

Phenology is a critical component of many aspects of human life, including agriculture, gardening, health, cultural events, and recreation. Phenology is also key to ecological relationships and processes, including pollination, predation, competition, as well as water and carbon cycles (Cleland and others, 2007; Noormets, 2009). An improved understanding of the causes and consequences of variation and change in phenology across taxonomic groups has great potential to improve agricultural production, human well-being, and ecosystem management (Schwartz, 2003; Enquist and others, 2014).

The goals of the USA National Phenology Network (USA-NPN, www.usanpn.org) are to advance science, inform decisions, and communicate and connect researchers, resource managers, educators and the public regarding phenology and species' responses to environmental variation and climate change (USA-NPN National Coordinating Office, 2014). The USA-NPN seeks to facilitate informed ecosystem stewardship and management by freely and openly providing phenological information (Box 1).

¹University of Arizona.

²USA National Phenology Network.

³U.S. Geological Survey.

Box 1: The USA National Phenology Network

The USA National Phenology Network, established in 2007 by the US Geological Survey in collaboration with other governmental and nongovernmental organizations, is a national-scale science and monitoring initiative focused on phenology. The organization seeks to improve our understanding of how plants, animals, and landscapes respond to environmental variation and change, primarily through the collection, storage, and dissemination of phenology data products. A diverse range of stakeholders contributes to and utilizes network knowledge, data, and tools, including researchers, resource managers, educators, and the public.

Observational phenology data are collected and submitted by professional and citizen scientists, primarily through the USA-NPN plant and animal phenology observing program, *Nature's Notebook* (Box 2). These data are submitted to the National Phenology Database (NPDb) and the USA-NPN observational database and serve as the backbone of all USA-NPN observational data products.

Box 2: *Nature's Notebook*

Nature's Notebook (www.nn.usanpn.org) is a national phenology program of the USA-NPN, in which professional and citizen scientists record long-term observations of plant and animal life cycle events. *Nature's Notebook* includes a website, training materials, mobile apps and visualization tools (Rosemartin and others, 2014). Launched in 2009, the program offers standardized observation protocols for more than 1,200 plant and animal taxa. Over 10,000 participants have contributed over 11 million phenophase status records via *Nature's Notebook*, as of 2017.



While many observers participate in *Nature's Notebook* independently, 52% of observers collect phenology observations in groups, as part of Local Phenology Projects. In addition to the tools provided by the NCO to all *Nature's Notebook* participants, leaders of these groups of observers (referred to as “partner groups” in data documentation) offer communication, training, and support, coordinated monitoring and locally-relevant driving questions. As a result, Local Phenology Projects typically contribute more frequent observation records than individual observers.

[Image: Observers at the Mission Garden Local Phenology Project, part of the Tucson Phenology Trail, in Tucson Arizona. Credit: Brian Forbes Powell]

The USA-NPN delivers four observational data products: Status and Intensity Data, Individual Phenometrics, Site Phenometrics and Magnitude Phenometrics. Each data type is suited for particular analyses or visualizations as detailed in individual sections in this report.

In this report we describe protocols for data collection, data access, processing and documentation relevant to all USA-NPN observational phenology data products, an overview of the four data types followed by descriptions of each of the four data types that are available from the NPDb:

1. Status and Intensity data,
2. Individual Phenometrics
3. Site Phenometrics, and
4. Magnitude Phenometrics.

We conclude with some recent applications of the data products for science and management.

Data Collection Protocols

Rigorous, scientifically vetted observation protocols form the foundation of *Nature's Notebook* and yield high quality, reliable data. The protocols are standardized within plant functional types and animal guilds, and feature phenophases that are easily observable, responsive to seasonal changes and accurately reflect species life histories (Denny and others, 2014). The observation protocols employ a status monitoring approach, in which observers visit a site at regular intervals to evaluate the phenological status of marked individual plants (or patches of plants) and animal species. Individual animals are not tracked; rather, animals are observed at the level of the site. Box 3 provides definitions for key terms used in USA-NPN protocols and data products.

Box 3. Terminology used for USA-NPN data products

- **Phenophase** is an observable stage or phase in the annual life cycle of a plant or animal that can be defined by a start and end point.
 - Examples: Period over which newly emerging leaves are visible, or the period over which open flowers are present on a plant.
- **Phenometrics** are broadly defined as derived metrics that characterize the expression of a phenophase at multiple temporal and spatial scales.
 - Examples: Onset date for flowering at one site, or weekly animal abundance for a region.
- **Intensity** is a categorical measure indicating the extent to which a phenophase is expressed for an individual plant or an animal species at a site, observed on a given visit.
 - Examples: “25-49%” of flowers are open on a plant, or a “Full chorus” of frog calling is heard.
- **Abundance** is the count of animals in a given phenophase at a site, observed on a given visit.
 - Example: 7 American robins “Feeding”.
- **Magnitude** is the extent to which a phenophase is expressed across multiple individuals or sites, during a specified time interval.
 - Example: The proportion of individual red maples in flower in the month of April.

Phenological status is reported by yes or no answers to a series of questions, for example, “Do you see leaves?” or “Do you see active individuals?”. In contrast to traditional monitoring of annual “first” events (for example, date of first leaf or first robin), this approach captures absence data when the phenophase is not occurring and repeat events (for example, a second wave of blooming after a desert rain). The approach allows for the quantification of uncertainty around the actual date of the beginning or end of a phenophase. In addition to “Yes” or “No,” observers may also report “?”, indicating that they are uncertain of the phenophase status. The beginning of a phenophase may have occurred prior to, but not after, the first “Yes” report. Thus, statistical methods that allow data users to estimate onset dates with missing data, including interval censor approaches, may be useful for analysis (Bjorkman and others, 2015; Pearse and others, 2017). In 2011, the USA-NPN added measures of intensity and abundance to the protocols, enabling observers to not only report on the presence or absence status of the phenophase during each site visit, but also the degree to which the phenophase is expressed on an individual plant, or for animals, at a site (for example, “95–100 percent” of a beech tree’s canopy is full with “Leaves,” or 12 “Active individual” robins are seen). Figure 1 provides an example of the protocols used to record two species in Tucson, AZ.

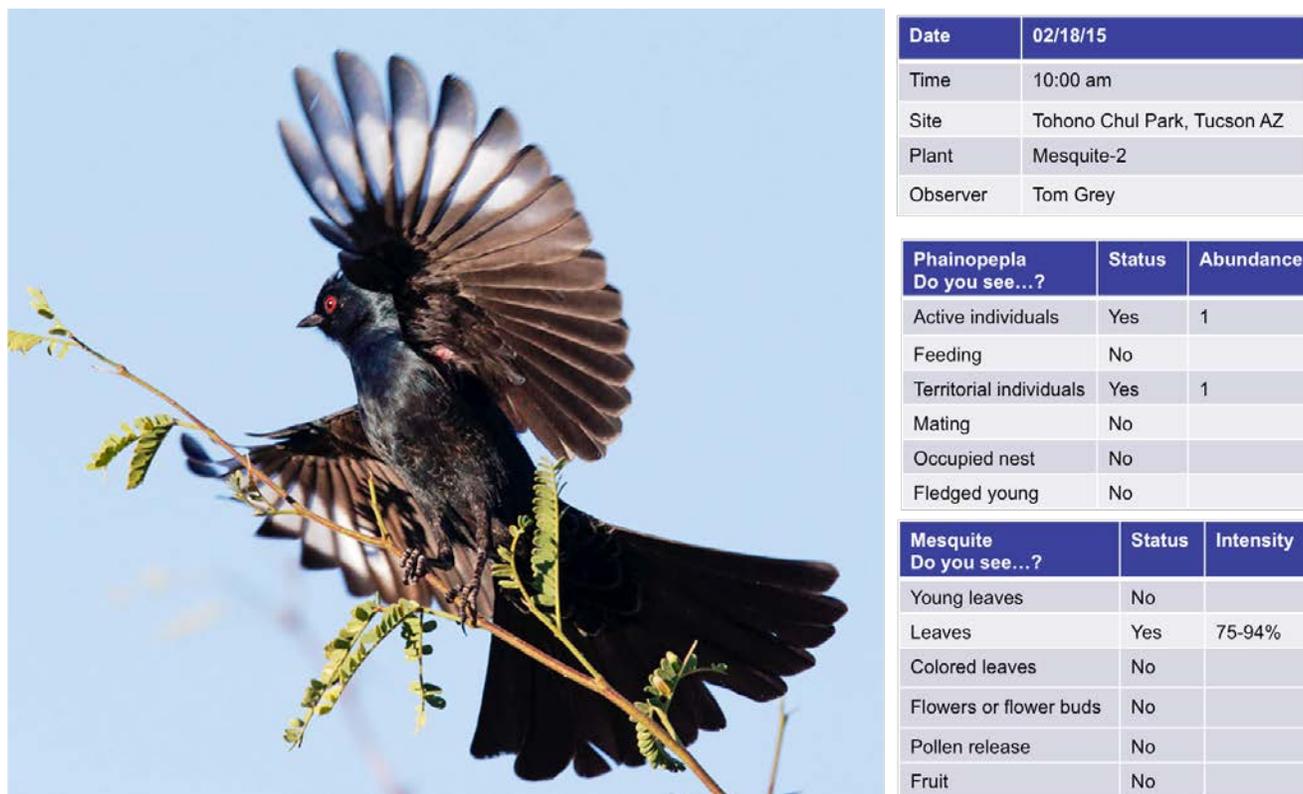


Figure 1. Example Observation Report. In this example observation using the USA-NPN protocols, an observer recorded data on a male phainopepla and a mesquite tree, in Tucson, Arizona in 2015. The observer answered “Yes” or “No” to a series of questions. Where he responded yes, he also provided measures of intensity and abundance for the phenophases. Photograph by Tom Grey.

To foster the creation of a robust, standardized, continental-scale data resource, the NCO encourages the collection of data following the USA-NPN protocols and the integration of that data into the NPDb. The USA-NPN protocols have been used independently of the *Nature’s Notebook*

infrastructure when adopted by USA-NPN collaborators (for example, National Ecological Observatory Network) and researchers (for example, Browning and others, 2015; Hallman and Arnott, 2015; Kern and others, 2015; Kharrassi and others, 2015; Elmendorf and others, 2016). Through direct entry, bulk upload, or web services, data resulting from these projects can and have been entered into the NPDb. In cases where the data collected using the USA-NPN protocols are housed by a collaborator, the Plant Phenology Ontology (<https://www.ebi.ac.uk/ols/ontologies/ppo>, Stucky and others, 2016) supports automated integration of datasets housed within and external to the NPDb. In addition, future iterations of the ontology and other efforts will facilitate integration of USA-NPN datasets with other data sources such as herbarium specimens, phenology cameras (phenocams), and satellite-derived land surface phenology.

Data Discovery, Access, Quality and Documentation

In the interests of making USA-NPN Observational Data widely available and usable, we enable users to find, access, trust, and appropriately analyze and interpret data. The data are readily discoverable and accessible (see Data Discovery and Access section). Several forms of ancillary data are available along with the primary data types (see Ancillary Data section). Data undergo extensive quality assurance and quality control processes (see Data Quality section). All data products are provided with technical documentation and metadata (see Data Documentation section).

Data Discovery and Access

Metadata, which describe the four observational data types (Status and Intensity, Individual Phenometrics, Site Phenometrics, and Magnitude Phenometrics) are available in several repositories to facilitate data discovery, including USGS ScienceBase, USGS Science Data Catalog, Data.Gov, and the Knowledge Network for Biocomplexity.

Observational phenology data are housed in the NPDb. Data users can customize data downloads based on data type, desired optional fields, time period, species, phenophase, partner group collecting the observations, and geographic location (by state or bounding box) using the Phenology Observation Portal (POP; <http://doi.org/10.5066/F78S4N1V>).

Users may alternatively opt to access these data products using the USA-NPN Application Program Interface (API; USA-NPN National Coordinating Office, 2016a). The API provides access to the same filtering, data types, fields, and ancillary data as the Phenology Observation Portal. Data are also available via the Phenology Visualization Tool (www.usanpn.org/data/visualizations), where users can explore the data by viewing custom phenology calendars, graphs, and maps.

Ancillary Data

Ancillary data that support interpretation and analysis are available with observational data. Twenty climate variables derived from Daymet (gridded, daily surface weather and climate data, <https://daymet.ornl.gov/overview.html>), including spatially explicit seasonal temperature variables and heat accumulation, are also available for download with Status and Intensity, Individual and Site Phenometrics data types. Climate data are made available with the observational data shortly after their release by Daymet, which is typically six months after the close of the calendar year. Every time a unique location is added to the database, the Daymet data services are queried for climate information about that location, and the aggregated climate values, including Accumulated Growing Degree Days, are calculated and stored in the database for future use. As new data are made available from Daymet

each year, new values are calculated for all unique locations. New versions of Daymet result in the updating of all stored values.

Users may download ancillary data files with additional information on observers (for example, if trained in person), sites (for example, if urban), individual plants (for example, if watered), and source datasets (for example, if submitted directly through *Nature's Notebook* or integrated from another source). Data are output from the POP as a comma-separated value (.csv) file, together with a file listing the search parameters and files with any user-selected ancillary data or metadata. Appendix 1 provides the full suite of default and optional fields for each data type, with data field definitions, as well as definitions for all ancillary data fields.

Data Quality

The NCO undertakes substantial effort to ensure quality of the observational phenology data housed in the NPDb. In the following sections we describe data quality assurance and quality control and results of research investigating broader aspects of data quality. Data users may also wish to consult the literature on spatial and temporal biases in large multi-contributor datasets (for example, Isaac and others, 2014; Bird and others, 2014).

Quality Assurance and Quality Control

Substantial quality assurance (QA; applied prior to data entry into the NPDb) and quality control (QC; applied prior to data output from NPDb) measures are in place for all observational data products (see appendix 2 for complete list).

QA measures limit species and phenophase misidentification and data entry errors. Extensive online training and print resources, such as the *Botany Primer* and *Phenophase Primer* (Guertin and others, 2015; Guertin and others, 2017) are in place to support observers in collecting high quality data. Online tools are designed with usability in mind, to prevent user error, and processing prevents illogical values (for example, dates in the future) from being submitted.

Quality Control measures assess data in terms of validation, reliability, and plausibility. Validation refers to assessing values relative to defined rules, for example, that coordinates are in North America. Reliability refers to dataset completeness and freedom from error and bias, for example, flagging of instances where multiple onsets occur in a season or when one or more observers have reported differing phenophase statuses for the same individual plant on the same date. Plausibility refers to the likely accuracy of the values. Our QC processes do not yet directly address plausibility, as a large number of observations would be required for each species and phenophase, spanning many sites and (or) years, a requirement that has yet to be met for most species. Instead, we support data users in exploring the likely accuracy of suspect records through comments fields and ancillary data about observers and site conditions when observations are made.

For the selection of QC measures tracked dynamically on our Quality Control Dashboard (<https://www.usanpn.org/data/quality>), error rates are relatively low (under 2 percent of all records). Nearly half of all first and last “Yes” observations have a preceding or following “No” observation within a week, enabling precise estimates of phenophase onset and end.

Results of Quality Control Investigations

Published explorations of the quality of USA-NPN observational data have found high levels of reliability and utility for their intended applications. For example, Fuccillo and others (2015) found a 91 percent overall concordance rate between trained and expert observers, with lower rates early in the

season. Further, the modern *Nature's Notebook* program is built on a lilac and honeysuckle monitoring program begun in the 1950s. In this long-term phenology dataset housed in the NPDb, researchers found that only 1.46 percent of first leaf or bloom records to be outliers within an individual plant's period of record for plants with at least 10 years of data (Rosemartin and others, 2015); similarly, in this dataset, only 2.09 percent of first leaf or bloom records were flagged as outliers relative to other onsets in similar geographic and climatic contexts (Mehdipoor and others, 2015).

Data Documentation

All existing and planned data products, which include modeled, integrated, and observational data types, are described in the USA-NPN Data Product Catalog (USA-NPN, 2016b). The USA-NPN Data Management Plan describes all data resources, software and tools (USA-NPN, 2016c).

The datasets described here are supported by multiple forms of technical documentation, including:

- FGDC-compliant metadata (<https://data.usanpn.org:3002/pop/fgdc>)
- Datafield definitions (appendix 1)
- Application Programming Interface documentation (USA-NPN, 2016a)
- *Nature's Notebook* User Interface Documentation (<https://www.usanpn.org/results/nndocumentation>)
- Protocol documentation (Denny and others, 2014)

Data Types

Data stored in the NPDb are output in four primary formats, or data products (figs. 2 and 3; USA-NPN, 2016b). Status and intensity data consist of presence/absence records for individual phenophases on individual plants or species of animals at a site on a single visit. These records also include intensity and abundance measures. Individual Phenometrics and Site Phenometrics, which are synthesized sequentially from Status and Intensity data, provide estimated phenophase onset and end dates. Individual Phenometrics are derived estimates of phenophase onset and end dates for organisms within a given period of interest. Site Phenometrics are summary metrics of the onset and end date of phenophase activity across multiple individuals of the same species at a site within a given period of interest. Magnitude Phenometrics provide measures of the extent to which a phenophase is expressed across multiple individuals or sites, for a given time interval. These metrics include several approaches for capturing the shape of seasonal activity curves. In *Nature's Notebook*, plants are marked and tracked through time, while animals are not, resulting in several key differences between the phenometric data types for plants (fig. 2) and for animals (fig. 3). Individual Phenometrics and Site Phenometrics are nearly identical for animals, while for plants the former provide data for individual plants and the latter aggregate data across plants of the same species at a site. Magnitude Phenometrics provide additional information on animals, including correcting abundance values by search time and search area, which is not relevant for plants. As additional observational phenology data types are created by the USA-NPN, they will be described at www.usanpn.org/data/new_data_products.

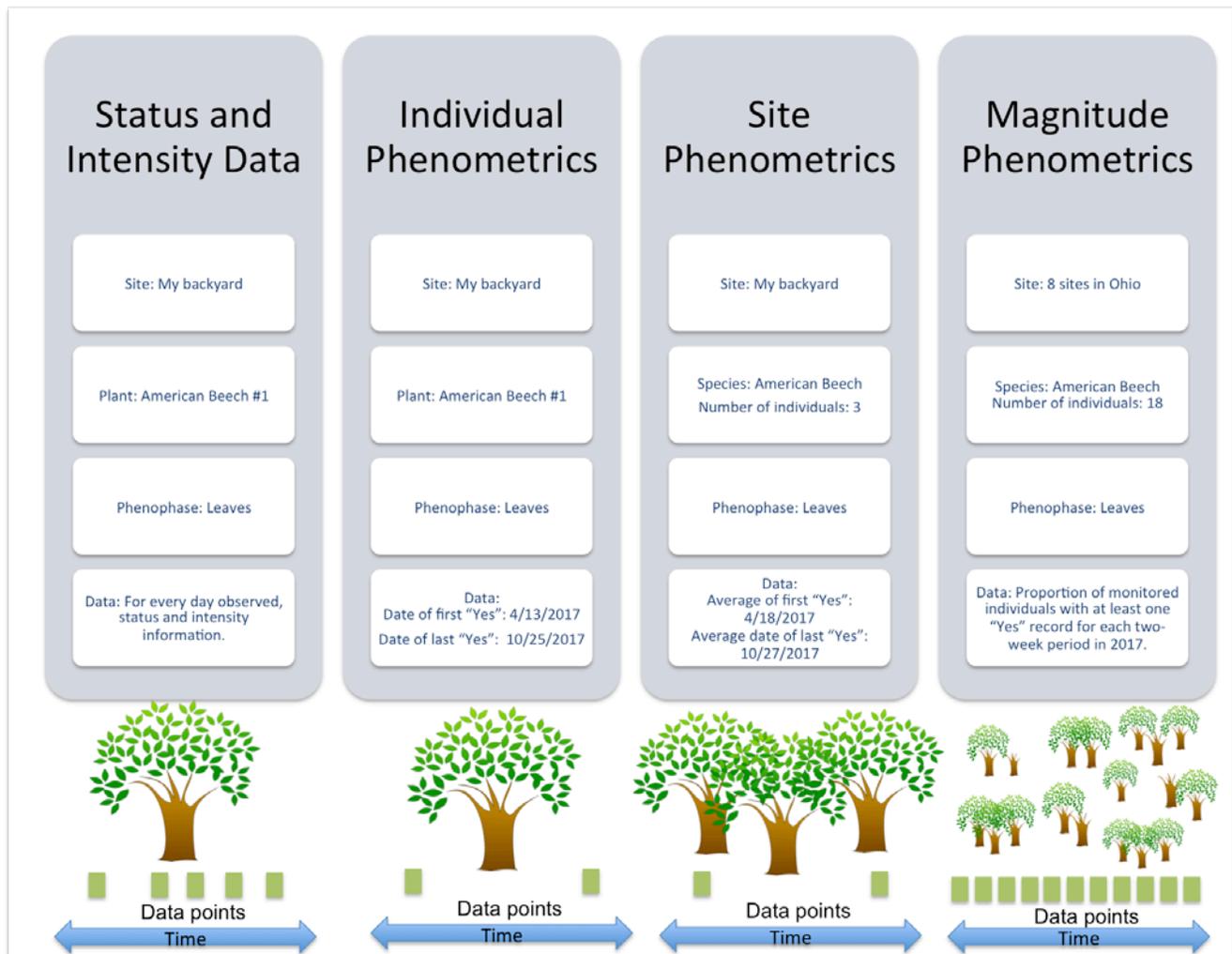


Figure 2. Data Type Comparison for Plants. This conceptual diagram compares Status and Intensity data, Individual Phenometrics, Site Phenometrics, and Magnitude Phenometrics for plants. From left to right the scale increases from an individual tree to three individuals trees monitored at a site, to eight sites in Ohio monitoring different numbers of individual American beech trees. Green squares illustrate that data points throughout the season are available in Status and Intensity data (left) and Magnitude Phenometrics (right), while only first and last "Yes" dates are available for Individual and Site Phenometrics (center).

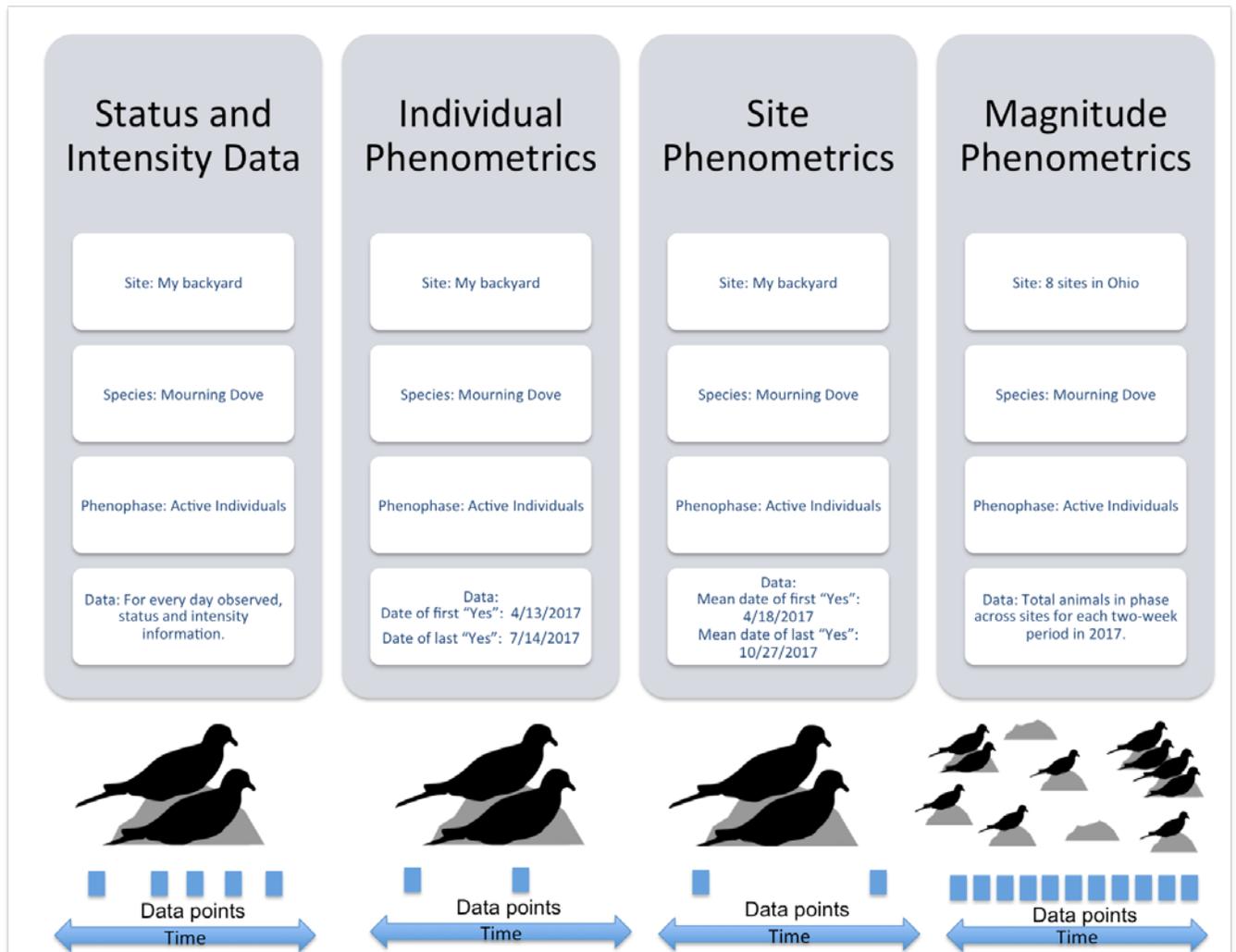


Figure 3. Data Type Comparison for Animals. This conceptual diagram compares Status and Intensity data, Individual Phenometrics, Site Phenometrics, and Magnitude Phenometrics for animals. Animals are monitored as species presence at a site (rather than as individual organisms), the first three panels represent a single site, and the right panel represents multiple sites. For animals, Individual and Site Phenometrics are conceptually equivalent, and differ only in that Site Phenometrics provides mean date for multiple individuals and only uses the initial first "Yes" and final last "No". Blue squares illustrate that data points throughout the season are available in Status and Intensity data (left) and Magnitude Phenometrics (right), while only first and last "Yes" dates are available for Individual and Site Phenometrics (center).

Status and Intensity Data

The Status and Intensity data type is the most direct presentation of the phenology data stored in the NPDb. Each row is comprised of a single record of the status (present/"Yes", absent/"No" or uncertain/"?") of a single phenophase on an individual plant or species of animal at a site on a single site visit, as well as the estimated intensity or abundance (for example, percent canopy fullness or number of individual robins observed respectively, see Box 3).

Status and Intensity datasets are appropriate for a variety of applications, including the development of phenology calendars. Figure 4 provides an example phenology calendar based on Status and Intensity data for American beech. This data type also supports applications such as validating remotely sensed phenology observations and comparison with model predictions (Euskirchen and others, 2013; Melaas and others, 2016). It can also be used to determine site occupancy rates for animals. The intensity and abundance data can be used to investigate the amount of ripe fruit available for frugivorous animals, or to find the peak date of yellow warbler arrival.

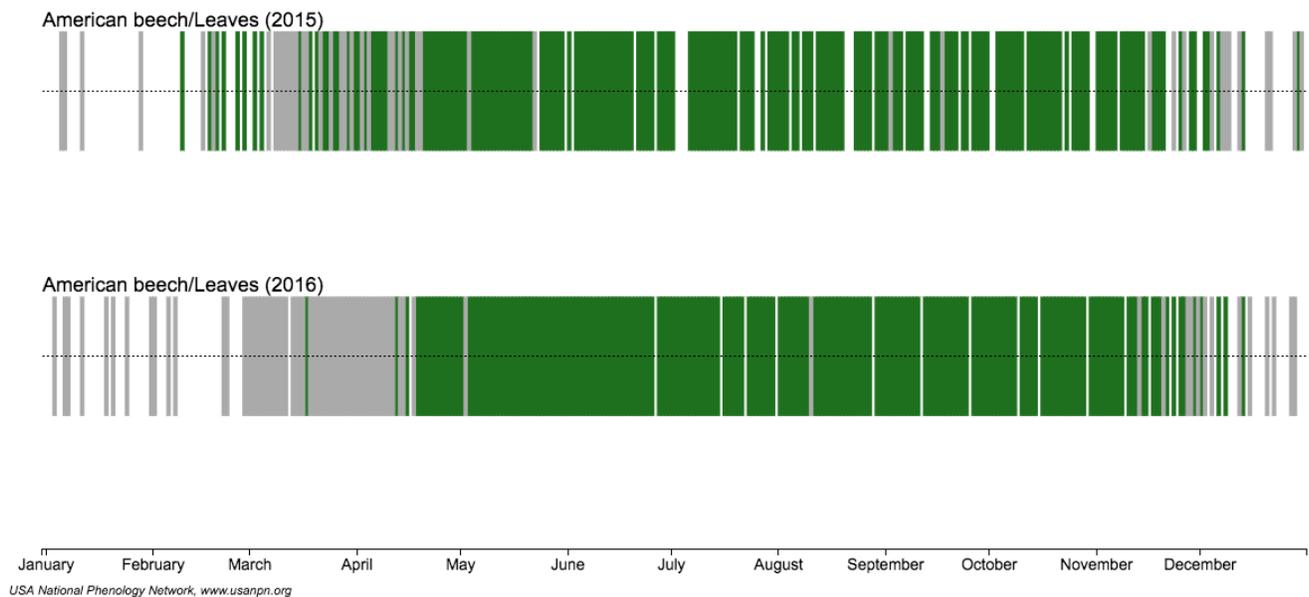


Figure 4. Status Data Phenology Calendars. These calendar displays from the USA-NPN visualization tool show status data for American beech, summarized across sites in the northeastern United States, for 2015 and 2016. Grey bars indicate dates when “No” records were reported for the “Leaves” phenophase and green bars indicates dates when “Yes” records were reported.

Status and Intensity Data Usage Notes

If one or more observers report conflicting phenophase status observations on the same date for the same plant or animal species at a site, the “Single/Multiple Observer Conflict Flag” is applied (as described under Reliability in appendix 2). The “Status_Conflict_Related_Records” field provides the “Observation_ID” numbers for all conflicting records, to enable further investigation of the conflict.

Individual Phenometrics

Individual Phenometrics are derived from phenophase status data and provide estimates of phenophase onset and end dates based on the first and last “Yes” status values for organisms within a specified season of interest. Each row in this data type is comprised of values that are derived from a string of consecutive “Yes” status reports without an intervening “No” status report for a single

phenophase for an individual plant or an animal species at a site, called a “series”. For plants, this data type provides information on the onset and end of a phenophase on an individual plant. For animals, it provides information on the onset and end of the presence of an animal species at a site. As animal presence at a site is much more likely to be interrupted by absence than the presence of a phenophase on a plant, Status and Intensity data or Site Phenometrics may be more appropriate for investigating animal phenology. However, we provide animal phenology in the same format as individual plants in this data type to allow users to readily compare individual plant phenology with animal activity.

Individual Phenometrics are appropriate for comparing onset and end dates across individual plants and comparing individual plant phenology to predictive models. This data type facilitates the creation of phenology calendars for individual plants. Figure 5 provides an example calendar, for an individual creosote plant, based on Individual Phenometrics. These data are also appropriate for quantifying how many episodes of phenophase activity occur and the duration of those episodes (for example, Gerst and others, 2017).

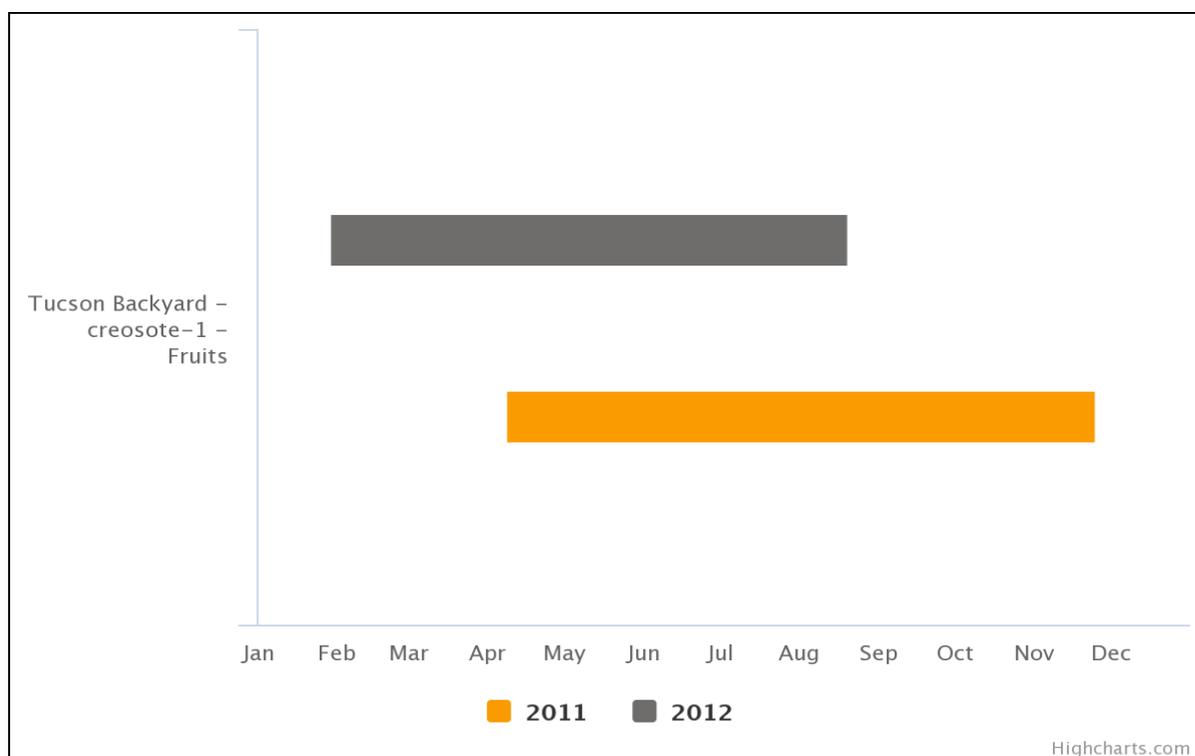


Figure 5. Individual Phenometric Calendar. This calendar visualization, available through the *Nature's Notebook* observation interface, enables observers to explore differences in individual plant phenology over time. In this example, fruiting periods for an individual creosote plant (Creosote-1, at the site “Tucson Backyard”), were similar in duration in 2011 and 2012, but began much earlier in 2012.

Individual Phenometrics Usage Notes

Phenophase status values of “?” (meaning the observer was uncertain of status) are not considered when calculating Individual Phenometrics.

Quality control flags, as described under Reliability (appendix 2), are applied to the row if any record in the series meets the criteria for a particular flag.

More than one series may exist for a given phenophase in an individual plant or animal species within a single growing season or year, this might occur in the case of leaf bud break followed by a killing frost and second round of breaking leaf buds. It could also occur at group sites where two or more observers are reporting on the same plant on sequential days but are not in agreement on phenophase status. This scenario can result in a pattern of yes-no-yes-no-yes reports. Both the “Multiple Observer” and the “Multiple First Yes” flags (as described under Reliability, appendix 2) can aid data users in identifying these issues and taking appropriate actions based on their assessment of whether the phenomena is likely real or an artifact.

Individual Phenometrics provide the date of first reported “Yes”, the number of days since the prior “No”, the date of last “Yes”, and number of days until the following “No”. Phenometrics are calculated with a temporal resolution of one day, the time of day an observation was collected, if reported, is not taken into account. The temporal precision, or “gap size”, reported as number of days between the first or last “Yes” and a “No” report, is provided to allow users to estimate the uncertainty in onset or end date. If both a “Yes” and a “No” were reported on the same day, at any point in the series, the “Yes” record is given precedence, and a “Single/Multiple Observer Status Conflict” flag is applied. The number of “Yes” reports in the series is also included, together with the number of days in the series to provide summary and quality information. For example, a series of 12 days with two “Yes” reports is less reliable for the purpose of estimating phenophase duration than a series of 12 days with eight “Yes” reports.

Individual phenometric dates are provided in three formats in the data files: (a) MM/DD/YYYY, (b) Year and Day of Year (DOY) in separate columns, and (c) Julian Date. Julian dates are calculated as the number of solar days since November 24, 4714 BC (based on the Gregorian Calendar), and are used primarily by astronomers. We provide phenometrics in Julian dates to simplify the calculation of phenophase duration, particularly in cases where a phenophase overlaps the turn of the calendar year. Julian dates use decimals to represent time of day. As we are providing phenometrics at 1-day resolution, we round Julian date to the nearest integer in all processing and output.

When downloading Individual Phenometrics via the POP, users specify both the span of years of interested (for example, 2010 to 2015), and their within-year period of interest to define the day that the phenometric calculations start. Filters allow selection of “Calendar Year” (January 1 to December 31), “Water Year” (October 1 to September 30), “Summer Year” (July 1 to June 30), and a custom option. For example, if “Calendar Year” is selected, series are calculated starting January 1 of each year and ending December 31 of the same year. As the date of first “Yes” depends on the within-year period, users must know the natural history of the species they are investigating. For instance, a researcher interested in flowering of winter annuals in the southwestern U.S. might select the “Water Year” rather than the “Calendar Year”. In the latter, calculations would initiate January 1, in the middle of the flowering season, potentially dividing a single series of “Yes” reports (representing a single continuous period of flowering) into two separate series, one before and one after January 1. If no “Yes” reports are available for the time period requested, no rows are output.

Note that the API does not allow a user to select a season of interest, only a start and end date. As observers do not always capture the preceding and following negative data around a phenophase, it is possible for a series in this data type, requested via API, to give a first “Yes” for the phenophase in one year and the last “Yes” a year or more later. To avoid this issue the user should create separate web service calls for each season of interest. For example, to explore spring leaf out dates by calendar year, from 2010 to 2015, a data user would create six web service calls, beginning January 1 and ending December 31 for each year. Changes are planned to the API to reduce the likelihood of this issue.

Site Phenometrics

Site Phenometrics, derived from Individual Phenometrics, provide summary metrics of the onset and end date of phenophase activity for a species at a site. Observers are directed to create sites that represent uniform habitat and are no larger than 15 acres. For plants, this metric is calculated as an average for all individuals of a species at the site. For animals, where individuals are not tracked, this metric represents the first and last recorded appearance of the species during the season of interest.

This data type is suitable for understanding how phenophase onset and end dates relate to latitude, elevation, and climate drivers, such as seasonal temperatures and precipitation, leading to the development of predictive pheno-climatic models (Mazer and others, 2015; Crimmins and others, 2017; Gerst and others, 2017). Figure 6 illustrates the relationship between leaf out date for American beech and latitude, using Site Phenometrics. Late season outliers seen in this figure are typical of the observational dataset, and would be explored and potentially excluded by the data user. This data type is also appropriate for the calibration of remotely sensed information, as it characterizes phenological patterns for a species at the level of a site and provides information on sample size and within-site variation.

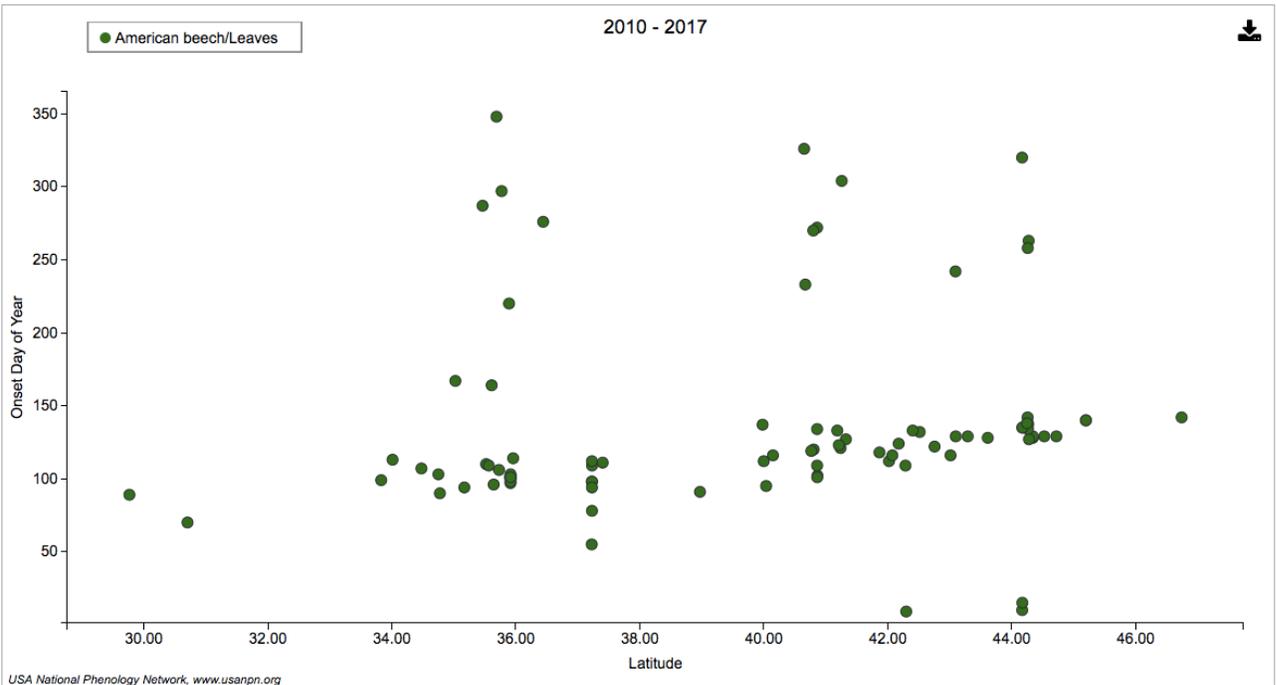


Figure 6. Leaf Out by Latitude. In this scatterplot display from the USA-NPN Visualization Tool, onset date based on Site Phenometrics, for “Leaves” in American beech is plotted against latitude across the eastern United States, for the years 2010–2017.

Site Phenometrics Usage Notes

Conceptually, and in terms of processing and output, this data type is the same as Individual Phenometrics, with the following exceptions. This data type provides only the average initial first “Yes” and the average final last “Yes” date of the phenophase for the period of interest because multiple within-year series (as calculated for Individual Phenometrics) are meaningless when averaged. Restated,

this data type returns only the mean of the first “Yes” dates for each individual of the species and the mean of the last “Yes” dates for each individual and ignores any ending and restarting of the phenophase in between the specified within-year start and end dates.

As the data user can define the within-year time frame of interest for which the data are summarized and returned (as described in the Individual Phenometrics section), the data can be summarized appropriately for any seasonal phenomenon of interest. For example, a user could isolate spring flowering of a plant species from a second round of late summer flowering in the same species by specifying a “Season Date Range” of January 1 to May 31.

As a result of the processing described above, the data user must be careful in selecting a custom “Season Date Range” on the POP. For example, if a user creates a custom date range that encompasses several years, there will be only one start and end date for that period of time, ignoring yearly phenophase onset and end. Conversely, if a user creates a custom data range of only one month within a long period of plant activity, the phenophase will inaccurately appear to start at the beginning of the month and end at the end of the selected month.

As with Individual Phenometrics, the API for Site Phenometrics does not allow a user to select a “Season Date Range”, but rather, only a start and end date. If a user wants to explore spring leaf out dates by calendar year from 2010 to 2015, he or she would create six web service calls, beginning January 1 and ending December 31 for each year.

To increase precision in the calculations of onset and end, this data type excludes first or last “Yes” values without a prior or following “No” during the selected time period. A filter in the POP allows the user to refine the desired temporal precision to exclude “Yes” values without a “No” within 30, 14, or 7 days. In the API, this filter may be set at any value.

Site phenometric dates are provided in the same three formats as described above for Individual Phenometrics (see Individual Phenometric Usage Notes section).

Additional fields provide the number of individual plants included in the mean onset or end date calculations, the standard error of the mean onset or end date, the mean number of days between the first or last “Yes” and the prior or following “No”, and the standard error of these “gap sizes.” Users may opt to include the minimum, maximum, and median day of year and standard deviation for each mean onset and end date.

Magnitude Phenometrics

Magnitude Phenometrics are a suite of eight metrics derived from Status and Intensity data. This data type provides information on the extent to which a phenophase is expressed across multiple individuals or sites, for a given set of sequential time intervals. The data user may select a weekly, bi-weekly, monthly, or custom time interval to summarize the metrics in either the POP or the API. Two metrics are available for both plants and animals, one metric is available for plants alone and five metrics are available for animals alone (table 1). Three of the five animal metrics correct animal abundance values for observer effort in time and space.

This set of data products serves a variety of applications including investigating patterns of synchrony between species and determining peak periods of activity. The data type is well-suited for creating phenological activity curves enabling comparison of the amount of overlap between related phenophases, such as presence of leaves and caterpillar activity. Figure 7 illustrates the temporal overlap between relevant monarch and milkweed life stages, using Magnitude Phenometrics. Future additions to this data type will include summarized intensity data (for example, percent canopy fullness).

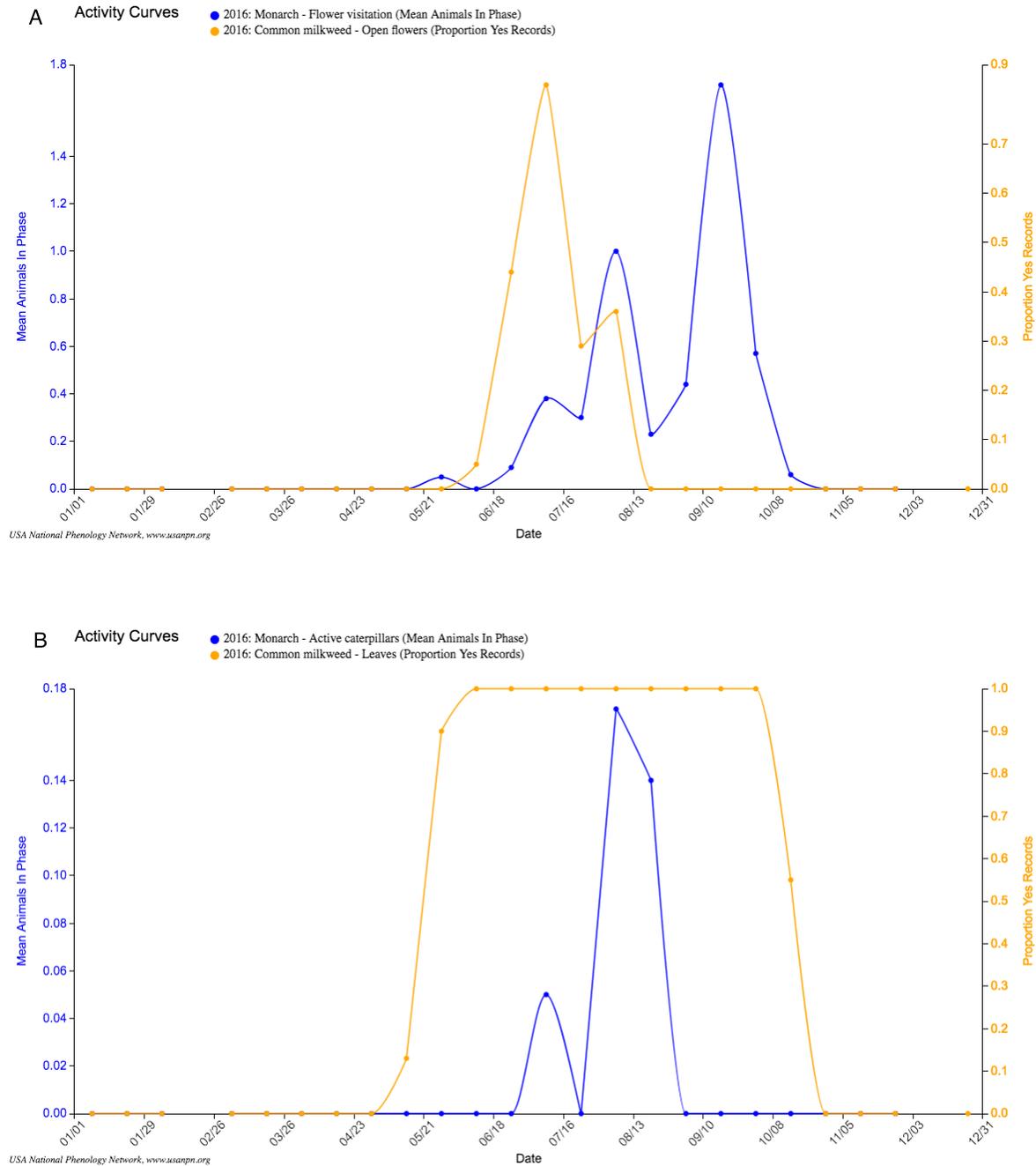


Figure 7. Phenological activity curves. This display from the USA-NPN Visualization tool is based on Magnitude Phenometrics and shows the degree to which a phenophase is expressed over time across individuals and (or) sites. A, The mean number of adult monarch butterflies in the “Flower visitation” phase relative to the proportion of “Yes” records for “Open flowers” in common milkweed, in the upper Midwest, summarized on a biweekly interval. This visualization suggests that milkweeds may be a nectar source for monarchs during only part of their adult phase. B, The mean number of “Active caterpillars” relative to the proportion of “Yes” records for “Leaves” in common milkweed, in the upper Midwest, summarized on a biweekly interval. This visualization suggests that milkweed leaves are available during the full caterpillar phase of the species.

Table 1. Magnitude Phenometrics. These eight metrics enable the quantification of phenological activity across one or more individuals and sites, for both plants and animals.

Metric	Kingdom	Source data	Description
Total Yes Records	Plants and Animals	Status	Sum of all “Yes” records.
Proportion Yes Records	Plants and Animals	Status	Number of “Yes” records divided by number of all “Yes” and “No” records.
Proportion Individuals with Yes Records	Plants	Status	Individual plants with at least one “Yes” record divided by all individual plants with at least one “Yes” or “No” record.
Proportion Sites with Yes Records	Animals	Status	Sites with at least one “Yes” record divided by all sites with at least one “Yes” or “No” record.
Total Animals In Phase	Animals	Abundance	Sum of all abundance values. ¹
Mean Animals In Phase	Animals	Abundance	Abundance value averaged across site visits.
Mean Animals In Phase Per Hour	Animals	Abundance	Abundance value is divided by time spent searching for animals at each site visit. Values for each site visit are then averaged. Includes only “Stationary”, “Walking” and “Area Searches” as the observation method, for site visits lasting under 3 hours.
Mean Animals In Phase Per Hour Per Acre	Animals	Abundance	Abundance value is divided by time spent searching and area searched. Values for each site visit are then averaged. Includes only “Area Searches” as the observation method, for site visits lasting under 3 hours.

¹ Abundance values represent the number of animals observed in a given phenophase. For all Magnitude Phenometrics based on abundance positive status records with no reported abundance are excluded.

Magnitude Phenometrics Usage Notes

Phenophase status values of “?” (meaning the observer was uncertain of status) are not considered in the calculation of Magnitude Phenometrics. Accordingly, if an individual tree has eight “No” records and three “?” records for a given time period, the denominator for “Proportion of Individuals In Phase” is 8, rather than 11.

In both the Phenology Observation Portal, and the API, users may select a weekly, bi-weekly, monthly or custom time interval for summarization. All summarization initiates on January 1, such that a weekly time interval selection would return values for 52 weeks (for example, 7-day intervals, January 1–7, 8–14, and so forth). The final 7-day interval in a non-leap year ends on December 30. Any data for December 31 is included in this final whole interval. A bi-weekly selection creates a 2-day remainder at the end of the year, and data for these two days are also included with the final whole 7-day interval. As monthly calculations are based on the calendar month, there is no remainder. When any custom time intervals are selected, the corresponding remainder is included in the final whole interval.

Magnitude phenometric datasets calculated at a spatial scale greater than a single site may have spatial or temporal biases, as some elevations or habitat types may be overrepresented or sampling frequency may not be evenly distributed over the course of the year. *Nature’s Notebook* site and observation density generally follows patterns of human population density.

Caution should be used in comparing “Total Yes Records” and “Total Animals In Phase” across multiple geographic regions or time periods that may have group sites contributing to the dataset. If

multiple observers report a “Yes” for the same phenophase on the same plant on the same day, or report abundance for the same group of animals, values will be artificially inflated.

For the metrics “Mean Animals in Phase per Hour,” and “Mean Animals in Phase per Hour per Acre” the animals per hour are calculated at the level of the site visit. When there are no animals observed at a site visit, the value is zero regardless of how long the search time was for a given visit. Thus the search time in the case of a phenophase that was not detected is not accounted for in our summarized metrics. For example, two “No” reports for a phenophase, with search times of 5 minutes and 60 minutes will both return a value of zero animals per hour. Further, these metrics are standardized by hour and acre regardless of total time and area searched, and therefore ignore variable detection rates over time and space.

Additional fields for this data type include sample size for each metric, and for metrics that are mean calculations, standard error and standard deviation are included.

Applications

The four observational data products described in this report, enable researchers to further the science of phenology. Over 400 users downloaded datasets from 2013 to 2017, and over 40 peer-reviewed publications (www.usanpn.org/pubs/results) have resulted from the data housed in the NPDb.

Data are shedding light on the climate drivers of phenology. The broad geographic scope of these data enables the identification of accumulated heat thresholds across species’ ranges (Crimmins and others, 2017). Understanding these drivers improves management, as in the case of Buffelgrass (*Pennisetum ciliare*), a critical threat to the native vegetation communities of the Sonoran desert, including Saguaro National Park. Using phenology observations collected via *Nature’s Notebook*, precipitation data, and satellite imagery, USGS scientists identified the sequence of rainfall events that trigger conditions that maximize treatment efficacy for this species while minimizing impacts to native vegetation (Wallace and others, 2016). With a similar intention, resource managers at seven National Parks in California collected data on shared species of interest using *Nature’s Notebook*, to better understand climate drivers, on elevational and latitudinal gradients, in water-limited systems (Mazer and others, 2015).

Data are also used in management directly; by tracking the phenology of invasive *Verbesina encelioides* using *Nature’s Notebook*, staff at Midway Atoll National Wildlife Refuge are able to determine when to schedule removal to efficiently eradicate remaining patches of this problematic plant on the Refuge (<https://fws.usanpn.org/midway-atoll-nwr>).

Data have been used to validate or improve models developed with other observational phenology datasets and from satellite-derived phenology, to better understand the ecological impacts of climate change (for example, Euskirchen and others, 2014; Jeong and Medvigy, 2014; Melaas and others, 2016). Data also serve in predicting species’ range limits (Chapman and others, 2014).

As these observational datasets continue to grow rapidly, they will be valuable for an increasing diversity of applications, including wildfire management, human health and allergy concerns, spread of disease, and more.

References Cited

- Bird, T.J., Bates, A.E., Lefcheck, J.S., Hill, N.A., Thomson, R.J., Edgar, G.J., Stuart-Smith, R.D., Wotherspoon, S., Krkosek, M., Stuart-Smith, J.F., Pecl, G.T., Barrett, N., and Frusher, S., 2014, Statistical solutions for error and bias in global citizen science datasets: *Biological Conservation*, v. 173, p.144–154, <https://doi.org/10.1016/j.biocon.2013.07.037>.
- Bjorkman, A.D., Elmendorf, S.C., Beamish, A.L., Vellend, M., and Henry G.H.R., 2015, Contrasting effects of warming and increased snowfall on Arctic tundra plant phenology over the past two decades: *Global Change Biology*, v. 21, no. 12, p. 4,651–4,661, <https://doi.org/10.1111/gcb.13051>.
- Browning, D.M., Rango, A., Karl, J.W., Laney, C.M., Vivoni, E.R., and Tweedie, C.E., 2015, Emerging technological and cultural shifts advancing drylands research and management: *Frontiers in Ecology and the Environment*, v. 13, no. 1, p. 52–60, <https://doi.org/10.1890/140161>.
- Chapman, D.S., Haynes, T., Beal, S., Essl, F., and Bullock, J.M., 2014, Phenology predicts the native and invasive range limits of common ragweed: *Global Change Biology*, v. 20, no. 1, p. 192–202. <https://doi.org/10.1111/gcb.12380>.
- Cleland, E.E., Chuine, I., Menzel, A., Mooney, H.A., and Schwartz, M.D., 2007, Shifting plant phenology in response to global change: *Trends in Ecology & Evolution*, v. 22, no. 7, p. 357–365, <https://doi.org/10.1016/j.tree.2007.04.003>.
- Crimmins, T.M., Crimmins, M.A., Gerst, K.L., Rosemartin, A.H., and Weltzin, J.F., 2017, USA National Phenology Network's volunteer-contributed observations yield predictive models of phenological transitions: *PLoS One*, v. 12, no. 8, art. e0182919, <https://doi.org/10.1371/journal.pone.0182919>.
- Denny, E.G., Gerst, K.L., Miller-Rushing, A.J., Tierney, G.L., Crimmins, T.M., Enquist, C.A.F., Guertin, P., Rosemartin, A.H., Schwartz, M.D., Thomas, K.A., and Weltzin, J.F., 2014, Standardized phenology monitoring methods to track plant and animal activity for science and resource management applications: *International Journal of Biometeorology*, v. 58, no. 4, p. 591–601, <https://doi.org/10.1007/s00484-014-0789-5>.
- Elmendorf, S.C., Jones, K.D., Cook, B.I., Diez, J.M., Enquist, C.A.F., Hufft, R.A., Jones, M.O., Mazer, S.J., Miller-Rushing, A.J., Moore, D.J.P., Schwartz, M.D., and Weltzin, J.F., 2016, The plant phenology monitoring design for The National Ecological Observatory Network: *Ecosphere*, v. 7, no. 4, art. e01303, <https://doi.org/10.1002/ecs2.1303>.
- Enquist, C.A.F., Kellermann, J.L., Gerst, K.L., and Miller-Rushing, A.J., 2014, Phenology research for natural resource management in the United States: *International Journal of Biometeorology*, v. 58, no. 4, p. 579–589, <https://doi.org/10.1007/s00484-013-0772-6>.
- Euskirchen, E.S., Carman, T.B., and McGuire, A.D., 2014, Changes in the structure and function of northern Alaskan ecosystems when considering variable leaf-out times across groupings of species in a dynamic vegetation model: *Global Change Biology*, v. 20, no. 3, p. 963–978, <https://doi.org/10.1111/gcb.12392>.
- Fuccillo, K.K., Crimmins, T.M., de Rivera, C.E., and Elder, T.S., 2015, Assessing accuracy in citizen science-based plant phenology monitoring: *International Journal of Biometeorology*, v. 59, no. 7, p. 917–926, <https://doi.org/10.1007/s00484-014-0892-7>.
- Gerst, K.L., Rossington, N.L., and Mazer, S.J., 2017, Phenological responsiveness to climate differs among four species of *Quercus* in North America: *Journal of Ecology*, v. 105, no. 6, p. 1,610–1,622, <https://doi.org/10.1111/1365-2745.12774>.

- Guertin, P., Barnett, L., Denny, E.G., and Schaffer, S.N., 2015, USA National Phenology Network Botany Primer: USA-NPN Education and Engagement Series 2015–001, accessed March 12, 2018, at https://www.usanpn.org/files/shared/files/USA-NPN_Botany-Primer.pdf.
- Guertin, P., Barnett, L., Denny, E.G., and Schaffer, S.N., 2017, USA National Phenology Network Phenophase Primer for Plants—Draft Version: USA-NPN Education and Engagement Series 2016–002, accessed March 12, 2018, at https://www.usanpn.org/files/npn/reports/USA-NPN-PhenophasePrimer_Section1-June2017.pdf.
- Hallman, C., and Arnott, H., 2015, Morphological and physiological phenology of *Pinus longaeva* in the White Mountains of California: Tree-Ring Research, v. 71, no. 1, p. 1–12, <https://doi.org/10.3959/1536-1098-71.1.1>.
- Isaac, N.J.B., van Strien, A.J., August, T.A., de Zeeuw, M.P., and Roy, D.B., 2014, Statistics for citizen science—Extracting signals of change from noisy ecological data: Methods in Ecology and Evolution, v. 5, no. 10, p. 1,052–1,060, <https://doi.org/10.1111/2041-210X.12254>.
- Jeong, S.J., and Medvigy, D., 2014, Macroscale prediction of autumn leaf coloration throughout the continental United States: Global Ecology and Biogeography, v. 23, no. 11, p. 1,245–1,254, <https://doi.org/10.1111/geb.12206>.
- Kern, A.L., Roehrig, G.H., Bhattacharya, D., Wang, J.Y., Finley, F.A., Reynolds, B.J., and Nam, Y., 2015, Drawing on place and culture for climate change education in native communities, in Mueller, M.P., and Tippins, D.J., eds., Ecojustice, citizen science and youth activism—Environmental discourses in science education: New York, Springer International Publishing, p. 121–138, https://doi.org/10.1007/978-3-319-11608-2_8.
- Kharrassi, Y.E., Mazri, M.A., Mabrouk, A., Nasser, B., and Mzouri, E.H.E., 2015, Flowering and fruiting phenology, and physico-chemical characteristics of 2-year-old plants of six species of *Opuntia* from eight regions of Morocco: Journal of Horticultural Science and Biotechnology, v. 90, no. 6, p. 682–688, <https://doi.org/10.1080/14620316.2015.11668731>.
- Mazer, S.J., Gerst, K.L., Matthews, E.R., and Evenden, A., 2015, Species-specific phenological responses to winter temperature and precipitation in a water-limited ecosystem: Ecosphere, v. 6, no. 6, art. 98, p. 1–27, <https://doi.org/10.1890/ES14-00433.1>.
- Mehdipoor, H., Zurita-Milla, R., Rosemartin, A., Gerst K.L., and Weltzin J.F., 2015, Developing a workflow to identify inconsistencies in volunteered geographic information—A phenological case study: PLoS One, v. 10, no. 10, art. e0140811, <https://doi.org/10.1371/journal.pone.0140811>.
- Melaas, E.K., Friedl, M.A., and Richardson, A.D., 2016, Multiscale modeling of spring phenology across deciduous forests in the eastern United States: Global Change Biology, v. 22, no. 2, p. 792–805, <https://doi.org/10.1111/gcb.13122>.
- Noormets, A., ed., 2009, Phenology of ecosystem processes—Applications in global change research: New York, Springer Science & Business Media.
- Pearse, W.D., Davis, C.C., Inouye, D.W., Primack, R.B., and Davies, T.J., 2017, A statistical estimator for determining the limits of contemporary and historic phenology: Nature Ecology & Evolution, v. 1, no. 12, p. 1,876–1,882, <https://doi.org/10.1038/s41559-017-0350-0>.
- Rosemartin, A.H., Crimmins, T.M., Enquist, C.A.F., Gerst, K.L., Kellermann, J.L., Posthumus, E.E., Denny, E.G., Guertin, P., Marsh, L., and Weltzin, J.F., 2014, Organizing phenological data resources to inform natural resource conservation: Biological Conservation, v. 173, p. 90–97, <https://doi.org/10.1016/j.biocon.2013.07.003>.
- Rosemartin, A.H., Denny, E.G., Weltzin, J.F., Marsh, R.L., Wilson, B.E., Mehdipoor, H., Zurita-Milla, R., and Schwartz, M.D., 2015, Lilac and honeysuckle phenology data 1956–2014: Scientific Data, vol. 2, art. 150038, <https://doi.org/10.1038/sdata.2015.38>.

- Schwartz, M.D., ed., 2003, *Phenology—An integrative environmental science*: New York, Springer Science and Business Media, 610 p., <https://doi.org/10.1007/978-94-007-6925-0>.
- Stucky, B. J., Deck, J., Denny, E.G., Guralnick, R.P., Walls, R.L., and Yost, J., 2016, W12–04: The Plant phenology ontology for phenological data integration: International Conference on Biomedical Ontology and BioCreative (ICBO BioCreative 2016), accessed March 12, 2018, at http://icbo2016.cgrb.oregonstate.edu/sites/default/files/W14-03_ICBO2016.pdf.
- USA-NPN National Coordinating Office, 2013, USA National Phenology Network Plant and Animal Observation Handbook: USA-NPN Education & Engagement Series 2013–001, 51 p., accessed March 12, 2018, at https://www.usanpn.org/files/shared/files/USA-NPN-HTOFull_8.27.13.pdf.
- USA-NPN National Coordinating Office, 2014, USA National Phenology Network Five Year Strategic Plan (F14-FY18): USA-NPN Programmatic Series 2014-001, 26 p., accessed March 21, 2018, at https://www.usanpn.org/files/shared/files/USA-NPN%20Five-Year%20Strategic%20Plan%20-%20January%202014_0.pdf.
- USA National Phenology Network Coordinating Office, 2016a, USA National Phenology Network Data Web Service API v1.2.1: USA-NPN web document, accessed March 12, 2018, at <https://tinyurl.com/hghrvml>.
- USA-NPN National Coordinating Office, 2016b, USA National Phenology Network data product development framework and data product catalog, v 1.1: USA-NPN Technical Series 2016–001, 46 p., accessed March 12, 2018, at https://www.usanpn.org/files/npn/reports/USA-NPN_Data-Product-Catalog1.1-2016Dec.pdf.
- USA-NPN National Coordinating Office, 2016c, USA National Phenology Network Data Management Plan: USA-NPN Technical Series, 10 p., accessed March 12, 2018, at https://www.usanpn.org/files/shared/USA-NPN_DMP_12-5-16.pdf.
- Wallace, C.S.A., Walker, J.J., Skirvin, S.M., Patrick-Birdwell, C., Weltzin, J.F., and Raichle, H., 2016, Mapping presence and predicting phenological status of invasive Buffelgrass in southern Arizona using MODIS, climate and citizen science observation data: *Remote Sensing*, v. 8, no. 7, 24 p., <https://doi.org/10.3390/rs8070524>.

Appendix 1: Datafield Definition Tables

Field names and descriptions for the primary dataset file for each data type, as well as for the ancillary data files are available as a set of tables. They may be downloaded in Excel format at the report index page, <https://doi.org/10.3133/ofr20181060>.

Appendix 2: Quality Assurance and Quality Control Measures

Quality Assurance Measures

Measures described below are organized by potential source of error. Measures not yet implemented appear in grey text.

1. Species identification

- A suite of “How to Observe” resources (<https://www.usanpn.org/nn/guidelines>) emphasize the importance of accurate species identification and direct observers to general identification resources.
- Species profile pages include a photo, a range map, in some cases a written description of the species, and links to resources with additional identification information.
- Messages in the interface ask observers to confirm species identification when an individual is registered at a location that is out of the species’ known geographic range.

2. Phenophase status evaluation

- The date and the presence/absence status of each phenophase (for example, “Open flowers”) are reported on each visit; observers are not asked to infer the date of an event (for example, date of “first flower”).
- Observers report presence/absence with a simple “Yes”/“No” and are given an “Uncertain” (“?”) option to reduce tendency for reporting false positives or negatives.
- Phenophase definitions are written precisely and in accessible language.
- Phenophase definitions are generalized and identical within phenological functional types for consistency. Species-specific additions to the general definitions more completely describe how the phenophase appears in a particular species.
- Frequently Asked Questions (FAQs) address tricky issues in phenophase status evaluation.
- Plant anatomy and phenophase evaluation are covered in a series of resources:
 - *Botany Primer* (Guertin and others, 2015)
 - *Phenophase Primer* (Guertin and others, 2017)
 - Webinars: <https://www.usanpn.org/nn/connect/WebinarArchive>
 - *Nature’s Notebook* Nuggets: <https://www.usanpn.org/nn/observe/questionsanswered>
 - Species-specific phenophase photo guides <https://www.usanpn.org/nn/groups/resources>

3. Data entry

- Datasheets mirror the online data entry form.
- Observers can reorder plant and animal lists such that the data entry form and datasheet printout mirror the order encountered at the field site.
- Species names and intensity measures are presented as pick lists in the online data entry interface, limiting the possible responses the user can provide.
- Phenophase definitions and intensity questions appear on rollovers in the data entry form.
- Site latitude and longitude can be provided a number of ways to facilitate the reporting of accurate coordinates. The user can place a marker on a Google Map, the user can geocode the location based on an address, or the user can manually provide a latitude and longitude. In every case, reverse geocoding is attempted as well to resolve the state in which the site is located.

- Elevation is automatically populated using Google Elevation Services but can be modified by the user.
 - Observers can review previously submitted observations in user interfaces (on data entry screens and via calendar visualizations) or via a downloaded Excel file, and can edit these observations.
 - Usability testing has been conducted on user interfaces to increase intuitiveness and reduce transcription errors.
 - When an observer deletes a plant from their list, a rationale for deletion is requested. Plant data are deleted from the database if the reason given is that the plant was a “test plant” and retained if the reason given is that the observer is “no longer observing”.
 - Mobile applications for data collection eliminate datasheet-to-interface transcription errors.
 - Training and FAQs address common data entry issues.
 - User interface validation measures include:
 - Controlled values are used whenever feasible, including for time, date, phenophase status, and categorical intensity measures.
 - The list of phenophases for each species is constrained to phenophases relevant to the species for which they are entering data.
 - Observers may not enter an intensity or abundance measure unless the phenophase status is set to “Yes” or “Uncertain”.
 - Date field is required; the default is for the observer to select the date using a calendar in the interface.
 - Dates in the future are not allowed.
 - Duplicate date and time values are not allowed.
 - Observers entering past observations are informed of changes in phenophase names and definitions through time.
4. Training
- Guidelines describing standardized field observing methods (selecting a site, selecting species, and making observations) are accessible via the suite of “How to Observe” resources (<https://www.usanpn.org/nn/guidelines>).
 - Detailed FAQs are available as context-specific help.
 - In-person and online workshops provide training opportunities for observers.
 - Local Phenology Leaders are supported by the NCO to provide on-the-ground training and sustained in-person support for observers.
 - The Local Phenology Leader Certification Program (<https://www.usanpn.org/nn/LPLCertification/>) provides virtual, in-person training to group leaders, with advanced training materials, to ensure observers are calibrated and sites are established according to USA-NPN guidelines.

Quality Control Measures

Measures described below are organized by error type. Measures not yet implemented appear in grey text.

1. Validation

- Sites with geographic coordinates outside of the United States, Canada, or Mexico are identified through an annual process that leverages Google’s Reverse Geocoder. These sites are manually

reviewed and, when possible, corrected (for example, when +/- signs preceding latitude and longitude are reversed).

- A nightly process corrects elevations below -90 meters or above 5000 meters, using Google's Digital Elevation Models for the site coordinates; user-defined elevation is retained in the database.
- The phenophase status reported by an observer is confirmed via expert review of images.

2. Reliability

- A nightly process flags instances where conflicting phenophase status records are present on the same date, regardless of time of day, for the same individual plant or species of animal at a site. These conflicting records can be submitted either by one observer and are flagged as a "One Observer Status Conflict", or by multiple observers and are flagged as a "Multiple Observer Status Conflict".
- Phenometric data provide information regarding the temporal precision of a phenophase transition by providing the number of days between a "Yes" and the preceding or following "No" report for a phenophase onset or end.
- Phenometric data are flagged as "Multiple Observers" when multiple observers contributed to a given series. This highlights potential issues with the series, as described in Individual Phenometrics Usage Notes section.
- Phenometric data are flagged as "Multiple First Yes" in instances of multiple onsets during the time period of interest.
- Detection bias in animal phenology reporting is exposed via observer reports of the time spent observing animals and their selection of an animal survey method from a pick list. Site area is provided in site-level ancillary data.

3. Plausibility

- Ancillary data at the site, plant, and site visit level (for example, land cover type for sites, watered status for plants) is available to enable exploration of outliers.
- Comments at the site, plant, and site visit level, submitted by observers are available to enable the exploration of outliers.
- Observers provide contact information to enable NCO-mediated communication regarding outliers or other issues.
- Self-reporting of training, skill, and experience level by observers is made available to data users (as of October 2016, 18 percent of active observers report these features), enabling exploration of outliers and prefiltering of data.
- A nightly process identifies and flags individual plants or animals registered at locations outside of the species' known geographic range.
- A nightly process identifies and flags phenophases reported in implausible order.
- A nightly process identifies and flags implausible changes in intensity measure values.
- An annual process flags outliers in phenophase onset and end dates relative to distribution of records that are similar geographically, climatically, and (or) taxonomically.