

## Development and analysis of spring plant phenology products: 36 years of 1-km grids over the conterminous US

Emma Izquierdo-Verdiguier<sup>a,1,\*</sup>, Raúl Zurita-Milla<sup>a,1</sup>, Toby R. Ault<sup>b</sup>, Mark D. Schwartz<sup>c</sup>

<sup>a</sup> Faculty of Geo-information Science & Earth Observation (ITC), University of Twente, Enschede, The Netherlands

<sup>b</sup> Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, USA

<sup>c</sup> Department of Geography, University of Wisconsin-Milwaukee, Milwaukee, USA

### ARTICLE INFO

#### Keywords:

Climate change  
Extended Spring Indices  
Google Earth engine  
Cloud computing  
Time series analysis

### ABSTRACT

Time series of phenological products provide information on the timings of recurrent biological events and on their temporal trends. This information is key to studying the impacts of climate change on our planet as well as for managing natural resources and agricultural production. Here we develop and analyze new long term phenological products: 1 km grids of the Extended Spring Indices (SI-x) over the conterminous United States from 1980 to 2015. These new products (based on Daymet daily temperature grids and created by using cloud computing) allow the analysis of two primary variables (first leaf and first bloom) and two derivative products (Damage Index and Last Freeze Day) at a much finer spatial resolution than previous gridded or interpolated products. Furthermore, our products provide enough temporal depth to reliably analyze trends and changes in the timing of spring arrival at continental scales. Validation results confirm that our products largely agree with lilac and honeysuckle leaf and flowering onset observations. The spatial analysis shows a significantly delayed spring onset in the northern US whereas in the western and the Great Lakes region, spring onset advances. The mean temporal variabilities of the indices were analyzed for the nine major climatic regions of the US and results showed a clear division into three main groups: early, average and late spring onset. Finally, the region belonging to each group was mapped. These examples show the potential of our four phenological products to improve understanding of the responses of ecosystems to a changing climate.

### 1. Introduction

Changes in climate are evident in observational weather and ecological records (Kerr and Ostrovsky, 2003). According to the Intergovernmental Panel on Climate Change (IPCC), there is strong evidence that human activities are behind most of these changes. A tangible impact of these modifications is the increasing frequency of temperature extremes. The spatial and temporal variability of temperature has a direct impact on the timing of recurrent biological events of plants and animals (bird migrations (Cohen et al., 2018), early appearance or early flowering of the plants (Thomson, 2010), for example). This, in turn, has a direct impact on land surface-atmosphere interactions and associated biogeochemical cycles. Therefore, it is essential to understand how terrestrial ecosystems are responding to climate change (Richardson et al., 2013). In recent decades, climate change research has increasingly involved remote sensing technologies (Rosenqvist et al., 2003) using satellite images to derive land surface phenology and

in situ measurements provided by weather stations. The former overcome interpolation problems but the latter supply measurements that are easier to relate to ground processes and observations (Mendelsohn et al., 2007). Satellite, airborne and meteorological sensors provide observations of the Earth's surface at global, regional and, local scales. All these measures are used to derive products to study the impact of climate change on our planet (Broich et al., 2015; Villoria et al., 2016).

Consistent climate change indicators are needed to better understand the different causes and impacts of climate change on our ecosystems. Phenological observations constitute one of the most sensitive indicators of climate change (Parry et al., 2007) because they contain information about the timing of recurrent biological events that are strongly linked to the local weather and climate of the area. Various phenological indicators have been derived using phenological models (Tucker, 1979; Gilbert et al., 2014). The simplest phenological models are Thermal Time models (Linkosalo et al., 2008). Among several meteorologically based measures of thermal time, Growing Degree Day

\* Corresponding author.

E-mail addresses: [e.izquierdo-verdiguier@utwente.nl](mailto:e.izquierdo-verdiguier@utwente.nl) (E. Izquierdo-Verdiguier), [r.zurita-milla@utwente.nl](mailto:r.zurita-milla@utwente.nl) (R. Zurita-Milla), [toby.ault@cornell.edu](mailto:toby.ault@cornell.edu) (T.R. Ault), [mds@uwm.edu](mailto:mds@uwm.edu) (M.D. Schwartz).

<sup>1</sup> Both authors have contributed equally to the work development and the manuscript.

(GDD) is suitable for modelling plant growth (Shaykewich, 1995). GDD is the basis of the Extended Spring Indices models (SI-x) (Schwartz, 1985; Ault et al., 2015b). The SI-x models are used to generate a *Start of Spring* indicator<sup>2</sup> which is included in the US Global Change Research Program. The *Start of Spring* indicator uses the accumulation of heat to predict the day of the year on which temperature-sensitive plants leaf out and start blooming. *Start of Spring* provides a direct connection between vegetation phenology effects and global warming.

Different studies have used the SI-x models to analyze variations in spring onset linked to climate change (Schwartz et al., 2013; Allstadt et al., 2015). Most of these studies are based on plant and weather observations stations at specific locations (Ault et al., 2013, 2015b) but current work broadens the analyses to gridded SI-x products. Originally available at relatively coarse spatial resolutions ( $1^{\circ}$  (Ault et al., 2015a) and 25 km (Wu et al., 2016)) and more recently at resolutions of about 15 (Allstadt et al., 2015) and 4 km (Crimmins et al., 2017), from which a resampled 2 km product is also derived.<sup>3</sup> High spatial resolution SI-x products have special relevance in highly variable territories such as North America, which present multiple and complex topographies. Thus, high spatial resolution phenological products could be used to obtain more realistic views of local phenology and to support regional ecological studies. Additionally, having long time series of high spatial resolution products helps to avoid drawing misleading conclusions based on short-term conditions and/or trends (Cohen et al., 2018).

Until now, technology has limited high spatial resolution phenological modeling to small areas due to the huge quantity of data that had to be processed. However, the advancement of Information and Communications Technologies (ICT) allows not only the visualization and analysis of climate data (Zhang et al., 2016; Arundel et al., 2016; Bradley et al., 2010) but also the development of new SI-x products using cloud computing (Broich et al., 2015). The increasing accessibility and lower costs of cloud computing have made it possible to study geographic phenomena at high spatial resolution, over long periods of time, and at continental to global scales. One example of an easily accessible and free cloud computing application is Google Earth Engine. This application is based on the well-known map-reduce paradigm introduced by Google,<sup>4</sup> which considerably speed up data processing and helps to scale up the required computations (Gorelick et al., 2017).

In this paper we present new spring onset gridded products based on the SI-x models. Our products, available at 1 km and for the period 1980–2015, consist of four variables (two primary, one observational and one derived; c.f. Section 2). The primary products are verified, validated and analyzed to evaluate their quality and to check their consistency with previous SI-x products and studies. Our analysis focuses on studying spatio-temporal patterns of spring onset, mapping trends and on the use of the SI-x to regionalize the conterminous US.

## 2. The Extended Spring Indices

The Extended Spring Indices (SI-x) are a suite of models developed by Schwartz et al. (2013) by removing the chilling requirements from the original of spring indices models (Schwartz, 1997). This allows the SI-x to have a wider geographic applicability and to model spring onset for the complete conterminous US (CONUS). The SI-x models are primary used to predict “Leaf” (LF) and “Bloom” (BL) indices for three indicator plant species (*Lilac* (*Syringa chinensis* “*Red Rothmagensis*”) and *Honeysuckle* (*Lonicera tatarica* “*Arnold Red*” and *Lonicera korolkowii* “*Zabeli*”)).

The SI-x models are based on Growing Degree Hours (GDH), which are calculated from daily minimum and maximum temperatures. These GDH are used to define accumulation of short- and long-term variables.

These variable are used in regression based models that predict LF and BL for each plant species. The regression coefficients were calculated by Schwartz et al. (2013):

- The LF index is the average of the first day of the year that fulfills:

$$\begin{aligned} \text{DDE2}^*0.201 + \text{DD57}^*0.153 + \text{SYNOP}^*3.306 + \text{MDS0} \\ *13.878 > = 1000(\text{Lilac}) \end{aligned}$$

$$\begin{aligned} \text{DD57}^*0.248 + \text{SYNOP}^*4.266 + \text{MDS0} \\ *20.899 > = 1000(\text{Arnold Red Honeysuckle}) \end{aligned}$$

$$\begin{aligned} \text{DDE2}^*0.266 + \text{SYNOP}^*2.802 + \text{MDS0} \\ *21.433 > = 1000(\text{Zabeli Honeysuckle}), \end{aligned} \quad (1)$$

where *DDE2* is the accumulated GDH from day  $t$  until day  $t + 2$ , *DD57* is the accumulated GDH from day  $t + 5$  until day  $t + 7$  with  $t$  being a temporal index from January 1st, *ASYNOP* is accumulative of the synop variable which is 1 when *DDE2*  $> 637$  and otherwise is 0 and *MDS0* is a counter that starts on January 1st.

- Likewise, the BL index is obtained by averaging the first day of the year that fulfills:

$$\begin{aligned} \text{ACGDH}^*0.116 - \text{MDS0}^*23.934 > = 1000(\text{Lilac}) \\ \text{ACGDH}^*0.127 - \text{MDS0}^*24.825 > = 1000(\text{Arnold Red Honeysuckle}) \\ \text{ACGDH}^*0.096 - \text{MDS0}^*11.368 > = 1000(\text{Zabeli Honeysuckle}). \end{aligned} \quad (2)$$

In this case, *ACGDH* is the accumulation of GDH from LF index and, *MDS0* is still a counter but it starts on the LF index date.

The SI-x models are not limited to predicting just these primary indices. The SI-x also output two derivative products: Last Freeze (LSF) and Damage Index (DI). The LSF is the last day of the year whose minimum temperature is lower than or equal to  $28^{\circ}\text{F}$  ( $\sim -2.22^{\circ}\text{C}$ ). The DI links the LF index and LSF to measure the risk of frost damage. That risk is quantified by the difference between the anomalies of LF index and the LSF. Thus, very negative DI values indicate a high probability of frost damage. For more information on the SI-x models see Ault et al. (2015b), Schwartz et al. (2006).

## 3. Scaling up the Spring Index models

The Google Earth Engine (GEE) platform was chosen to scale up the calculation of the SI-x models. GEE<sup>5</sup> is a free and cloud-based application that specializes in geospatial processing. However, GEE cannot run the original Fortran and Matlab codes (Ault et al., 2015b) so these had to be restructured to exploit the parallel processing environment of the cloud. In this section, we first describe the data used in this work, then the restructure of the model to cloud computing, and finally, the process of verifying, validating and analyzing.

### 3.1. Data

The data belongs to two main groups: the data used to obtain the SI-x products at 1 km (environmental data) and the data applied in the evaluation and analysis of the products (ancillary data).

- Environmental data:

Daily surface weather data (Daymet) version 2 is available in GEE. Daymet is a continuous surface dataset available at a spatial resolution of 1 km for the CONUS (Thornton et al., 2014). Daymet data is available between January 1st, 1980 and January 1st, 2016 and it covers the following spatial range: latitudes between 10 and 53° and longitudes between  $-133.5$  and  $-49.9^{\circ}$ . This means that

<sup>2</sup> <http://www.globalchange.gov/explore/indicators>.

<sup>3</sup> [https://www.usanpn.org/data/spring\\_indices](https://www.usanpn.org/data/spring_indices).

<sup>4</sup> <https://developers.google.com/earth-engine/>.

<sup>5</sup> <https://developers.google.com/earth-engine/>.

Daymet provides data for a long period of time (36 years) and over a wide geographical extend (2700 by 6700 pixels). This allows characterizing spring onset over the CONUS as well as analyzing phenological changes. Daymet includes 7 variables although the SI-x models only requires: duration of the daylight period (s), daily maximum 2-m air temperature (°C) and daily minimum 2-m air temperature (°C). All Daymet variables were obtained by interpolating daily weather values. These values were derived from meteorological station data, taking into account elevation, the desired spatial resolution and a land value mask. For more information of Daymet see Thornton et al. (2000).

- Ancillary data: Three databases are used as ancillary data to confirm the high quality of the obtained products and their analysis.
  - Volunteered phenological observations: The database is provided by U.S. Geological Survey (USGS)<sup>6</sup> for 134 ground phenological observation (GPO) sites spread around the CONUS over the period 1980–2015 (Fig. 1). This database contains the latitude and longitude where the First Leaf and Bloom were observed, the DOY and the event type (e.g. Leaf or Bloom). A total of 2550 GPO were collected over the study period. The number of LF and BL observations depends on the year and the location because few observation sites contain data for the complete study period. The locations of the GPO were employed to verify our implementation of the SI-x models whereas both location and observed DOY were used to validate the SI-x products.
  - Global Historical Climatology Network Daily (GHCN-D): This dataset<sup>7</sup> consists of daily weather summaries from land surface stations across the CONUS. It consists of the latitude and longitude of the weather stations and the daily maximum and minimum temperature for the temporal range of this study. This database was used to both verify and validate the GEE SI-x version models.
  - Climatic regions: The National Centers for Environmental Information have identified nine climatically consistent regions within the CONUS (Karl and Koss, 1984). The climate regions were used to analyze the temporal variability of the SI-x products.

### 3.2. The Extended Spring Index models on the cloud

The parallel-processing nature of GEE motivated rewriting and adapting the calculation of SI-x to obtain products at high spatial resolution. The implementation of the SI-x models was adapted to utilize the *function mapping* capabilities of GEE, replacing the use of traditional “**for**” loops.

The computation of SI-x is divided into several steps (Fig. 2). First the Daymet collection is used to compute *GDH* for each of the first 250 days of each year. Next, the predictor variables, *DDE2*, *DD57* and *MDS0* and the *SYNOP* threshold are computed for each day. A cumulative sum is then applied to this daily collection to produce the *ASYNOP* variable and the three inequalities of Eq. (1) are evaluated each day. In each daily image, locations where Eq. (1) is true are assigned the value of the day of year (DOY) and the final LF index image is obtained by reducing this collection of images to find the minimum DOY at each location for each year.

The bloom step also starts by calculating *MDS0* and, *GDH* values to obtain accumulated *GDH* (*AGDH*). Unlike the LF index process, both measures (*AGDH* and *MDS0*) begin counting from the date of LF index instead of January 1st. Both predictors are linearly combined to generate a time series of images whose values are the DOY when the conditions of Eq. (2) are met.

Finally, the derivative products are calculated. The LSF is obtained from the minimum daily temperature in the Daymet dataset and the difference between LF and LSF produced the Damage Index. JavaScript



Fig. 1. Ground phenological observation places in the CONUS from 1980 to 2015.

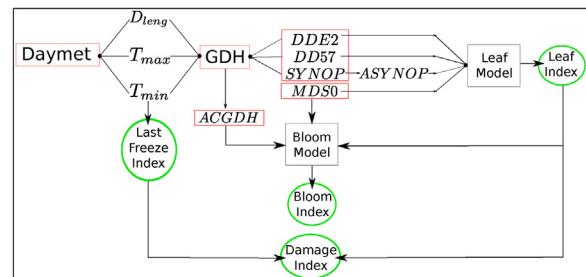


Fig. 2. Scheme of SI<sub>x</sub> on GEE for processing one year.  $D_{length}$ ,  $T_{max}$ ,  $T_{min}$  are the day length and daily temperature bands through which the Growing Degree Hours (GDH) are obtained. The accumulation of temperature between two previous days and the day (DDE2), between seven previous days and five previous days (DD57) and accumulative *SYNOP* event (ASYNOP) are used together with *MDS0* (day of the year less one) to predict the Leaf index, using a linear regression model. Bloom index is obtained by linear regression of accumulative *GDH*, *MDS0* and Leaf index. DI is calculated from Leaf index and Last Freeze index. The latter uses  $T_{min}$  for its calculation.

code is available for the interested reader: LF index<sup>8</sup>, BL index<sup>9</sup> and to obtain the LSF.<sup>10</sup>

### 3.3. Performance evaluation and analysis of the SI-x models on GEE

The models' performance was evaluated by verification (comparison between local and cloud model versions at GPO positions) and validation (comparing the predictions and GPOs). The SI-x analysis focuses on the spatial patterns and trends.

#### 3.3.1. Performance evaluation

The GEE-based model products were compared against a legacy implementation (Matlab MATLAB, 2010a toolbox) provided in Ault et al. (2015b). The comparison of both outputs was done at GPO sites (Section 3.1). The verification focused on the Pearson correlation and root mean square error between the outputs obtained by GEE-API and Matlab toolbox.

The validation used the DOY observed by the volunteers, which are regarded as ground truth. We compared the SI-x products using Daymet and GHCN-D (Section 3.1) with the GPOs. The comparison was accomplished as follows:

1. For each GPO (lat, long and observation year), we extracted the Daymet temperatures of the corresponding grid cell and the GHCN-D temperatures of the nearest weather station in distance and in elevation difference.
2. Using the temperatures, we calculated the LF and BL indices using GEE-API for Daymet dataset and Matlab toolbox for GHCN-D stations.

<sup>8</sup> <https://code.earthengine.google.com/532a5e99f776eba9b106d070b5a885ca>.

<sup>9</sup> <https://code.earthengine.google.com/640aeab2b2b9040e162e578400f21209>.

<sup>10</sup> <https://code.earthengine.google.com/8e7ef4d4e9c0b020b93cd27ee35af384>.

### 3. The LF and BL index products were evaluated against the observed DOY (GPO).

Finally, we analyzed the differences in the output when using different input data (GHCN-D vs Daymet). The LF and BL products generated through GEE were compared against the ones obtained when running the same models with GHCN-D temperature data. This comparison was done to study the precision of the models [through the root-mean-square error (RMSE) and the mean absolute error (MAE)], the bias [through the mean error (ME)] and the goodness of fit [through the Pearson's correlation coefficient ( $R$ )] of the models. Furthermore, we used scatter plots to visualize possible differences between Daymet- and GHCN-D-based model outputs.

#### 3.3.2. Spatial-temporal analysis

The spatial patterns were analyzed by the temporal average of all years (1980–2015), its standard deviation during that period and the trends of the all SI-x products. The temporal patterns were estimated by fitting a linear regression (least squares approach) on the full time series. The statistical significance ( $p$ -value) of these trend was analyzed and mapped to show areas with clear phenological changes. Lastly, the temporal averages of LF and BL indices were analyzed for the nine climate regions (Section 3.1) within the CONUS.

## 4. Results and discussion

### 4.1. Performance evaluation

The Pearson's correlation coefficient between the LF index obtained when using the GEE and MATLAB implementations is equal to 1, and the RMSE is equal to 0. The same results were obtained for the BL index. This verifies our SI-x cloud-based implementation and confirms that we now have a way to generate phenological grids at high spatial resolution for long periods of time. These results represent a step forward in phenological modelling because previous studies either delivered data using coarse spatial resolution grids or at individual points (typically the locations of weather stations). Furthermore, our GEE-based implementation is very fast; the complete temporal series was processed in around 12 h.

Errors between predictions (using Daymet and GHCN-D) and the GPOs are shown in Table 1. Results show that leaf RMSE errors for GHCN-D and Daymet are similar while bloom errors are almost two days smaller when using the Daymet database to calculate SI-x products. Furthermore, the bloom errors are higher than leaf errors for both comparisons. This was as expected because in our implementation we coupled the BL and FL indices (i.e. BL depends on FL index predictions). Comparison of the indices produced by GHCN-D and Daymet shows that both databases lead to highly correlated outputs ( $R$  values above 0.95) but also show that there are differences in the actual values (RMSE values of around 7 days).

**Table 1**

Estimated results for SI-x model using Daymet dataset (GEE-API) compared with voluntary observations, GHCN-D stations using Matlab compared with voluntary observations and Daymet compared with GHCN-D stations results. Notation: ME: mean error, RMSE: root mean square error, MAE: mean absolute error,  $R$ : Pearson coefficient.

Type pred.	GHCN-D vs Obs.		Daymet vs Obs.		Daymet vs GHCN-D	
	Leaf	Bloom	Leaf	Bloom	Leaf	Bloom
# samples	1309	1241	1309	1241	1309	1241
ME	2.645	11.493	0.337	9.575	2.310	1.920
RMSE	12.129	14.091	12.309	12.737	7.521	6.129
MAE	8.599	11.877	8.403	10.547	3.650	3.395
$R$	0.884	0.929	0.874	0.925	0.955	0.965

We represented as scatter-plots the LF and BL index predictions to analyze their differences (Fig. 3a and b). As the figures show, LF prediction is balanced between over/under predictions for both databases while for bloom, Daymet produces more balanced predictions than the GHCN-D dataset, which over predicts BL index. That is confirmed by the ME values (Table 1). Fig. 3c and d shows the cross plots between both predictions. The predictions made with GHCN-D and Daymet are highly correlated notwithstanding that there are GPO sites with great differences (some exceed more than 3 standard deviations) which are analyzed in Fig. 4. The differences between databases predictions are higher in LF than BL predictions. That is due to slight variation in the inputs of the model causing greater variations in LF than BL predictions. The slope of the LF linear combination (Eq. (1)) is more gentle than the BL linear combination (Eq. (2)) and that implies, small variations in the LF linear combination produce larger day prediction increments. The steeper slope of the BL is due to high dependency on GDH accumulation ( $AGDH$ ) requiring further increase of the constant to change the prediction day. These results show the sensitivity of SI-x products with respect to the models' input variables.

The spatial distribution of GPO sites with high differences between Daymet and GHCN-D predictions are shown in Fig. 4a for LF and in Fig. 4b for BL. We have split the differences into three cases: between 1 and 2 standard deviation (gray), between 2 and 3 standard deviations (green), and higher than 3 standard deviations (red). In both figures, the size of dots represents the number of times (predictions in different years) when the difference between the predictions is higher than the corresponding threshold. Note that most of the GPO sites have small differences, only one site at north center for LF and four GPO sites at center for BL are large. Thus, large differences are spread out around the conterminous US. These results indicate which GPO sites might require additional sampling to improve the SI-x model output. Focusing on the temporal dimension, the main differences between GHCN-D and Daymet predictions occur in the first decade of the temporal series, although a few outliers are found later (Fig. 4c). The quantitative results between both datasets are presented in Table 1.

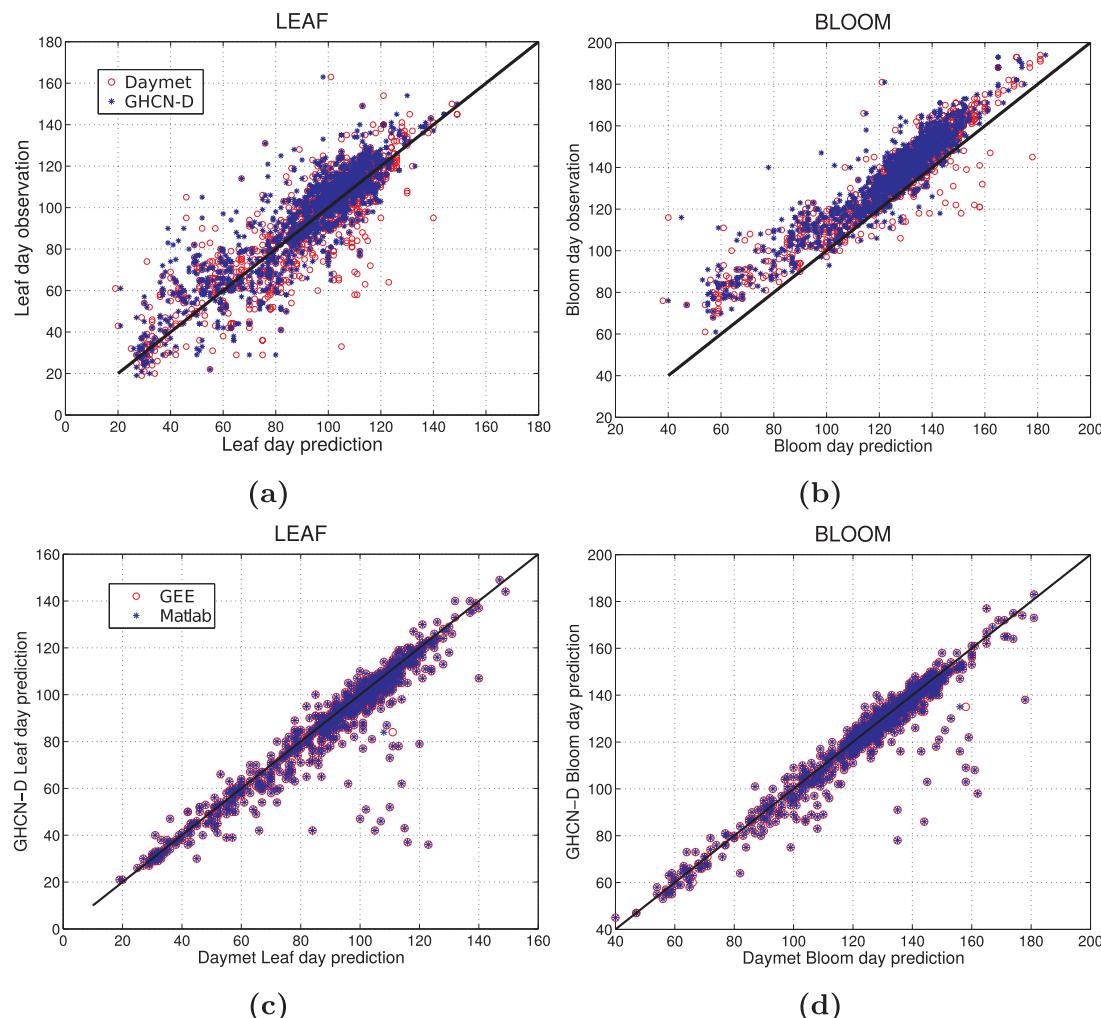
### 4.2. Analysis of SI-x model

#### 4.2.1. Spatial-temporal analysis of SI-x models

The spatial and temporal average of the LF and BL indices indicate that spring onset takes place on days 75 (March 16) and 108 (April 18). Nevertheless, as Fig. 5a and d shows, these indices exhibit a clearly noticeable gradient, with low values in the south and high values in the north. BL occurs 40 days after LF in most parts of the country, except in the west where the difference increases to almost 100 days.

The standard deviations of the LF and LB indices are generally low and spatially homogeneous, although there are some notable points in the West Coast and Pacific regions. The LF index exhibits less variations along the Southern coast of the country than observed in the BL index. The trends (Fig. 5c and f) show a slight tendency towards delayed LF and BL index from northwest to north central regions, whereas in the western region, the trend is reversed. Statistical significance (range of  $p$ -values) is high in the western region for both the LF (Fig. 6a) and BL (Fig. 6b) indices. However, the BL index also shows high values in the north central, north west and around the Great Lakes. The advancement of the BL in the latter region (i.e. Great Lakes) is consistent with results obtained in previous work using a one-degree spatial resolution grid (Ault et al., 2015b). However, our results present finer spatial details.

The LSF (Fig. 5g) has a similar latitudinal gradient to that of LF and BL. However, in this case the gradient is less horizontal because LSF is directly affected by minimum temperature values. The spatial and temporal average of the LSF is equal to 88 (March 29). Considering the average LF index (March 16), there is a clear possibility of plants being damaged by frost. The range of LSF varies between day 120 (April 30) and 0, for areas where it never freezes. The standard deviation (Fig. 5h) shows higher variability in the Pacific, Southwest and Southeast coasts



**Fig. 3.** Scatter plots of leaf (a) and bloom (b) index predictions using Daymet and GHCN-D versus the observations and, leaf (c) and bloom (d) index days predicted by GEE and Matlab versus GHCN-D predictions ( $n = 2250$ ). Statistical results of the fitting are presented in Table 1.

than in the North Central and Northeast and, it does not exceed more than 20 days. The trend analysis shows that LSF are generally getting earlier and earlier each year.

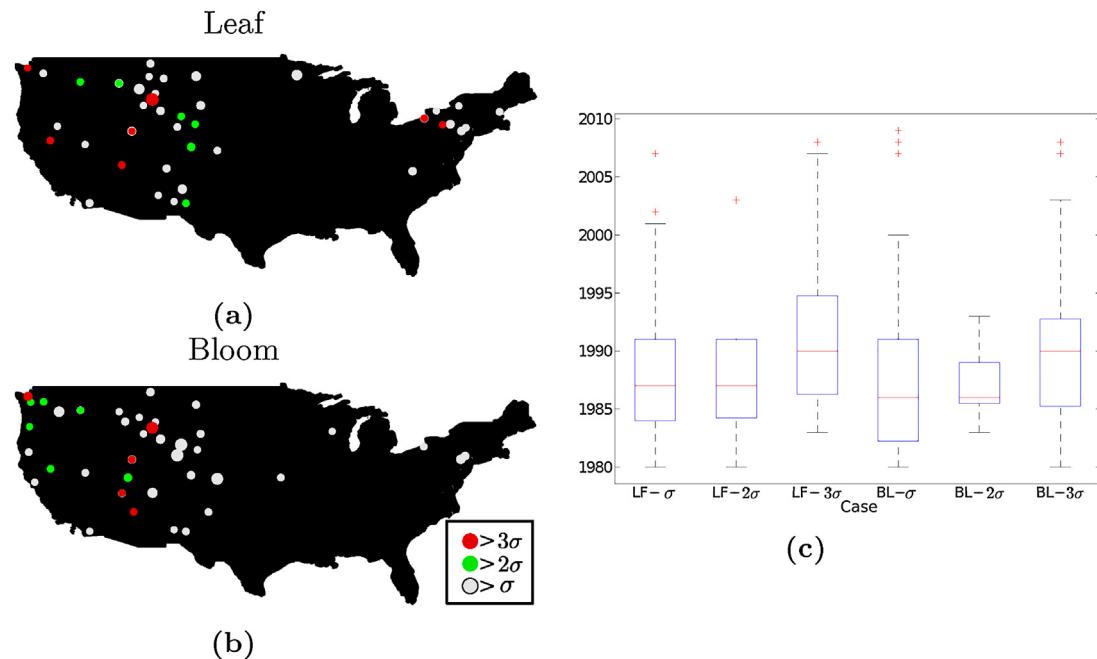
Fig. 5j and k also shows the average DI and its standard deviation. The average DI is mostly negative although there are a few small regions with positive values at the south. Negative DI values mean that LSF happened relatively later than the LF index; this may mean damage to sensitive plants, depending on the temporal differences between LSF and LF. The spatial patterns of DI standard deviation are similar to those of LSF. This shows the high incidence of the LSF over the DI in the study period. The analysis of the trend shows that DI is advancing in most parts of the country, with higher advancement rates in the west than elsewhere. This is largely due to earlier LSF. For both the LST and the DI, the statistical significance maps show low values where their trends have extreme values (Fig. 6c and d).

The temporal variability of the LF and BL indices was analyzed for the nine climatic regions defined by the National Centers (Fig. 7c). The average LF and BL indices for each region are shown in Fig. 7a and b, respectively. Both figures clearly show that the study area can be divided into three main spring onset groups: early (LF DOY 40 and BL DOY 70), average (LF DOY 80 and BL DOY 110) and late (LF DOY 100 and BL DOY 140). These averages also show that spring onset is advancing in recent years. This advancement is visible in all regions except the Northeast (climatic region 5 in Fig. 7c). The high temperatures of 2012 (Karl et al., 2012) lead to a very early spring onset in all regions except 1 and 9. In particular, the high temperatures of 2012 resulted in

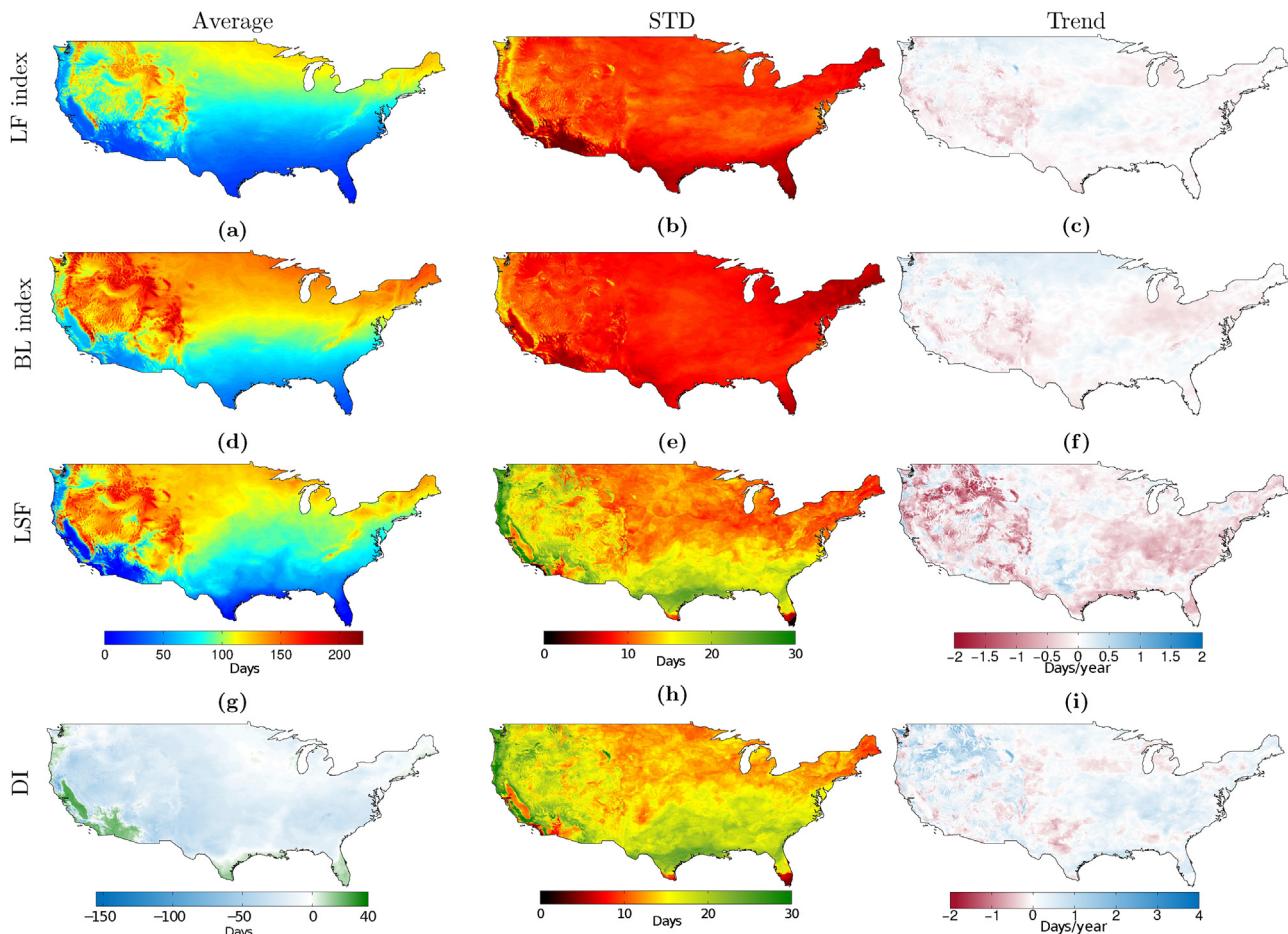
earlier spring onsets in regions 2–5 (North Central, Great Lakes and Northeast regions) but not in regions 1 and 9, which had LF and BL values similar to their long term averages. However, in 2015 (one of the warmest years on record (NOAA, 2015)) regions 1 and 9 (West region) show a considerable advancement of spring onset.

## 5. Conclusions

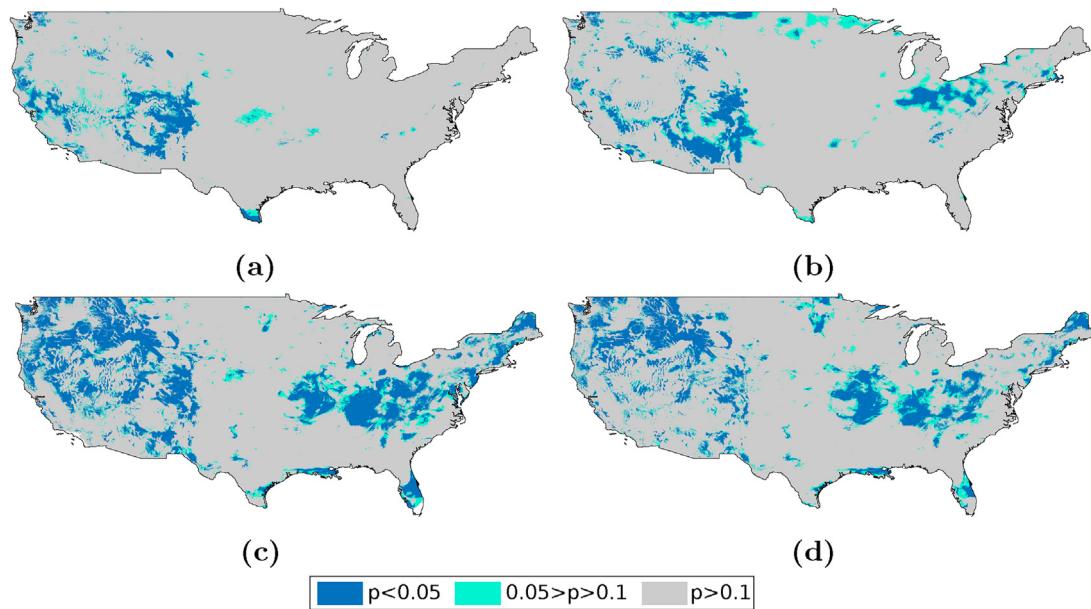
This paper describes the development of new long-term (1980–2015) and high spatial resolution (1 km) phenological products for the conterminous US. These products were created by adapting the Extended Spring Indices (SI-x) models to a cloud computing platform (Google Earth Engine). This adaptation allowed us to run the SI-x models with a gridded daily weather dataset (Daymet) to generate two primary outputs: the leaf (LF) and bloom (BL) indices. These indices were successfully verified and validated by comparing our cloud-based results with previous implementations of the SI-x, with results obtained at locations with historical weather stations (GHCN-D network), and with ground phenological observations. The temporal analysis of the LF and BL indices shows that spring onset has, in general, advanced. The rate of advancement depends on the geographic location, and it is modulated by the topography of the US. Relatively large advancements rates are evident in high altitude regions, through rates are lower in the eastern portion of the country. The adaptation of the SI-x models to a cloud environment also allowed us to derive two additional metrics: the day of last freeze (LSF) and the damage index (DI). Both metrics,



**Fig. 4.** Left: Spatial distribution of the sites where differences in leaf (a) and bloom (b) predictions obtained from Daymet and GDHC-D are larger than  $\sigma$  (in gray),  $2\sigma$  (in green) and,  $3\sigma$  (in red). Note that the size of points is proportional to the number of times the difference between predictions is higher than the threshold. The largest is equivalent to 14 and 12 times for leaf and bloom predictions, respectively. Box plot of years (c) for the three  $\sigma$  thresholds. (For interpretation of the references to color in this legend, the reader is referred to the web version of the article.)



**Fig. 5.** Averages, standard deviation and trends for the leaf index (LF; a–c), the bloom index (BL; d–f), last freeze (LSF; g–l) and the damage index (DI; j–l). All values are calculated for the complete study period (i.e. 1980–2015).

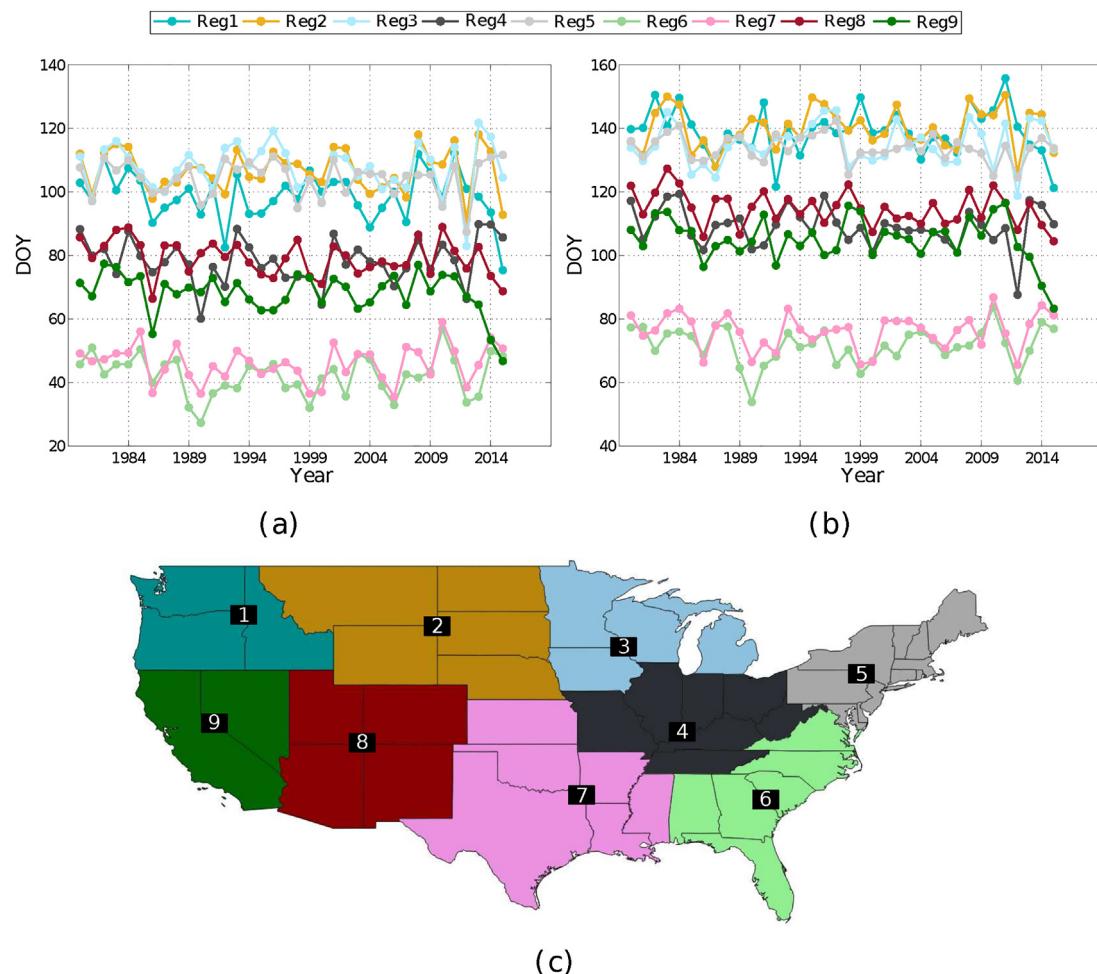


**Fig. 6.** Statistical significance (range of  $p$ -values) for the trends of the leaf index (a), bloom index (b), last freeze (c) and damage index (d) illustrated in Fig. 5.

presented here for the first time as gridded products, were used to study the likelihood of frost damage.

Our gridded products provide phenological information at more actionable scales and, as such, they have the potential to support

management and decision-making activities. For instance, phenological networks could use our products to both inform citizens on the arrival of spring at their neighborhoods or to direct the data collection efforts. Furthermore, the availability of phenological information at 1 km



**Fig. 7.** Mean leaf (a) and bloom (b) temporal variability from 1980 to 2015 of the nine climate regions (c) in the USA.

allows capturing finer surface heterogeneities. It also facilitates the comparison and integration of temperature-based phenological information with other phenological data sources such as Earth observation satellites or the PhenoCam network.<sup>11</sup> This could lead to an improved understanding of, for instance, crop seasonality at regional scales or to more robust ecological monitoring and management systems. Finally, having long time series of spring onset gridded products at high spatial resolution means that we can analyze phenological changes at multiple spatial scales. Such analyses will hopefully lead to new insights and to an improved understanding of the responses of ecosystems to a changing climate.

## Acknowledgements

This work has been partially supported by the “Green-wave” project (Google Faculty Award). We thank the broad Google Earth Engine community for their help and comments. Special thanks to Dr. David Thau and Noel Gorelick from Google Inc. for actively supporting this project.

## References

Allstadt, A.J., Vavrus, S.J., Heglund, P.J., Pidgeon, A.M., Thogmartin, W.E., Radeloff, V.C., 2015. Spring plant phenology and false springs in the conterminous US during the 21st century. *Environ. Res. Lett.* 10, 104008.

Arundel, J., Winter, S., Gui, G., Keatley, M., 2016. A web-based application for bee-keepers to visualise patterns of growth in floral resources using MODIS data. *Environ. Model. Softw.* 83, 116–125.

Ault, T.R., Henebry, G.M., de Beurs, K.M., Schwartz, M.D., Betancourt, J.L., Moore, D., 2013. The false spring of 2012, earliest in North American record. *EOS Trans. Am. Geophys. Union* 94, 181–182.

Ault, T.R., Schwartz, M.D., Zurita-Milla, R., Weltzin, J.F., Betancourt, J.L., 2015a. Trends and natural variability of spring onset in the conterminous United States as evaluated by a new gridded dataset of spring indices. *J. Clim.* 28, 8363–8378.

Ault, T.R., Zurita-Milla, R., Schwartz, M.D., 2015b. A Matlab® toolbox for calculating spring indices from daily meteorological data. *Comput. Geosci.* 83, 46–53.

Bradley, E., Roberts, D., Still, C., 2010. Design of an image analysis website for phenological and meteorological monitoring. *Environ. Model. Softw.* 25, 107–116.

Broich, M., Huete, A., Paget, M., Ma, X., Tulbure, M., Coupe, N.R., Evans, B., Beringer, J., Devadas, R., Davies, K., Held, A., 2015. A spatially explicit land surface phenology data product for science, monitoring and natural resources management applications. *Environ. Model. Softw.* 64, 191–204.

Cohen, J.M., Lajeunesse, M.J., Rohr, J.R., 2018. A global synthesis of animal phenological responses to climate change. *Nat. Clim. Change* 8, 224–228.

Crimmins, T.M., Marsh, R.L., Switzer, J.R., Crimmins, M.A., Gerst, K.L., Rosemartin, A.H., Weltzin, J.F., 2017. USA National Phenology Network Gridded Products Documentation, Technical Report. US Geological Survey.

Glibert, P.M., Icarus Allen, J., Artioli, Y., Beusen, A., Bouwman, L., Harle, J., Holmes, R., Holt, J., 2014. Vulnerability of coastal ecosystems to changes in harmful algal bloom distribution in response to climate change: projections based on model analysis. *Glob. Change Biol.* 20, 3845–3858.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27.

Karl, T.R., Koss, W.J., 1984. Regional and National Monthly, Seasonal, and Annual Temperature Weighted by Area, 1895–1983. Historical Climatology Series 4-3.

Karl, T.R., Gleason, B.E., Menne, M.J., McMahon, J.R., Heim, R.R., Brewer, M.J., Kunkel, K.E., Arndt, D.S., Privette, J.L., Bates, J.J., Groisman, P.Y., Easterling, D.R., 2012. U.S. temperature and drought: recent anomalies and trends. *EOS Trans. Am. Geophys. Union* 93, 473–474.

Kerr, J.T., Ostrovsky, M., 2003. From space to species: ecological applications for remote sensing. *Trends Ecol. Evol.* 18, 299–305.

Linkosalo, H.K., Lappalainen, T., Har, P., 2008. A comparison of phenological models of leaf bud burst and flowering of boreal trees using independent observations. *Tree Physiol.* 28, 1873–1882.

MATLAB, version 7.10.0 (R2010a). The MathWorks Inc., Natick, MA.

Mendelsohn, R., Kurukulasuriya, P., Basist, A., Kogan, F., Williams, C., 2007. Climate analysis with satellite versus weather station data. *Clim. Change* 81, 71–83.

NOAA National Centers for Environmental Information, 2016. State of the Climate: Global Analysis for Annual 2015. Published online January 2016, retrieved from <http://www.ncdc.noaa.gov/sotc/global/201513> (04.12.16).

Parry, M., Canziani, O., Palutikof, J., van der Linden, P.C.H. (Eds.), 2007. Chapter 1: Assessment of Observed Changes and Responses in Natural and Managed Systems. Working Group II: Impacts, Adaptation, and Vulnerability, Technical Report.

Richardson, A.D., Keenan, T.F., Migliavacca, M., Ryu, Y., Sonnentag, O., Toomey, M., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agric. For. Meteorol.* 169, 156–173.

Rosenqvist, A., Milne, A., Lucas, R., Imhoff, M., Dobson, C., 2003. A review of remote sensing technology in support of the Kyoto protocol. *Environ. Sci. Policy* 6, 441–455.

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

Schwartz, M.D., Ault, T.R., Betancourt, J.L., 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.

Schwartz, M.D., 1985. The Advance of Phenological Spring Across Eastern and Central North America (PhD dissertation). Department of Geography, University of Kansas, Lawrence.

Schwartz, M.D., 1997. Spring Index models: an approach to connecting satellite and surface phenology. *Phenol. Seas. Clim.* 1 23–28.

Shaykewich, C., 1995. An appraisal of cereal crop phenology modelling. *Can. J. Plant Sci.* 75, 329–341.

Thomson, J.D., 2010. Flowering phenology, fruiting success and progressive deterioration of pollination in an early-flowering geophyte. *Philos. Trans. R. Soc. B: Biol. Sci.* 365, 3187–3199.

Thornton, P.E., Hasenauer, H., White, M.A., 2000. Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: an application over complex terrain in Austria. *Agric. For. Meteorol.* 104, 255–271.

Thornton, P., Thornton, M., Mayer, B., Wilhelmi, N., Wei, Y., Devarakonda, R., Cook, R., 2014. Daymet: Daily Surface Weather Data on a 1-km Grid for North America. Version 2.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Rem. Sens. Environ.* 8, 127–150.

Villoria, N.B., Elliott, J., Müller, C., Shin, J., Zhao, L., Song, C., 2016. Rapid aggregation of global gridded crop model outputs to facilitate cross-disciplinary analysis of climate change impacts in agriculture. *Environ. Model. Softw.* 75, 193–201.

Wu, X., Zurita-Milla, R., Kraak, M.J., 2016. A novel analysis of spring phenological patterns over Europe based on co-clustering. *J. Geophys. Res. Biogeosci.* 121, 1434–1448.

Zhang, T., Li, J., Liu, Q., Huang, Q., 2016. A cloud-enabled remote visualization tool for time-varying climate data analytics. *Environ. Model. Softw.* 75, 513–518.

<sup>11</sup> <http://phenocam.sr.unh.edu/>.