Leaf and Bloom Dates

Identification

1. Indicator Description

This indicator examines the timing of first leaf dates and flower bloom dates in lilacs and honeysuckle plants in the contiguous 48 states between 1900 and 2013. The first leaf date in these plants relates to the timing of events that occur in early spring, while the first bloom date is consistent with the timing of later spring events, such as the start of growth in forest vegetation. Lilacs and honeysuckles are especially useful as indicators of spring events because they are widely distributed across most of the contiguous 48 states and widely studied in the peer-reviewed literature. Scientists have very high confidence that recent warming trends in global climate have contributed to the earlier arrival of spring events (IPCC, 2014).

Components of this indicator include:

- Trends in first leaf dates and first bloom dates since 1900, aggregated across the contiguous 48 states (Figure 1)
- A map showing changes in first leaf dates between 1951–1960 and 2004–2013 (Figure 2)
- A map showing changes in first bloom dates between 1951–1960 and 2004–2013 (Figure 3)

2. Revision History

April 2010: Indicator posted
December 2011: Updated with data through 2010
December 2013: Combined original Figures 1 and 2 (leaf and bloom date time series) and updated with data through 2013; added new maps (Figures 2 and 3)

Data Sources

3. Data Sources

This indicator is based on leaf and bloom observations that were compiled by the USA National Phenology Network (USA-NPN) and climate data that were provided by the U.S. Historical Climatology Network (USHCN) and other databases maintained by the National Oceanic and Atmospheric Administration's (NOAA's) National Climatic Data Center (NCDC). Data for this indicator were analyzed using a method described by Schwartz et al. (2013).

4. Data Availability

Phenological Observations

This indicator is based in part on observations of lilac and honeysuckle leaf and bloom dates, to the extent that these observations contributed to the development of models. USA-NPN provides online access to historical phenological observations at: www.usanpn.org/?q=data_main.

Temperature Data

This indicator is based in part on historical daily temperature records, which are publicly available online through NCDC. For example, USHCN data are available online at: www.ncdc.noaa.gov/oa/climate/research/ushcn/#access, with no confidentiality issues limiting accessibility. Appropriate metadata and "readme" files are appended to the data so that they are discernible for analysis. For example, see: ftp://ftp.ncdc.noaa.gov/pub/data/ushcn/v2/monthly/readme.txt. Summary data from other sets of weather stations can be obtained from NCDC at: www.ncdc.noaa.gov/oa/ncdc.html.

Model Results

The processed leaf and bloom date data set is not publicly available. EPA obtained the model outputs by contacting Dr. Mark Schwartz at the University of Wisconsin–Milwaukee, who developed the analysis and created the original time series and maps. Results of this analysis have been published in Schwartz et al. (2013) and other papers.

Methodology

5. Data Collection

This indicator was developed using models that relate phenological observations (leaf and bloom dates) to weather and climate variables. These models were developed by analyzing the relationships between two types of measurements: 1) observations of the first leaf emergence and the first flower bloom of the season in lilacs and honeysuckles and 2) temperature data. The models were developed using measurements collected throughout the portions of the Northern Hemisphere where lilacs and/or honeysuckles grow, then applied to temperature records from a larger set of stations throughout the contiguous 48 states.

Phenological Observations

First leaf date is defined as the date on which leaves first start to grow beyond their winter bud tips. First bloom date is defined as the date on which flowers start to open. Ground observations of leaf and bloom dates were gathered by government agencies, field stations, educational institutions, and trained citizen scientists; these observations were then compiled by organizations such as the USA-NPN. These types of phenological observations have a long history and have been used to support a wide range of peer-reviewed studies. See Schwartz et al. (2013) and references cited therein for more information about phenological data collection methods.

Temperature Data

Weather data used to construct, validate, and then apply the models—specifically daily maximum and minimum temperatures—were collected from officially recognized weather stations using standard meteorological instruments. These data have been compiled by NCDC databases such as the USHCN and TD3200 Daily Summary of the Day data from other cooperative weather stations. As described in the methods for an earlier version of this analysis (Schwartz et al., 2006), station data were used rather than gridded values, "primarily because of the undesirable homogenizing effect that widely available coarse-resolution grid point data can have on spatial differences, resulting in artificial uniformity of processed outputs..." (Schwartz and Reiter, 2000; Schwartz and Chen, 2002; Menzel et al., 2003). Ultimately, 799 weather stations were selected according to the following criteria:

- Provide for the best temporal and spatial coverage possible. At some stations, the period of record includes most of the 20th century.
- Have at least 25 of 30 years during the 1981–2010 baseline period, with no 30-day periods missing more than 10 days of data.
- Have sufficient spring–summer warmth to generate valid model output.

For more information on the procedures used to obtain temperature data, see Schwartz et al. (2013) and references cited therein.

6. Indicator Derivation

Daily temperature data and observations of first leaf and bloom dates were used to construct and validate a set of models that relate phenological observations to weather and climate variables (specifically daily maximum and minimum temperatures). These models were developed for the entire Northern Hemisphere and validated at 378 sites in Germany, Estonia, China, and the United States.

Once the models were validated, they were applied to locations throughout the contiguous 48 states using temperature records from 1900 to 2013. Even if actual phenological observations were not collected at a particular station, the models essentially predict phenological behavior based on observed daily maximum and minimum temperatures, allowing the user to estimate the date of first leaf and first bloom for each year at that location. The value of these models is that they can estimate the onset of spring events in locations and time periods where actual lilac and honeysuckle observations are sparse. In the case of this indicator, the models have been applied to a time period that is much longer than most phenological observation records. The models have also been extended to areas of the contiguous 48 states where lilacs and honeysuckles do not actually grow—mainly parts of the South and the West coast where winter is too warm to provide the extended chilling that these plants need in order to bloom the following spring. This step was taken to provide more complete spatial coverage.

This indicator was developed by applying phenological models to several hundred sites in the contiguous 48 states where sufficient weather data have been collected. The exact number of sites varies from year to year depending on data availability (the minimum was 297 sites in 1901; the maximum was 771 sites in 1991).

After running the models, analysts looked at each location and compared the first leaf date and first bloom date in each year with the average leaf date and bloom date for 1981 to 2010, which was established as a "climate normal" or baseline. This step resulted in a data set that lists each station along with the "departure from normal" for each year—measured in days—for each component of the indicator (leaf date and bloom date). Note that 1981 to 2010 represents an arbitrary baseline for comparison, and choosing a different baseline period would shift the observed long-term trends up or down but would not alter the shape, magnitude, or statistical significance of the trends.

Figure 1. First Leaf and Bloom Dates in the Contiguous 48 States, 1900–2013

EPA obtained a data set listing annual departure from normal for each station, then performed some additional steps to create Figure 1. For each component of the indicator (leaf date and bloom date), EPA aggregated the data for each year to determine an average departure from normal across all stations. This step involved calculating an unweighted arithmetic mean of all stations with data in a given year. The aggregated annual trend line appears as a thin curve in each figure. To smooth out some of the year-to-year variability, EPA also calculated a nine-year weighted moving average for each component of the indicator. This curve appears as a thick line in each figure, with each value plotted at the center of the corresponding nine-year window. For example, the average from 2000 to 2008 is plotted at year 2004. This nine-year average was constructed using a normal curve weighting procedure that preferentially weights values closer to the center of the window. Weighting coefficients for values 1 through 9, respectively, were as follows: 0.0076, 0.036, 0.1094, 0.214, 0.266, 0.214, 0.1094, 0.036, 0.0076. This procedure was recommended by the authors of Schwartz et al. (2013) as an appropriate way to reduce some of the "noise" inherent in annual phenology data.

EPA used endpoint padding to extend the nine-year smoothed lines all the way to the ends of the period of record. Per the data provider's recommendation, EPA calculated smoothed values centered at 2010, 2011, 2012, and 2013 by inserting the 2009–2013 average into the equation in place of the as-yet unreported annual data points for 2014 and beyond. EPA used an equivalent approach at the beginning of the time series.

Figures 2 and 3. Change in First Leaf and Bloom Dates Between 1951–1960 and 2004–2013

To show spatial patterns in leaf and bloom changes, Figures 2 and 3 compare the most recent decade of data with the decade from 1951 to 1960 at individual stations. The 1950s were chosen as a baseline period to be consistent with the analysis published by Schwartz et al. (2013), who noted that broad changes in the timing of spring events appeared to start around the 1950s. To create the maps, EPA calculated the average departure from normal during each 10-year period and then calculated the difference between the two periods. The maps are restricted to stations that had at least eight years of valid data in both 10-year periods; 561 stations met these criteria.

For more information on the procedures used to develop, test, and apply the models for this indicator, see Schwartz et al. (2013) and references cited therein.

Indicator Development

The 2010 edition of EPA's *Climate Change Indicators in the United States* report presented an earlier version of this indicator based on an analysis published in Schwartz et al. (2006). That analysis was referred to as the Spring Indices (SI). The team that developed the original Spring Indices subsequently

developed an enhanced version of their algorithm, which is referred to as the Extended Spring Indices (SI-x). EPA adopted the SI-x approach for the 2012 edition of *Climate Change Indicators in the United States*. The SI-x represents an extension of the original SI because it can now characterize the timing of spring events in areas where lilacs and honeysuckles do not grow. Additional details about the SI-x are discussed in Schwartz et al. (2013).

For the 2014 edition of this indicator, EPA added a set of maps (Figures 2 and 3) to provide a more robust depiction of regional variations. These maps were published in Schwartz et al. (2013) and have since been updated with more recent data.

7. Quality Assurance and Quality Control

Phenological Observations

Quality assurance and quality control (QA/QC) procedures for phenological observations are not readily available.

Temperature Data

Most of the daily maximum and minimum temperature values were evaluated and cleaned to remove questionable values as part of their source development. For example, several papers have been written about the methods of processing and correcting historical climate data for the USHCN. NCDC's website (www.ncdc.noaa.gov/oa/climate/research/ushcn) describes the underlying methodology and cites peer-reviewed publications justifying this approach.

Before applying the model, all temperature data were checked to ensure that no daily minimum temperature value was larger than the corresponding daily maximum temperature value (Schwartz et al., 2006).

Model Results

QA/QC procedures are not readily available regarding the use of the models and processing the results. These models and results have been published in numerous peer-reviewed studies, however, suggesting a high level of QA/QC and review. For more information about the development and application of these models, see Schwartz et al. (2013), McCabe et al. (2012), and the references cited therein.

Analysis

8. Comparability Over Time and Space

Phenological Observations

For consistency, the phenological observations used to develop this indicator were restricted to certain cloned species of lilac and honeysuckle. Using cloned species minimizes the influence of genetic differences in plant response to temperature cues, and it helps to ensure consistency over time and space.

Temperature Data

The USHCN has undergone extensive testing to identify errors and biases in the data and either remove these stations from the time series or apply scientifically appropriate correction factors to improve the utility of the data. In particular, these corrections address changes in the time-of-day of observation, advances in instrumentation, and station location changes.

Homogeneity testing and data correction methods are described in more than a dozen peer-reviewed scientific papers by NCDC. Data corrections were developed to specifically address potential problems in trend estimation of the rates of warming or cooling in the USHCN. Balling and Idso (2002) compare the USHCN data with several surface and upper-air data sets and show that the effects of the various USHCN adjustments produce a significantly more positive, and likely spurious, trend in the USHCN data. In contrast, a subsequent analysis by Vose et al. (2003) found that USHCN station history information is reasonably complete and that the bias adjustment models have low residual errors.

Further analysis by Menne et al. (2009) suggests that:

> …the collective impact of changes in observation practice at USHCN stations is systematic and of the same order of magnitude as the background climate signal. For this reason, bias adjustments are essential to reducing the uncertainty in U.S. climate trends. The largest biases in the HCN are shown to be associated with changes to the time of observation and with the widespread changeover from liquid-in-glass thermometers to the maximum minimum temperature sensor (MMTS). With respect to [USHCN] Version 1, Version 2 trends in maximum temperatures are similar while minimum temperature trends are somewhat smaller because of an apparent overcorrection in Version 1 for the MMTS instrument change, and because of the systematic impact of undocumented station changes, which were not addressed [in] Version 1.

USHCN Version 2 represents an improvement in this regard.

Some observers have expressed concerns about other aspects of station location and technology. For example, Watts (2009) expresses concern that many U.S. weather stations are sited near artificial heat sources such as buildings and paved areas, potentially biasing temperature trends over time. In response to these concerns, NOAA analyzed trends for a subset of stations that Watts had determined to be "good or best," and found the temperature trend over time to be very similar to the trend across the full set of USHCN stations (www.ncdc.noaa.gov/oa/about/response-v2.pdf). While it is true that many stations are not optimally located, NOAA's findings support the results of an earlier analysis by Peterson (2006) that found no significant bias in long-term trends associated with station siting once NOAA's homogeneity adjustments have been applied.

Model Results

The same model was applied consistently over time and space. Figure 1 generalizes results over space by averaging station-level departures from normal in order to determine the aggregate departure from normal for each year. This step uses a simple unweighted arithmetic average, which is appropriate given the national scale of this indicator and the large number of weather stations spread across the contiguous 48 states.

9. Data Limitations

Factors that may impact the confidence, application, or conclusions drawn from this indicator are as follows:

1. Plant phenological events are studied using several data collection methods, including satellite images, models, and direct observations. The use of varying data collection methods in addition to the use of different phenological indicators (such as leaf or bloom dates for different types of plants) can lead to a range of estimates of the arrival of spring.

2. Climate is not the only factor that can affect phenology. Observed variations can also reflect plant genetics, changes in the surrounding ecosystem, and other factors. This indicator minimizes genetic influences by relying on cloned plant species, however (that is, plants with no genetic differences).

10. Sources of Uncertainty

Error estimates are not readily available for the underlying temperature data upon which this indicator is based. It is generally understood that uncertainties in the temperature data increase as one goes back in time, as there are fewer stations early in the record. However, these uncertainties are not sufficient to mislead the user about fundamental trends in the data.

In aggregating station-level "departure from normal" data into an average departure for each year, EPA calculated the standard error of each component of Figure 1 (leaf date and bloom date) in each year. For both components, standard errors range from 0.3 days to 0.5 days, depending on the year.

Uncertainty has not been calculated for the individual station-level changes shown in Figures 2 and 3.

Schwartz et al. (2013) provide error estimates for the models. The use of modeled data should not detract from the conclusions that can be inferred from the indicator. These models have been extensively tested and refined over time and space such that they offer good certainty.

11. Sources of Variability

Temperatures naturally vary from year to year, which can strongly influence leaf and bloom dates. To smooth out some of the year-to-year variability, EPA calculated a nine-year weighted moving average for each component of this indicator in Figure 1, and EPA created the maps in Figures 2 and 3 based on 10-year averages for each station.

12. Statistical/Trend Analysis

Statistical testing of individual station trends within the contiguous 48 states suggests that many of these trends are not significant. Other studies (e.g., Schwartz et al., 2006) have come to similar conclusions, finding that trends in the earlier onset of spring at individual stations are much stronger in Canada and parts of Eurasia than they are in the contiguous 48 states. In part as a result of these findings, Figure 1 focuses on aggregate trends across the contiguous 48 states, which should be more

statistically robust than individual station trends. However, the aggregate trends still are not statistically significant ($p<0.05$) over the entire period of record, based on a simple t-test.

References

Balling, Jr., R.C., and C.D. Idso. 2002. Analysis of adjustments to the United States Historical Climatology Network (USHCN) temperature database. Geophys. Res. Lett. 29(10):1387.

IPCC (Intergovernmental Panel on Climate Change). 2014. Climate change 2014: Impacts, adaptation, and vulnerability. Working Group II contribution to the IPCC Fifth Assessment Report. Cambridge, United Kingdom: Cambridge University Press. www.ipcc.ch/report/ar5/wg2.

McCabe, G.J., T.R. Ault, B.I. Cook, J.L. Betancourt, and M.D. Schwartz. 2012. Influences of the El Niño Southern Oscillation and the Pacific Decadal Oscillation on the timing of the North American spring. Int. J. Climatol. 32:2301–2310.

Menne, M.J., C.N. Williams, Jr., and R.S. Vose. 2009. The U.S. Historical Climatology Network monthly temperature data, version 2. Bull. Am. Meteorol. Soc. 90:993-1107. ftp://ftp.ncdc.noaa.gov/pub/data/ushcn/v2/monthly/menne-etal2009.pdf.

Menzel, A., F. Jakobi, R. Ahas, et al. 2003. Variations of the climatological growing season (1051–2000) in Germany compared to other countries. Int. J. Climatol. 23:793–812.

Peterson, T.C. 2006. Examination of potential biases in air temperature caused by poor station locations. Bull. Am. Meteorol. Soc. 87:1073–1080. http://journals.ametsoc.org/doi/pdf/10.1175/BAMS-87-8-1073.

Schwartz, M.D., and X. Chen. 2002. Examining the onset of spring in China. Clim. Res. 21:157–164.

Schwartz, M.D., and B.E. Reiter. 2000. Changes in North American spring. Int. J. Climatol. 20:929–932.

Schwartz, M.D., R. Ahas, and A. Aasa. 2006. Onset of spring starting earlier across the Northern Hemisphere. Glob. Chang. Biol. 12:343–351.

Schwartz, M.D., T.R. Ault, and J.L. Betancourt. 2013. Spring onset variations and trends in the continental United States: Past and regional assessment using temperature-based indices. Int. J. Climatol. 33:2917–2922.

Vose, R.S., C.N. Williams, Jr., T.C. Peterson, T.R. Karl, and D.R. Easterling. 2003. An evaluation of the time of observation bias adjustment in the U.S. Historical Climatology Network. Geophys. Res. Lett. 30(20):2046.

Watts, A. 2009. Is the U.S. surface temperature record reliable? The Heartland Institute. http://wattsupwiththat.files.wordpress.com/2009/05/surfacestationsreport_spring09.pdf.