Trends and Natural Variability of Spring Onset in the Coterminous United States as Evaluated by a New Gridded Dataset of Spring Indices

Toby R. Ault
Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, New York

Mark D. Schwartz
Department of Geography, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin

Raul Zurita-Milla
Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, Netherlands

Jake F. Weltzin

Julio L. Betancourt
U.S. Geological Survey, Reston, Virginia

(Manuscript received 29 October 2014, in final form 22 June 2015)

ABSTRACT

Climate change is expected to modify the timing of seasonal transitions this century, impacting wildlife migrations, ecosystem function, and agricultural activity. Tracking seasonal transitions in a consistent manner across space and through time requires indices that can be used for monitoring and managing biophysical and ecological systems during the coming decades. Here a new gridded dataset of spring indices is described and used to understand interannual, decadal, and secular trends across the coterminous United States. This dataset is derived from daily interpolated meteorological data, and the results are compared with historical station data to ensure the trends and variations are robust. Regional trends in the first leaf index range from 2.0 to 2.1 days decade⁻¹, while first bloom index trends are between −0.4 and −1.2 for most regions. However, these trends are modulated by interannual to multidecadal variations, which are substantial throughout the regions considered here. These findings emphasize the important role large-scale climate modes of variability play in modulating spring onset on interannual to multidecadal time scales. Finally, there is some potential for successful subseasonal forecasts of spring onset, as indices from most regions are significantly correlated with antecedent large-scale modes of variability.

1. Introduction

Understanding the effects of climate change on seasonality requires metrics that can clearly and consistently identify seasonal transitions through time and across large spatial scales. Indicators of phenological events may be particularly useful in this regard because they tend to integrate weather and climatic information through time (Lieth 1974; Schwartz et al. 2012). Phenology is linked to primary productivity, crop yield, insect emergence, and bird migration; it is also a powerful indicator of climate change impacts (Rosenzweig et al. 2007; Parmesan 2006). Phenological events can further act as proxies for seasonal transitions that occur in other biological and physical variables alike (e.g., Schwartz 1996; McCabe et al. 2011, 2013).
Phenological observations are collected from individual sites by professional scientists (Cook et al. 2012), across well-organized regional and national networks by volunteers (e.g., van Vliet et al. 2003; Schwartz et al. 2012), and by passive and active sensors mounted on Earth-observing and Earth-orbiting satellites (e.g., White et al. 2009; de Beurs and Henebry 2010; Zurita-Milla et al. 2013). Although these observational records are invaluable, they are seldom widespread, continuous, or long enough on their own to identify and evaluate interannual variations and decadal trends, much less understand their climate drivers; that is, some observer-based records may be quite long, but they are restricted to a relatively small geographic scale, whereas satellite data are continuous in space but only cover the last few decades. Given the incomplete empirical data, a more suitable approach to understanding long-term trends and interannual variations in seasonality may be to apply weather-based phenological models to spatially complete gridded meteorological data.

A few such weather-based phenological indices have been developed and employed to understand the patterns and causes of variations in spring onset (e.g., Jolly et al. 2005; Schwartz et al. 2006). Such indices have the advantage of being calculable anywhere that meteorological data are available, enabling researchers to use them across much wider geographic domains and longer periods than those available from observational phenology alone. One such suite of indices, the “spring indices” (Schwartz et al. 2006), has been particularly useful in this regard because the algorithms require only daily temperature minima ($T_{\text{min}}$) and maxima ($T_{\text{max}}$) as inputs, which are widely available from many meteorological stations throughout the world for most of the twentieth century (e.g., Klein Tank et al. 2002). The spring indices use models of lilac (Syringa chinensis, “Red Rothomagensis”) and honeysuckle (Lonicera tatarica, “Arnold Red”; and L. korolkowii, “Zabeli”), validated with phenological observations, as proxies for the beginning of the growing season at mid- to high latitudes (e.g., Schwartz and Reiter 2000). Spring indices calculated from station data have recently been employed to understand hemispheric-scale changes in spring onset (Schwartz et al. 2006), the relationship between spring onset and snowmelt in the western United States (McCabe

![FIG. 1. Regions used to produce average time series.](image1)

![FIG. 2. Schematic diagram of how the seasonally appropriate windows are identified from spring indices.](image2)
et al. 2013), the interannual-to-decadal predictability of spring onset across North America (McCabe et al. 2011), and large-scale patterns of trends and interannual variability in spring onset and "false springs" across North America (Ault et al. 2011; Schwartz et al. 2013; Ault et al. 2013; Peterson and Abatzoglou 2014).

Our primary goal in this study is to document regional and continental-scale trends and variability in the timing of spring using spring indices calculated from both station data and a new gridded $T_{\text{min}}/T_{\text{max}}$ product, the Berkeley Earth dataset (Mueller et al. 2013; http://berkeleyearth.org/), for the coterminous United States. By focusing on gridded data with long temporal coverage (1920–2013), we hope to gain better insight into the regional patterns and trends, as well as establish new benchmarks for climate models, especially those used to understand and prepare for climate change (e.g., Taylor et al. 2012).

2. Methods and data

The original spring indices (SI-o) were developed as phenological models of lilac and honeysuckle leafing and blooming (Schwartz et al. 2002, 2006). As such, they incorporated "chilling requirements," or the minimum exposure (measured in chilling hours) that some plants need to initiate their dormant period in winter. Subtropical climates rarely satisfy such chilling requirements, precluding the calculation of SI-o across all of the coterminous United States. Recently, however, Schwartz et al. (2013) lifted the chilling requirements and found that the resulting "extended" spring indices (SI-x) were equally valid in regions where SI-o had been used previously while expanding coverage to subtropical areas of the United States. Accordingly, Schwartz et al. (2013) suggested that the SI-x offer insight into both large-scale climate processes and local phenological responses. Here we extend this approach further by calculating SI-x from gridded observational datasets and a much denser network of station data than employed previously (e.g., as in Schwartz et al. 2013; Ault et al. 2013), and we document the regional patterns of interannual-to-decadal variability and long-term trends in these datasets.

A detailed explanation of the underlying assumptions, algorithm, and code for calculating SI-x is provided in Ault et al. (2015). Briefly, the indices are based on averages of phenological models for three plants (one species of lilac and two species of honeysuckle), validated with widespread phenophase observational data, and require only daily $T_{\text{min}}$ and $T_{\text{max}}$ as input (as well as latitude to compute day length). As such, they can be...
calculated across broad geographic domains and provide a consistent metric of the same event (e.g., the emergence of spring foliage) through time. The two primary variables comprising SI-x are the “first leaf index” (the average date when the first leaves appear on the three plants) and “first bloom index” (the average date when blossoms first appear). Conceptually, the leaf index indicates the first phase of the transition from winter to spring, typically when shrubs start to become active but before deciduous trees put on leaves. The bloom index occurs later and is indicative of the window of time when lilacs (and other shrubs) first start blossoming and deciduous trees “leaf out.” Because the bloom index lags the leaf index by several weeks, it is necessarily more indicative of the weather patterns later in the growing season.

Here we calculate SI-x from two observational sources. First, we used daily station data from 2858 sites in the Global Historical Climatology Network (GHCN) (Klein Tank et al. 2002). As in Schwartz et al. (2006), sites were used if they had sufficient daily data from each year (no more than 30 days total missing during any given spring; no gaps longer than 10 consecutive days) to calculate SI-x from at least 25 years out of the 30-yr reference period from 1981 to 2010.

Second, to produce spatially complete maps of spring onset, we employed the gridded daily $T_{\text{min}}/T_{\text{max}}$ Berkeley Earth data product with 1° spatial resolution (Mueller et al. 2013). Although daily temperatures from this product are available going back to 1880, only the 1920–2013 time interval is used because of the relatively high density of instrumental data going into the product over this time as compared with earlier periods. As with other global gridded data products, the process for constructing the Berkeley Earth dataset entails statistically interpolating a sparse, uneven network of stations (GHCN sites) to a rectangular (latitude × longitude) grid. The details of this process, described in Mueller et al. (2013), differ from other similar efforts primarily in how discontinuities in the underlying data are treated. Rather than homogenizing the underlying station data, it treats periods of individual stations separated by discontinuities as distinct records. It also uses a greater number of total stations and generally exhibits less somewhat lower levels of uncertainty than other similar products.

The chief advantage of the Berkeley Earth dataset for our purposes is that it is one of just a very small handful of gridded daily $T_{\text{min}}/T_{\text{max}}$ products with the dual advantages of affording relatively high spatial resolution
and long temporal coverage. It is also updated regularly, allowing our analysis to diagnose trends and long-term variations in spring onset until relatively recently. Other similar products include reanalysis data (e.g., Kalnay et al. 1996) and high-resolution gridded products such as PRISM (Daly et al. 1994) and DAYMET (Thornton et al. 1997). These datasets generally cover a shorter period of time (in most cases only from the satellite era onward). Long datasets with monthly resolution are inadequate for our purposes because the underlying SI-x models require daily \( T_{\text{min}}/T_{\text{max}} \) as input.

The most similar type of product to the Berkley Earth data used here is the Hadley Centre Global Historical Climatology Network–Daily (HadGHCND) dataset (Caesar et al. 2006), but even this only covers the period from 1950 onward, and the spatial resolution is \( 2.5^\circ \times 3.75^\circ \) rather than \( 1^\circ \times 1^\circ \) (we nonetheless computed SI-x from this dataset and found the results to be consistent with the ones presented here). Hence, for diagnosing long-term trends and interannual variations over the twentieth century, we argue that the Berkeley Earth dataset reflects a good compromise between long-term coverage, spatial resolution, and uncertainty introduced by the underlying data. Moreover, it is produced at spatial scales (hundreds of kilometers) that are quite comparable to those of existing global climate models, making it possible to compare the SI-x benchmarks we establish here to model output.

Monthly indices of large-scale modes of climate variability were obtained from the National Oceanic and Atmospheric Administration’s (NOAA) Physical Sciences Division (PSD) of the Earth System Research Laboratory (ESRL) “climate indices” webpage (http://www.esrl.noaa.gov/psd/data/climateindices/list/). The indices considered here include 1) the Niño-3.4 time series, as an indication of the state of the El Niño–Southern Oscillation (ENSO), defined as average sea surface temperature (SST) from the Pacific Ocean between \( 5^\circ S \) and \( 5^\circ N \) and \( 120^\circ W \) and \( 170^\circ W \); 2) the Pacific decadal oscillation (PDO) (Zhang et al. 1997; Mantua et al. 1997), defined as the first principal component of Pacific SST poleward of \( 20^\circ N \); 3) the Atlantic multidecadal oscillation (AMO), computed as the detrended, area-weighted average of SST in the Atlantic between \( 0^\circ \) and \( 70^\circ N \) (Enfield et al. 2001); 4) the Pacific–North American pattern (PNA), identified from 500-mb height anomalies in NCEP reanalysis (http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/month_pna_index2.shtml); and finally 5) the North Atlantic Oscillation (Hurrell 1996), defined again using the 500-mb height anomalies and methods used to define the PNA (see http://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/history/method.shtml for more information).

We used several simple metrics to compare SI-x station data with gridded data from the same region over
the reference period of 1981–2010. First, we estimated the bias of gridded mean values by identifying each station encompassed by each grid cell, then we subtracted the grid cell average from each station’s average. Second, we calculated root-mean-square errors between time series of each station and each underlying cell. Third, we correlated station data and gridded data at each station within each cell. Fourth, linear trends were computed over the 1981–2010 period and removed to compute standard deviations of detrended station and gridded data.

The gridded first leaf and bloom data for the period from 1955 to 2013 were grouped into regional domains (Fig. 1) similar to the nine NCDC U.S. climate divisions (Karl and Koss 1984), shown in Fig. 1. These time series are used to identify a “seasonally appropriate” temporal window for each region for each year to diagnose climate influences on spring onset; that is, for a given year and region, we are interested in understanding what aspects of regional climate are most anomalous right around the beginning of spring. Therefore we identify an index of days for a window of time 10 days before and 5 days after a given region’s leaf index average. This is shown schematically for a single time series in Fig. 2. Seasonally appropriate windows are then used to connect year-to-year variations in spring onset with 250-mb geopotential height and wind fields, which are obtained from the National Centers for Environmental Prediction (NCEP) daily reanalysis product (Kalnay et al. 1996). In total, nine of these seasonally adjusted fields were produced (one for each region) by identifying the temporal window surrounding the first leaf date for each region during each year.

Each regional average is also correlated with the large-scale climate indices described earlier. As most of these indices are defined at monthly resolution, the month prior to the mean onset date for that region is used. This decision is in part necessary because only climatic information leading up to an event can be relevant to it, but it also allows us to explore the potential predictability of spring onset on subseasonal and longer lead times.

3. Results

Station-based versus gridded spring indices

Average first leaf and first bloom index values from station (GHCN) and gridded (Berkeley Earth) data are shown in Fig. 3. First leaf index averages for both datasets follow elevational and latitudinal gradients in the western United States, with greater mean values (later dates) occurring at higher latitudes and altitudes. In the eastern United States, the gradients are primarily dominated by latitude. The first bloom index exhibits a similar geographic pattern, albeit with greater values...
(later dates) because blooming occurs after leafing in the lilac and honeysuckle species comprising the SI-x models. The underlying gridded data and the station data exhibit similar spatial patterns.

Comparisons between the GHCN and the gridded datasets reveal biases in mean dates in the gridded data that are generally greater in the mountainous West than in the East (Fig. 4). Specifically, throughout the Cascade and Rocky Mountain ranges north of about 38°N, many of the stations exhibit first leaf and first bloom dates that are earlier than the underlying grid. In the mountain ranges south of 38°N, the biases are reversed, with station dates being later than the underlying grid. Systematic biases farther east are small and generally less than 5 days, meaning that they are within the uncertainty of the original models (e.g., Schwartz et al. 2006).

Root-mean-square error (RMSE) values between the time series of each station and each corresponding grid are illustrated in Fig. 5. Again, the greatest discrepancies between the gridded product and the station data are found in the mountainous regions of the western United States, with values between 20 and 30 days of error. Despite the biases and larger RMSE values between some stations and the gridded data, temporal correlations between the two products are quite strong throughout the domain (Fig. 6), although they are still lowest in mountainous western regions.

Detrended standard deviations (Fig. 7) are lower for the gridded data than for the raw station data, especially for the first leaf index between 45° and 50°N. Nonetheless, both products exhibit broadly similar geographic patterns, with higher standard deviations found in a belt stretching diagonally from the Pacific Northwest to the central Atlantic coast. Standard deviations are lower in California and throughout the southern portion of the domain. First bloom index standard deviations are similar to those of the first leaf index but slightly lower (Fig. 7).

The patterns of trends in the first leaf and first bloom indices are very similar for the station and gridded data (Fig. 8). In the northwest Cascades, trends are positive (toward later onset dates), whereas they are negative throughout most of the rest of the domain. The area of greatest negative trend is centered on the Great Lakes for the first leaf index, and it is displaced slightly to the south of the Great Lakes for the first bloom index.

Overall, the geographic patterns of the means, standard deviations, and trends are similar between the gridded and station-based SI-x. Likewise, the interannual variations of both products covary quite strongly—so much so that most correlation values are close to unity, implying that any biases in mean values (leading also to high RMSE values) do not influence unduly the interannual variations of either product on regional scales.
Averages of the first leaf and bloom indices are plotted in Fig. 9 for each of the nine regions. Statistics summarizing the means, standard deviations, and trends of each region are provided in Table 1. Table 2 summarizes the results of correlating each regional first leaf index time series with antecedent large-scale modes of climate variability, and Table 3 reports the first bloom index results. Significant negative correlations (positive climate index values with early spring) are found between the Northwest and Niño-3.4, the PDO, and the PNA for both the first leaf index and first bloom index time series. Conversely, the first leaf and first bloom indices from the South and Southeast are positively correlated with these same three modes and negatively correlated with the North Atlantic Oscillation (NAO). Moreover, the NAO correlates negatively with the West, Southwest, and Central regions.

Correlations between each region’s average leaf index and 250-mb geopotential heights of seasonally appropriate anomalies (see methods) are shown in Fig. 10. In this figure, negative correlations associate high geopotential height anomalies (indicative of ridging) with early spring, or low heights (troughs) with late spring. Positive correlations relate low height anomalies (troughs) with late spring or high height anomalies (ridges) with early spring. In the Northwest, for example, interannual variations in spring onset are negatively correlated with geopotential heights regionally as well as with those of the central subtropical Pacific and Caribbean sectors (red and orange, Fig. 10). They are positively correlated with heights in the central North Pacific and eastern half of the continent (blue colors, Fig. 10). The pattern for the West exhibits just one center of significant correlation, while the Southwest is negatively correlated with regional height anomalies and positively correlated with anomalies farther south. Correlation patterns for the West North Central, East North Central, and Northeast regions are similar to each other, with early spring associated with high heights throughout the central part of the continent and to a lesser extent near the Hawaiian Islands. The correlation patterns for the Central, South, and Southeast regions are similar to one another, with local heights negatively correlated with spring timing, and positive correlations farther south in the case of the South and Southeast.

4. Discussion

a. Consistency in gridded and station-based indices

Results from calculating spring indices from GHCN station data as well as a gridded $T_{\text{min}}/T_{\text{max}}$ dataset yield a qualitatively similar picture of the means, trends, and variability in the timing of spring in the coterminous United States. Spring onset (as indicated by the first leaf
Index) occurs between January and February for the most southern regions, March for the West Coast and much of the central part of the continent (Fig. 3), and April for the Great Lakes, Northeast, and Rockies. Mean onset dates also follow meridional and elevational gradients, with later dates [April–June (AMJ)] occurring for both the first leaf and first bloom index at greater elevations and higher latitudes. Biases in the gridded product, however, show systematic regional differences, with most grid cells indicating dates that are later than the station data for the mountainous West north of 38°N (Fig. 4). South of that latitude, for the mountainous West north of 38°N (Fig. 4).

![Fig. 9. Time series of mean annual first leaf (green) and first bloom (red) index dates for each region shown in Fig. 1. Note that the values of the y axes differ among panels, but they all cover a range of 90 days.](image)

**Table 1.** Summary statistics for each of the nine regions’ averages in Fig. 9. The first column is the region name from Fig. 1. The next three columns report the mean (integer value), standard deviation, and trend for the first leaf index. The next three columns report the mean, standard deviation, and trend for the first bloom index averages for each region. Units of the mean and standard deviation values are in days, while the units of the trends are in days decade⁻¹. All values are computed over the 1955–2013 period.

<table>
<thead>
<tr>
<th>Region</th>
<th>Leaf index</th>
<th>Bloom index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std dev</td>
</tr>
<tr>
<td>Northwest (42°–49°N, 125°–112°W)</td>
<td>98</td>
<td>6.7</td>
</tr>
<tr>
<td>West (32.5°–42°N, 125°–114°W)</td>
<td>55</td>
<td>4.3</td>
</tr>
<tr>
<td>Southwest (31.25°–41°N, 114°–102.5°W)</td>
<td>76</td>
<td>5.9</td>
</tr>
<tr>
<td>West North Central (41°–49°N, 112°–96.5°W)</td>
<td>108</td>
<td>6.6</td>
</tr>
<tr>
<td>East North Central (42°–49°N, 96.5°–82.5°W)</td>
<td>112</td>
<td>7.2</td>
</tr>
<tr>
<td>Central (35°–42°N, 91°–80.5°W)</td>
<td>76</td>
<td>7.9</td>
</tr>
<tr>
<td>South (27°–38°N, 102.5°–91°W)</td>
<td>42</td>
<td>6.0</td>
</tr>
<tr>
<td>Northeast (39°–47.5°N, 80.5°–67°W)</td>
<td>106</td>
<td>6.2</td>
</tr>
<tr>
<td>Southeast (29°–35°N, 91°–76°W)</td>
<td>33</td>
<td>7.0</td>
</tr>
</tbody>
</table>
TABLE 2. Correlation coefficients between each regional leaf index time series and several well-known indices of large-scale climate variability for the previous month. The previous (prev) month is used here to investigate the potential for predictability on subseasonal time scales. The month used from each index is indicated in parentheses in the second column. Correlations that are significant at the 95% confidence limit are shown in bold.

<table>
<thead>
<tr>
<th>Region</th>
<th>Climate index</th>
<th>Prev month</th>
<th>Niño-34</th>
<th>PDO</th>
<th>AMO</th>
<th>NAO</th>
<th>PNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northwest</td>
<td>(Apr)</td>
<td>-0.44</td>
<td>-0.45</td>
<td>-11</td>
<td>-0.25</td>
<td>-0.46</td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>(Feb)</td>
<td>-0.01</td>
<td>-0.19</td>
<td>0</td>
<td>-0.32</td>
<td>-0.22</td>
<td></td>
</tr>
<tr>
<td>Southwest</td>
<td>(Mar)</td>
<td>0.35</td>
<td>0</td>
<td>0.04</td>
<td>-0.32</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>West North</td>
<td>(Apr)</td>
<td>-0.08</td>
<td>-0.16</td>
<td>-0.05</td>
<td>-0.22</td>
<td>-0.06</td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>(Apr)</td>
<td>-0.09</td>
<td>-0.24</td>
<td>-0.21</td>
<td>0.08</td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>(Mar)</td>
<td>0.15</td>
<td>0.23</td>
<td>0.14</td>
<td>-0.45</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>(Feb)</td>
<td>0.34</td>
<td>0.24</td>
<td>-0.04</td>
<td>-0.43</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>(Apr)</td>
<td>-0.11</td>
<td>-0.24</td>
<td>-0.07</td>
<td>-0.17</td>
<td>-0.17</td>
<td></td>
</tr>
<tr>
<td>Southeast</td>
<td>(Feb)</td>
<td>0.33</td>
<td>0.35</td>
<td>-0.02</td>
<td>-0.47</td>
<td>0.47</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 3. Correlation coefficients between each regional bloom index time series and several well-known indices of large-scale climate variability for the previous month. The previous month is used here to investigate the potential for predictability on subseasonal time scales. The month used from each index is indicated in parentheses in the second column. Correlations that are significant at the 95% confidence limit are shown in bold.

<table>
<thead>
<tr>
<th>Region</th>
<th>Climate index</th>
<th>Prev month</th>
<th>Niño-34</th>
<th>PDO</th>
<th>AMO</th>
<th>NAO</th>
<th>PNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northwest</td>
<td>(May)</td>
<td>-0.54</td>
<td>-0.59</td>
<td>-0.1</td>
<td>-0.41</td>
<td>-0.32</td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>(Mar)</td>
<td>0</td>
<td>-0.42</td>
<td>0.12</td>
<td>-0.35</td>
<td>-0.27</td>
<td></td>
</tr>
<tr>
<td>Southwest</td>
<td>(Apr)</td>
<td>0.27</td>
<td>0.07</td>
<td>0.13</td>
<td>-0.34</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>West North</td>
<td>(May)</td>
<td>-0.05</td>
<td>-0.24</td>
<td>0.02</td>
<td>-0.54</td>
<td>-0.22</td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>(May)</td>
<td>0.02</td>
<td>-0.08</td>
<td>-0.09</td>
<td>-0.26</td>
<td>-0.18</td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>(Apr)</td>
<td>0.12</td>
<td>0.04</td>
<td>0.07</td>
<td>-0.36</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>(Mar)</td>
<td>0.37</td>
<td>0.31</td>
<td>0.16</td>
<td>-0.51</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>(May)</td>
<td>-0.08</td>
<td>-0.09</td>
<td>-0.08</td>
<td>-0.21</td>
<td>-0.17</td>
<td></td>
</tr>
<tr>
<td>Southeast</td>
<td>(Feb)</td>
<td>0.44</td>
<td>0.33</td>
<td>0.1</td>
<td>-0.49</td>
<td>0.32</td>
<td></td>
</tr>
</tbody>
</table>

High-elevation sites tend to indicate earlier dates than the underlying grid. East of the Rockies, biases are small (within 5 days) and are in fact within the uncertainty of the original first leaf and first bloom models (Schwartz et al. 2006). Despite these biases (see also Fig. 5), correlations between station data and local grid points are quite strong and almost universally positive and close to one (Fig. 6). Hence, interannual variations in spring onset are very similar between the station data and the grid even when they are offset in time from one another. The patterns of interannual variability (Fig. 7) and trends (Fig. 8) between 1981 and 2010 likewise support this interpretation; hence, we focus the remainder of our discussion on results obtained from the gridded SI-x data.

b. Climate indices

The results in Tables 2 and 3 lend themselves to a relatively straightforward interpretation, especially in light of the correlation results in Fig. 10. The Niño-3.4, PDO, and PNA time series are connected with each other, although they also vary independently (Wallace and Gutzler 1981; Thompson and Wallace 1998; Zhang et al. 1997; Mantua et al. 1997). Positive values in any of these climate indices are indicative of anomalous high pressure over western Canada and lower-than-average pressure in the Southeast. Such conditions favor more warm-air intrusion earlier into the Northwest, connecting high Niño-3.4, PDO, or PNA index values with early spring (e.g., as also noted in Cayan et al. 2001; Ault et al. 2011; McCabe et al. 2011). The positive phases of these modes also increase the number of cold-air outbreaks in the Southeast, prolonging spring onset (e.g., Schwartz et al. 2013). Anomalies of the opposite sign occur during the negative phases of these modes, favoring anomalously late spring in the Northwest and early spring in the South and Southeast.

The role of the NAO in modulating spring onset is similar to that of the PNA, although the regions affected by it are different. In particular, high index phases favor a poleward jet in the eastern United States (e.g., Thompson and Wallace 2001) and consequently earlier spring. Low index phases of the NAO also play a major role because they are indicative of a more meridional jet stream (Thompson and Wallace 2001), which in turn allow for more frequent cold outbreaks and delayed spring in the Southeast. Hence, the negative correlations between regional first leaf index time series, and almost all first bloom index time series, with the NAO suggests that cold outbreaks during its low index periods can delay spring, while the absence of such events during its warm phase favors earlier spring.

c. Regional trends

While regional and hemispheric variations in the atmosphere help us to interpret interannual variations in spring onset, the length (and quality) of reanalysis data before the satellite era make it difficult to conduct comparable analyses over longer time horizons with the same level of detail (e.g., Schneider et al. 2013). Importantly, however, the patterns of trends in the gridded SI-x data are quite different for longer time horizons than they are for the relatively short period of the satellite era. Over the period from 1979 to 2013, when satellite data could be
available, trends in the conterminous United States show an apparent dipole pattern with “centers of action” in the northwest Cascade and Great Lakes regions (boxes in Fig. 11). The maps for the longer 1920–2013 period, on the other hand, are very different for both the first bloom and first leaf indices with more uniform warming across the continent. There is also evidence of a “warming hole” (Pan et al. 2004; Kunkel et al. 2006; Meehl et al. 2012; Schwartz et al. 2013), a region where spring onset was either delayed, or did not advance as fast as elsewhere in the United States, in the Southeast and the Atlantic coast and Texas and the Gulf Coast.

To investigate these trends further, we show time series from each of the two centers of action identified above (Fig. 12). The northwest Cascade time series of average first leaf and first bloom index values show pronounced multidecadal variability, with relatively early dates occurring in the 1920s, 1930s, as well as the 1980s and 1990s. In contrast, the Great Lakes trends toward earlier values start sometime in the late 1950s and continue through 2013. Spring onset dates continue to decrease steadily with time, with the first bloom index dates exhibiting trends that are slightly more negative than those of the first leaf index. These results suggest that the mechanisms for generating the dipole pattern seen in Fig. 11 (and also Fig. 8) may be different in the northwest Cascades and Great Lakes regions (e.g., they may not be driven by the same climate process).

The PDO’s influence on spring onset has been documented and discussed elsewhere (McCabe et al. 2011, 2013). It is worth mentioning here insomuch as it helps explain the trends in the northwest Cascades, a region with spring temperature variations strongly correlated to the PDO, especially during the spring (Fig. 13). Correlations with the PDO are negative (see also Table 2), indicating the positive phase of the PDO favors early spring, while the negative phase favors later spring (e.g., McCabe et al. 2011, 2013; Schwartz et al. 2013). As the PDO was in a negative phase during the last part of the record analyzed, the positive (later spring) trends in the Northwest appear to be following suit.

Variability in the Pacific Ocean does not appear as strongly linked to the Great Lakes region (Fig. 13), nor do the first leaf and first bloom time series exhibit
multidecadal variations like those in the Northwest time series (Fig. 12). Therefore, the negative (earlier spring) trends in the Great Lakes region cannot be readily explained by multidecadal variations in the PDO (nor can they be explained by the Atlantic multidecadal oscillation; not shown). Instead, given the steady, post-1950s trend evident in Fig. 12, we interpret changes in this region to reflect the secular long-term warming trend. Hence, we suggest that the apparent dipole in North America is not driven by a single mode of climate variability but instead the combined effects of two different processes. In the west, the PDO appears to suppress, or even counteract, the role of the warming trend. In the Great Lakes and eastern United States, the secular trend toward earlier spring appears more steady after 1950, albeit with prominent interannual variations superimposed. The trends for the Great Lakes region (2–3 days decade\(^{-1}\)) are quite comparable to Schwartz et al. (2006), whereas the trends in the Northwest appear to have reversed since Cayan et al. (2001) published their results from analysis of observed lilac and honeysuckle phenology. However, these are not apples-to-apples comparisons because the latter study only focused on observational data from the western United States and covered a shorter period that ended in the 1990s. There are also elevation effects that make reconciling the gridded product with the station-based data difficult, particularly in the West. Nonetheless, the apparent reversal in the sign of spring onset trends from Cayan et al. (2001) to the present study does suggest that natural variability, as opposed to global warming per se, plays an important role in pacing spring onset in that region. We note also, however, that at the time of this writing (June 2015), the West has experienced an extremely early spring this year. Hence, it cannot be assumed that trends will continue to be toward later spring in the Northwest for any appreciable time into the future simply on the basis that they were positive through 2013.

5. Conclusions and future work

We have described continental and regional means, variability, and trends in spring onset for the conterminous United States using extended spring indices (SI-x) calculated from a dense network of historical station data and a spatially complete (gridded) daily \(T_{\text{min}}/T_{\text{max}}\) product (Mueller et al. 2013). We found that both the station data and the gridded product provide a consistent view of the means, variances, and trends in spring onset across the continent. However, there are some key differences between the two that could have important ramifications for certain areas of inquiry. For example, the gridded product is generally biased toward later values in the mountainous western regions. It is also necessarily smoother (in
space) than the underlying station data, and consequently the interannual variability (as measured by detrended standard deviations) is lower in the gridded SI-x than the station-based values. Nonetheless, the spatial patterns of standard deviations, as well as trends and means, are all very similar between the two types of underlying data.

Despite the considerable warming experienced globally and across the coterminous United States during recent decades (e.g., Melillo et al. 2014), spring onset is not advancing toward uniformly earlier values as might be expected. Instead, interannual-to-decadal variations appear to pace regional trends and the overall spatial patterns. Specifically, the coterminous United States is characterized by a dipole of spring onset trends with “centers of action” in the northwest Cascades and Great Lakes regions from 1979 to 2013, with later dates seen in the former and earlier ones in the latter. Over the longer 1920–2013 period, the trends are more similar to each other throughout the domain, but they still exhibit some geographic differences. We interpret these results to reflect primarily the PDO’s influence on the climate of the Northwest, which drives trends toward later dates most prominently since the late 1970s. The rest of the domain is experiencing earlier dates but not in connection to any one mode or combination of modes. Hence, the differences between the Northwest and Great Lakes are likely to be more reflective of the suppression of early spring onset in the Northwest by the PDO, and the influence of climate change elsewhere. These results apply to both the first leaf and first bloom indices and throughout much of the eastern United States.

Other regional trends are between \(-0.6 \text{ and } -1.7\) days decade\(^{-1}\) for the first leaf index and \(-0.2 \text{ and } -1.4\) days decade\(^{-1}\) on average from 1955 to 2013. These estimates are quite consistent (same sign and within \(\pm 50\%\) of the trend) with those published earlier for continental- and global-scale metrics, and phenologies, of “shrubs” that respond to early spring (Cayan et al. 2001; Schwartz et al. 2006; Parmesan 2006; Schwartz et al. 2013). Interestingly, the trends of the first leaf index, which responds to early spring warmth, outpace those of the first bloom index in most areas. This finding could have certain consequences for agriculture and ecosystem dynamics, if the period of spring warming is becoming longer as a consequence of climate change.

Our findings also highlight some of the unique aspects of working with data for which dependent values are days of the year instead of physical quantities. As such, interpreting the interannual variability of SI-x requires standard methods in the atmospheric and climate sciences to be used somewhat differently than is typical. Correlations in Fig. 10, for example, are between spring index time series and seasonally adjusted geopotential height anomalies, which we employ because the time window of variability of the atmosphere relevant to spring onset changes from one year to the next. Other studies using similar kinds of metrics (e.g., “center of timing” or “center of mass” in hydrology) could apply similar seasonal adjustments to identify locally important climate drivers of variability and change. These considerations are especially important in light of efforts to develop and understand climate change “indicators” in support of future national assessments (e.g., Kenney and Janetos 2014).
Finally, there is some potential opportunity for exploring the predictability of spring onset a month or so in advance (Tables 2 and 3), as was also suggested by Cayan et al. (2001). Significant correlations between spring indices and large-scale modes of variability a month in advance occur in most regions. Hence, as suggested in McCabe et al. (2013) and Ault et al. (2011), early spring in the Northwest should be more likely during warm phases of ENSO, the PDO, or the positive phase of the PNA. Likewise, nearly all regions appear influenced by the NAO, which likely acts to delay spring by increasing the number of cold outbreaks at the end of winter (Thompson and Wallace 2001).

In addition to exploring the predictability of spring onset on subseasonal and longer time scales, future work could also investigate the atmospheric dynamics and thermodynamics linked to early and late spring in each of the regions identified here. Likewise, the SI-x could be readily calculated from global climate models to characterize and anticipate future changes in climate. Similar analysis could also be conducted at the hemispheric or global scale. By publishing the first, to our knowledge, gridded spring indices dataset, this work opens the door to new systematic comparisons of spring onset timing in observations and climate model simulations with data represented at comparable scales.

Acknowledgments. This project was facilitated in large part by the USA National Phenology Network (USA-NPN; https://www.usanpn.org/). We are especially grateful to Alyssa Rosemartin and the National Coordinating Office of the USA-NPN for providing logistical support and feedback. The SI models were developed using phenological data that are now available from the National Phenology database at the USA National Phenology Network. The project described in this publication was supported by Grant/Cooperative Agreement G13AC00248 from the U.S. Geological Survey.

REFERENCES


