A novel analysis of spring phenological patterns over Europe based on co-clustering

Xiaojing Wu1, Raul Zurita-Milla1, and Menno-Jan Kraak1

1Department of Geo-Information Processing, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands

Abstract The study of phenological patterns and their dynamics provides insights into the impacts of climate change on terrestrial ecosystems. Here we present a novel analytical workflow, based on co-clustering, that enables the concurrent study of spatio-temporal patterns in spring phenology. The workflow is illustrated with a long-term time series of first leaf dates (FLD) over Europe, northern Africa, and Turkey calculated using the extended spring index models and the European E-OBS daily maximum and minimum temperatures (1950 to 2011 with a spatial resolution of 0.25°). This FLD dataset was co-clustered using the Bregman block average co-clustering with l-divergence (BBAC_l), and the results were refined using k-means. These refined co-clusters were mapped to provide a first spatially-continuous delineation of phenoregions in Europe. Our results show that the study area exhibits four main spatial phenological patterns of spring onset. The temporal dynamics of these phenological patterns indicate that the first years of the study period tend to have late spring onsets and the recent years have early spring onsets. Our results also show that the study period exhibits 12 main temporal phenological patterns of spring onset. The spatial distributions of these temporal phenological patterns show that western Turkey tends to have the most variable spring onsets. Changes in the boundaries of other phenoregions can also be observed. These results indicate that this co-clustering based analytical workflow effectively enables the simultaneous study of both spatial patterns and their temporal dynamics and of temporal patterns and their spatial dynamics in spring phenology.

1. Introduction

The spatio-temporal inhomogeneity of climate change is well documented [Intergovernmental Panel on Climate Change, 2013]. For instance in Europe, the increase of surface air temperature in high latitudes is different than that in other places [European Environment Agency, 2012; Haylock et al., 2008] and the decadal increase of average temperature over the period of 2002–2011 is higher than over the period of 1850–1899 [Brohan et al., 2006; Hansen et al., 2010; Smith et al., 2008]. As a result, the impacts of climate change on terrestrial ecosystems are also inhomogeneous across space and time [Menzel et al., 2006]. In this respect, it is important to study dynamics in phenological patterns from both spatial and temporal dimensions as they provide insights into the inhomogeneous impacts of climate change on terrestrial ecosystems [Parmesan and Yohe, 2003; Walther et al., 2002].

Phenology is the science that deals with the study of life cycle phases in plants and animals driven by environmental factors [Lieth, 1974; Schwartz and Chen, 2002]. Several studies have demonstrated that plant phenology is one of the most reliable bioindicators of climate change [Gordo and Sanz, 2010; Menzel et al., 2006; Schwartz et al., 2006]. Phenological spring events (e.g., first leaf appearance) are reported to be more sensitive to climate change than events occurring in other seasons [Doi and Katano, 2008; Gordo and Sanz, 2010; Matsumoto et al., 2003]. Consequently, indices that measure the onset of spring are ideal biological indicators of climatic variability in space and time [Schwartz, 1998; Schwartz et al., 2006].

Several methods are available to monitor spring phenology. One of such methods relies on time series of remotely sensed satellite images to derive the so-called start of spring season (SOS). However, there is no universally accepted method to characterize SOS from satellite data and results depend on the method and/or sensor used [Schwartz et al., 2002; White et al., 2009]. Besides, satellite data is only available since the 1980s, and thus, it is insufficient for studying long-term phenological responses to climate change. Another method uses ground phenological observations of selected species and biological events (e.g., leafing or blooming). Long-term phenological records exist, but they exhibit a poor spatial coverage [Menzel et al., 2006;
Studer et al., 2007; White et al., 2005]. A third method relies on the use of phenological models. Often these models are calibrated using ground phenological observations and allow predicting spring onset in regions where no observations exist. In this study, we use the extended spring index models (section 2.2) [Schwartz et al., 2013] to consistently characterize spring onset over large areas and for long time periods.

One of the most popular methods to analyze patterns in spatio-temporal data is clustering, which identifies groups of similar data elements and enables analysts to consider them at a higher level of abstraction [Andrienko et al., 2009]. In environmental studies, clustering is particularly useful because regional analysis for a given time period (e.g., a season) is generally more informative than analyzing a collection of observations made at a few locations and timestamps [Zirlewagen and Von Wilpert, 2010]. Thus, several studies can be found on the analysis of phenological patterns using clustering [Ahas and Aasa, 2003; Ahas et al., 2007; Gu et al., 2010; Kumar et al., 2011; Mills et al., 2011; White et al., 2005; Zhang et al., 2012; Zurita-Milla et al., 2013].

Ahas and Aasa [2003] used clustering to group phenological events by years to identify early and late years in terms of seasonal rhythm. Ahas et al. [2007] used clustering analysis to group years in terms of human activities’ rhythm to characterize urban and rural dynamics. White et al. [2005] used an iterative k-means clustering to identify phenologically and climatically similar clusters in space, which they termed phenoregions. Based on this latter work, Kumar et al. [2011] and Mills et al. [2011] developed and implemented a parallel k-means clustering to identify phenoregions in conterminous United States (CONUS). Gu et al. [2010] applied a widely used clustering technique, ISODATA, to generate another map of phenoregions for CONUS. Zhang et al. [2012] combined principal component analysis (PCA) and k-means++ to generate a phenoregions map for a smaller geographical region (the Upper Colorado River Basin), and Zurita-Milla et al. [2013] used self-organizing maps to cluster time series of satellite data to identify the main phenological patterns in the Kruger National Park (South Africa).

All these phenological clustering studies can be categorized into two types (refer to Figure 1): (1) spatial clustering for spatial pattern analysis (Figure 1b) and (2) temporal clustering for temporal pattern analysis (Figure 1c). In spatial clustering, the locations are regarded as objects and each timestamp as an attribute. Clustering results are locations with similar values of phenological variable(s) along all timestamps (the thick rectangle in Figure 1b is an example of a spatial cluster). In the temporal clustering, the timestamps are regarded as objects and each location as an attribute. Clustering results are timestamps with similar values of phenological variable(s) along all locations (the thick rectangle in Figure 1c is an example of a temporal cluster). However, patterns identified by only using spatial clustering lack the ability to describe the time-varying behavior present in the data and vice versa [Deng et al., 2011]. For instance in Figure 1b, values in location-cluster1 are more similar in timestamp-1 and timestamp-2 rather than along all timestamps, while in Figure 1c, values in timestamp-cluster1 are more similar in location-1 and location-2 rather than along all locations. This deficiency demands a clustering method that allows the simultaneous discovery of spatial and temporal patterns. Co-clustering enables this type of analysis.
Co-clustering, unlike one-way clustering such as $k$-means, treats locations and timestamps equally [Han et al., 2012]. This is done by mapping locations to location-clusters and timestamps to timestamp-clusters at the same time (Figure 1d). Co-clustering identifies homogeneous spatio-temporal co-clusters (co-clusters for short), formed by intersecting each location- and time-stamp-cluster. The values of phenological variables in timestamps at locations that belong to the same co-cluster are similar. The thick rectangle in Figure 1d shows an example of a co-cluster.

Each co-cluster contains elements that are similar to each other. However, as pointed out by Wu et al. [2015], the values in different co-clusters might still be similar due to the need to arbitrarily predefine the number of location- and timestamp-clusters before applying the algorithm to the dataset. Also, co-clustering algorithms assign complete rows/columns to location- and year-clusters. Consequently, complex spatio-temporal patterns might not be fully captured. To deal with this, co-clustering could be run with a relatively large number of co-clusters that could be subsequently grouped to create axis-parallel irregular (i.e., nonrectangular) co-clusters.

Fully considering the above listed issues, this study proposes a novel analytical workflow based on co-clustering that enables the exhaustive analysis of complex spatial and temporal phenological patterns. Briefly, we first generate time series of date of first leaf using the extended spring index models. Then, we co-cluster the time series and refine the results using $k$-means, and finally, we map the results to reveal the main spring phenological patterns as well as their dynamics.

2. Materials and Methods

2.1. Materials

The European E-OBS dataset (v10.0; http://www.ecad.eu/download/ensembles/download.php [Haylock et al., 2008]), which is a product of the European Climate Assessment and Dataset (ECA&D) project, is used in this study. This gridded dataset runs from 1950 to 2013 and covers the whole Europe as well as northern Africa and Turkey. For this study, we downloaded geo-referenced daily maximum (TX) and minimum (TN) temperatures with a spatial resolution of 0.25° x 0.25°.

We examined the completeness of TX and TN as the extended spring index models (section 2.2) require daily temperature records and will not calculate the date of first leaf if there are more than 20 missing TX or TN values per month for the first 6 months of the year. Thus, these records were excluded. Also, grid cells for which the spring phenological index could not be calculated for more than 20 years were excluded from further analysis. The application of the criteria resulted into a new dataset that contains 28,225 grid cells (98.5% of all the cells covered by the original E-OBS dataset) covering the period of 1950 to 2011 (62 out of 64 years). The spatial extent of the filtered dataset is depicted in Figure 2 that, as an example, shows the spatial distribution of TX on an arbitrary day (21 March 2010).

2.2. Extended Spring Index Models

The so-called spring indices (SI) are a suite of regression-based models that can be used to characterize spring onset [Schwartz, 1997; Schwartz and Reiter, 2000; Schwartz et al., 2006]. These models predict the first leaf dates (FLD) and first bloom dates for three key species: lilac (Syringa chinesis ‘Red Rothomagensis’) and honeysuckles (Lonicera tatarica ‘Arnold Red’ and Lonicera korolkowii Zabeli). The model inputs are daily maximum and minimum temperatures for a site as well as its latitudinal information—used as a proxy for day

![Figure 2. The spatial extent of the filtered dataset used in this study, depicted by the spatial distribution of the maximum temperature for an arbitrary day (21 March 2010) as an example.](image)
length. The SI have been extensively validated in the Northern Hemisphere to generate reliable phenological timings for the above-mentioned species [Schwartz et al., 2006]. However, the original SI can only produce outputs for locations where the chilling and warmth requirements are satisfied [Schwartz et al., 2013]. This is a typical characteristic of sequential phenological models [Chuine, 2000] where the model should first accumulate chilling degrees and then start accumulating temperature to predict leafing, blooming, or other phenological phases. To be able to cover large areas, here we use the recently developed extended spring indices (SI-x) where chilling is not required anymore [Schwartz et al., 2013]. The SI-x models have been validated in continental United States with extensive ground phenological observations, and they were found as effective as the original SI models [Schwartz et al., 2013]. Thus, the SI-x must perform well in Europe too, given the fact that the SI are already validated in this continent [Schwartz et al., 2006]. Besides, the three SI species are also widely present in Europe and monitored by Europe-wide programs, such as the European Phytology Network [van Vliet et al., 2003] and the International Phenological Gardens in Europe [Menzel et al., 2006].

In this study, we focus on the FLD model output as it is the earliest spring bioindicator. This index from the SI-x models refers to the average of the first leaf dates of aforementioned three key species. Nevertheless, FLD has been proven to have a larger scope and found relevant also for the general phenological onset of grasses and shrubs growth [Schwartz and Chen, 2002; Schwartz et al., 2006] and even fruit trees [Schwartz et al., 2013]. The FLD is calculated as follows:

$$FLD_{ij} = SI-x(TX_{ij}, TN_{ij}, lat_i), \; i=1^{28225} \; \text{and} \; j=1^{62}$$

where TX is the daily maximum temperature, TN is the daily minimum temperature, and lat represents the latitude of the location. The index i symbolizes the locations (i.e., the 28,225 grid cells in the study area) and j the years (62 years). The TX and TN are used to accumulate degree-days and synoptic events (i.e., warm peaks in temperature) from the first of January of each year. Notice that TX and TN need to be transformed from Celsius to Fahrenheit before passing them to the model because the SI-x was calibrated using this unit. Latitudinal information is used in this model as a proxy for day length to account for photoperiodic changes [Basler and Körner, 2012]. For a detailed explanation of the SI-x models and for the code to calculate the FLD, please see Ault et al. [2015].

### 2.3. Co-clustering Analysis

Co-clustering algorithms have been used for pattern analysis in many fields [Banerjee et al., 2007; Cho et al., 2004; Dhillon et al., 2003; Wu et al., 2015]. Dhillon et al. [2003] proposed the information theoretic co-clustering (ITCC) algorithm, which uses the l-divergence metric to find optimum co-clusters in the input matrix while preserving the row, column, and co-cluster averages during the process. They applied the ITCC for word-document analysis to find groups of interrelated documents and words. Cho et al. [2004] proposed the minimum sum-squared residual co-clustering algorithm, which employs the squared Euclidean distance for optimization, to obtain subsets of genes and conditions with similar expression values. In 2007, Banerjee et al. developed the so-called Bregman co-clustering algorithm, which is a meta co-clustering algorithm with the above two algorithms as special cases. In their application of the Bregman co-clustering for word-document analysis, Banerjee and colleagues empirically proved the superiority of the l-divergence metric. Recently, Wu et al. [2015] applied the Bregman block average co-clustering algorithm with l-divergence (BBAC_I), which is a special case of the Bregman co-clustering that preserves co-cluster averages, to time series of temperature values and successfully identified observations with similar temperatures along both the spatial and the temporal dimensions. This latter algorithm is also employed in this study to identify co-clusters with similar FLD values along locations and years. The BBAC_I algorithm allows to co-cluster positive data matrices with real-valued elements which represent co-occurrences or joint probability between two random variables [Wu et al., 2015]. It regards co-clustering as an optimization issue in information theory where mutual information measures the amount of shared information between two variables, and yields the optimal co-clusters by minimizing the loss in mutual information between the original data matrix and the co-clustered one.

The FLD data can be regarded as a co-occurrence matrix (OFLD) between a spatial variable taking values in 28,225 locations and a temporal variable taking values in 62 years. The pseudocode of the BBAC_I algorithm (Figure 3) briefly illustrates the iterative process to optimize the partitions of the FLD data matrix to the co-clusters. The MATLAB code for implementing the algorithm is provided in the supporting information.
The first step of the BBAC_I is to perform an initial random mapping of the locations to location-clusters and of the years to year-clusters. This produces the co-clustered matrix \( \hat{O}_{\text{FLD}} \). Then, in the second step, the loss in mutual information between the original and the co-clustered matrices is measured by applying the equation listed in this step, where \( D_1(\cdot \| \cdot) \) denotes the I-divergence between two matrices. In the third step, the algorithm starts an iterative process to update the mapping from locations to location-clusters and years to year-clusters to minimize the loss function. This function has been proven to be monotonically decreasing after each iteration [Banerjee et al., 2007]. The iterations cease when the loss function converges to a local minimum; i.e., the change in the loss is below a predefined threshold. Since a global minimum cannot be guaranteed, this co-clustering process is repeated with several random mappings to maximize the likelihood of finding the global minimum and yield the optimal co-clustering results. Finally, the rows and columns of the \( \text{FLD} \) matrix are re-ordered so that locations/years belonging the same location-/year-cluster are arranged together. In particular, the co-clusters are arranged according to their average \( \text{FLD} \) values: low \( \text{FLD} \) values are positioned at the bottom left and high \( \text{FLD} \) values at the top right corner of the re-ordered matrix. Re-ordered matrices typically exhibit a checkerboard pattern because of the assignment of full rows and columns to clusters in co-clustering algorithms. To better capture the spatio-temporal patterns in the \( \text{FLD} \) dataset, the well-known k-means algorithm was used to regroup the results into k axis-parallel non-rectangular co-clusters, also named irregular co-clusters. The mean and variance of the \( \text{FLD} \) values in each regular co-cluster were used as input features for k-means, and the number of clusters (i.e., k) was optimized using various runs (to obtain optimal solutions) and the Silhouette method [Rousseeuw, 1987]. In this respect, it is important to notice that both clustering algorithms are local optimization ones. This means that the stability of the results must be carefully considered. Especially because the BBAC_I results feed into k-means, using multiple runs is therefore essential to obtain stable and optimal results.

2.4. Spatio-temporal Phenological Patterns

The main spatio-temporal phenological patterns in the \( \text{FLD} \) dataset were explored by visually analyzing the irregular co-clusters from a spatial and a temporal perspective. For this, we visualized the irregular co-clusters using three sets of small multiples (i.e., a series of geographic maps) and two sets of linear timelines (i.e., a linear line that visualizes events in a chronological order). The first set of small multiples was used to show the spatial distribution of each irregular co-cluster or phenoregions, which indicate self-similar clusters in terms of \( \text{FLD} \) in this study. The second set of small multiples was used to display the unique spatial patterns found in the study area (one map per pattern). These spatial patterns were extracted from the irregular co-clusters by combining grid cells with the same variation over the study area. In other words, these
spatial patterns were created by combining phenoregions falling into the same years. A linear timeline was used to show the temporal dynamics of these spatial patterns over the whole study period. The percentage of each phenoregion per spatial pattern as well as the number of years that exhibit each spatial pattern were also calculated to characterize the main spatial patterns and support the study of their temporal dynamics. The third set of small multiples was used to display the spatial extent of unique temporal patterns found during the whole study period. These temporal patterns were extracted from the irregular co-clusters by combining years with the same variation along the study period and arranging them chronologically. Each timeline is used to show each temporal pattern, and thus, the number of timelines is equal to the number of unique temporal patterns.

3. Results

3.1. First Leaf Dates (FLD)

The FLD values were calculated for all valid grid cells and years in the E-OBS temperature datasets. As an example, Figure 4 shows their spatial distribution over the study area in 2010. The greener the color, the earlier the first leaf date is. As expected, the FLD increases from the south to north of Europe. The range of variation is fairly wide with very early spring dates (end of January and beginning of February) for Southern Europe, Turkey, and northern Africa. Late to very late spring dates (end of June and the start of July) occur in Scandinavia and northern Russia as well as in a few “cold islands” corresponding to the Alps and the Caucasus Mountains. It is important to note that although the spatial patterns of FLD (Figure 4) resemble those of TX (Figure 2), these patterns are fundamentally different. Not only because the former displays the spring phenological pattern of one year while the latter shows the patterns of a particular day in that year but also because the Si-x model captures non-linearity in the accumulation of temperature. Therefore, the FLD is more than an accumulation of degree-days.

3.2. Regular and Irregular Co-clusters

The BBAC_I co-clustering algorithm was applied to the FLD matrix, which was created by reorganizing the input FLD maps so that all valid locations appear as rows and the yearly FLD values as columns. The number of year-clusters was set to 4 after testing values from 4 to 10 in the step of one because the BBAC_I algorithm always returned this number of year-clusters. This means that the loss function of BBAC_I achieves its minimum with this value for this particular dataset. The number of location-clusters was set to 45 after testing values in the range of 20 to 60 with an interval of 5. Again, the reason for choosing this value is that it minimizes the value of the loss function. We empirically observed that setting the threshold for convergence of the loss function to $10^{-6}$ and the number of iterations to 2000 was sufficient to guarantee stable results. Finally, the number of random mapping initializations was set to 200 to help find a global minimum of the loss function.

This BBAC_I parameterization produced the $45 \times 4$ regular co-clusters displayed as a heatmap in Figure 5. The values of x axis are the 62 years covered by the dataset arranged according to their membership to the year-clusters, which are sorted from early to late FLD values. Most of the years belonging to year-cluster1 are recent years, whereas the ones in year-cluster4 correspond to the first years of the study period. The values of y axis are the 28,225 grid cells arranged from location-cluster1 to location-cluster45 from top to bottom.
The higher the number of location-cluster, the earlier the FLD. Thus, based on the arrangement of location- and year-clusters along the axes, the regular co-cluster location-cluster45/year-cluster1 has the earliest FLD values, while the regular co-cluster location-cluster1/year-cluster4 has the latest FLD values. Figure 5 also shows several co-clusters with almost the same FLD values, for example, location-cluster45/year-cluster1 and location-cluster45/year-cluster2. This problem, which tends to occur with any co-clustering algorithms where the number of clusters is specified a priori, requires the refinement of the co-clusters. Here this is done by regrouping the co-clusters using k-means. To do this, we calculated the mean and the variance of each co-cluster and used them as input features for k-means (Figure 6). As expected, the mean of FLD values increases from the first to the last co-cluster. However, the FLD variance presents many fluctuations, with a peak at the 115th co-cluster. The most probable reason for these fluctuations are anomalous TX and/or TN values in the original dataset. To cope with this, we named one of k irregular co-clusters as “abnormal”.

The value of k was optimized by testing values from 3 to 10 in the step of one and using the Silhouette method. Results showed that the optimum number of irregular co-clusters is five, and therefore, besides the abnormal co-cluster, we defined “very early”, “early”, “late”, and “very late” co-clusters according to their average FLD values, which (rounded) correspond to 37th, 67th, 105th, and 152nd days of the year. These irregular co-clusters discretize the whole FLD dataset and hence, we refer to them as the main FLD categories. Figure 7 shows the composition of the irregular co-clusters using the same format as that for the regular co-clusters (Figure 5). Figure 7 also shows that some regular co-clusters (e.g., location-cluster39/year-cluster1) belonging to one FLD category (e.g., very early FLD) are embedded in another category (e.g., early FLD). This is because BBAC_I re-orders locations and years into co-clusters using the average value of a complete location and year (i.e., full row or column of the original FLD matrix). For instance, the average value of one location can be higher than another when considering the whole time series but lower when considering a subset of years.

### 3.3. Spatiotemporal Patterns in FLD

The spatial distribution of each FLD category is displayed in the small multiples shown in Figure 8. It shows that the phenoregions for “late” FLD located mostly in Russia and Eastern Europe have the largest extents while those for “very early” FLD located mostly in Southern Europe and northern Africa have the smallest extents. It also shows that there are at most four phenoregions (as this is the number of year-clusters) for each
category due to different spatial extents. For instance, for "very late" FLD located mostly in Northern Europe, the first two phenoregions have the same spatial extent, while for "very early" FLD, the last three phenoregions have the same extent. For both "early" FLD located mostly in Western Europe and Turkey and "late" FLD, there are four different phenoregions, and thus, there are more spatial variations for this two categories. We also observe that the "abnormal" FLD category only happens in a very small area of southern Iceland (for all 62 years) and the borders between Western and Eastern Europe, Scandinavian countries, and northern UK in recent years.

Figure 6. The mean and variance of the FLD values within each of the 45 × 5 regular co-clusters. The number is indexed from location-cluster45 to location-cluster1 and from year-cluster1 to year-cluster4. The left y axis shows the mean of FLD values, and the right y axis shows the variance of FLD values.

Figure 7. The five irregular FLD co-clusters named "very late", "late", "early", "very early" and "abnormal". The composition of irregular co-clusters from regular ones is also displayed in this heatmap by preserving the co-cluster arrangement in Figure 5.
Grid cells with the same variation over the whole study area in Figure 7, that is, phenoregions falling into the same year-cluster in Figure 8, compose four unique spatial patterns over Europe. These spatial patterns, which were named SPattern1 to SPattern4, were shown in the small multiples in Figure 9a, and their temporal dynamics from 1950 to 2011 was shown in the linear timeline in Figure 9b. The percentages of phenoregions

**Figure 8.** (a–e) Phenoregions (self-similar clusters in terms of FLD) for each of the five FLD categories. For instance, the map for VL_phenoregion1 displays the phenoregion for “very late” FLD in accordance with year-cluster1. For categories with less than four, four maps were still used for consistency.
for each FLD category and the number of years taken by these patterns were shown in Table 1. Among the four spatial patterns, SPattern1 has the highest percentages of phenoregions for “very early” and “early” FLD and lowest for “late” FLD, while SPattern4 has the highest of phenoregions for “very late” FLD and lowest for “very early” and “early” FLD. Thus, variations from SPattern1 to SPattern2 or SPattern3 result mainly in 4.9% decrease in “very early” phenoregions located in the Iberian Peninsula, western France, and southern UK.
and also 4% decrease in “early” phenoregions in the Balkan Peninsula and northern UK. Thus, such variations indicate increasingly late FLD and spring onsets and so do variations from SPattern2 or SPattern3 to SPattern4. Besides aforementioned decrease in “very early” phenoregions, variations from SPattern1 to SPattern4 also result in 6% decrease in “early” phenoregions located in Germany and Austria and also 9% increase in “very late” phenoregions in Scandinavia. Therefore, such variations indicate the trend for the very late FLD and spring onsets. The timeline in Figure 9b shows that there is a general trend toward earlier FLD values from early to recent years and from north to south of Europe. It shows that more than two thirds (11 out of 15) of SPattern1 occurred after 1987, while others of SPattern1 did not occur until 1960. It also shows that few (1 out of 16) SPattern4 occurred after 1987 in 1996. Thus, the period of 1950–1960 had very late spring onsets and the period of 1988–2011 had very early spring onsets except 1996. The first big variation of spring onsets during the study period occurred from 1965 (SPattern4) to 1966 (SPattern1), where large areas in Western and Southern Europe, UK, and Scandinavia experienced increasingly very early FLD. The biggest variation of spring onsets occurred in the period of 1975–1977 where these areas experienced increasingly very late and then very early FLD.

Table 1. Percentages of Phenoregions per Spatial Pattern and the Number of Years That Exhibit Each Spatial Pattern

<table>
<thead>
<tr>
<th>Spatial Patterns</th>
<th>VL_phenoregions</th>
<th>LA_phenoregions</th>
<th>EA_phenoregions</th>
<th>VE_phenoregions</th>
<th>Number of Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPattern1</td>
<td>24.4%</td>
<td>46.4%</td>
<td>16.7%</td>
<td>9.1%</td>
<td>15</td>
</tr>
<tr>
<td>SPattern2</td>
<td>24.2%</td>
<td>57.9%</td>
<td>13.7%</td>
<td>4.2%</td>
<td>18</td>
</tr>
<tr>
<td>SPattern3</td>
<td>28.4%</td>
<td>54.9%</td>
<td>12.5%</td>
<td>4.2%</td>
<td>13</td>
</tr>
<tr>
<td>SPattern4</td>
<td>33.4%</td>
<td>51.7%</td>
<td>10.7%</td>
<td>4.2%</td>
<td>16</td>
</tr>
</tbody>
</table>

Years with the same variation over the whole study period in Figure 7 compose 14 unique temporal patterns. In Figure 10 each timeline shows one temporal pattern with variations among FLD categories over the time period of 1950–2011, and the map below shows the geographical extent of that temporal pattern. Among the temporal patterns except those related to “abnormal” FLD, there are four temporal patterns with one FLD category for all years, indicating very late (1st pair), late (6th pair), early (12th pair), and very early (14th pair) spring onsets separately. The maps show that very late and late spring onsets are prevalent in Northern and Eastern Europe, respectively, and the geographical extent of the latter is the largest among all. Other eight temporal patterns are with variations between different FLD categories, indicating changes between very late and late spring onsets (3rd and 4th pairs), late and early spring onsets (7th to 11th pairs), early and very early spring onsets (13th pair). The maps show that Western Europe and the Balkan Peninsula, which predominately have variations between late and early FLD, have the most complex spring onsets. In general, the more south the regions and the more recent the years are, the more are the variations biased toward early spring onsets. It also shows that western Turkey has the most intensive variations of spring onsets because FLD varied 32 times between late and early categories over the study period.

These identified spatio-temporal phenological patterns will be discussed in the following session by comparing them with those found in previous studies. Besides this, possible reasons for these patterns will also be discussed.

4. Discussion

Previous studies on satellite-derived land surface phenology and ground phenological observations have provided information on spring dynamics in Europe using a wide variety of datasets. Through the analysis of advanced very high resolution radiometer normalized difference vegetation index time series for the period of 1982–2001, Stöckli and Vidale [2004] found that the period of 1985–1987 had late spring onsets and that the years 1989, 1990, 1994, and 1995 had early ones. Our results confirm those findings. As shown in Figure 9, the period of 1985 to 1987 indeed had late sprint onsets (SPattern3 and Spattern4) and the four years had very early spring onsets (SPattern1). These early spring onsets could be attributed to an increase in temperatures, as reflected by the fluctuation of atmospheric CO2 assimilation by vegetation [Keeling et al., 1996]. Jeong et al. [2011] and Fu et al. [2014] extended the study period till 2008 and 2011 separately and found that spring onset (start of the growing season in their work) advanced significantly in the period of 1982–1999 and that “earlier springs” weakened in the periods of 2000–2008 and 2000–2011, respectively. Although with a different later changing point at 1988, our study also shows that the pattern in the recent
Figure 10. Each linear timeline to display each of 14 unique temporal patterns with variations among FLD categories over the time period of 1950–2011 and the map below to show the geographical extent of the temporal pattern.
period 1988–2011 (stable very early spring onsets) except 1996 is different with that in previous years. This weakened advance of spring onsets might be caused by asymmetric warming patterns of springs between the two periods. As stated by Fu et al. [2014], the reduced spring warming in the latter period might have led to a retarded advance of spring onsets. The difference in the changing point might be caused by differences in the study period; most of the years in the period of 1950–1981 have late spring onsets, and this might explain why we found an earlier changing point. Using ground phenological records of leaf unfolding from 1951 to 1996, Menzel [2000] concluded that the increasing temperature mostly took place in the late 1980s and 1990s. Our study shows consistent spring onset dynamics that spatial patterns varied from SPattern4 to SPattern1 from 1987 to 1989 (Figure 9), indicating increasingly very early spring onsets. However, our study observed stable early spring onsets in the late 1990s. This discrepancy might be caused by our extended study period and spatial coverage and/or differences in the indicator plants. Recall that the FLD index indicates the spring onset of lilacs and honeysuckles (which are linked to grasses, shrubs, and fruit trees), whereas Menzel [2000] used phenological records from various species. Chmielewski and Rötzer [2001] analyzed the ground observational values of the timing of leaf unfolding for the period of 1969–1998 and reported that spring onsets were generally early in 1990s except 1996, due to the long and strong winter of 1995/1996. These results agree well with our findings. Chmielewski and Rötzer [2001] also analyzed regional patterns of the timing of leaf unfolding and found weak trends of early spring onsets in northern Scandinavia and of late ones in south-eastern Europe. We also found these regional trends (1st and 7th pairs in Figure 10). The mildly decreasing temperature observed in the last few years of their study period may explain the trends in these regions. Ahas et al. [2002] extended the study period of this latter study to 1951–1998 and reported that first third of the study period mostly has late spring onsets and the last third has early spring. Such trends are can also be observed in Figure 9 in this work. All of these results confirm the effectiveness of co-clustering and SI-x based analysis to capture the spatio-temporal phenological patterns.

In contrast to satellite-derived phenological metrics and ground observations as cited above, phenological models have the advantage of predicting phenological metrics of spring onsets over large regions and for long time periods. For instance, Schwartz et al. [2006] analyzed European FLD patterns from 1955 to 2002 using the original spring index (SI) models and temperature data from meteorological stations. The stable spring onsets in Eastern Europe reported in their study are consistent with this study (6th pair in Figure 10). In both cases, this stability might be caused by the weak spring warming observed in this region [Jones and Moberg, 2003]. However, we found variations between late and early spring onsets, with a bias toward the latter in recent years, in parts of Germany. This is not consistent with the conclusions of Schwartz et al. [2006], which indicate that central Europe (mainly Germany) exhibits the strongest trend of earlier spring onset. This inconsistency might be caused by two reasons. The first one is the difference in the study periods. The second is the different sources of datasets. Schwartz et al. [2006] used temperature data from individual meteorological stations and were only able to study trends in isolated and unevenly distributed stations. In our case we use interpolated grid cells, and thus, the trends are spatially-continuous. Nevertheless, the temporal dynamics of four spatial phenological patterns explored in current study is confirmed by the similar temporal variation of departure from long-term average shown in their work by a simple curve. However, more detailed spring phenological patterns over Europe are described in this work by the four spatial patterns shown in maps (Figure 9).

Even though SI-x do not consider phenological strategies such as chilling requirements [Polgar and Primack, 2011], they retain the utility and accuracy of SI which take above strategies into account [Schwartz et al., 2013]. This fact enables the extension of the analysis into areas such as northern Africa in this work or those analyzed by Zhang et al. [2007], where chilling requirements are not met [Schwartz et al., 2013]. Thus, together with the use of the E-OBS temperature datasets, SI-x allowed us to delineate spatially contiguous phenoregions in Europe, northern Africa, and Turkey. These phenoregions are more informative than the analysis of temporal dynamics based on unevenly distributed locations (meteorological stations or ground observations). Moreover, our approach allows visualizing the boundaries of the main phenoregions and studying their changes in time. To the best of our knowledge, this study reports the first spatially-continuous phenoregions in Europe (northern Africa and Turkey) while several phenological regionalizations already exist for CONUS [Gu et al., 2010; Kumar et al., 2011; Mills et al., 2011; Zhang et al., 2012]. Our results also confirm that SI-x work well for large-scale analysis. However, one potential caveat of the SI-x models is that they do not consider other environmental variables than temperature and, indirectly, day length. Therefore, the SI-x models might not fully capture the phenological dynamics of areas or species that are driven by, for instance, precipitation.
Finally, in contrast to previous phenological studies, that analyzed spring phenology from a separate view, this is the first study that applies co-clustering to the exploration of spring phenological patterns. Co-clustering allows the simultaneous analysis of the spatial and temporal dimensions of the data. This resulted in the simultaneous identification of not only the main spatial patterns in the study area and their temporal dynamics but also the identification of the main temporal patterns over the study period and their spatial distributions.

5. Conclusion

In this paper, we have presented a novel analytical workflow that allows to simultaneously study spatio-temporal patterns over large areas and long time periods. In more details, our analysis first used the extended spring index models to calculate time series of FLD from the E-OBS daily maximum and minimum temperature datasets over Europe, northern Africa, and Turkey from 1950 to 2011. Then the FLD dataset was co-clustered by applying the Bregman block average co-clustering algorithm with I-divergence (WBAC_I) and k-means. This co-clustering based approach identified irregular co-clusters that contain similar FLD values along both locations and years. Finally, these irregular co-clusters were visually explored to reveal the main spatio-temporal patterns in spring phenology.

The developed analytical workflow allowed the delineation of spatially-continuous phenoregions in Europe for the first time. Besides, our analysis identified four unique spatial patterns and showed that the early years of the study period (1950–1960) have very late spring onsets, especially in Northern Europe and Russia. The period of 1961–1987 has many variations in the timing of spring onset. The largest variation took place in the years 1975 to 1977 where the spring onset occurred from very early to very late and back. We also saw that recent years (from 1988 onward) tend to have stable early springs, especially in the Iberian Peninsula and northern Africa. One exception is the year 1996 that had a very late spring onset. Our results also identified 12 unique temporal phenological patterns as well as their spatial distributions. Four of these temporal patterns indicate stable spring onsets and eight patterns indicate variable spring onsets. The analysis of their spatial distribution showed that very late and early spring onsets prevail in Northern and Eastern Europe while variations between late and early spring onsets prevail in Western Europe and the Balkan Peninsula, with the most intensive variations located in western Turkey.

For the future work, we anticipate the identification of the driving forces behind the patterns found in this study. This could be done by linking the patterns (e.g., phenoregions and temporal dynamics) with environmental variables. More generally, further work could investigate the role of decadal and interannual climatic variabilities, e.g., from the influence of North Atlantic Oscillation in the variations of spatial and/or temporal spring onset patterns.

Acknowledgments

We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu). The E-OBS dataset is available at http://www.ecad.eu/download/ensembles/download.php. We also thank David G. Rossiter and Toby R. Ault (both from Cornell University) and Julio Betancourt (USGS) for their insightful comments on this work.

References

Dhillon, I. S., S. Mallela, and D. S. Modha (2003), Information-theoretic co-clustering, paper presented at The 9th International Conference on Knowledge Discovery and Data Mining (KDD).