Geocomputational Workflows for Analysing Spring Plant Phenology in Space and Time

Hamed Mehdi Poor

ITC

UNIVERSITY OF TWENTE.

GEOCOMPUTATIONAL WORKFLOWS FOR ANALYSING SPRING PLANT PHENOLOGY IN SPACE AND TIME

Hamed Mehdi Poor

GEOCOMPUTATIONAL WORKFLOWS FOR ANALYSING SPRING PLANT PHENOLOGY IN SPACE AND TIME

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof.dr. T.T.M. Palstra,
on account of the decision of the Doctorate Board,
to be publicly defended
on 30 January 2019 at 16.45 hrs

by

Hamed Mehdi Poor

born on 15 January 1986

in Sirjan, Iran

This thesis has been approved by
Prof.dr. R. Zurita-Milla, supervisor
Dr.ir. P.W.M. Augustijn-Beckers, co-supervisor

Cover designed by Job Duim
Printed by ITC Printing Department

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Graduation committee:

Chairman/Secretary
Prof.dr.ir. A. Veldkamp   University of Twente

Supervisor
Prof.dr. R. Zurita-Milla   University of Twente

Co-supervisor
Dr.ir. P.W.M. Augustijn-Beckers   University of Twente

Members
Prof.dr. M.J. Kraak   University of Twente
Prof.dr. A.D. Nelson   University of Twente
Prof.dr. A. Wytzisk   Bochum University of Applied
                                      Sciences
Dr. K. Hufkens   Ghent University

To friendship between Iranian, Dutch and American people

Acknowledgements

Table of Contents

List of figures

List of tables

Chapter 1 Introduction

1.1   Background

Among the various research questions raised by climate change, the question: "how does climate change affect vegetation seasonality?" is crucial because changes in vegetation seasonality have global and substantial implications for our planet (Bakkenes et al., 2002; Walther et al., 2002; Parmesan and Yohe, 2003; Inouye, 2008). For instance, several studies have shown that changes in vegetation seasonality are affecting the distribution and productivity of natural and agricultural plants (Chmielewski et al., 2004; Park et al., 2005). Vegetation seasonality information is also needed for a wide range of applications such as food security (Anyamba and Tucker, 2005; Vintrou et al., 2012), nature management (van Rooijen et al., 2015), and public health (Luber and Lemery, 2014; MacDonald, 2018). Moreover, vegetation seasonality controls the global biochemical circles, including water and carbon cycles (Keenan et al., 2014; Yuan et al., 2018).

Exploring changes in vegetation seasonality in space and time is a pre-requisite to understand the impact of climate change (and of inter-annual weather variability) on our planet. It also helps to design climate change adaptation strategies. This chapter provides an overview of the research problem addressed in this thesis. Section 1.1 describes the study of seasonal plant life cycle events. Section 1.2 focuses on volunteered observations of the events. Section 1.3 reviews phenological modelling approach. Section 1.4 describes the impact of observations consistency on phenological studies. Section 1.5 reviews the type of models, as well as the source and scale of inputs used to estimate the timing of the events. Section 1.6 provides an overview of the application of geocomputation workflows for large-scale analysis of vegetation seasonality.

Spring plant phenology

Phenology is the science that studies periodic plant and animal life cycle events (phenophases) and how annual and inter-annual variations in weather and environmental conditions affect them (Lieth, 1974; Kramer, 1996). Examples of plant phenophases are first leaf, first flower, full leaf, full flower and first fruit (Figure 1.1). The start of a phenophase can be pinpointed to a single day of year (DOY). This PhD thesis focuses on spring plant phenology (SPP) because not all plant phenophases are equally useful for studying the impact of climate change on vegetation seasonality. In particular, the first leaf and first flower phenophases that occur after winter dormancy are sensitive to climate variability (Cayan et al., 2001; Schwartz et al., 2006, 2013; Post et al., 2018). Moreover, the impact of climate change is typically greater in spring, and more phenological observations are available for plant phenophases in spring than in any other season (Bonsal et al., 2001; Robeson, 2004).
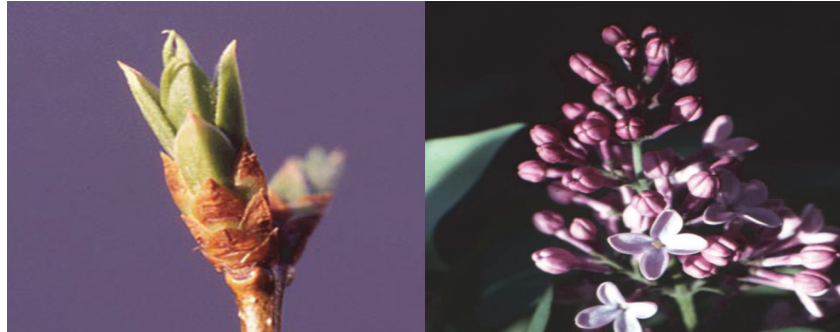
Figure 1.1 Lilac first leaf (left, credit: Mark Schwartz) and first flower (right, credit: Elisabeth Beaubien).

Increases in global temperature, particularly at the end of the winter season, have resulted in earlier spring onsets throughout the Northern Hemisphere (Schwartz et al., 2006, 2013; Allstadt et al., 2015). When plants leaf and bloom earlier than usual, pollinators and herbivores have to adjust their life cycle events (Marra et al., 2005; Miller-Rushing et al., 2010; Gornish and Tylianakis, 2013; Broussard et al., 2017). Earlier leafing and flowering can dry out soils and advance, and even exacerbate the wildfire season (Abatzoglou and Williams, 2016). Advancements in flowering can cause frost damage to fruit crops (Gu et al., 2008; Ault et al., 2011; Munson and Sher, 2015; Chen, 2017). As a result, several studies have analysed and modelled the spatial and temporal variation of these two phenophases. Different species are used to study SPP, ranging from natural to agricultural plants (Schwartz, 1999; Schwartz and Chen, 2002). Some experiments rely on cloned individuals to highlight the effect of climate change and weather variability over genetic variability. For example, cloned lilacs (*Syringa chinensis 'Red Rothomagensis'*) have been widely used to study SPP for more than half a century across the Northern Hemisphere.

The collection of timely phenological observations on selected species is the first step in the study of SPP. The analysis and modelling of phenological observations provide valuable insights into the influence of weather and climate dynamics on plant growth (Studer et al., 2005; Chmielewski, 2013; Zurita-Milla et al., 2015). Phenological observations often contain the geographic location and the DOY of a particular phenophase for a given species. This data helps to study vegetation seasonality in space and time (Schwartz and Reiter, 2000; Wu et al., 2016).

1.2   Volunteered geographic information for SPP

For centuries, volunteers have contributed to the production of information about geographic phenomena such as the impacts of climate change on our

planet (Bock and Root, 1981). However, progress in online information communication and mobile location-aware technologies have dramatically increased the amount of geographic information that can be collected by volunteers (Beaubien and Hamann, 2011; Ferster and Coops, 2013). The development of global positioning systems enabled volunteers to efficiently georeference their information (Gouveia and Fonseca, 2008). Further, evolutions in web 2.0 have allowed volunteers to register their own information and to share it via the internet (van Vliet et al., 2003; Wiersma, 2010). These developments have led to the emergence of volunteered geographic information (VGI).

In particular, phenology has benefited from VGI because it offers a practical approach to acquire timely and detailed information at low cost across a variety of spatial and temporal scales (Goodchild and Li, 2012; Comber et al., 2013). These developments have contributed to the low number and small extent of observations, which were the main limitations of most of the ecological and phenological studies (Dickinson et al., 2010; Rosemartin et al., 2015). Phenological VGI containing the geographic location and DOY of the observed phenophases are hereafter referred to as volunteered phenological observations (VPOs). These observations open new opportunities for the study of spatial patterns and temporal trends of vegetation seasonality from both a spatial and a temporal perspective (Sparks et al., 2008; Beaubien et al., 2011; Zurita-Milla et al., 2013). Accordingly, worldwide efforts to collect, monitor, and synthesize VPOs enable scientists to obtain a new perspective on how global change is affecting organisms across a wide range of spatial scales (Brunsdon and Comber, 2012; Fuccillo et al., 2015). The national phenology networks are using various platforms (e.g., web and mobile application, etc.) and protocols to collect and store VPOs of a wide range of species. Phenological networks have large sets of VPOs of spring phenophases because these events are fairly simple to observe for volunteers and are promoted by scientists who study climate change effect.

Although VPOs provide valuable phenological information, they are not ideal. For instance, VPOs tend to be unevenly distributed (e.g., clustered around cities where most of the volunteers live); also because some locations are unreachable for volunteers. Therefore, it is necessary to develop and use alternative approaches to generate spatially continuous phenological information. In this respect, the use of VPOs to calibrate and validate models that estimate the DOY of phenophases from contextual environmental information is a scientifically interesting alternative.

4

1.3 *Phenological models*

For many years, geographers have already used modelling approaches to fix the lack of geographic information in either space or time. As a result, there is a range of models that can be used to estimate the location and time of geographic phenomena (Giorgi and Mearns, 1991; Boyd and Doney, 2002; Fowler et al., 2007; Fitchett et al., 2015; Fraga et al., 2016). These models allow the extrapolation of in-situ phenological observations such as VPOs to unvisited areas (Schwartz, 1994; Chmielewski et al., 2014; Jochner et al., 2014). Phenological models (PMs) are designed and calibrated to estimate the DOY of the phenophase at variuos geographic locations (Worner, 1992; Ault et al., 2015). The outputs of PMs are used to discover patterns and trends in plants phenophases.

PM derived information has a large potential for different environmental applications such tacking the rhythm of seasons (Morisette et al., 2009), estimating carbon sequestration potential of forests (Leinonen and Kramer, 2002), agriculture and natural resource management (Schwartz et al., 2013; Gerstmann et al., 2016; Nissanka et al., 2017). Moreover, phenological model outputs are used to reconstruct and qualify ground- (Chuine et al., 2004; Menzel, 2005) and satellite-based (Schwartz et al., 2002; Macbean et al., 2015) time-series of VPOs, and to estimate species-specific phenology (Krinner et al., 2005; Chuine et al., 2013). PMs use environmental geo-information such as weather parameters to estimate the DOY of phenophases. This geo-information is typically available over a larger area and longer time periods than in-situ phenological observations (Schwartz et al., 2000; Chuine et al., 2013; Richardson et al., 2013).

Among the weather parameters which are used to calibrate PMs, temperature has been found to be crucial. De Réaumur, who was an entomologist, commenced plant phenology modelling in 1735 (Puppi, 2007). He explained the differences between years and locations by differences in the summation of daily temperature from an arbitrary date to the date of the phenophase; something that is now known as degree-day summation. This summation of daily temperatures has been recognized as a significant factor reflecting inter-annual variation in plant phenology (Chuine et al., 2013). Later, Adanson (1750) modified de Reaumur's model introducing the concept of the thermal threshold by which the summation was calculated excluding temperatures below a specific degree. SPP models use degree-days and other predictors to estimate the DOY of occurrence of plants events (Schwartz and Marotz, 1988; H. Wang et al., 2015). These statistical models can generalize the phenology of a wide variety of plants to make predictions national and continental scales and over several decades (Allstadt et al., 2015). The Extended Spring Indices, the Thermal Time and the Photothermal Time are examples of widely used SPP

models in the Northern Hemisphere (Linkosalo et al., 2008; Hufkens et al., 2018).

Although PMs provide valuable information to explore patterns and trends in SPP, little research has been conducted on the effect of the consistency of VPOs that are used to calibrate these models (Mendoza et al., 2017). In these respects, the next section elaborates further on the checking of VPO consistency, and its impact on the modelling of trend and synchrony of VPOs.

1.4   Consistency and synchrony of VPOs

Spatial and temporal uncertainties in the actual location and time of VPOs are an inseparable part of these observations as only volunteers decide where and when to observe (Schaber et al., 2010). Volunteers are non-professionals and have different levels of expertise in recognizing specific phenological events or, the target species (Brunsdon et al., 2012). Moreover, volunteers may also perform observations at locations with environmental conditions that are not representative of the phenological events being monitored (e.g., they might report data for an individual plant growing under a special micro-climate). Further, there is often no prescribed scientific experimental approach for the collection of VPOs and there are changes in VPO collection protocols over time that negatively affect the consistency of VPOs (Yanenko and Schlieder, 2012; Schwartz, 2013). As a result, there are VPOs anomalously early or late in relation to their associated environmental conditions; these observations are called inconsistent VPOs in this thesis.

Inconsistent VPOs might affect phenological studies that can be supported by volunteered observations. Among the various phenological studies that can be supported by VPOs, is the analysis of phenological synchrony, defined here as the temporal dispersion of a phenological event across individuals of the same species (Sparks et al., 2008; Mihorski et al., 2012). Analysis of phenological synchrony is sensitive to inconsistent observations. Phenological synchrony is often quantified by the standard deviation of DOY of all the observations collected in a given area and year (Henderson et al., 2000; Gordo and Sanz, 2010; C. Wang et al., 2016). Phenological synchrony is particularly interesting because changes in phenological synchrony have ecological consequences for individual survival and ecosystem stability (Ims, 1990; English-Loeb and Karban, 1992; Both and Visser, 2001). For example, low flowering synchrony can hamper the expected random mating pattern because early bloomers are more likely pollinated by early plants, and late plants by late plants (Weis and Kossler, 2004). Phenological synchrony is strongly controlled by annual weather variability in regions with a marked seasonality (Both et al., 2001; Gordo et al., 2010). Thus, checking the consistency of VPOs is necessary to

investigate phenological synchrony and its inter-annual variations, which increases our understanding of the impact of climate change on species.

Consistency checks of VPOs primarily rely on human review, or simple statistical deviation from an expected probability distribution. Human-dependent workflows can be costly and time-consuming. The purely statistical checks assume that the majority of the observations are consistent and, therefore, can be used to identify inconsistent VPOs. In 2010, Schlieder and Yanenko proposed a consistency check in which observations in close spatial and temporal proximity confirm each other as a criterion. Their method introduced a graph in which observations are modelled as nodes. Edges connect nodes to each other creating a so-called confirmation graph. For each edge, there is a value (positive or negative) that shows the extent to which connected observations confirm or deny each other. Then, a value shows the degree of consistency of each node or observation.

Although using locations of VPOs to check the consistency is an added value of current methods, the methods do not use independent sources of information from the environmental context of the VPO. Besides, environmental contextual informants, such as temperature, are widely used to build different PMs. In addition to VPO consistency, the type of PMs and the source and scale of their inputs might affect the study of phenological patterns and trends. The next section provides an overview of the effect of these latter factors in more detail.

1.5   Types and inputs of SPP models

Weather-driven SPP models are based on different statistical and/or ecological assumptions. Some SPP models assume that changes in plant phenology are only (directly or indirectly) driven by daily temperature while other models use both daily temperature and photoperiod to model plant phenology (Capiro, 1993; Schwartz et al., 2012). SPP model parameters range from simple accumulations of degree days to advanced counting of high-energy synoptic events (Chuine et al., 2013). Some SPP models use the same parameters but apply different mathematical formulations. For example, some SPP models define forcing temperatures (i.e., temperatures at which the plant develops) using linear and non-linear formulas. Further, ground-based phenological observations that are used to calibrate SPP models vary (Wolfe et al., 2005; Chmielewski, 2013; Hamunyela et al., 2013). As a result, outputs of SPP models and patterns and trends which are derived from these models might differ significantly.

In addition to different model parameters, mathematical formulations and calibration datasets, SPP models also use different sources of input temperature data to estimate DOY. In particular, we focus on gridded

temperature time-series (GTT) in this PhD thesis, which are available and used more than ever to study SPP (Ault et al., 2015; Izquierdo-Verdiguier et al., 2018). GTTs-driven model's outputs are widely used to support management decisions that support the adaptation of the ecological and agricultural system to global change (Enquist et al., 2014; Gerst et al., 2016). Several studies have used GTTs to generate and to analyse patterns and trends in spring phenology of plants (Ault, 2015; Melaas et al., 2016; Izquierdo-Verdiguier, 2018; L. E. Parker and Abatzoglou, 2018). This is because SPP models can provide continuous phenological information using these data (Izquierdo-Verdiguier et al., 2018). GTTs are generated from varying ground-based daily measurements and interpolation models. Further, they are available at different spatial resolutions. These differences in GTT might affect outputs of SPP models, and consequently the patterns and trends based on these data.

It is necessary to analyse the effect of model type, data source and data scale on the phenological patterns and trends derived from SPP models, especially at large spatial and long temporal scales. Current evaluations of SPP models are divided into the calibration and the validation of the model. The calibration phase is used to find the values of the model parameters that minimize the error of the model. The validation phase is used to assess the error of the model using an independent input dataset. Calibration and validation of SPP models over a large area are now possible for two reasons: wide availability of new gridded temperature time series and of contemporary VPOs. At large spatial and long temporal scales, such evaluations require the implementation of steps which are computationally efficient and reproducible. The next section provides and overview of workflows that overcome the limitations of computational intensiveness and reproducibility.

1.6   Geocomputational workflows

Technological advancements and their general adoption have led to a tighter integration of the geosciences with computer science. This, in turn, has led to geocomputational approaches, which help to process and integrate massive amounts of geographic information to solve complex spatio-temporal problems (Ehlen et al., 2002; Heppenstall and Harland, 2014; Batty, 2017). Geocomputation has improved analytical methods by going beyond classical statistical and spatial analytical approaches, and reaching out more advanced methods such as data-driven and distributed computing (J. Liu et al., 2015; Thill and Dragicevic, 2018). Data-driven methods such as machine learning and data mining are getting more and more popular in scientific research and these methods can be integrated with the geographic information system (GIS) and Earth Observation data to solve non-linear and nonparametric problems (Thill et al., 2018). Data-driven methods do not require specific distributions or other constraints over input variables. This explains the

impact of novel regression and supervised and unsupervised classification tasks in many (ecological) studies. Data-driven methods reduce computation time and tend to improve model performance (Rodriguez-Galiano et al., 2016; Talbert et al., 2017).

Large-scale distributed computing such as cloud computing has scaled up the storage and data processing of spatio-temporal data (Guo et al., 2010). This development enables analysis and modelling of geographical phenomena at national and continental scales. Something that was not possible in the past. Cloud-based approaches also allow the development of highly customizable geoprocessing tools (Karimi et al., 2011; Haynes et al., 2018; Huang et al., 2018). The CyberGIS Gateway and Geospatial Building Blocks (GABBs) are examples of such tools (Y. Liu et al., 2015; Song et al., 2016). Moreover, cloud-based Geo-platforms often offer data and computation together. This empowers researchers who can now focus on their work without having to deal with technical issues. For example, Google Earth Engine, based on its millions of servers around the world, has a large catalog of Earth observation data that enables the scientific community to work on gridded and vector data in an intrinsically parallel way (Gorelick et al., 2017). Thus, cloud computing should be integrated with data-driven approaches in scientific researches.

Scientific workflows are based on rich and diverse data resources while they provide a systematic way of describing the processing steps needed and provide the interface between scientists and computing infrastructures (Atkinson et al., 2017; Cohen-Boulakia et al., 2017; Yenni et al., 2018). These workflows improve the reproducibility of evidence which supports scientists to take responsibility for the quality of their results and findings. The reproducibility of a study does not necessarily mean that the results are scientifically correct, but ensures computational transparency in the result (Stodden, 2010; Yin et al., 2017). Reproducible workflows allow researchers to test the findings, as well as to use the methods which are developed by other researchers (Morisette et al., 2013; Cohen-Boulakia et al., 2017). Hence, there is no doubt that reproducible geocomputational workflows are ideal for scientists who study geographic phenomena such as SPP at large scales. However, reproducible geocomputational workflows are not addressed in large-scale phenological studies.

It is not always clear what source of data and what interconnection and order of steps are used in phenological studies. This is because phenological studies often explain their input data and processing steps in plain (e.g., English) text. Reading the same text might result in various interpretations, which might produce different results (Gil et al., 2007; Piekielek et al., 2015). There is a lack of geocomputational workflows that analyse the effect of varying source and type of VPOs, weather data and models in large-scale phenological studies.

In this PhD thesis, we designed and illustrated such geocomputational workflows that access and retrieve data from data repositories that provide and keep evolving datasets. In the next section, we describe the main research objective and research questions.

1.7   Research objective and questions

To the best of our knowledge, there is no comprehensive study that analyses at large spatial and long temporal scales the effect of VPO consistency as well as type and input of SPP models on vegetation seasonality. Hence, the main objective of this PhD thesis is:

"*To design novel geocomputational workflows to explore vegetation seasonality at large scale and over long periods using volunteered information and phenological models*"

This main objective is operationalized by splitting it into two sub-objectives, which are achieved by answering four research questions:

Sub-objective 1: "*To check the consistency of volunteered phenological observations using contextual geo-information and domain knowledge*"

> Q1. How to use environmental contextual information to check the consistency of volunteered phenological information?

> Q2. How to integrate domain information (i.e., phenological synchrony) with contextual information to check the consistency of volunteered phenological observations?

Sub-objective 2: "*To analyse the impact of the type of phenological model as well as of its input data sources and their spatial resolution on the patterns and trends derived from the model*"

> Q3. How to analyse the impact of the type of phenological model on the patterns and trends that can be derived from it?

> Q4. How to analyse the effect of using various gridded model inputs and of their spatial resolution on the patterns and trends derived from a phenological model?

1.8   Thesis outline

This thesis has six chapters including the introduction and synthesis. The core chapters have been published, or are submitted to, peer-reviewed journals. After this Introduction:

Chapter 2 presents a workflow to check VPO consistency applying dimensionality reduction, model-based clustering and outlier detection methods on weather information and volunteered observations. The workflow is demonstrated using Daymet data and highlights inconsistent VPOs from the USA National Phenology Network (USA-NPN).

Chapter 3 describes a workflow to check the consistency of VPOs while taking phenological synchrony into account. The workflow, based on network graphs, regression modelling and constraint satisfaction methods, is tested using temperature data from the Royal Netherlands Meteorological Institute and phenological observations from the Dutch national phenological network.

Chapter 4 illustrates a cloud-computing based workflow to assess and compare the effect of using various kinds of phenological models on phenological patterns and trends over the coterminous United States. The workflow uses simulated annealing and regression modelling to calibrate models and to assess their outputs using historical and contemporary VPOs from USA-NPN and Daymet data.

Chapter 5 illustrates a cloud-computing based workflow to validate and compare the gridded phenological patterns and trends generated from high resolution gridded weather data over the coterminous United States. The workflow uses cloud-computing and regression modelling to access the effect of the data and to model long-term pattern and trends.

Chapter 6 summarizes the main findings from chapters 2 to 5, includes a research reflection, answers the research questions, presents the main contributions of this PhD thesis, and provides recommendations for future research.

Chapter 2 Developing a workflow to identify inconsistencies in volunteered geographic information: a phenological case study*

2.1 Introduction

The contribution of volunteers to the production of information about geographic phenomena, such as the impacts of climate change, is not new. For example, the Christmas Bird Count has studied the impacts of climate change on the spatial distribution and population trends of selected bird species in North America since 1900 (Butcher and Niven, 2007). However, improvements in online information communication and mobile location-aware technologies have led to a dramatic increase in the amount of volunteered geographic information (VGI) in recent years (Gouveia, 2008; Feick and Roche, 2013; C. J. Parker, 2014).

VGI, a term coined by Goodchild (2007), refers to "the harnessing of tools to create, assemble, and disseminate geographic data provided voluntarily by individuals". VGI is a practical approach to acquire timely and detailed geographic information at low cost across a variety of spatial and temporal scales (Goodchild et al., 2012). Because of this, VGI is used to understand and manage important emerging problems in many fields such as conservation biology (Newell et al., 2012), urban planning (Brabham, 2009), disaster management (Goodchild and Glennon, 2010) and earth observation (van Vliet et al., 2003; Mayer, 2010; Ferster et al., 2013).

Despite the wide applicability and acceptability of VGI in science (Dickinson et al., 2010; Feick et al., 2013) many studies argue that the quality of the observations provided by volunteers remains a concern (Elwood, 2008; Flanagin and Metzger, 2008; Goodchild, 2009; Coleman et al., 2009; Matyas et al., 2011; Galindo et al., 2011; Goodchild, 2012; Elwood et al., 2013; Bimonte et al., 2014). This is because VGI does not often follow scientific principles of sampling design, and levels of expertise vary among volunteers (Brunsdon et al., 2012; Comber et al., 2013). Moreover, unlike traditional authoritative geographic information, VGI typically lacks automated quality checking mechanisms (Kelling et al., 2011, 2012; See et al., 2013).

Among the different data quality aspects, consistency of VGI is considered key for most studies, where inconsistent VGI are observations that are implausible regarding the conditions, geographic location or time they were obtained. Such inconsistent observations can bias analysis and modelling results because they are not representative for the variable studied, or because they decrease the ratio of signal to noise. Hence, the identification of inconsistent observations would clearly benefit VGI-based applications and provide more robust datasets to the scientific community.

The approaches to check VGI quality can be categorized into three main types (Goodchild et al., 2012; Elwood et al., 2013): 1) crowdsourcing where

volunteers validate and thus refine the quality of observations by themselves, 2) social which relies on a hierarchy of trusted people who act as moderators, and 3) geographic, where given the location of the volunteered observations, one can use certain geographic rules to assess quality, e.g., Tobler's "first law of geography" which states that "all things are related, but nearby things are more related than distant things" (Tobler, 1970). The geographic approach is more readily machine-automated than the other two approaches (which rely on human subjectivity), and is therefore the focus of this study (Goodchild et al., 2012).

As an example, eBird, a popular VGI-based initiative for bird monitoring, uses the geographic approach to automatically verify new observations, using historical observations, prior to human moderation (Sullivan et al., 2009). The eBird quality filter relies on substantial prior knowledge about a given organism, geography or time (e.g., a measure of how frequently a species is reported in a region during a specific time period), as well as information about volunteer expertise levels (Kelling et al., 2012). Such information is not always available for VGI-based initiatives.

Schlieder and Yanenko (2010) used spatiotemporal proximity and social distance (i.e., the distance between the observers in the social network of observers on the web) to define constraints for checking the inconsistency of observations. The hypothesis was that spatiotemporally and socially close observations presumably referred to the same event so would more likely be consistent. Their workflow was used to formulate general rules and to find observations that have low confirmation. This workflow was further developed using constraint satisfaction approach to produce more sophisticated results (Yanenko et al., 2012). However, the improved workflow still uses spatial distance as the only criterion to connect observations. Moreover, this workflow is useful only when a sequential order of volunteered observations is available at a given location.

Yet another geographic workflow was proposed by Ali and Schmid (2014) based on machine learning for identifying wrongly-categorized Open Street Map observations. These authors trained a classifier using contributed entities and their associated class labels (e.g., park or garden). However, their model was only concerned with the inconsistency of areal entities (i.e., extended geometric entities such as buildings) regarding administrative boundaries and semantic classifications.

There is a lack of standardized workflows that address VGI inconsistency. Current inconsistency workflows primarily rely on human review, or simple statistical deviation from an expected probability distribution. Human-dependent workflows can be costly and time-consuming, and are impracticable

in some situations, e.g., in cases where events persist only for short periods of time. The statistical workflows assume that the majority of the observations are consistent and, therefore, that these can be used to check for inconsistency. Moreover, existing workflows do not optimally use environmental contextual data. This raises the question of how to address inconsistency using a more objective, efficient and automated workflow.

This paper describes a novel automated workflow to identify inconsistency in VGI. A robust identification of inconsistent observations allows testing their potential impact on VGI-based studies. The workflow relies on the availability of contextual information and is built using a combination of dimensionality reduction, clustering and outlier detection techniques and it was illustrated using observations on the timing of the first flower of lilac plants collected by volunteers. While some inconsistent observations may reflect real, unusual events, here we demonstrate that these observations bias the trends (advancement rates) of the date of lilac flowering onset. This shows that identifying inconsistent observations is a pre-requisite to study and interpret the impact of climate change on the timing of life cycle events (Ault et al., 2013; Schwartz et al., 2013).

2.2 Materials and methods

Phenological VGI

Phenology is the science of the study of periodic plant and animal life cycle events and how seasonal and inter-annual variations in climate affect them. Phenological studies are important to understand the impact of global change in our planet (Schwartz, 1990; Cleland et al., 2007; Barr et al., 2009; Keatley and Hudson, 2010). Worldwide, several VGI-based initiatives collect or have collected phenological data (Schwartz, 2003; Koch, 2010). One VGI-based initiative, the USA National Phenology Network (USA-NPN; www.usanpn.org), has recently released a curated dataset of lilac leafing and flowering observations across the continental United States for the period 1956 to 2014 (Rosemartin et al., 2015). From this dataset we extracted flowering records for common lilac (*Syringa vulgaris*) and cloned lilac (*S. x chinensis 'Red Rothomagensis'*). Considering data completeness and the availability of environmental contextual data, we concentrated our analyses on flowering onset dates for the period 1980 to 2013, for cloned lilacs (with 2174 observations) and common lilacs (with 2682 observations) separately.

Widespread and readily observable, lilac plants have been observed across the continental United States since the 1950's, as a complement to cooperative weather data collection (Schwartz et al., 2012). Observations of lilac leafing, flowering and fruiting have been used for a variety of applications, including

understanding trends and variations in the onset of spring and tracking the impacts of climate change on natural resources (Schwartz et al., 2006). Although lilacs are ornamental plants, their phenology and response to climate have been shown to closely track native species and crops (Schwartz et al., 2013).

The following attributes were used to check inconsistency for cloned and common lilac flowering dates: 1) a unique ID for each record, 2) the year when the flowering occurred, 3) the day of the year (DOY) when the flowering occurred and 4) geographic location where the phenological phase was reported (latitude, longitude and elevation). It is important to note that since 2009, volunteers report the status of each phenological phase with "Yes" when it is visible and "No" when it is not visible (Denny et al., 2014). This status monitoring approach allows for the quantification of uncertainty in flowering onset DOYs (i.e., number of days between the "Yes" and the preceding "No"). Thus, the status monitoring provides additional information on the occurrence of multiple flowering events in a year for individual plants. When a "Yes" report was followed by at least one "No" report and then a subsequent "Yes" record was present on an individual plant, all corresponding DOYs to "Yes" reports were flagged and stored as multiple "Yes" observations in the dataset.

Environmental contextual data

The proposed workflow requires environmental contextual data to characterize observation locations. In phenology, cumulative climatic parameters are the most relevant contextual datasets, because most phenological processes are driven by climate conditions (Barr et al., 2009; Ranta et al., 2010; Schwartz, 2013). Therefore, we extracted climate parameters for the period 1980 to 2013 from Daymet, a dataset that provides 1 by 1-km gridded estimates of daily climatic parameters for North America (Thornton et al., 2014).

Cumulative climatic variables were created for each geographic location by summing parameter values from the 1 January for the year of the observation to the reported DOY of flowering. Cumulative variables calculated include: maximum daily temperature (degrees C), minimum daily temperature (degrees C), daily precipitation (mm/day), daily water vapor pressure (Pa), daily solar radiation (W/m2), daily day-length (s/day) and daily snow water equivalent (kg/m2). In addition, using the daily maximum and minimum temperatures, we calculated daily average temperatures and cumulative average daily temperature (degrees C). Thus, a total of 11 contextual variables (i.e., 8 cumulative climatic variables and the 3 geographic variables of latitude, longitude and elevation) were associated with each phenological observation expressed as DOY (Table *2.1*).

Table 2.1 Mean and standard deviation of the geographic and climatic parameters for cloned and common lilacs.

Variable	Cloned lilac		Common lilac	
	Mean	Standard deviation	Mean	Standard deviation
DOY (of flowering)	133	21	123	22
Latitude	42.42	2.46	42.13	4.15
Longitude	-79.38	9.73	-105.60	12.99
Elevation	255.86	252.79	917.51	1051.14
Cumulative maximum daily temperature	934.59	535.02	1058.95	406.49
Cumulative minimum daily temperature	-547.59	465.42	-522.19	469.43
Cumulative average daily temperature	542.89	372.66	506.69	207.62
Cumulative daily day length	5512631	1044335	5000749	1083965
Cumulative daily precipitation	348.47	151.50	222.39	205.54
Cumulative daily solar radiation	43261.2	8504.76	40775.14	11737.74
Cumulative daily snow water equivalent	5321.42	6187.48	3106.81	7462.61
Cumulative daily water vapor pressure	69509.64	31508.84	52258.41	19946.21

The context-aware workflow

The proposed context-aware inconsistency check workflow builds upon elements from existing workflows. More precisely, it relies on the wide availability of contextual (environmental and geographic) information, enabling us to characterize complex differences between observation locations in space and time. When this characterization results in a high-dimensional dataset, the data are mapped to a low-dimensional space to facilitate the subsequent analysis of the data and the visualization of the results. Next, observations are clustered into contextually homogenous subsets. Finally, inconsistent observations are identified by analysing the outliers present in each cluster.

Dimensionality reduction

The t-distributed stochastic neighbour embedding (t-SNE) algorithm (Van der Maaten and Hinton, 2008) was selected to reduce the dimensionality of the contextual information. This algorithm maps the data to a low-dimensional space, typically two or three dimensions, so that data visualization is possible. It retains the local structure of the data which means that similar objects are mapped to nearby points in the low-dimensional space. Moreover, the model-based clustering step of the workflow has limited ability to deal with high-dimensional data, which further justify the use of the t-SNE algorithm.

The t-SNE defines a probability distribution over pairs of data points in the high-dimensional space so that similar ones have a high probability of being selected. Next, the t-SNE defines a similar distribution over the data points in the low-dimensional space in such a way it minimizes the information lost when such distribution is used to approximate the distribution in high-dimensional space. In particular, t-SNE uses the Kullback–Leibler divergence (Kullback and

Leibler, 1951) which quantifies the difference between the two probability distributions (in this case, those of the original and of the low dimensional data points).

The t-SNE algorithm requires the definition of the perplexity value, which is a smooth measure of the effective number of neighbours used to define the probability distribution in the high- and low-dimensional spaces. However, typical perplexity values are located in a limited interval (between 5 and 50) so optimizing its value is relatively easy. We used the "t-SNE" R package to perform all calculations in this study (Donaldson, 2010).

Model-based clustering

Model-based clustering (Banfield and Raftery, 1993; Fraley and Raftery, 2002) was selected to cluster the contextual information because it automatically identifies the number, shape and size of the clusters present in a dataset. This increases the objectivity of the analysis by reducing the need for human intervention and facilitates its use for multiple applications. The automated identification of cluster characteristics is realized by sequentially fitting several mixture models (Rasmussen, 1999) to the dataset and selecting the one that maximizes the Bayesian Information Criterion or BIC (Biernacki et al., 2000). We calculated the BIC values for ten Gaussian mixture models currently available in the R package, "mclust" (Fraley et al., 2012).

The uncertainty of the clustering was calculated (by subtracting the probability of the most likely group for each data point from one) and analysed to determine its impact on the identification of inconsistent observations. Data points with an uncertainty value of more than 0.5 were ignored as they could be either an inconsistent or a mis-clustered observation.

The model-based clustering method implemented in "mclust" uses the expectation maximization (EM) algorithm (Dempster et al., 1977). The EM, an iterative method, is used to find maximum likelihood parameters of a mixture model, specifying the mixture component to which each data point belongs. This algorithm is relatively robust but its efficiency is negatively affected by the dimensionality of the input data because the number of parameters that need to be estimated is proportional to the dimensionality of the data (Fraley et al., 2012).

Intra-cluster outlier detection

The identification of inconsistent observations requires defining objective and easily automatable rules. Here we used the Tukey boxplot as a main tool to highlight inconsistent observations (Frigge et al., 1989). The boxplot is a hybrid

non-parametric method that displays variation and outliers in numerical data by visually indicating its degree of dispersion and skewness in the data (Figure 2.1). The bottom and top of the box represent the first (Q1) and third (Q3) quartiles of the data respectively, and the band inside the box represents the second quartile (the median).
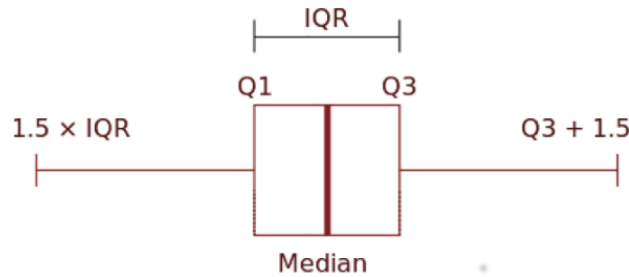
Figure 2.1 The Tukey boxplot.

In the Tukey boxplot the whiskers cover 150% of the interquartile range (i.e., 1.5 x IQR). If the numerical data are normally distributed, points larger or smaller than the values represented by the whiskers are 0.7% of the data and are typically considered outliers (Frigge et al., 1989). In this study, these outliers are highlighted as inconsistent observations. The outlier detection is also done using the built-in function of boxplot in the R software package to create an automated and clean workflow that can be re-used for multiple applications.

Impact of inconsistent observations

To investigate the impact of the inclusion of inconsistent observations in an analysis of phenological patterns, we used linear regression to model the trend in the flowering onset DOY–with and without inconsistent observations–over the complete study period. Regression models were developed for pooled observations of cloned and common lilacs, and separately for each type of lilac. Finally, we used analysis of covariance (Logan, 2010) to test the effect of the inconsistency of observations (i.e., consistent and inconsistent) on flowering onset DOY while controlling for the effect of the year of observations. This analysis is used to statistically test for differences in slopes among regression models. The regression modelling and the covariance analysis were done using built-in functions of the R software package.

2.3 Results and discussion

The eleven-dimensional data space that characterizes the phenological observation was transformed to a two-dimensional space (

Figure *2.2*) while testing several perplexity values (5 to 50 in steps of 5 units). The optimal perplexity value was chosen as the one that maximizes clustering (i.e., the one that better "spreads" and "separates" the observations into distinct groups). For both datasets, the perplexity value equalled 35, which led to the maximum number of clusters that the EM algorithm could identify.

A visual inspection of the transformed data space in
Figure *2.2* shows that the environmental conditions of the observation sites for cloned lilac are similar to each other, as the majority of points formed a cloud shape. It also shows that the observation sites for the common lilac are more clustered, indicating that these observations are made in more contrasting environments (Cayan et al., 2001) relative to the cloned lilacs (Schwartz, 1994). This is consistent with the fact that cloned lilacs were only observed in the Eastern U.S. (Frigge et al., 1989), which is characterized by less environmental variability than the Western U.S. (Table *2.1*).
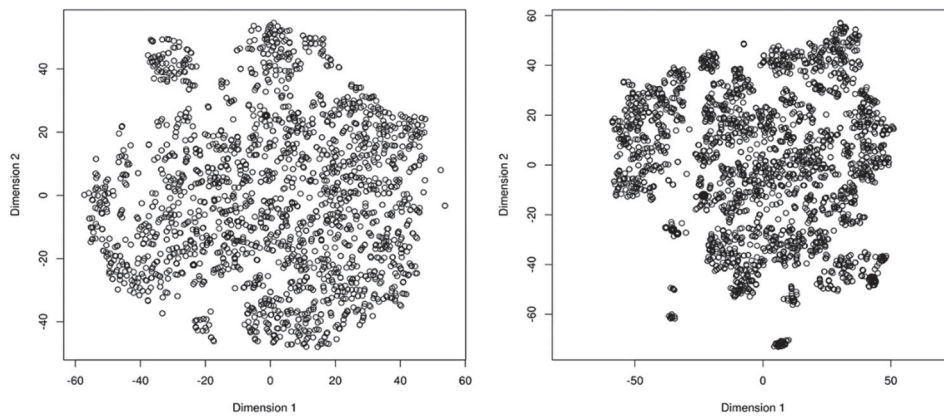
Figure 2.2 The results of applying t-SNE on contextual information.

As expected from the t-SNE results, the number of clusters for the common lilac (47 clusters) is larger than for the cloned lilac (12 clusters). These results (Figure *2.3*) demonstrate that a diagonal Gaussian mixture distribution—with equal shape, variable volume and coordinate axes orientation—fits best the contextual information for both cloned and common lilacs (Table 2.2).

21

Table 2.2 The fitted mixture models currently in the "mclust" package and their corresponding BIC values.

Distribution	Volume	Shape	Orientation	BIC of cloned lilac	BIC of common lilac
Spherical	Equal	Equal	-	-40096	-49152
Spherical	Variable	Equal	-	-39977	-48801
Diagonal	Equal	Equal	Coordinate axes	-40130	-49124
Diagonal	Variable	Equal	Coordinate axes	**-39905**	**-48768**
Diagonal	Equal	Variable	Coordinate axes	-40091	-49307
Diagonal	Variable	Variable	Coordinate axes	-40082	-48950
Ellipsoidal	Equal	Equal	Equal	-40119	-49172
Ellipsoidal	Equal	Equal	Variable	-40115	-49203
Ellipsoidal	Variable	Equal	Variable	-40130	-48960
Ellipsoidal	Variable	Variable	Variable	-40036	-49268
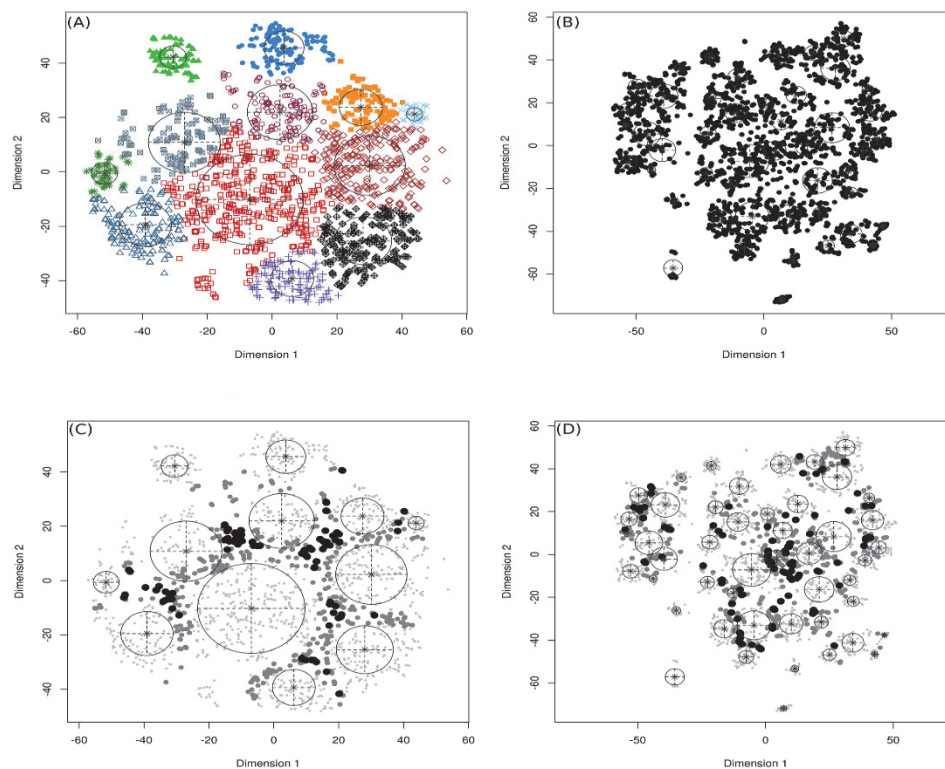
Figure 2.3 The results and uncertainty of model-based clustering. Clusters of the transformed contextual information about (A) cloned lilac and (B) common lilac. The uncertainty in clustering of transformed contextual information about (C) cloned lilac and (D) common lilac. In uncertainty plot, the symbols have the following meaning: large filled symbols, 95% quantile of uncertainty; smaller open symbols, 75–95% quantile; small dots, first three quartiles of uncertainty.

22

The phenological observations belonging to each cluster were projected into the geographic space to study their geographic distribution (Figure *2.4* and Figure *2.5*). For both types of lilac, the observation sites that belong to the same cluster are often spatially clustered (i.e., clusters tend to be compact). Nevertheless, there are some sparse clusters (e.g., cluster 7 and 10 of cloned and clusters 29, 31, 32, 36 and 40 of common lilac) that indicate geographically distant observation sites with similar climatic context.

Figure 2.4 The geographic distribution of the clusters in context condition of cloned lilac.

Figure 2.5 The geographic distribution of the clusters in context condition of common lilac.

The variability across the interquartile ranges and median values of the clusters for common lilacs is greater than for cloned lilac (Figure *2.6*). The greater variability in observations on common lilac reported from the Western U.S. was expected based on the clusters described above, and has been noted in other studies (Schwartz et al., 2000; Brunsdon et al., 2012). The outliers identified by the boxplots were highlighted as inconsistent phenological observations in this study.
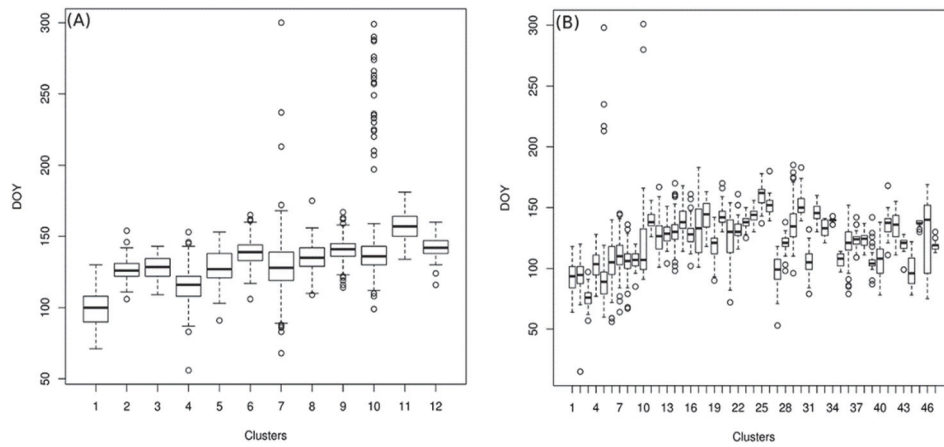
Figure 2.6 Intra-cluster boxplot of DOYs that lilac started flowering. Boxplots of corresponding DOYs in clusters of transformed contextual information for (A) cloned lilac and (B) common lilac. Hollow circles represent intra-cluster outliers.

Inconsistent observations were found in both pre- and post-2009 phenological observations (Figure *2.7*). For both types of lilacs, the highlighted inconsistencies accounted for about 3% of phenological observations (3.1% and 2.9% of phenological observations on cloned and common lilac respectively). 53% of the inconsistent observations on cloned lilacs have greater than one week uncertainty (>7 days between the prior "No" and the first "Yes" observation) whereas less than 15% of inconsistent observation on common lilac have greater than one week uncertainty in the estimated onset DOYs. Moreover, 41% of the inconsistent observations of cloned lilac and 50% of the common lilacs are associated with sites that report multiple flowering in a year (post 2009, when reports of repeat flowering were allowed, e.g., to account for flowering activity after frosts).
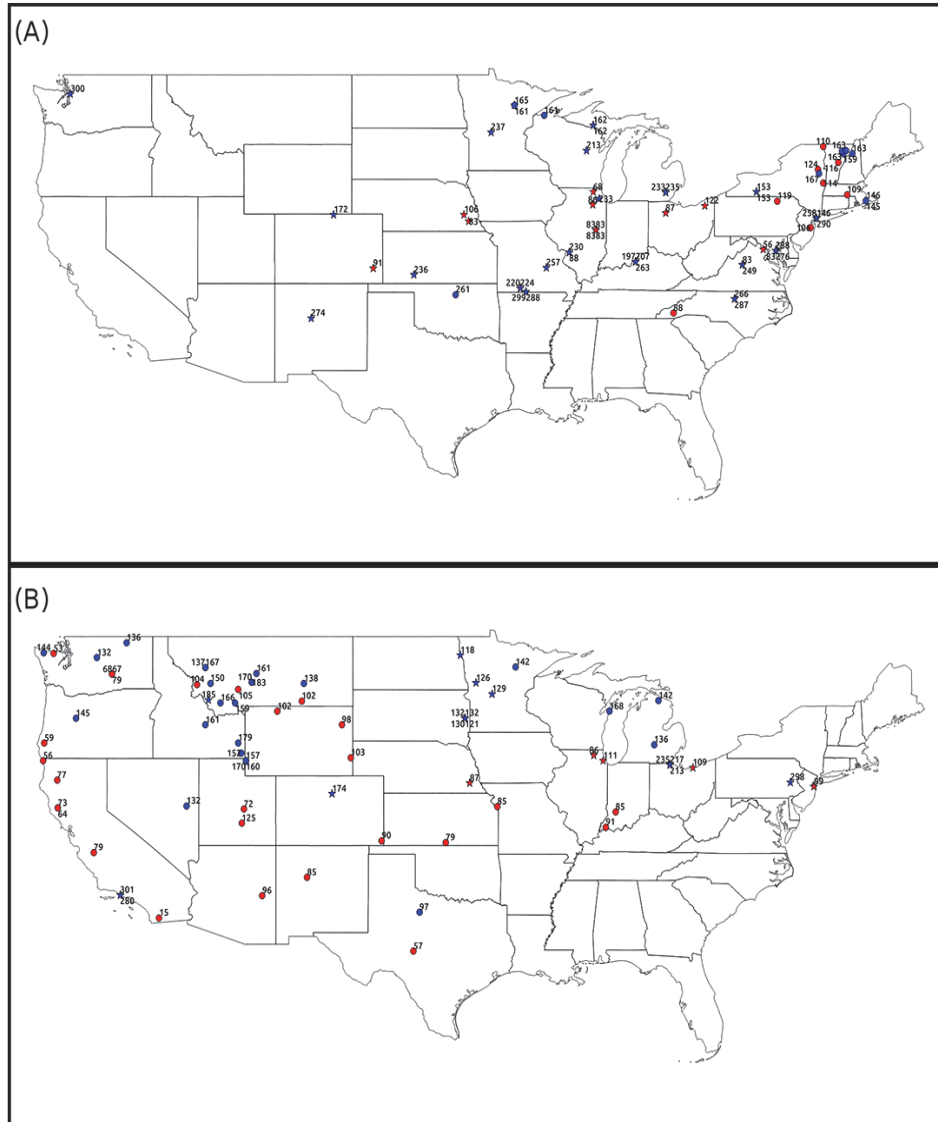
Figure 2.7 Plot of inconsistent phenological observations through study area. Inconsistent volunteered observations on flowering onset DOY of (A) cloned lilac and (B) common lilac. Red points show unusually early while blue ones show unusually late phenological observation. Circles show that phenological observations from historical initiatives whereas stars show phenological observations from contemporary initiatives. Inconsistencies were labelled with the day of year that lilac started flowering.

The unusually late "Yes" observations are not necessarily a result of erroneous data collection, because lilacs can also flower in the autumn (which may be associated with different environmental factors). In addition, unusually early "Yes" reports preceded by a second consistent "Yes" spring record might point

to mild winter in which lilacs start flowering early, experience frost, and then set flower again. For example, in 2012 in Charlottesville, Virginia, first flowering of a cloned lilac shrub was reported in February (i.e., early relative to other observations at the site). The flowering of the shrub was also reported later, on April 7th, which is more consistent, as determined by the workflow.

For cloned lilacs, the rate of change in flowering onset DOY (i.e., the slope of the regressions) significantly ($P < 0.001$) changed from -0.19 to -0.37 when inconsistent observations were excluded. In other words, using the cleaned dataset for the trend analysis resulted in two days additional advancement per decade in flowering onset of cloned lilac compared to the raw dataset. Likewise, for common lilacs, excluding inconsistent observations affected the regression slope, but to a lesser degree (from 0.12 to 0.9; $P = 0.06$) than in the cloned lilacs. For the pooled observations, the slope changed from -0.02 to -0.12 ($P < 0.001$) when the inconsistent observations were removed, resulting in one additional day advancement per decade in flowering onset across the US. Thus, the inclusion of inconsistent observation underestimates the rate of acceleration of the lilac onset dates over the period 1980–2013 (Figure *2.8*). These results are in agreement with previous studies that found a gradual advance in the flowering onset DOYs (Brunsdon et al., 2012; Ault et al., 2013).
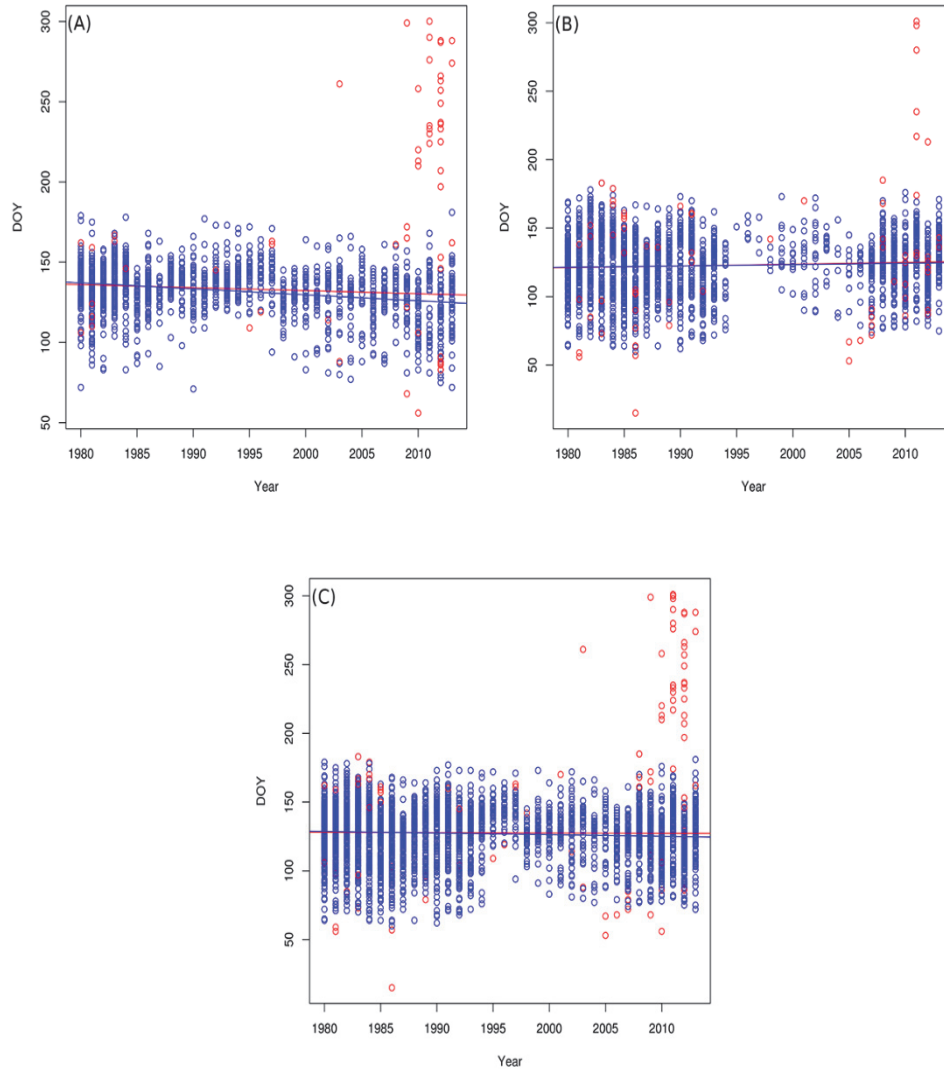
Figure 2.8 Comparison of the linear modelling of the original phenological observations and the consistent phenological observations. Temporal trends in the flowering onset DOY of (A) cloned lilac, (B) common lilac, and (C) pooled observations of cloned and common lilac.

2.4   Conclusions

The identification of inconsistent observations is a pre-requisite for any kind of analysis or modelling effort. In this paper, using a phenology case study, we present and demonstrate a computational workflow that has potential to automate the identification of inconsistencies in data collected by VGI-based initiatives. The workflow relies on environmental data as critical context that

affects the variability in the observational datasets, and consists of a sequence of dimensionality reduction, model-based clustering and outlier detection.

The workflow demonstrated that we can highlight unusually early or late observations of the flowering onset DOYs for lilacs. The identified inconsistencies should be further analysed using more granular climate data or expert knowledge to determine if they are likely observation or transcription errors or represent truly anomalous events, due to microclimate, or genetic variation, in the case common lilacs. Overall low inconsistency rate (about 3%) indicates that volunteer collected observations are a valuable source of information for the study of phenology.

Phenological VGI has greatly contributed to our understanding of seasonal spatial and temporal patterns for plants and animals across the globe. Given that phenology has been recognized as an important indicator of climate change and has emerged as a vibrant area of research at multiple ecological scales, analyses that increase data quality and usability will greatly benefit the fields of climate research, ecology, and natural resource management. We envision that this workflow will greatly increase the reliability of, and potential for scientific contribution from, spatially and temporally rich VGI datasets.

Focusing subsequent analysis on the inconsistent observations identified by our workflow reduces human checks, which saves money and time. Moreover, unlike existing workflows, the proposed workflow uses relevant contextual information for the phenomena under study (as climate drives phenological events). Therefore, we recommend that initiatives collecting volunteered geographic information use the proposed automated workflow and relevant contextual information to check inconsistency in order to improve data quality. This workflow could be applied to volunteered meteorological data (Council, 1998) to, for instance, highlight unusually high or low temperature reports because daily weather data has a long history and is increasingly available (Menne et al., 2012).

Chapter 3 Checking the consistency of volunteered phenological observations while analysing their synchrony*

3.1 Introduction

Because of the progress in information, communication and mobile location-aware technologies, the use of volunteered geographic information (VGI) has dramatically increased in recent times (Beaubien et al., 2011; Ferster et al., 2013). This geographic information is very useful since it permits volunteers (non-experts) to act as human sensors, which contribute, for instance, to environmental research. However, the quality of the information provided by volunteers remains a concern (Ballatore and Zipf, 2015; Senaratne et al., 2017). Inconsistent VGI, which conflict associated contextual conditions, is one of the critical quality issues (Yanenko et al., 2012; Comber et al., 2013). The lack of consistency can affect the results of environmental studies such as phenology (Mehdipoor et al., 2015; Rosemartin et al., 2015). Phenology is the study of periodic plant and animal life cycle events (phenophase) and how seasonal and inter-annual variations in weather conditions affect them (van Vliet et al., 2003; Mayer, 2010; Chmielewski, 2013). Phenology benefits from VGI developments to acquire observations that support climate change studies at various scales (Doi and Katano, 2008; Gordo, 2010; Zurita-Milla et al., 2017).

National and regional phenological networks promote phenological research and curate ever-increasing collections of Volunteered Phenological Observations (VPOs) collected by large crowds of volunteers (Beaubien et al., 2011; Ferster et al., 2013). VPOs are available at fine spatial scales and provide timely and low-cost information (Devictor et al., 2010; Rosemartin et al., 2015; Ault et al., 2015). Hence, VPOs support novel phenological studies that study the impact of climate change on plants and animals (Soroye et al., 2018). Although VPOs support climate change studies, the consistency of the observations provided by volunteers remains a concern. Phenological studies are sensitive to anomalously early or late VPOs regarding their associated environmental conditions (Sparks et al., 2008; Mihorski et al., 2012). These so-called inconsistent VPOs can bias the results of trend analysis and modelling because they are not representative and decrease the ratio of signal to noise (Mehdipoor et al., 2015).

Inconsistent VPOs are expected because volunteers have different levels of expertise in recognizing specific phenophase or, the target species (Brunsdon et al., 2012). Moreover, volunteers may also perform observations at locations with environmental conditions that are not representative of the phenophases being monitored (e.g., they might report data for an individual plant growing under a special micro-climate). Consistency checks for VPOs often rely on applying a threshold to identify outliers.

For example, the commonly used Tukey boxplot (Frigge et al., 1989) uses 1.5 times the absolute value of the difference between the first and third quartiles of the annual DOYs to highlight outliers. These checks assume that most of the observations are consistent and can be used to estimate the distribution of the reported dates, and to recognize outliers. However, this assumption is not always true when working with volunteered observations. Contextual environmental information (e.g., temperature) play a key role in many phenological studies. Yet, this information is not used in consistency checking (Hochachka and Fink, 2012). This contextual information has the potential to check the consistency of VPOs by integrating assumptions about the effect of context on VPOs.

The analysis of the temporal dispersion of a phenophase across individuals of the same species, or phenological synchrony, could provide information about how context derives the timing of phenophase. Changes in phenological synchrony have ecological consequences for individual survival and ecosystem stability (Ims, 1990; English-Loeb, 1992; Bolmgren and Eriksson, 2015). For example, low synchrony in plant flowering can hamper the expected random mating pattern because early bloomers will likely get pollinated by early plants, and late plants by late plants (Weis et al., 2004). In regions with a marked seasonality, synchrony is strongly controlled by temperature variability (Both et al., 2009; Gordo et al., 2010). Phenological synchrony is typically quantified by the standard deviation of Day of Year (DOY) of all the observations collected in a given area and year (Henderson et al., 2000; Menzel et al., 2006; Gordo et al., 2010; C. Wang et al., 2016). This quantification method is sensitive to inconsistent VPOs (Sparks et al., 2008; Mihorski et al., 2012).

This study checks the consistency of VPOs while taking the effect of inconsistent observation on phenological synchrony into account. In the next section, we describe a geocomputational workflow that uses the geographic location and DOY of VPOs, and the associated daily temperature at the locations where the observations were made. Here we tested the workflow using VPOs from the Dutch phenological network and daily temperature time-series from the Dutch national weather service. The added value of our workflow is evaluated by comparing it with the results of a more classical outliers identification method, namely the Tukey boxplot.

3.2 Materials and methods

VPOs and temperature datasets

This study is illustrated with VPOs from the Dutch phenological network Natuurkalender[1] (Nature's Calendar). This volunteer-based network was established in 2001 to monitor phenophases for a wide range of species in the Netherlands. Considering the spatio-temporal coverage of the observations, spring phenophases of four plants species were selected for this study: flowering of lesser celandine (*Ficaria verna Huds*), wood anemone (*Anemone nemorosa L.*), and cow parsley (*Anthriscus sylvestris (L.) Hoffm*), and leafing of pedunculate oak (*Quercus robur L.*).

Lesser celandine and wood anemone are herbs that grow in forests, grasslands, and next to roads and waterways. Both herbs are intensively studied in Europe as an indicator of global environmental changes on the distribution and abundance of forest understorey (Baeten, 2010). Lesser celandine starts to flower early in March and wood anemone flowers around mid-March. Cow parsley is also a herb, which grows in rich, moist soils along roadsides and forest paths. It flowers from late March onwards. Pedunculate oak is a tree, often planted on sandy soils, which starts leafing around late April. Cow parsley and pedunculate oak have been found to be indicators of climate change by studies performed in Europe. For the period 2003-2015, the Natuurkalender database has a total of 3,042, 1,298 and 1,811 flowering observations for lesser celandine, wood anemone and cow parsley, and 586 leafing observations for pedunculate oak. These observations only contain the geographic location provided by a volunteer (in the Dutch National Coordinate System; EPSG: 28992) and the date that a phenophase was first observed in a given year.

In the Netherlands and other temperate regions, temperature is the main driver of plant phenophases (Jolly et al., 2005; De Frenne et al., 2009). Hence, the annual temperature regimes of each observation were characterized using 1-km gridded estimates of daily average temperature, as downloaded from the Royal Netherlands Meteorological Institute (KNMI[2]). These grids, created by interpolating daily temperature records were collected by 150 meteorological stations distributed across the country, and were used to calculate Growing Degree Days (GDDs) (McMaster and Wilhelm, 1997) by adding up the daily average temperature above zero degrees Celsius from the first of January of the year of the observation until the observed date of each VPO. According to Lappalainen and Heikinheimo (1994), GDDs calculated from 1 January, using 0 or 5 °C as the base temperature, provide a valid tool for predicting the timing

---

[1]   www.natuurkalender.nl
[2]   https://data.knmi.nl/datasets

of leafing and flowering and their regional variation for several species in Western Europe. This has been confirmed by other spring phenology studies (Zavalloni et al., 2006; Cook et al., 2012).

Analysing consistency and synchrony of VPOs

The proposed workflow consists of four steps (Figure 3.1) and uses the geographic location of the VPOs, their observed DOY and the GDDs accumulated until that DOY. The first three steps are used to check annual VPOs consistency. These steps connect, model and compare VPOs that are near to each other to identify sets of inconsistent observation. The fourth step focuses on modelling synchrony and optimizing the identification of inconsistent VPOs. The next paragraphs describe each of the steps in more detail.
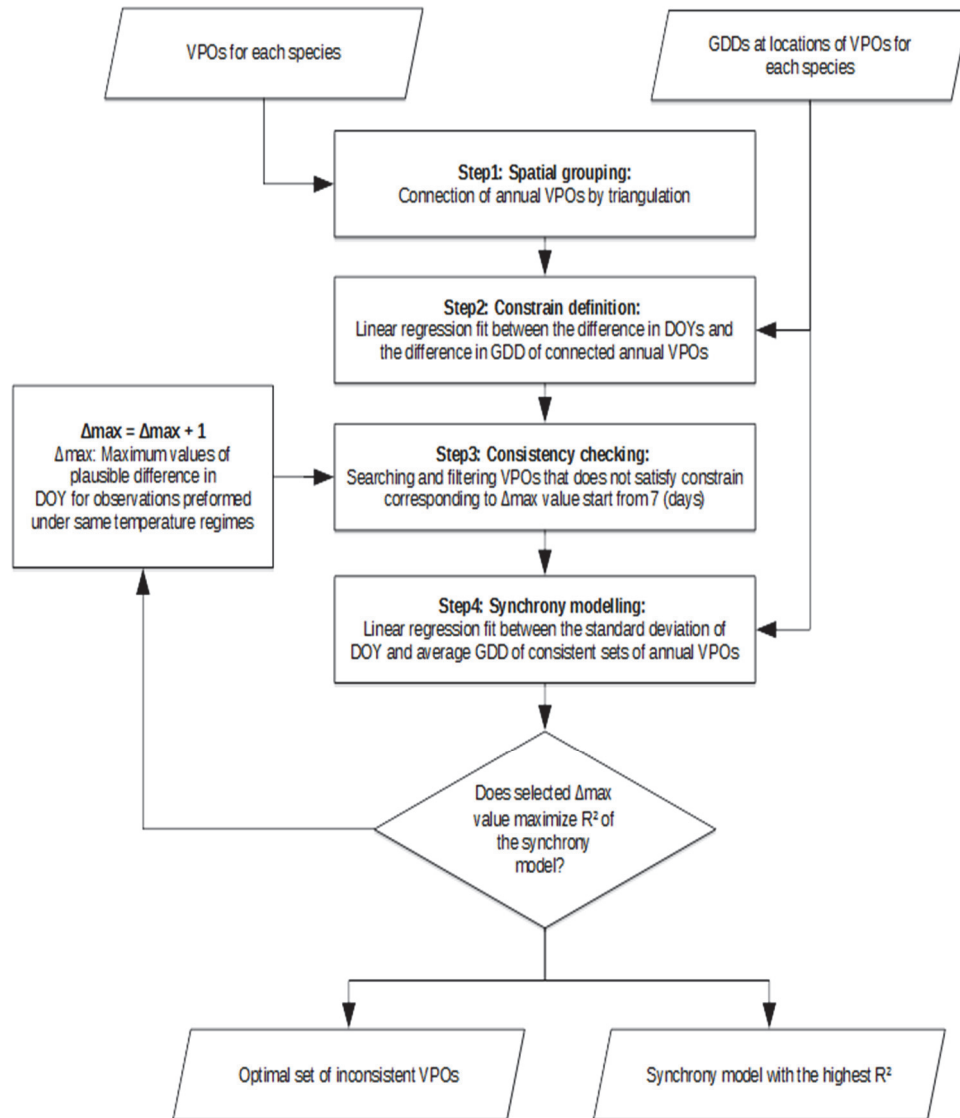
Figure 3.1 Flowchart of the proposed workflow that uses the coordinates of the volunteered phenological observations, their observed date and the GDDs accumulated until that date to check their consistency and to model phenological synchrony.

In the first step, the geographic locations of the annual VPOs were used to spatially group observations by connecting observations that are near to each other. This creates a "network graph". This approach was chosen because near observations locations are more likely to have similar weather and observations that are further apart. The spatial variability of meteorological parameters, other than temperature, is small across the Netherlands because of the relatively homogeneous topography (maximum difference in elevation of 300m), and a gradual and smooth steepness (van Vliet et al., 2014). These annual graphs were made using a Delaunay triangulation and edges longer than a given threshold distance were pruned (Chew, 1989). This triangulation method was chosen because it is computationally efficient and avoids long edges (Yanenko et al., 2012; Shi et al., 2016). The pruning distance was set to 100-km because this value ensured good connectivity (each observation was connected to at least two other observations).

In the second step, the difference in the observed DOY of connected observations was modelled using their corresponding difference in GDDs. For each year, the Pearson product-moment correlation coefficient was calculated and a linear regression model was fitted (De Frenne et al., 2009; Gordo et al., 2010). The slopes of the regression lines (which represent the rate of spatial change in DOY per unit of GDDs) are used to define consistency constraints. More precisely, constraints were defined by fixing a maximum difference in DOY (ΔMax) for observations performed under similar environmental conditions. In practical terms, this indicates that for a given difference in GDDs, the difference in DOY should not exceed ΔMax, otherwise, the two reported DOYs are inconsistent to each other.

In the third step, the consistency constraint was checked by using a range of ΔMax values. Since ΔMax cannot be known a priori, the chosen range varies from one week to one month in steps of one day. The minimum ΔMax was set to one week to take genetic variation into account because two individuals of the same species that grow at the same location (i.e., under the same temperature regimes) may not start flowering (or leafing) at the same DOY. For each species and year, the inconsistent observations corresponding to each ΔMax value (i.e., VPOs refuted by more than one connected VPO) were highlighted and excluded. This resulted in 24 sets of consistent VPOs for each species. The optimal ΔMax for each species, and consequently the final set of inconsistent VPOs, was selected after performing the phenological synchrony analysis that takes place in the last step (step 4 – Figure *3.1*) of the workflow.

In the fourth step, the synchrony of consistent VPOs was analysed and modelled using the inter-annual variability in temperature as an explanatory variable. Wang et al. (2016) and Thackeray et al. (2016) have proved that

warmer days in late winter and early spring lead to less synchronous flowering and leafing onsets than colder transitions. They showed that the annual standard deviations of flowering and leafing onset are significantly higher in years with a lower rate of change of temperature from winter to spring. In such years, the number of days prior to flowering and leafing onset that have high temperature increases. This results in a larger average GDD than in years with a higher rate of temperature change. For each set of consistent VPOs, the Pearson product-moment correlation coefficient between the annual standard deviation of the DOYs and the annual average of their GDD values was calculated. For each species, the optimal ΔMax corresponds to the value that maximizes the correlation. A linear regression to the annual standard deviation of the DOYs and average GDD of the final set of consistent VPOs was also fitted to model synchrony.

The fitness of this model was evaluated by its coefficient of determination (R-squared). The R-squared values of the synchrony model driven with consistent VPOs were compared with the R-squared values produced by the models fitted using: 1) the original and 2) the outlier-free (boxplot outliers excluded) VPOs. This comparison helps to understand the added value of the proposed workflow. The percentage of annual inconsistent and outlier VPOs was also calculated to compare the amount of data lost through the cleaning of observation by the proposed workflow and by the outlier-filtering method (based on the boxplot).

3.3 Results

The pruned Delaunay triangulation of VPO locations results in annual graphs of VPOs locations. For lesser celandine (Figure *3.2*), wood anemone (Appendix A.1) and cow parsley (Appendix A.2), these graphs cover the whole of the Netherlands. However, the VPOs of pedunculate oak (Appendix A.3) are mostly located in the centre of the country. For all species, a large dispersion of VPO locations from year to year was observed, which is intrinsic to volunteer monitoring network data. For example, the spatial variability found for lesser celandine (Figure *3.2*) shows that volunteers observed this species almost everywhere in the Netherlands however, in some years (e.g., 2008) observations are more clustered in the western part of the country.
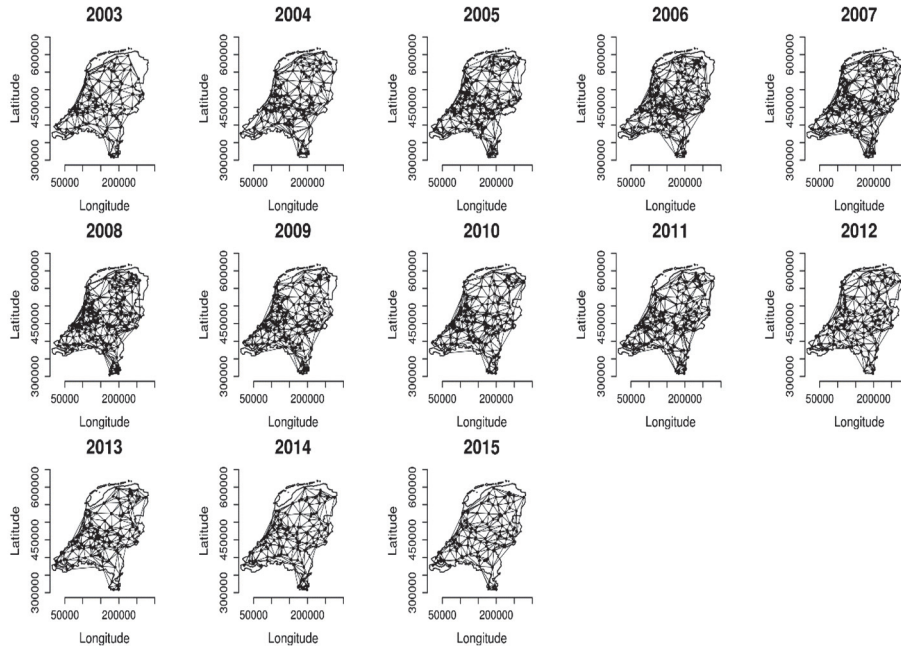
Figure 3.2 Annual graphs for lesser celandine flowering onset observations. The graphs were made using a Delaunay triangulation and edges longer than 100-km were pruned.

The correlation between the difference in DOYs and the difference in GDDs was significant for all species. The average correlation coefficient was 0.91 for lesser celandine (Figure *3.3*), 0.9 for wood anemone (Appendix A.4), 0.93 for cow parsley (Appendix A.5) and 0.9 for pedunculate oak (Appendix A.6). This indicates the considerable influence of the accumulated daily temperature on the timing of flowering. The slopes of the fitted regression lines show the rate of change of the difference in reported DOYs per unit of difference in GDD (i.e., the steeper the slope, the larger the difference in the timing of the phenophase). The comparison of the annually modelled and reported difference in DOYs is used to identify 24 sets of inconsistent observations.
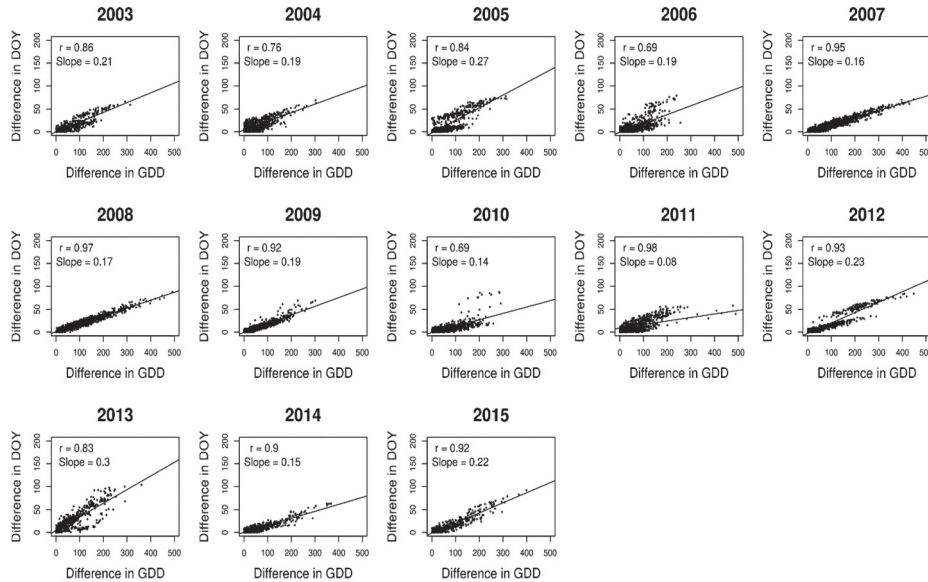
Figure 3.3 Linear regression lines fitted between the difference in observed flowering DOY (number of days) and the difference in modelled GDD (number of degree days) for pairs of spatially connected observations of lesser celandine flowering onset (see Figure 3.2). Correlation coefficient and rate of spatial change in DOY per unit of GDDs (i.e., slope of regression line) are given for each observation year.

The correlation coefficient between the annual standard deviation of consistent DOY corresponding to each ΔMax value and the annual average GDD were calculated for lesser celandine (Figure *3.4*), wood anemone, cow parsley and pedunculate oak (Appendix A.7). The ΔMax that lead to the phenological synchrony model with the largest coefficient of determination was 13 days for lesser celandine, 20 days for wood anemone, seven days for cow parsley and 15 days for pedunculate oak. For example, the optimal ΔMax value indicated that the maximum difference in the timing of flowering of two lesser celandine plants (located less than 100-km apart and growing under similar temperature regimes) was 13 days. The small value of ΔMax for cow parsley flowering, a plant with a shallow root system, might be caused by the impact of other environmental parameters such as soil moisture and light intensity (Jie et al., 2004).
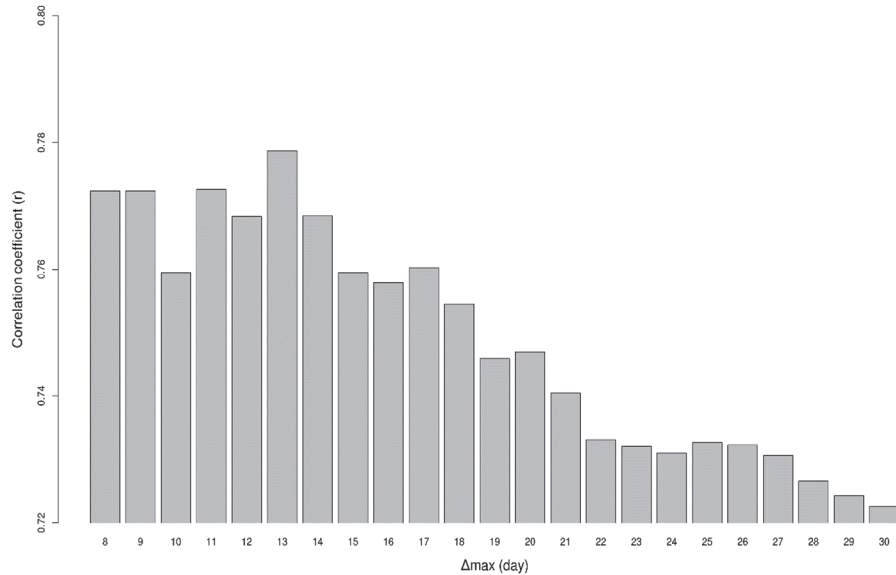
Figure 3.4 Correlation coefficient (r) between the standard deviation of the reported DOYs for the flowering onset of lesser celandine and the average GDD for ΔMax values varying from one week to 30 days.

For all species, optimal inconsistent VPOs show a large difference in the reported DOY compared to their surrounding VPOs (Figure *3.5*). For all species, except cow parsley, the annual percentages of VPOs highlighted as possible inconsistent observations are smaller than the annual percentages of boxplot outliers (Figure *3.6* and Appendix A.8). Inconsistent VPOs refer to unusually early or late DOYs with respect to the regional temperature regime of the observation site whereas boxplot outliers only identify very early or late observations for a given set of annual observations and, in consequence, do not consider the effect of regional contextual information. The annual boxplots for lesser celandine (Figure *3.7*d), wood anemone (Appendix A.9), cow parsley (Appendix A.10) and pedunculate oak (Appendix A.11) show that outliers are mostly located below the lower whisker of the boxplot, meaning that the distribution of observed DOYs is not normally (Gaussian) distributed as the boxplot assumes.
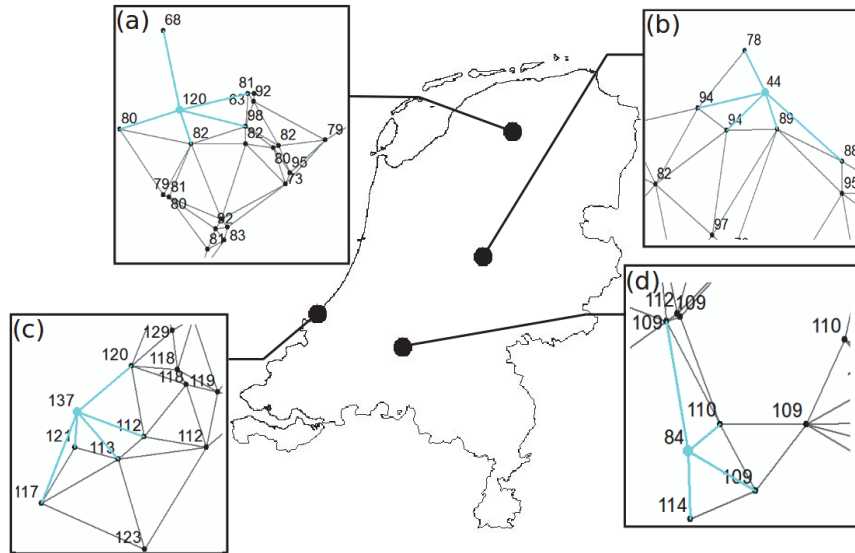
Figure 3.5 Examples of inconsistent observations for the flowering onset of lesser celandine. The highlighted observation is refuted by more than one other observation. Given the difference in GDDs between the highlighted observation and those connected to it, their difference in DOY exceeds the predicted difference by about two weeks. The numbers next to each observation show the reported DOYs.
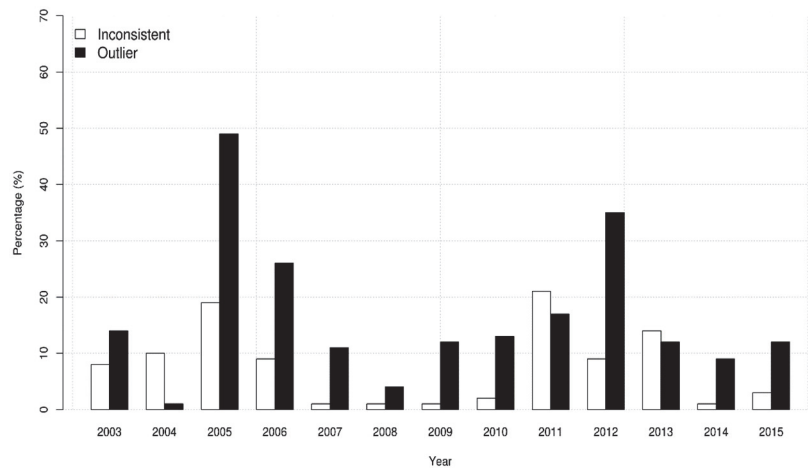
Figure 3.6 Annual percentages of inconsistent observations and of boxplot outliers in volunteered observations of lesser celandine flowering onset. Inconsistent observations are unusually early or late DOYs with respect to the regional temperature regime of the observation sites, while the outliers are only very early or late DOY.

The synchrony analysis resulted in models that predict the standard deviation of the DOY of the selected phenophases using the annual average GDD. For consistent VPOs, the correlation coefficient between the standard deviation of the DOY and the annual average GDD was 0.78 for lesser celandine, 0.63 for wood anemone, 0.61 for cow parsley and 0.6 for pedunculate oak. These results suggest that the timing of flowering and leafing onsets for the species under study is more synchronous in cold late winters and early springs than in warm ones. The comparison between the synchrony models, made from original, outlier-free and consistent VPOs, shows that using boxplots negatively impacts the quality of the model (Figure *3.7* and Appendix A.9, A.10 and A.11). For all species, the R-squared value of the models based on consistent VPOs is larger than that of the models based on outlier-free data (Table *3.1*). Moreover, the models based on consistent VPOs are more in line with the models made using original VPOs. Removing a large number of outliers by using the Tukey boxplot method leads to strongly distorted models at large geographical scales (Figure *3.7*b). The application of our workflow improves the quality of the model (notice the improved R-squared values in Figure *3.7*c).

Table 3.1 The coefficient of determination (R-squared) of phenological synchrony models driven from original, outlier-free and consistent observations.

	Original	Outlier-free	Consistent
Lesser celandine	0.54	0.00	0.61
Wood anemone	0.43	0.28	0.40
Cow parsley	0.12	0.07	0.37
Pedunculate oak	0.35	0.16	0.36
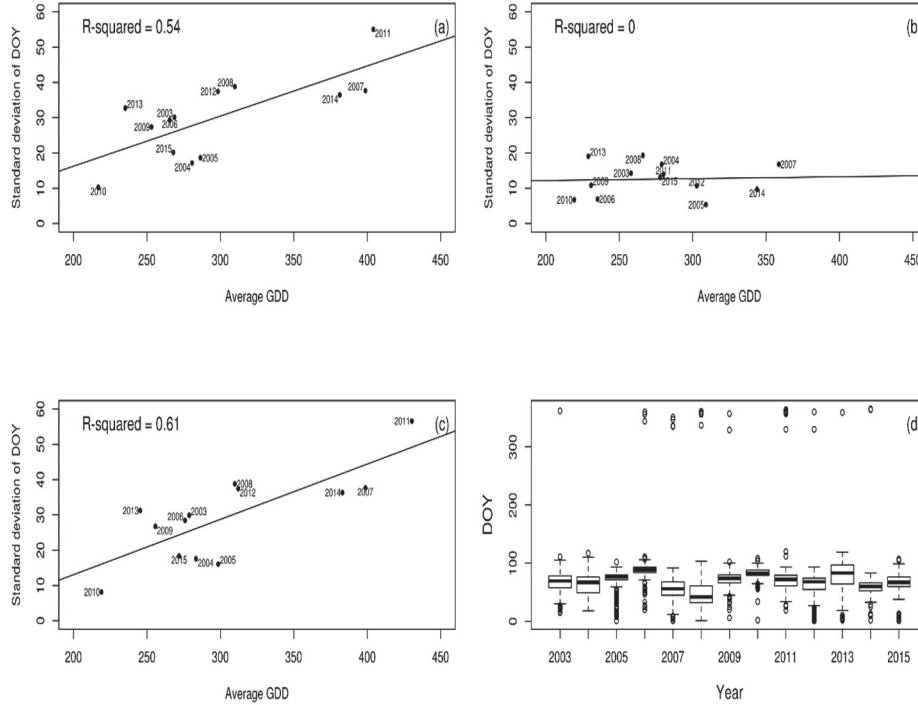
Figure 3.7 Lesser celandine flowering onset synchrony models for (a) original, (b) outlier-free and (c) consistent observations. Panel (d) shows annual boxplots of the reported DOYs for the original observations.

3.4 Discussion

This study presents a workflow to check the consistency of VPOs. Unlike purely statistical methods, our workflow uses the geographic location and the corresponding accumulation of daily temperature as independent sources of information. The workflow defines and evaluates consistency constraints based on the correlation between VPOs synchrony and the rate of change of temperature from winter to spring. We used the workflow to filter out phenological observations that do not provide regionally representative species-specific flowering and leafing DOYs and labels them as "inconsistent".

Inconsistent VPOs can be caused by either species and/or phenophase misidentification or they can be true observations influenced by a microscale temperature regime that is hard to model using 1 by 1-km temperature data. In either case, inconsistent observations are not representative of the phenology of this species in the Netherlands. The high correlation found between the difference in DOYs and the difference in GDD of near observations on flowering and leafing onsets indicates that daily temperature is indeed

relevant for the analysis of the selected species and phenophases. This is in line with the fact that daily temperature is a dominant factor for plant phenology in temperate and boreal regions (Shen et al., 2015). This highlights the importance of storing metadata about volunteered observations to improve the temporal consistency in phenological databases.

Considering ΔMax as a proxy for the spatial variability of the timing of a phenophase under similar temperature conditions helps to constrain the temporal window in which the occurrence of a VPO is consistent. As ΔMax takes into account the geographic context, a quality control mechanism of VPOs based on this metric outperforms alternative methods solemnly based on data distributions. Given the increasing popularity of citizen science networks, we expect to get more accurate estimates of ΔMax in the near future. The ΔMax metric can also help to understand the phenology of species. For instance, the relatively small value of ΔMax for cow parsley (seven days) indicates that the flowering onset of this species is controlled more strongly by temperature as opposed to the other selected species or phenophases. In consequence, this species shows the highest correlation between the difference in DOYs and the difference in GDD of VPOs.

In this study, we assume that GDD accumulations drive the synchrony of the selected phenophases within 100-km distance. In the Netherlands, the level of spatial connectivity of VPOs using this distance threshold is high, however, this might not be the case when analysing larger areas. In such areas, the analysis might be hampered by annual graphs with a low level of spatial connectivity. Other temperature-driven metrics than GDD such as the average temperature, could be tested with our workflow. For example, Calinger et al. (2013) showed the suitability of average monthly temperatures during the month of the phenophase and some number of months prior to the event to model phenological responses to temperature across many species. Moreover, GDD accumulations may not be the only or main driver of flowering and leafing onset in other study areas. The length of the chilling period (Sogaard et al., 2008; Laube et al., 2014), photoperiod (Laube et al., 2014; Ault et al., 2015), and precipitation and elevation (Penuelas et al., 2002) might also drive flowering onset in spring. For example, Studer et al. (2005) used a multivariate regression to model the timing of wood anemone flowering as a function of temperature and precipitation. A similar model could be used during the consistency check phase of our workflow because it is generic enough to accommodate other phenological drivers.

Our workflow works for events that are synchronized, yet, this is the case for several ecological phenomena. In citizen science, several networks are monitoring environmental events that are weather-driven. Examples of such types of monitoring are the reporting of tick bites and mosquito's and migrating

bird observation. For these types of phenomena, different types of weather data can be used to find inconsistent observations. The developed workflow can also be useful in these domains.

3.5   Conclusions

VGI has greatly contributed to phenological studies, leading to an improved understanding of plant and animal seasonality across the globe. In this respect, checking the consistency of volunteered phenological observations or VPO is a pre-requisite to ensure the validity and representativeness of VPO-based results. In this paper, we present a workflow designed to use geographical and contextual information associated with phenological observations to check the consistency of observations while analysing their synchrony. This workflow was used to improve our knowledge on the local impact of inter-annual temperature variations on the consistency and synchrony of VPOs from various plant species in the Netherlands.

Our results reveal that the most common method (boxplot) to filter outliers in VPOs substantially bias synchrony analysis of the timing of the spring flowering and leafing. Our results indicate that climate change and inter-annual weather variability determine changes in the synchrony of spring plant phenology. Given that several national and international initiatives facilitate and actively support the collection of VGI for ecological studies and that the open data movement is resulting in more contextual environmental information becoming available, the proposed workflow provides a unique opportunity to check the consistency of volunteered observations. Considering its general character, we think that this geocomputational workflow could be adapted to other kinds of VGI, hence contributing to the curation of this interesting source of geospatial data.

Chapter 4 Exploring differences in spatial patterns and temporal trends of phenological models at continental scale using gridded temperature time-series*

4.1   Introduction

Climate change is influencing the timing of key biological events. For example, warm springs are advancing the time of leaf onset of plants (Ellwood et al., 2013; Schwartz et al., 2013) and the migration of animals (Marra et al., 2005; Ault et al., 2011). Monitoring and analysing the timing of plants and animal development events is therefore essential to better understand the system Earth and define climate change adaptation strategies (Briske et al., 2015; Gerst et al., 2016; Labe et al., 2017). In this respect, phenology is one of the most reliable proxies to study climate change (Doi et al., 2008; Gordo et al., 2010; Zurita-Milla et al., 2017). Phenology is the science that studies periodic plant and animal life cycle events (phenophases) and how seasonal and inter-annual variations in environmental conditions affect them (Lieth, 1974; Chmielewski, 2013).

Phenological models are ideal tools for studying plants responses to both climate change and inter-annual weather variability (Badeck et al., 2004; Schwartz et al., 2006; Allstadt et al., 2015). Phenological models can be used to reconstruct and qualify ground observations (Chuine et al., 2004; Menzel, 2005), to estimate species-specific phenology (Chuine et al., 2000), and species performance (Basler, 2016). Phenological models are often calibrated using ground and weather observations. Spring plant phenology (SPP) models are particularly interesting because climate change effects are more visible in this season (Bonsal, 2001; Robeson, 2004; Polgar and Primack, 2011). SPP models are widely used to support management decisions and to design adaptation strategies for ecological and agricultural systems (Enquist et al., 2014; Gerst et al., 2016). These uses require phenological models of high quality.

The process of quality control of models starts with their calibration and continues with their validation. Calibration is used to find the values of the model parameters that minimize the error of the model. These values are often derived using specific phenological and environmental datasets. This makes the comparison of calibrated models challenging. Validation is used to check the error of the calibrated model. The calibration and validation of phenological models over large areas are now possible because we have access to continental-scale gridded weather time series such as daily temperature and large amounts of contemporary volunteered phenological observations (VPOs; (Rosemartin et al., 2015)). Gridded temperature time-series are key inputs to SPP models; they help to generate spatially-continuous data from which patterns and temporal trends can be extracted to evaluate the influence of climate change on plant development (Mehdipoor et al., 2018; Izquierdo-Verdiguier et al., 2018).

SPP models using various parameters range from simple accumulation of degree days to advance counting of high-energy synoptic events. Differences in the structure of SPP models and fine tuning of parameters of SPP models may lead to various estimates of DOY for the same species at the same location. Different SPP model formulations can be found in the literature (Hufkens et al., 2018). Exploring the effects of using varying phenological models improves our understanding of the outputs from SPP models, and consequently the management decisions based on the outputs. However, few studies have comprehensively explored the effects of using one or another model on phenological patterns and trends (Abatzoglou, 2013). This could be because the computational power required for such an exploration was not available until recently. The enhancements in large-scale distributed computing and cloud computing facilitate qualification of SPP models using higher spatial resolution, gridded input data (Guo et al., 2010; Mehdipoor et al., 2017).

The quality of phenological models is often assessed using the root-mean-square error (RMSE) between the estimated and observed day of the year (DOY) of a phenophase at a number of observations sites. The RMSE is an average model performance measure and several studies have shown that it overestimates average model error (Dhar et al., 2017). Although the RMSE is based on the values of individual observations, it expresses the quality of the model as a whole and local variations are therefore not detectable. Yet, several studies have shown that phenological patterns and trends derived from various kinds models may vary considerably (Janssen et al., 1995; García de Cortázar-Atauri et al., 2009; Basler, 2016; Chuine et al., 2016).

This study evaluates various SPP models and explores the impact of using one or another model on the phenological patterns and trend that can be extracted by running these models at continental scales. The workflow uses VPOs and gridded time-series temperature over the conterminous US (CONUS). We illustrate the workflow exploring the effect of using the Extended Spring Indices (SI-x; (Schwartz et al., 2013)), Thermal Time (TT; (Cannell and Smith, 1983)) and Photothermal- Time (PTT; (Masle et al., 1989; Črepinšek et al., 2006)) groups of models on the estimation of patterns and trends in DOY of lilac leafing.

4.2   Materials and methods

Volunteered observations and temperature data

Two groups of data were used to calibrate and validate the SPP models. For calibration, we collected historical lilac (*Syringa Chinensis 'Red Rothomagensis'*) VPOs and their corresponding temperature data from Prof.

Schwartz who developed the Extended Spring Indices (SI-xLM) (Schwartz, 1997; Ault et al., 2015). These VPOs contain the geographic location and DOY of first leaf (FL). Lilac FL phenophase provide a standard reference of spring plant phenology that can be compared to different location and years (Caprio, 1974; Santer, 1985). This phenophase responds directly to changes in temperature and day-length as opposed to changes in other environmental cues (Capiro, 1993; Schwartz et al., 2012). The VPOs dataset has a total of 2321 observations collected across 193 sites (known as phenological stations) across the continental US over the period 1961-1994. The corresponding daily maximum and minimum temperatures, and a day-length dataset include records from the nearest weather stations to the VPO sites. These weather stations are part of the Global Historical Climatology Network (GHCN).

To validate the selected SPP models and to explore their patterns and trends, we used VPOs from the USA National Phenology Network (USA-NPN) and Daymet gridded temperature time-series, from 2000 to 2014. The USA-NPN dataset contains 899 lilac FL observations that were checked for consistency (Mehdipoor et al., 2015; Rosemartin et al., 2015). The gridded daily maximum and minimum temperature and day-length from Daymet are available at 1-km for most of North America since 1980 (Daly et al., 2008; Thornton et al., 2014). Daymet uses spatially-referenced surface measurements of daily maximum and minimum temperature and precipitation from the Global Historical Climatology Network (GHCN) as the major input (Daly et al., 2008; Thornton et al., 2014). We used Daymet daily temperatures to generate and map spatial and temporal trends of the calibrated models.

SPP models used

We selected SPP models based on temperature and photoperiod because plant phenophases often respond to changes in daily value of these variables (Capiro, 1993; Schwartz et al., 2012). Our selection includes SI-xLM, Spring Warming (SW), UNIFORC, Photothermal SW and Photothermal UNIFORC models. The SI-x models are widely used to study the timing of plant leafing and its changes in the Northern Hemisphere (Linkosalo et al., 2008; Mehdipoor et al., 2016; Wu et al., 2016; Belmecheri et al., 2017; Hufkens et al., 2018). The output of the SI-x model, the estimated DOY of FL and first flower of indicator plants such as lilac, are used as an official indicator of climate change in the U.S. (Schwartz et al., 2006, 2013; Crimmins et al., 2016). SI-xLM was first calibrated about 25 years ago, using the original VPOs and daily minimum and maximum temperature (Schwartz, 1997; Ault et al., 2011; Schwartz et al., 2013; Ault et al., 2015).

SI-xLM has short- and long-term estimators based on growing degree hours, referred to as the sum of the hourly temperatures above 31 °F (-0.5556 °C).

Estimators of SI-xLM include days since January 1st (MDS0), accumulation of the number of high-energy synoptic events, which is defined as three-day accumulations of degree-hour higher than 637, (SYNOP), 5 - 7 day degree-hour accumulations (DD57) and 0 - 2 day degree-hour accumulations (DDE2). The calibration of SI-xLM fits a regression of the form equation 4.1, using estimators' values at observed DOYs. The coefficients of SI-xLM estimators (A_1, A_2, A_3, and A_4) were calibrated, and for prediction, inequalities of inequality 4.2 is operationally checked on a daily basis, starting on the 1st of January. For underlying assumptions and a more detailed definition of the SI-x models see (Ault et al., 2015).

$$\frac{1000}{\text{DOY}_{\text{Observed}}} = \frac{A_1 \times \text{MDS0}}{\text{DOY}_{\text{Observed}}} + \frac{A_2 \times \text{SYNOP}}{\text{DOY}_{\text{Observed}}} + \frac{A_3 \times \text{DD57}}{\text{DOY}_{\text{Observed}}} + \frac{A_4 \times \text{DDE2}}{\text{DOY}_{\text{Observed}}} \qquad 4.1$$

$$3.306 \times \text{MDS0} + 13.787 \times \text{SYNOP} + 0.201 \times \text{DDE2} + 0.153 \times \text{DD57} \geq 1000 \qquad 4.2$$

SW and UNIFORC phenology models are the second group of SPP models used in this study. These models consider only the role of the forcing temperatures (i.e., temperatures at which the plant develops). These models assume that phenophases such as FL phenophase occur when a critical state of forcing is reached. The state of forcing is modelled as the sum of the daily rate of forcing (R_f), which is a function of temperature only. We calibrated these models using equation 4.3 and equation 4.4:

$$R_{f_{\text{SW}}}(T) = \begin{pmatrix} 0 & \text{if } T < T_b \\ T - T_b & \text{if } T \geq T_b \end{pmatrix} \qquad 4.3$$

$$R_{f_{\text{UNIFORC}}}(T) = \begin{pmatrix} 0 & \text{if } T < T_b \\ \dfrac{1}{1 + e^{d(T-e)}} & \text{if } T \geq T_b \end{pmatrix} \qquad 4.4$$

where the T_b is the base temperature (i.e., minimum temperature required for plant development), T is daily average temperature, d and e correspond to the slope at the inflection point (width) and the temperature of the mid-response (centre) of the sigmoidal function. For both temperature response functions, the summation of the daily rate of forcing was calculated from the 1st of January (DOY=1) as in inequality 4.5:

$$S_f(t_s) = \sum_1^{t_s} R_f(T) \geq F^* \qquad 4.5$$

where t_s represent the DOY where FL is reached, and F* shows the amount of heat that is needed to be accumulated by the plant to reach that state of forcing $S_f(t_s)$.

Photothermal SW and Photothermal UNIFORC are the third group of models used in this study. These models are based on the average temperature during day-length (Burghardt et al., 2015). The summation of R_f was converted to photothermal units by adding a photoperiod variable to enhance the biological meaning of equation 4.5 (Črepinšek et al., 2006). The daily rate of forcing for the PTT models (R_{fPPT}) is defined as the multiplication of the light period as a proportion of a day to R_f of Spring Warming and UNIFORC (equation 4.6). The PTT models, Photothermal SW and Photothermal UNIFORC, apply the same approach to check if the plant received the amount of heat that is needed to reach that state of forcing, equation 4.3 and 4.4, where L_i is daily day-length in hour.

$$R_{fPTT} = \frac{L_i}{24} R_f \qquad\qquad 4.6$$

Exploring patterns and trends of models

We use the workflow presented in Figure *4.1* to analyse and compare the abovementioned models. First, we searched for the optimal set of parameters for the SW, UNIFORC, Photothermal SW and Photothermal UNIFORC models using the VPOs, temperature and day-length datasets used to calibrate SI-xLM, from 1961 to 1996. In particular, in the first step, we used simulated annealing (SA) to find the models parameters (i.e., T_b, e, d and F*). SA is a probabilistic optimization algorithm that performs a search in large multidimensional space (Aarts and Korst, 1988). This algorithm is robust to entrapment in local optima (van Laarhoven and Aarts, 1987) so we used it to find the optimal set of coefficients that minimize the objective function. Here, our objective function is the RMSE between the observed and estimated DOY from the calibration dataset. SA uses an initial random set of parameters for the objective function to start its search, then it takes steps within predetermined ranges. In this case, we used a range between [0, 10] for T_b, [-2, 0] for d, [0, 1000] for F* and [0, 5] for e.
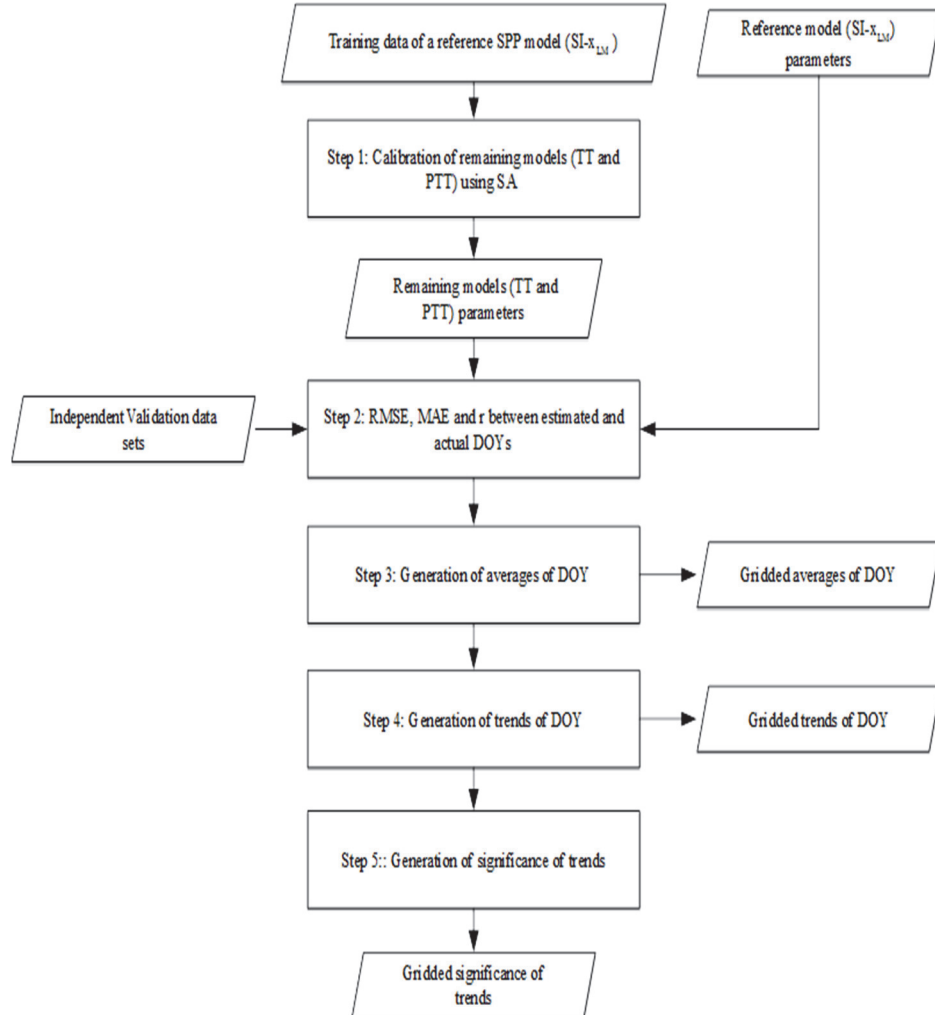
Figure 4.1 The main analysis steps for generating the spatio-temporal trends.

Next, in the second step, SI-xLM and the calibrated SW, UNIFORC, Photothermal SW and Photothermal UNIFORC models were validated using contemporary VPOs and Daymet data over the period 2000-2014. This period was selected because consistent VPOs were only available for these years. Daily temperature and day-length were extracted at the location of the VPOs. These data were used to estimate DOY at VPOs locations. The RMSE of all models was calculated comparing the observed and estimated DOY. Scatterplots of the observed and estimated DOYs were generated to explore the effect on the RMSE of the model qualitatively. Moreover, the mean absolute error (MAE) and the Pearson correlation coefficient (r) were calculated to quantify potential errors in the models. The scatterplots and correlations between the model errors (i.e., subtraction of observed DOY from estimated

ones) and geographic coordinates (i.e., Latitude and longitude) were generated to analyse the model output further and to understand better how model errors propagate over space.

In the third step, the average annual rate of change of DOY was calculated for each grid cell across CONUS to reveal differences in spatial and spatio-temporal trends between different geographic locations. Spatially continuous model outputs were only obtained for the two most accurate models to compare the effect of these models on the estimation of their average rate of change. Because we need to calculate for so many spatial locations (1-km resolution for the complete CONUS), we used the Google Earth Engine (GEE) cloud computing platform. The implementation of the SI-xLM model in the GEE (Gorelick et al., 2017; Izquierdo-Verdiguier et al., 2018) was used to generate annual and average outputs for the models. This helped to explore and compare regional variations between gridded model outputs. The histograms of the differences were plotted to provide more information about their distribution.

In the third steps, the annual difference between the two model outputs was spatio-temporally clustered to provide an overview on the effect of the models on spatial and temporal trend in DOY. The annual differences are the pairwise subtraction of grid cell values (i.e., 1-km by 1-km temperature-driven DOY) of the SW product from the SI-xLM product, from 2000 to 2014. We applied k means clustering in GEE, which is a widely used clustering technique that seeks to minimize the average squared distance between data in the same cluster. For underlying assumptions and details of the GEE k means clustering, see Arthur and Vassilvitskii, 2007. Since the number of clusters cannot be known a priori, we set it to seven as an example. Clusters define regions for which the difference between the estimates of the two models was similar throughout the years. Clusters were mapped and explored to understand the spatial variation of the regions.

Finally, for both models, the temporal trend in the gridded products was calculated and compared. For each grid cell, the temporal trend was obtained by fitting a linear regression line to the annual products from 2000 to 2014. The slope of the line is the rate of change of the models per year. We calculated the difference between the trends, by subtracting grid cell values of the two models. This highlighted regions where the estimated rates of change were highest. The statistical significance (p-value) of these trends were analysed and mapped to show areas with clear phenological changes. We applied the 2-sided p-value test to see if the estimated trend is significantly greater than 0 and if the mean significantly less than 0.

4.3   Results and discussion

The SA calibrated parameters of the TT and PTT models are shown in Table *4.1*. The train RMSE of the models is similar to what Hunter and Lechowicz (1992) found. The difference in model parameters indicate that models can be parameterized to provide good predictions, however, their parameters are biologically "meaningless" (Hunter et al., 1992). The calibrated parameters result in similar minimum values of the RMSE objective function (12 days). These values are also similar to the RMSE value calculated using the SI-xLM (11 days). Such similarities indicate that the selected phenological models can fit the historical data equally well.

Table 4.1 Calibrated parameters for the SW, UNIFORC, Photothermal SW, Photothermal UNIFORC models, and the corresponding RMSE of the calibration dataset.

Model Name	Model parameters				RMSE
	T_b	d	e	F*	
SW	2.34			220	11.48
UNIFORC	7.22	-0.35	4.75	28	11.92
Photothermal SW	0.26			151	12.57
Photothermal UNIFORC	6.23	-0.75	3.04	14	12.63

The results of the validation (Figure *4.2*) shows that the SI-xLM is two to three days more accurate in its estimates than the other models. The MAE also highlighted the similar difference in the error of the models. The visual exploration of scatterplots of observed and estimated DOY from UNIFORC, Photothermal SW and Photothermal UNIFORC estimations are more biased to later DOYs than SW and SI-xLM. Although there was no significant correlation between the model's error and geographic gradients (i.e., latitude and longitude), correlation coefficients are higher for TTs than other leaf models. Figure *4.2*e shows that the SW model has a higher error in the northern and southern CONUS compared to the centre. This is because for high and low latitudes, the temperature, which is the base of the SW model, might not be the only driver of lilac FL. However, the correlation between the estimated and observed DOY is similar ($r \sim 0.77$), which indicates a similar goodness of fit of the models for the estimation of the timing of lilac FL.
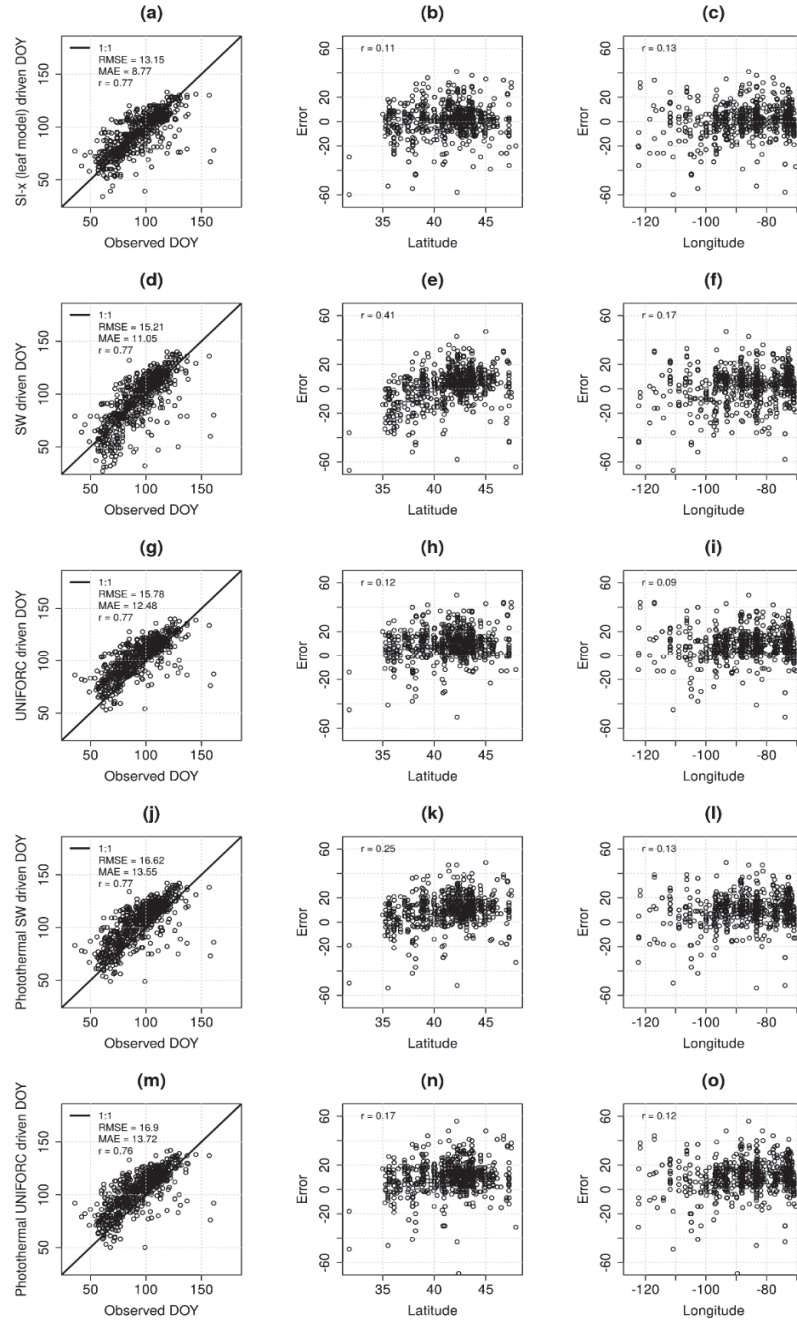
Figure 4.2 Scatterplots between observed and predicted by SPP models (first column), latitude and model error (second column) and longitude and model error (third column). Errors calculated by subtracting observed from the estimated DOYs.

The average DOY from SI-xLM and SW were mapped, generalizing grid cell values into half-months of DOYs (Figure *4.3*). For both models, DOYs range from January to June across CONUS. Moreover, there are latitudinal patterns in the eastern and elevational patterns in the western CONUS. The visual exploration of the generated products shows that DOYs from the SI-xLM model (Figure *4.3*a) are different from the results of the SW model (Figure *4.3*b). In the eastern CONUS, estimated DOYs from SW are earlier in the south, and later in the north than those estimated from SI-xLM. For example, SW mostly estimates DOY in early May while SI-xLM estimates it to be in late April. Similarly, in the western CONUS, DOYs driven from SW are earlier in low-altitudes, and later in high-altitude regions, compared to DOYs driven from SI-xLM. SW estimates DOY in early January for most locations in California, however, SI-xLM estimates this in early February.
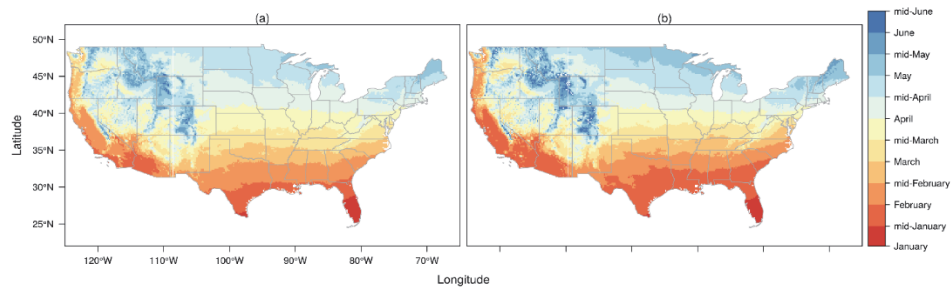
Figure 4.3 Average DOY of lilac FL generated from (a) SI-xLM and (b) SW, from1980 to 2014.

The histogram of difference in DOY for SI-xLM and SW (Figure *4.4*a) indicates that for most locations in CONUS the estimated average DOY is 11 days different. In these locations, SW estimates are earlier than SI-xLM ones. Although the RMSE of SI-xLM and SW were only two days different, the estimates can show up to one month difference in the West and North West of the CONUS (Figure *4.4*b). For example, in California and Washington State, SW estimates are about one month earlier. Or, in the Rocky Mountains, SW estimates are even until one month later. These follow the fact that SW uses only Daymet temperature, for which the interpolation method uses elevation as a key covariant (Daly et al., 2008). Moreover, SW considers only long-term effects of temperature while SI-xLM considers both the short-term and long-term effect and day-length.
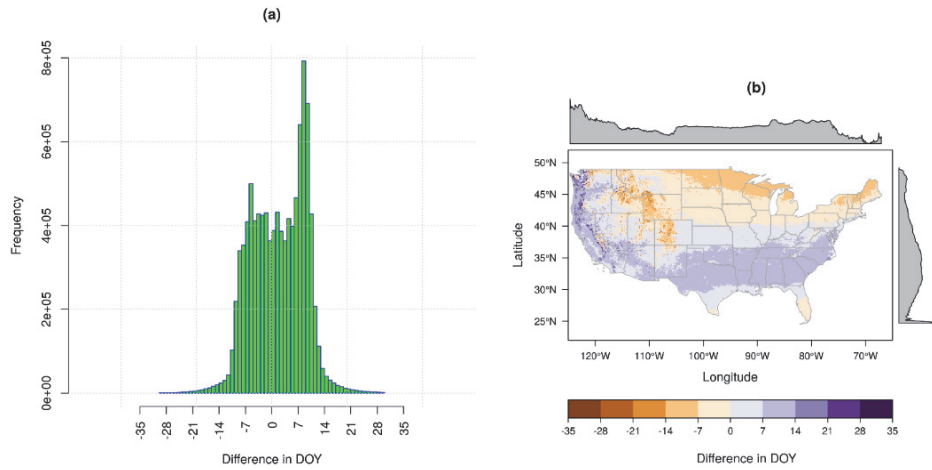
Figure 4.4 Histogram and map of the difference between products generated from SI-xLM and from SW. The differences are the pairwise subtraction of grid cell values of the SW product from the SI-xLM product. The two marginal plots are row and column summaries of the difference map, the average of grid cell values in each row and column.

Spatio-temporal clustering of annual difference in DOY between SI-xLM and SW grouped seven regions that have a similar difference over space and time (Figure *4.5*). The variability of the cluster type is larger in the eastern CONUS than the western CONUS, which shows that SI-xLM and SW perform substantially different in the eastern CONUS. In the eastern CONUS, the elevation gradient and consequently the temperature gradient are lower than in the western CONUS. Thus, compared to the western CONUS, the importance of day-length that changes with latitude is higher in the eastern CONUS. As a result, the variability of cluster types is higher in the eastern CONUS and the clusters have a latitudinal pattern. This is because SW is only based on temperature while SI-xLM takes the influence of both temperature and day-length into account.
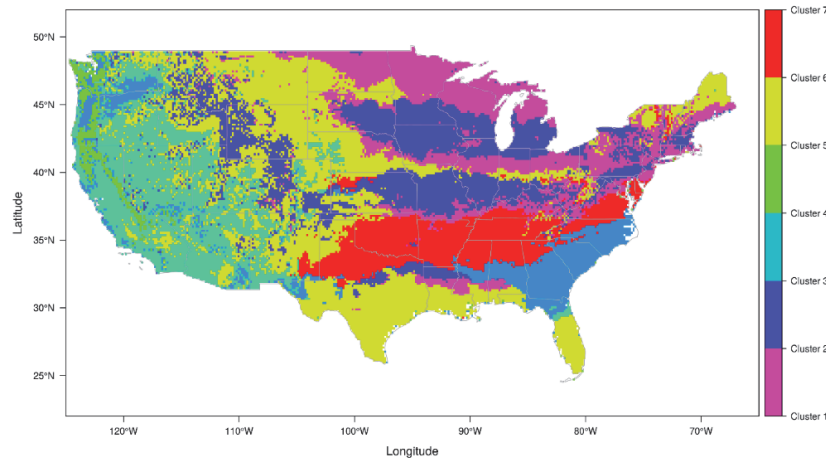
Figure 4.5 Clustered regions for which the difference between estimates of SI-xLM and SW is similar over the years.

The regression line fitted to the annual outputs of SI-xLM and SW, helps to explore spatial variations in the temporal trend (Figure *4.6*). The slope of the regression lines was mapped by generalizing them in 0.7 steps, which indicates about one week change per decade. Both models show advancement in DOY in the most western CONUS, ranging from one day to one month from 1980 to 2014. Temporal trends (Figure *4.6*a) driven from SI-xLM show both advancements and delay in the eastern CONUS while SW driven trends show mostly delay in this part of the CONUS. As referred to by other studies such as Schwartz et al. (2013), the delay in DOY happens only in the south-eastern CONUS.
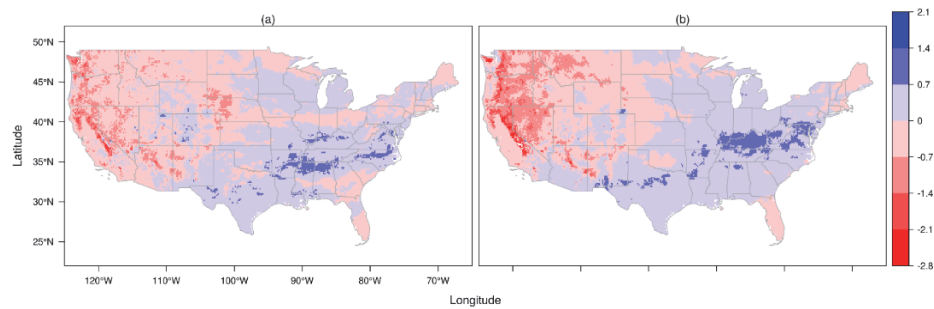
Figure 4.6 Trend maps of DOY of lilac FL from (a) SI-xLM and (b) SW. The trend values show the rate of change of DOY per year.

There is a significant trend variation between the outputs from SI-xLM and SW (Figure *4.7*). SW shows larger areas with a significant trend in the west coast and south central CONUS compared to SI-xLM. For example, SW shows significant advancement for most locations in Oregon, while SI-xLM does not

show such advancement in this state. Or, SW estimates significant delay in the southern and western part of Texas. The explanation for this is that SW uses average temperature for which annual variation is higher in the above-mentioned regions. In the eastern part of the US, significant trends between SI-xLM and SW do not match. Outputs from these models show a completely different spatial pattern in estimated delay over the period 2000-2014. However, both models highlight areas in the south eastern CONUS, close to the 'warming hole' region, where the secular trend during the past century has been towards later DOY (Meehl et al., 2012; Schwartz et al., 2013).
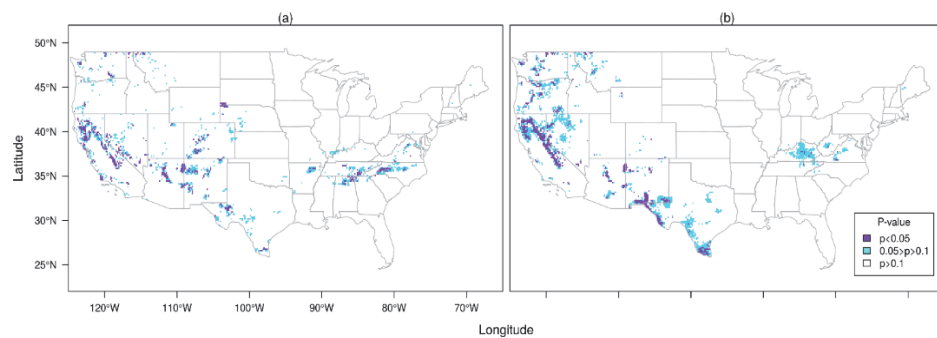
Figure 4.7 Statistical significance (range of p-values) for the trends in DOY of lilac FL from (a) SI-xLM (b) SW, illustrated in Figure 4.6.

The difference between the trends in SI-xLM and SW estimates can be up to a week per decade (Figure *4.8*). Difference values are the pairwise subtraction of grid cells values of the trend in SI-x estimated from the trend in SW estimates. The positive differences are more dominant in the western CONUS where the SI-x trend product shows less advancement in DOY compared to the SI-x trends. However, the negative difference values are more concentrated in the eastern CONUS where trends estimated from SW products are often showing a delay compared to estimated trends from SI-x.
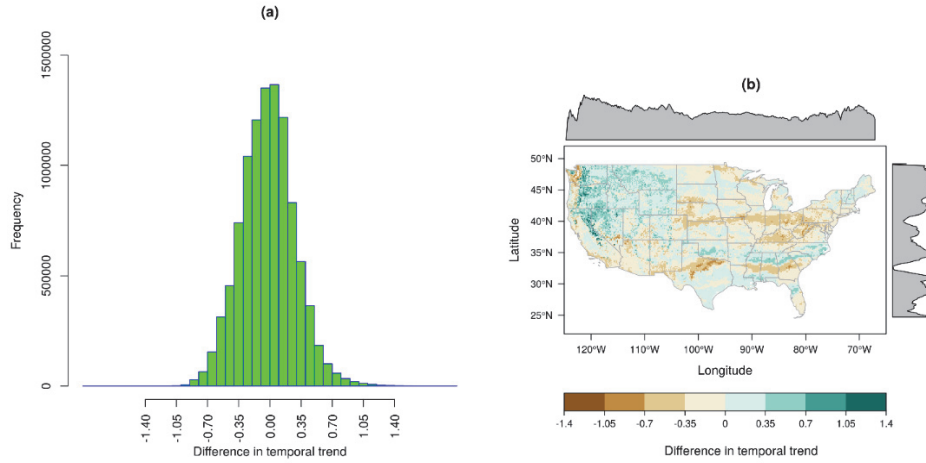
Figure 4.8 Histogram and map of the difference between the trends in SI-xLM and SW estimates. The differences are the subtraction of the trend in SI-x estimates from the trends in SW estimates. The two marginal plots are average of grid cell values in each row and column.

4.4 Conclusions

The analysis of spatial patterns and temporal trends in phenological model outputs is a necessary step to validate them and to obtain more reliable model predictions that can be used to study climate change and to design management and adaptation strategies for ecological and agricultural systems. This paper not only analyses patterns and trends at observational sites but also analyses spatio-temporal patterns for the complete spatial conterminous US and at a very fine spatial resolution (1-km). Due to the large volume of data, our workflow uses cloud computing and gridded temperate time-series to study phenology at continental scales. Volunteered phenological observations are also used to compare and validate the average and rate of change of DOY across CONUS.

Our results show that errors of SI-xLM and SW models are similar, and that these models are two days more accurate than those provided by other spring phenology models. However, patterns of DOY which are derived from SI-xLM and SW models are 11 days different in the CONUS. Given that the period only contains 15 years, this difference is considerable. The spatial variability of the SI-xLM and SW models is higher in the eastern CONUS. The results also indicate that the estimated rate of change in DOY from SI-xLM and SW can be up to one week per decade different across the CONUS. Moreover, our results show that the significance of the rate of change from SI-x and Spring Warming are spatially variable in the CONUS.

Therefore, current approaches for validating phenological models based on global statistics such as RMSE cannot be used to get information about the variability of patterns and trends in different regions. Studies using phenological models and gridded input data to study climate change impact on plant seasonality, and the eventual consequences on other living organisms, should check both the spatial and temporal variability at large-scale. Using a model that is found less valid across the study area than another one (i.e., with a "worse" RMSE) may still provide more realistic patterns and trends when compared with large-scale phenological data and/or information. Hence, we recommend applying our workflow to check the reliability of phenological models calibrated to be used at large scale. The workflow presented here can be applied to other phenophases, species and models to explore spatial phenological patterns and trends and to better understand the impact of climate change on the Earth system.

Chapter 5 Influence of source and scale of gridded temperature data on modelled spring onset patterns in the conterminous US*

5.1 Introduction

Climate variability and change affect the timing of plant development, most conspicuously after winter dormancy breaks in early spring (Cayan et al., 2001; Schwartz et al., 2006, 2013; Post et al., 2018). For example, increases in global temperature, particularly in the cool half of the year, have resulted in earlier spring onsets of leafing and flowering throughout the Northern Hemisphere (Schwartz et al., 2006, 2013; Allstadt et al., 2015). Such shifts have significant ecological, hydrological, and economic consequences. When plants leaf and bloom earlier than normal, for example, pollinators and herbivores, have to adjust their life cycle events (Marra et al., 2005; Miller-Rushing et al., 2010; Gornish et al., 2013; Broussard et al., 2017). Earlier leaf-out can dry out soils and advance, and even exacerbate, the wildfire season (Abatzoglou et al., 2016). And, early spring onsets can cause frost damage to fruit crops, when the last spring frost date fails to advance at the same pace as flowering (Gu et al., 2008; Ault et al., 2011; Munson et al., 2015; Chen, 2017).

Phenology is the science that deals with the study of annual life cycle events (phenophases) in plants and animals, and how variation in environmental conditions affect the timing of these events (Lieth, 1974). In particular, plant spring phenology is important for understanding the influence of weather and climate on plant growth, as well as a key indicator of climate change and its ecological and societal impacts (Schwartz, 1994; Root et al., 2003; Schwartz et al., 2012). It is commonly expressed in terms of day of year (DOY) for key phenophases, following established observation protocols such as First Leaf (LF) emergence and First Bloom (FB) emergence (Schwartz, 1998; Wolfe et al., 2005). DOY's for FL and FB for many early spring species are easy to observe and record, and tend to exhibit regional synchrony in response to temperature variations. FL and FB for these species are largely driven by the accumulation of warmth following the break in winter dormancy; and can be estimated to varying degrees for different species using photoperiod (day-length) and daily temperature series (Askeyev et al., 2005; Basler and Körner, 2012; Shen, 2015; Ault, 2015; Basler, 2016).

The exact response of FL and FB to these climate parameters varies among plant species (Polgar et al., 2011), however, phenological models generally capture the behaviour of a wide variety of plants in natural and agricultural systems (Wolfe, 2005; Schwartz and Hanes, 2010; Schwartz, 2013; Allstadt, 2015). A suite of statistical models referred to as Extended Spring Indices (SI-x) successfully generalize the DOY at regional to continental scales (Schwartz et al., 2006, 2013; Ault et al., 2015) for a wide range of species sensitive to the accumulation of warmth, especially in the early spring (Schwartz, 1990, 1993). SI-x indices have been promoted as official indicators of climate change

in the U.S., using station and gridded daily temperature data as model inputs (Crimmins et al., 2016). Moreover, these indices are widely used in the Northern Hemisphere to estimate patterns and trends in spring onset (Wu et al., 2016; Belmecheri et al., 2017; Zhu et al., 2017).

Gridded weather data are key because they support spatially continuous indices that can be used to explore spatio-temporal variability and trends in spring onset at the continental scale. The recent improvement in interpolation algorithms and computational technologies have led to gridded daily weather datasets such as Daymet, available at 1-km resolution and gridMET, available at ~4-km (1/24th degree) resolution. These two products are generally the most frequently-used datasets overall, and they are the most accurate gridded products available to calculate SI-x in the US. Daymet and gridMET are generated using ground-based daily temperature measurements, using two very different models, and are widely used for ecological, environmental, meteorological and atmospheric applications. Cross-validation errors have been calculated for Daymet (Thornton et al., 1997) and gridMET (Daly et al., 2007), but their impact on SI-x indices have not been evaluated. Such evaluation requires extensively-distributed reference phenological observations and adequate computational power.

Advances in online information communication and mobile location-aware technologies have dramatically increased the amount of phenological observations collected by volunteers (Ferster et al., 2013; Fuccillo et al., 2015; Mehdipoor et al., 2015). These volunteered phenological observations (VPOs) are timely observations at long temporal and large spatial scales (Rosemartin et al., 2015). Worldwide, several networks collect VPOs to model changes in the timing of plants phenological events in spring and what those changes imply (van Vliet et al., 2003; Mayer, 2010; Beaubien et al., 2011; Donnelly et al., 2014). Moreover, enhancements in large-scale distributed computing paradigm such as cloud computing facilitate processing of higher spatial resolution, gridded data (Guo et al., 2010).

This study analyses the effects of Daymet and gridMET data and their spatial resolution on gridded SI-x products in the conterminous US. SI-x indices using Daymet at and gridMET are used to estimate and compare annual and longer-term-average SI-x products using different spatial resolutions. These products also were validated using VPOs. In addition, temporal trend in SI-x indices using Daymet and gridMET were estimated and compared at four different resolutions.

5.2   *Materials and methods*

Temperature data and phenological observations

The data used in this study consist of gridded daily temperatures to generate the SI-x indices at 1, 4, 35 and 100-km, and VPOs to validate generated SI-x indices. Temperature data are daily maximum and minimum temperature extracted from Daymet and gridMET. Both datasets use a Digital Elevation Model (DEM), ground weather station data and local climate-elevation regression to generate gridded daily meteorological parameters estimates. However, the interpolation model uses different sources of input data and spatial resolution in each case.

Daymet dataset uses spatially-referenced surface measurements of daily maximum and minimum temperature and precipitation from the Global Historical Climatology Network (GHCN) and land and water mask, the NASA SRTM 30 arc second DEM as the major input (Daly et al., 2008; Thornton et al., 2014). Weather stations are weighted and filtered using a truncated Gaussian filter based on distance from the estimation point, where distance is a function of the concentration of stations in the estimation region. Daymet uses the chosen stations to assess the local relationship between temperature and elevation, and estimates temperature at the location of interest by using a weighted least-squared regression (Thornton et al., 1997; Hasenauer et al., 2003). Daily maximum and minimum temperature and day-length from Daymet and SI-x Daymet products (Izquierdo-Verdiguier et al., 2018) are available at 1-km spatial resolution for North America, since 1980. In this study, SI-x Daymet products also were generated at 4-km, 35-km and 100-km spatial resolution.

GridMET dataset combine gridded weather data from two sources including 1) the North American Land Data Assimilation System Phase 2 (NLDAS-2; see Mitchell et al., 2004), and 2) the Parameter-elevation Regressions on Independent Slopes Model (PRISM; see Daly et al., 2008) to create high-resolution gridded, surface meteorological data over the continental United States since 1979. The NLDAS-2 uses or assimilates surface measurements data to produce weather parameters at hourly time scales and 1/8th degree (~12-km) resolution over North America. PRISM preforms local regressions of station data to physiographic elements using an extensive knowledge base of spatial weather parameters to provide high spatial resolution (~800-m) weather parameter at monthly time scales.

First, gridMET downscales NLDAS-2 and upscales PRISM data to 4-km grid using a bilinear interpolation and an area-weighted average, respectably. In particular, gridMET data define the daily maximum temperature as the daily

maximum temperature from NLDAS-2 plus the difference of the monthly average maximum temperature from PRISM and the monthly average maximum temperature from NLDAS-2. GridMET also calculates the daily minimum temperature in a similar way. For underlying assumptions and detail definition of the gridMET see Mitchell et al., 2004. Daily maximum and minimum temperature from gridMET are available at 1/24th degree (4-km) across the contiguous United States, since 1979. And, they are used to generate the SI-x gridMET products at 4-km, 35-km. and 100-km.

The VPOs were used as reference observations to assess the accuracy of the different SI-x products. The most long-term and continentally-extensive VPOs available for CONUS are for phenophases of lilacs (*Syringa vulgaris 'common lilac'* and *S. x chinensis 'Red Rothomagensis'*) and honeysuckles (*Lonicera tatarica 'Arnold Red'* and *L. korolkowii 'Zabelli'*). Historical VPOs, for these species, which have been used extensively to both develop and validate SI-x products, can be uploaded directly from the Phenology Observation Portal of the USA National Phenology Network (USA-NPN; see Rosemartin et al., 2015). They contain the geographic location and the DOY that FL and FB for lilac and honeysuckle were first observed by volunteers in a given year. VPOs from USA-NPN include observations collected via the Nature's Notebook phenology program from 2009 to 2016, and additional integrated datasets, such as historical lilac and honeysuckle data from 1955 to 2016.

Extended Spring Indices

The SI-x is widely used to demonstrate that spring onset has generally shifted earlier in the Northern Hemisphere and that there is considerable variation in the magnitude of temporal trends (Schwartz et al., 2006, 2013; Allstadt et al., 2015). It estimates the DOY of FL and FB obtained by the time average of the FL and FB of three models from leafing and flowering of the four lilac and honeysuckle species. The models consider long- and short-term influence of temperature on FL and FB in spring, which makes the models unique compared to other spring phenology models.

Growing Degree Hours (GDH) forms the basis for calculating the SI-x regressors. Leaf model (equation 5.1) requires the regressors: day of the year since January 1 ($MDS0_{leaf}$), accumulation of number of high-energy synoptic events (SYNOP), 5-7 day degree-hour accumulations (DD57) and 0 - 2 day degree-hour accumulations (DDE2). The two regressors of bloom model (equation 5.2) are the day of the year ($MDS0_{bloom}$) and accumulated growing degree hours (AGDH) since the estimated DOY of FL (for underlying assumptions and detail definition of the regressors see (Schwartz et al., 2006; Ault et al., 2015). Once daily regressors values are calculated, the three inequalities of equation 5.1 are checked on daily basis. The mean of earliest

DOYs for which the inequalities are true, are DOY of FL. Once the DOY of FL is known, and applying equation 5.2, the mean of earliest DOYs for which the inequalities are true, are the DOY of FB.

$$3.306 \times MDS0_{leaf} + 13.787 \times SYNOP + 0.201 \times DDE2 + 0.153 \times DD57 \geq 1000$$ Lilac

$$4.266 \times MDS0_{leaf} + 20.899 \times SYNOP + 0.00 \times DDE2 + 0.248 \times DD57 \geq 1000$$ Arnold Red honeysuckle 5.1

$$2.802 \times MDS0_{leaf} + 21.433 \times SYNOP + 0.266 \times DDE2 + 0.00 \times DD57 \geq 1000$$ Zabeli honeysuckle

$$-23.934 \times MDS0_{bloom} + 0.116 \times AGDH \geq 1000$$ Lilac

$$-24.825 \times MDS0_{bloom} + 0.127 \times AGDH \geq 1000$$ Arnold Red honeysuckle 5.2

$$-11.368 \times MDS0_{bloom} + 0.096 \times AGDH \geq 1000$$ Zabeli honeysuckle

Exploring the effect of the scale and input data

The sensitivity of the SI-x indices respect to the spatial resolution and input data was analysed using Google Earth Engine (GEE) cloud computing platform (Figure *5.1*), in the following steps. First, Daymet dataset was aggregated to the gridMET spatial resolution and projection. Both datasets were used to generate annual and long-term-average SI-x indices at 1 (only in the case of Daymet), 4, 35 and 100-km in the contiguous US from 1980 to 2016. Next, these annual and long-term SI-x indices were compared spatially and temporally. They were then validated using annual and pooled VPOs, and the results were compared. Finally, rate of change in SI-x indices (per year) was calculated and compared. The next paragraphs describe each of the steps in more detail.
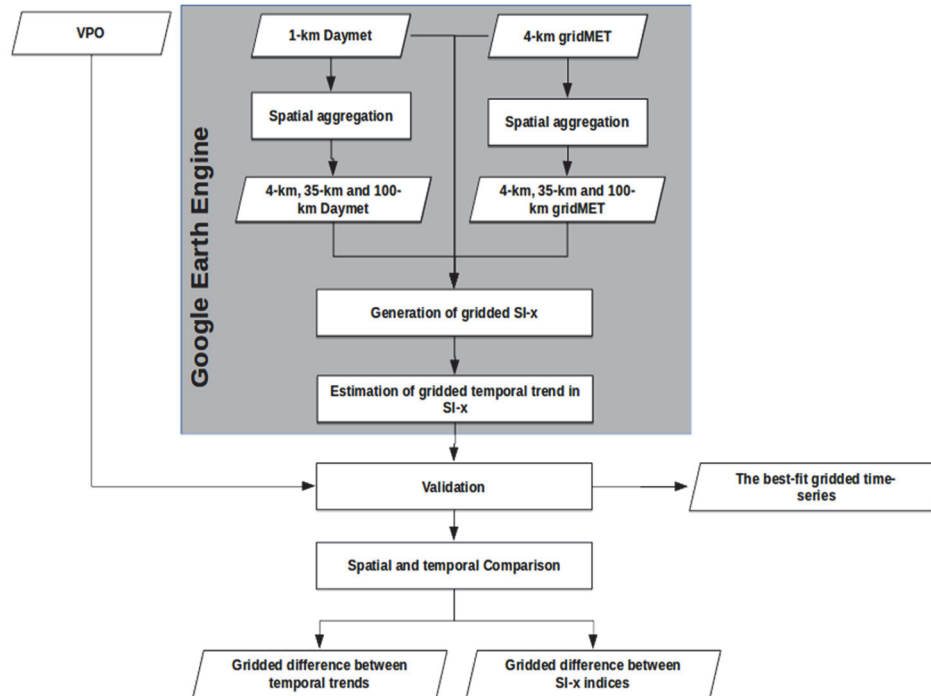
Figure 5.1 The diagram of workflow that analyses the sensitivity of SI-x indices respect to Daymet and gridMET using VPOs and GEE.

In the first step, daily maximum and minimum temperature and day-length from Daymet were spatially aggregated from 1-km to 4, 35 and 100-km. Similarly, gridMET datasets were spatially aggregated from 4-km to 35 and 100-km. These help to explore the effect of changing spatial resolution on SI-x estimates. For each year, in the second step, gridded SI-x indices were generated from 1-km, 4-km, 35-km, and 100-km Daymet and 4-km, 35-km, and 100-km gridMET using the SI-x indices code developed on GEE (Izquierdo-Verdiguier et al., 2018). Moreover, the long-term average of SI-x was calculated for indices driven for all spatial resolution cases. The long-term average provides an overview of SI-x products. The annual and long-term difference between the SI-x products were calculated to help explore and compare regional variations in the difference between gridded SI-x products. The differences are the pairwise subtraction of pixels values of 4-km, 35-km. and 100-km Daymet and gridMET. The histograms of the differences also were plotted to provide more information about the distribution of the difference.

In the third step, the accuracy of annual and long-term gridded SI-x products were assessed using the VPOs. The SI-x values associated with each VPO position and year were extracted from SI-x products using geographic location and year of VPOs. The scatter plots of the observed (VPO) and estimated (SI-x values) measures were generated to qualitatively explore potential bias in

SI-x products. A linear regression model was fitted and the root mean square error (RMSE), mean absolute error (MAE) and the Pearson correlation coefficient (r) were calculated. This helped to quantify and model potential errors in the products. For further analysis, we also plotted the daily SI-x regressors values versus the DOY for sample locations to explore the difference between them taking into account the datasets used.

In the last step, the temporal trend in the SI-x products were calculated and compared. For each pixel, the trend was obtained by fitting a linear regression line to annual SI-x indices, from 1980 to 2016. The slope of the line is the rate of change of SI-x indices per year. We calculated the difference between the trends subtracting pixels values of the pairwise SI-x products. This highlights regions where the estimated rates of change in SI-x indices are highest. The statistical significance (p-value) of these trends was analysed and mapped to show areas with clear phenological changes. We applied the 2-sided p-value test to see if the estimated trend is significantly greater than 0 and if the mean significantly less than 0.

5.3 Results and discussion

For both annual and long-term SI-x indices, the SI-x products generated from 1-km, 4-km, 35-km and 100-km Daymet (SI-x Daymet) are substantially different from SI-x products generated from 4-km, 35-km and 100-km gridMET (SI-x gridMET). However, visual exploration of SI-x Daymet in 1-km, 4-km, 35-km and 100-km and SI-x gridMET at 4-km, 35-km and 100-km resolutions shows that change in spatial resolution of SI-x input has no influence on spatial pattern of SI-x outputs. These products were mapped generalizing pixel values into month of DOYs (Figure *5.2* and Appendix B.1, B.2, B.3, B.4, B.5 and B.6). The long-term SI-x gridMET show later FL and FB compare to SI-x Daymet. For example, SI-x gridMET estimates FL to occur across Kansas in April and May, while SI-x Daymet estimates it in March. In the mountainous western U.S., the spatial pattern in SI-x Daymet tracks elevation SI-x gridMET, on the other hand, poorly tracks elevation. For annual variations in SI-x products (Appendix B.7, B.8, B.9 and B.10), we highlight large anomalies in 2012 in FL for SI-x products, with more moderate anomalies in FB.
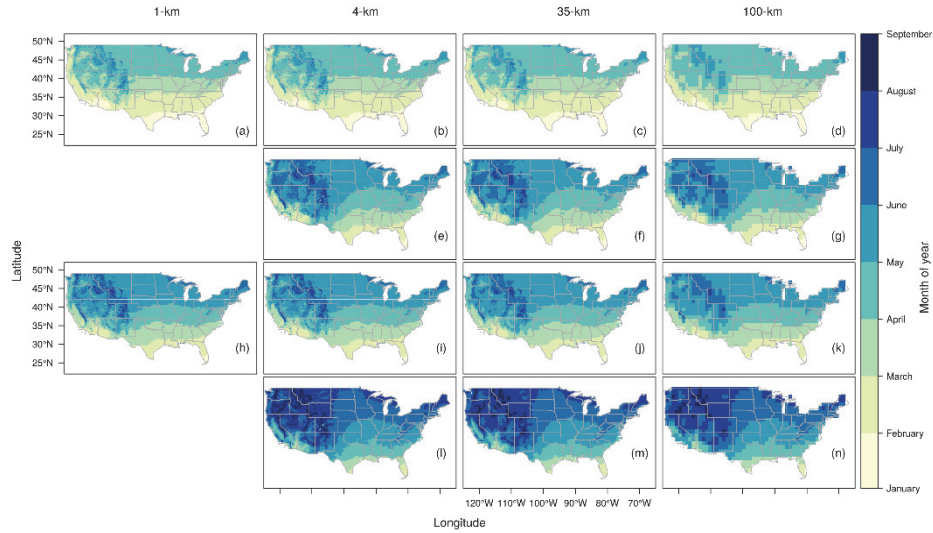
Figure 5.2 Long-term SI-x FL from (the first row: a–d) Daymet and (the second row: e–g) gridMET, and long-term SI-x FB from (the third row: h–k) Daymet and (the forth row: l–n) gridMET, from1980 to 2016.

The long-term differences between SI-x Daymet and SI-x gridMET are substantial for 4, 35 and 100-km (Figure *5.3* where the differences were plotted in steps of one week). The differences are five weeks for FL, and between five to six weeks for FB in most of the pixels (Figure *5.4* and Appendix B.11 and B.12). This is caused by the known high variation in accuracy for Daymet and gridMET in areas with steep elevation gradients (Scully, 2010). These differences are larger in the southwestern U.S. (e.g., Arizona and New Mexico) than elsewhere. These results indicate that SI-x indices are sensitive to input data that affect the estimated DOY of FB more than FL; the change in spatial resolution of input data does not affect SI-x indices in the less topographically-complex Eastern U.S., only slightly in the mountainous Western U.S.
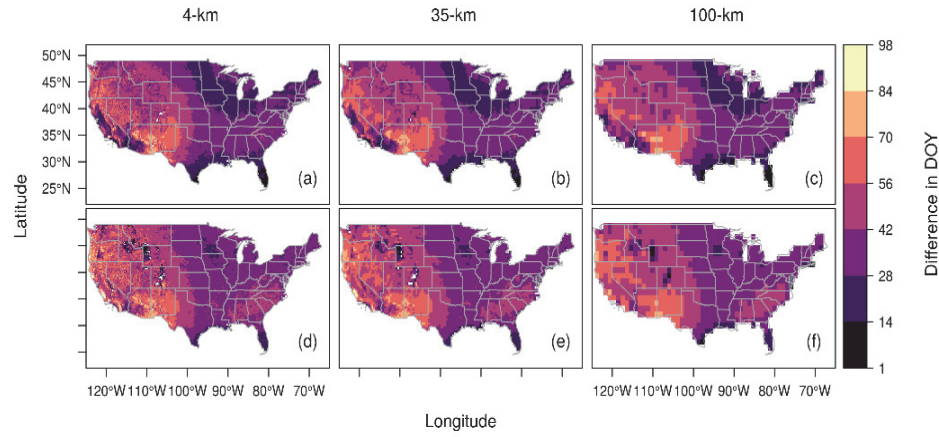
Figure 5.3 Maps of the difference between (the first row: a–c) SI-x FL products and (the second row: d–f) SI-x FB products.
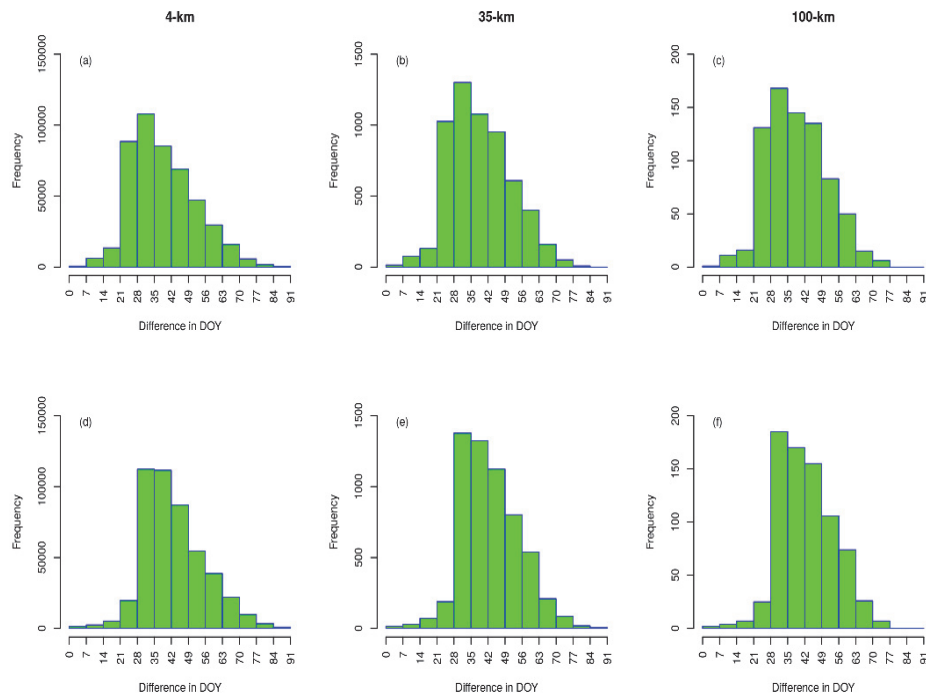
Figure 5.4 Histograms of the differences between (the first row: a–c) SI-x FL products and (second row: d–f) SI-x FB products.

The validation of SI-x Daymet and SI-x gridMET products are accomplished using annual VPO (Appendix B.13, B.14, B.15, B.16, B.17 and B.18) and pooled VPO (Figure *5.5*). For both the FL and FB indices, RMSE and MAE between

observed and SI-x Daymet as well as observed and SI-x gridMET change only about 2 days from 4-km to 100-km products. However, RMSE and MAE between observed and SI-x Daymet are smaller than SI-x gridMET at 4, 35 and 100-km spatial resolution for both the FL and FB indices. In terms of RMSE, SI-x Daymet products are about three weeks earlier than SI-x gridMET products for FL, whereas the difference increases up to about four weeks for FB. MAE results follow a similar pattern, showing SI-x Daymet are about four weeks for FL and five weeks for the FB earlier than 4-km SI-x gridMET. In terms of correlation, SI-x Daymet products are strongly correlated with VPOs, also indicated by the scatter plots (Figure *5.5*: the first and third rows). The SI-x gridMET values are overestimated by up to three and four months difference between modelled and observed FL and FB (Figure *5.5*: the second and fourth rows).

At sample locations, daily values of SI-x regressors driven from Daymet and gridMET differently evolve by DOY. For FL, the values driven by Daymet increase in higher rate in early year than those driven by gridMET. SI-x Daymet count larger numbers of high-energy synoptic events than SI-x gridMET (Figure *5.6*: the first column). They also take greater accumulations of GDH during the days prior to DOY of FL into account than SI-x gridMET (Figure *5.6*: the second and third columns). However, SI-x Daymet and SI-x gridMET highlight a similar evolution pattern for the accumulated growing degree hours since the estimated DOY of FL (Figure *5.6*: the fourth column).
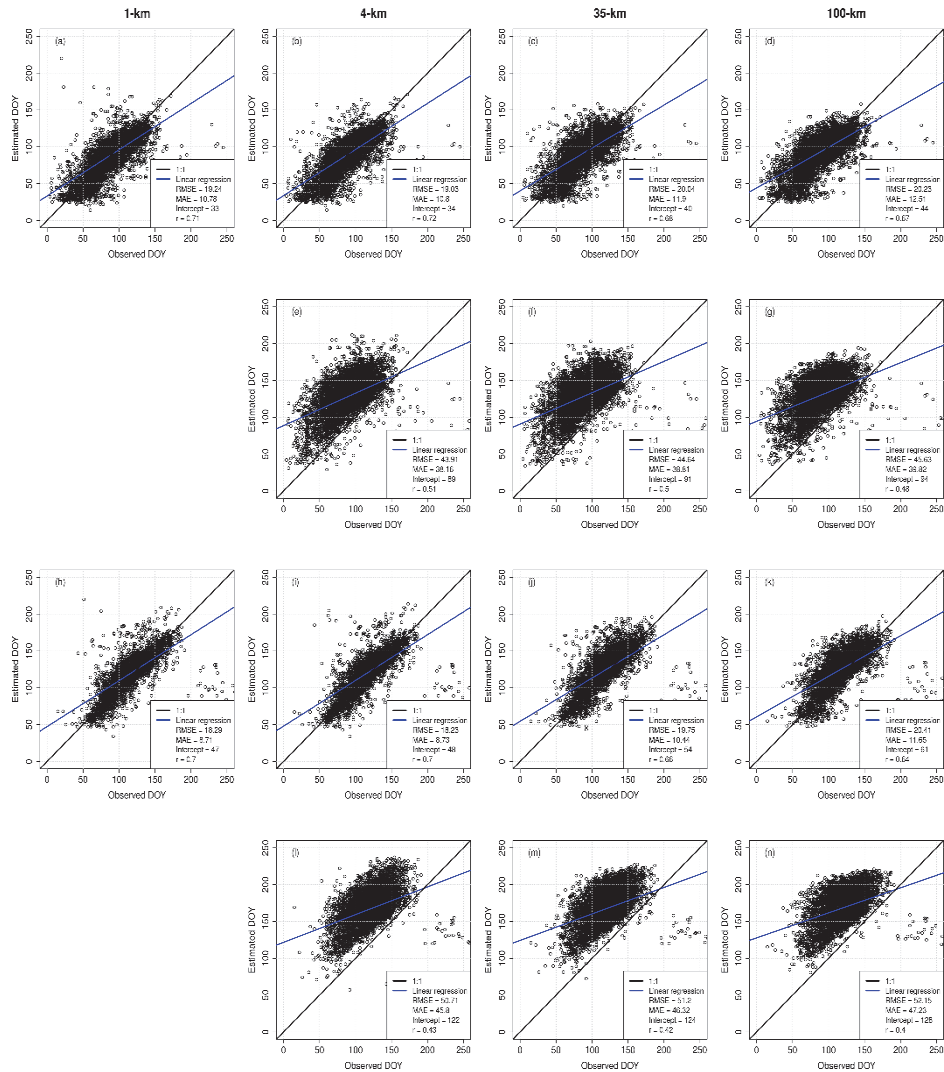
Figure 5.5 Scatter plots of volunteered observations versus (the first row: a–d) modelled FL index and (the third row: h–k) modelled FB index generated from Daymet; (the second row: e–g) modelled FL index and (the fourth row: l–n) modelled FB index generated from Daymet. The RMSE, MAE, correlation, slope of regression line (blue) fit between observed and estimated and identity line (black) are also included in scatter plots.
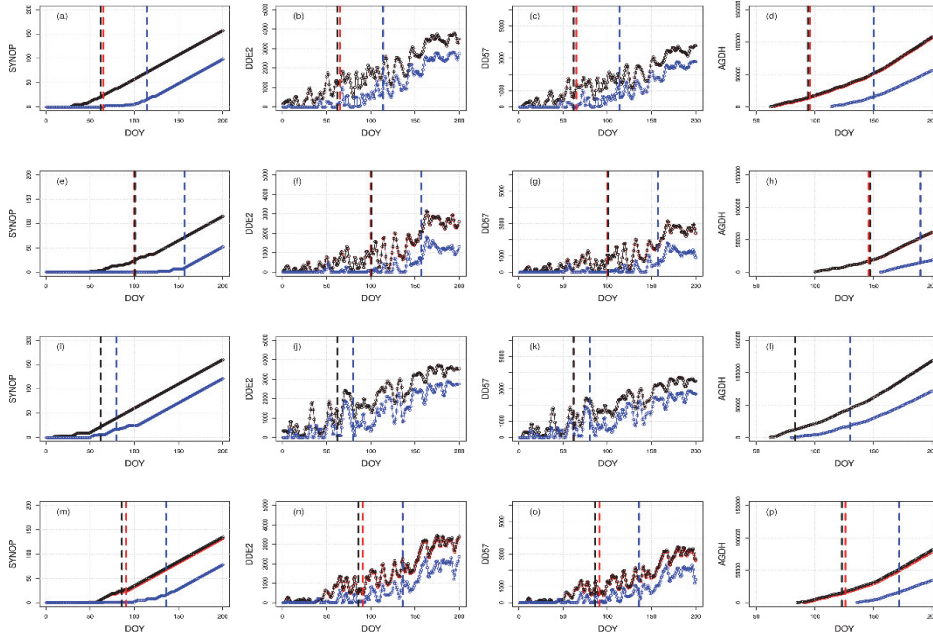
Figure 5.6 Illustration of daily values of SI-x regressors, driven from original Daymet (red), aggregated Daymet (black) and gridMET (blue), that change as function of time SYNOP, DDE2, DD57 and AGDH. In all panels, DOY of FL and FB is denoted by the dashed vertical lines (for these examples are from grid cells located at (a–d) latitude 37.64 N and longitude 99.28 W, (e–h) latitude 44.06 N and longitude 103.66 W, (i–l) latitude 44.06 N and longitude 103.66 W and (m–p) at a single year, 2016.

The regression line fitted to annual SI-x indices, from 1980 to 2016, helps explore temporal variations in SI-x Daymet and SI-x gridMET across the US. The slope of regression lines were mapped by generalizing in 0.2 steps, which indicate about one week change per the study period (Figure *5.7*). For both Daymet and gridMET, the maps of temporal trend in SI-x products show the similar spatial pattern in selected spatial resolutions. This is because these products exhibit no substantial differences. Both SI-x Daymet and SI-x gridMET show advancement in DOY of FL and FB for most of the locations in the US, ranging from one to five weeks from 1980 to 2016. However, the advancement in FB (Figure *5.7*: the second and fourth rows) covers larger areas than the advancement in FL (Figure *5.7*: the first and third rows). For FL, the advancement is higher in the Western U.S. (especially in California, Arizona and Colorado).
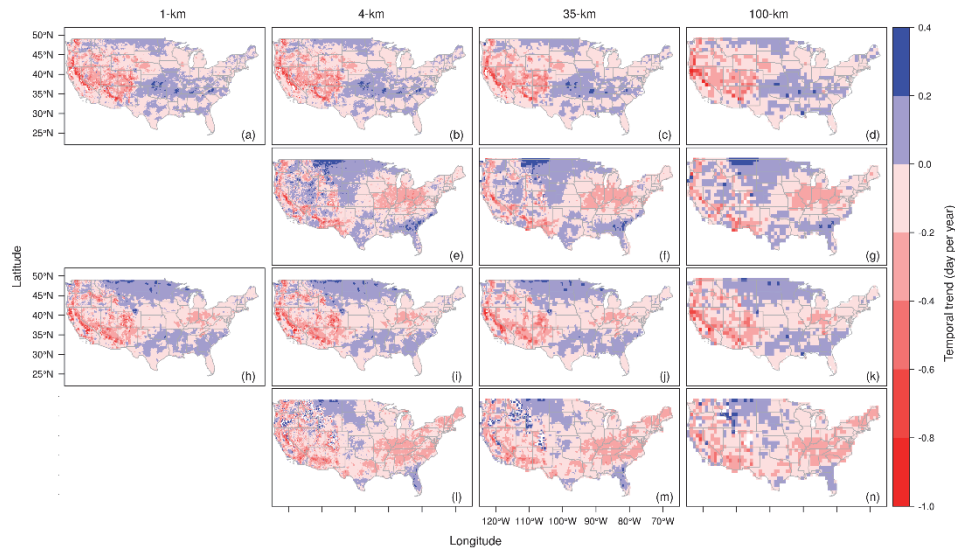
Figure 5.7 Trend maps of SI-x FL from (the first row: a–d Daymet and the second row: e–g) gridMET, and trend maps SI-x FB from (the third row: h–k) Daymet and (the fourth row: l–n) gridMET.

Significant trend in SI-x Daymet does not match those in SI-x gridMET (Figure *5.8*). The areas with significant trends in spring onset depend on the temperature product used to calculate the SI-x. For instance, the Daymet-based FL index has more significant areas in the West that the gridMET-based product. The opposite is found in for the BL product. In the East, the two indices derived from gridMET data show large areas with a significant advancement of spring. This is not in line with previous results (Schwartz et al., 2013; Crimmins et al., 2016), which found a delay in the timing of LF and BL. Trends are contradictory is some regions. For example, in Colorado and Utah there are areas with significant trends but Daymet-based trends point towards an advancement of spring whereas gridMET-based trends indicate that spring is getting delayed.
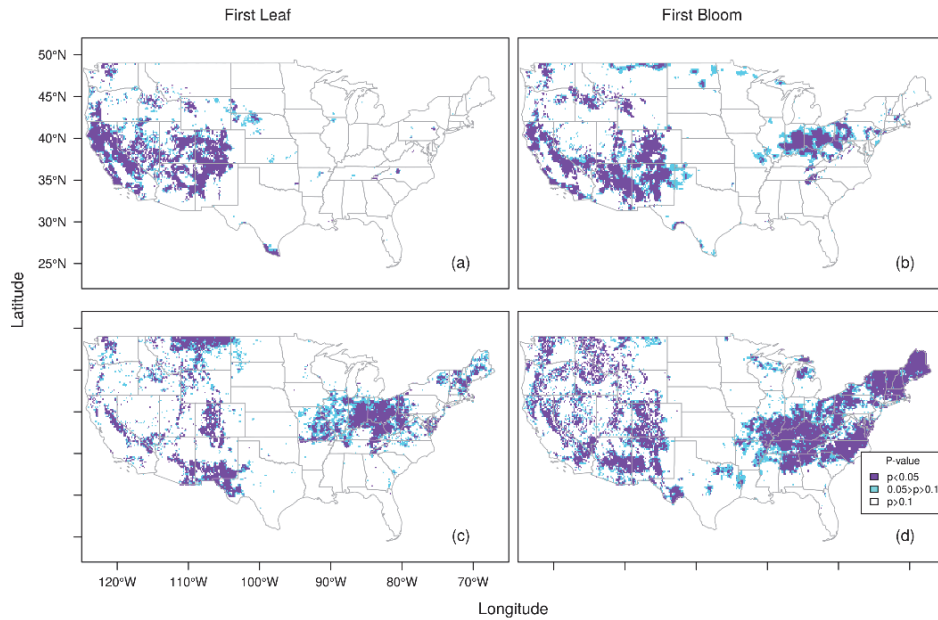
Figure 5.8 The statistics of the significance of temporal trend in (a) SI-x FL and (b) SI-x FB from, 4-km Daymet, and in (c) SI-x FL and (d) SI-x FB from, 4-km gridMET.

These differences in trends from Daymet and gridMET are highlighted in Figure *5.9*, which shows trend differences between 4, 35 and 100-km SI-x Daymet and SI-x gridMET. The trend in SI-x Daymet and SI-x gridMET can be up to 10 days per year different (Figure *5.9*). For both FL and FB indices, trend differences are higher in the Western U.S. than elsewhere. The trends in SI-x Daymet show smaller delays or larger advancement than those in SI-x gridMET in areas where the difference is positive, namely the highest elevations in the Western USA. The trends in SI-x Daymet show larger delay or smaller advancement than those in SI-x gridMET in areas with negative difference value. For FL, the difference value is positive over most of CONUS, especially in the West, while the negative values are centred in the Midwest and Southwest (Figure *5.9*: the first row). However, the differences for FB are negative over most of CONUS (Figure *5.9*: the second row).
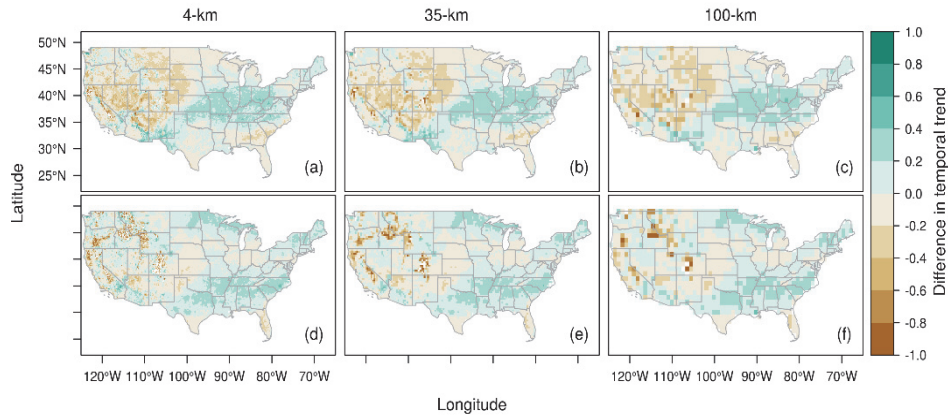
Figure 5.9 The difference between trends (day/year) in (the first row: a–c) SI-x FL and (the second row: d–f) SI-x FB from Daymet and gridMET.

5.4 Conclusions

The analysis of the effect of gridded time series input on Extended Spring Indices (SI-x) is necessary because these indices are being used increasingly as national and official indicators of climate change in the US. This paper presents an exploratory workflow analysing the effect of spatial resolution and input data on SI-x indices at the continental scale. The workflow utilized cloud computing and volunteered phenological observations to generate, compare and validate SI-x using Daymet and gridMET at selected spatial resolutions. We also analysed the impact of spatial resolution and input data on estimation of temporal trend in SI-x driven from these datasets, from 1980 to 2016.

Our results show that changing spatial resolution does not significantly affect the annual and long-term SI-x indices or the temporal trend in these products. However, the change in input data affects SI-x indices and, hence, the temporal trends in spring onset. The SI-x indices generated from Daymet are about three weeks for leaf index and about four weeks for bloom index more accurate than those driven from gridMET. The SI-x indices generated from gridMET are significantly biased toward later days of year, which can be expected to about three and four months for the FL and FB, respectively. The results also indicate that SI-x indices generated from both datasets exhibit the highest variation in FL and FB in the western US, as might be expected for areas that are more mountainous. Daymet and gridMET datasets show advances in FL and FB indices over most of the U.S. Difference between trends calculated from the different datasets, however, can be up to five weeks and even contradictory in some regions. In particularly, gridMET does not reflect the status of SI-x, and SI-x sub-models might need to be recalibrated to fit this dataset.

76

The proposed workflow in this paper can be applied to explore the effect of other high-resolution gridded time-series inputs on phenological models at the continental scale. By checking the translation of gridded input data into information, this workflow also can support other environmental and ecological studies being used to investigate the impact of climate change at local scales.

Chapter 6 Synthesis

This PhD thesis focuses on the design of novel geocomputational workflows to analyse the effects of using (inconsistent) volunteered phenological observations (VPOs), and of working with various phenological models, model inputs and spatial scales on the spatial patterns and temporal trends of spring plant phenology (SPP). Understanding such effects is crucial because SPP information is used to support policy making and the management of natural resources. In this chapter, I reflect on the link among the work presented in chapters 2 to 5. This reflection is articulated as a synthesis of knowledge that focuses on the two sub-objectives stated in chapter 1:

a. *To check the consistency of volunteered phenological observations using contextual geo-information and domain knowledge.*
b. *To analyse the impact of the type of phenological model as well as of its input data sources and their spatial resolution on the patterns and trends derived from the model.*

After the reflection (sections 6.1 and 6.2), I answer the research questions (section 6.3). Then, I summarize the main contribution of this PhD thesis (section 6.4). Finally, I outline various research avenues for future work (section 6.5). The setup of this chapter is illustrated in Figure 6.1.
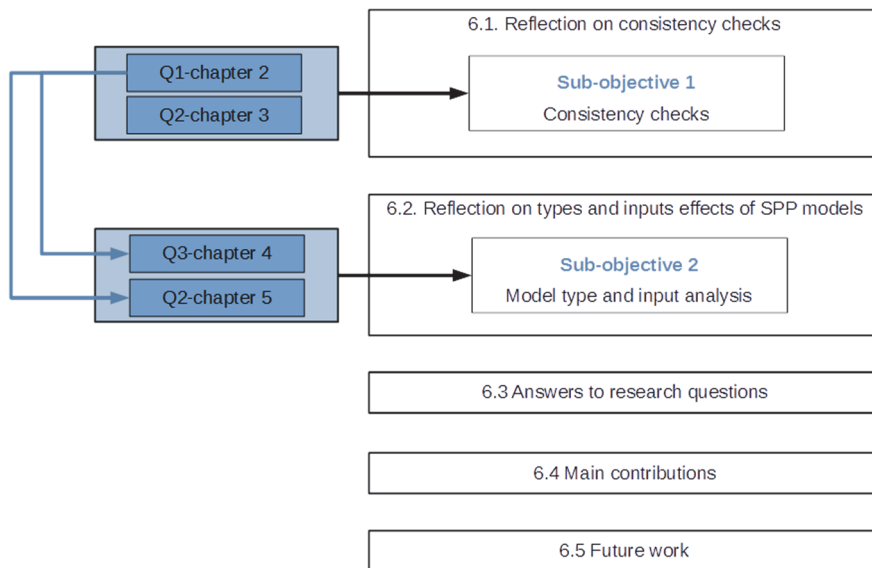

Figure 6.1 Overview of Chapter 6.

6.1 Reflection on consistency checks

The workflows presented in chapters 2 and 3 check VPOs consistency. Both workflows identify inconsistent VPOs. Inconsistency could be attributed to

transcription or volunteer errors or might be caused by non-representative microclimatic conditions or by unaccounted genetic variability. In the remaining of this section, I reflect and compare both workflows. Unlike existing consistency checks that only consider the statistical distribution of the VPOs distribution, our workflows use contextual and phenological information to provide robust consistency checks.

The workflows are based on different assumptions. The first workflow (Chapter 2) assumes that observations are collected in an area with a strong environmental gradient (in our case, the conterminous US (CONUS)). This assumption is crucial for grouping VPOs into similar locations so that we can identify anomalous observations. In areas where weather variables do not substantially change (e.g., the Netherlands), the dimensionality reduction and model-based clustering steps of the workflow might result in a small number of clusters or even in a single cluster. This hampers the context-aware outlier detection steps.

For the second workflow (Chapter 3), we assume that the key environmental variable that drives phenology does not change substantially in the neighbourhood of the observations and that there are at least two other observations in that neighbourhood. We also assume some domain knowledge, namely that the selected key variable drives the synchrony of the observations over the area. These assumptions form the basis for defining, searching and optimizing consistency constraints.

The first workflow can use as many environmental contextual (weather) variables as are available for the VPO sites. This makes the workflow flexible for various species and phenophases. However, using variables that do not, or to a lesser extent, drive the phenophase may lead to meaningless clusters. Unlike the first workflow, the second workflow uses only the variable that is related to domain knowledge of VPOs. Domain knowledge is often available about geographic phenomena such as spring onset.

Domain knowledge may be derived from more than one variable. In such a case, the constraint defined in chapter 3 should be adapted. However, the second workflow cannot be applied when no domain knowledge is available. Moreover, the second workflow needs a high level of connectivity in the network graph that connects the VPOs sites. As a result, in a large area with a low density of observations, the workflow fails to check the consistency of VPOs that are disconnected from other observations. A summary of the comparison between the workflows is presented in Table *6.1*.

The first workflow was used to identify inconsistent historical and contemporary VPOs from the USA National Phenology Network (USA-NPN).

This work was published as "Lilac and honeysuckle phenology data 1956–2014" (Rosemartin et al., 2015). The second workflow was used to check VPOs from the Dutch National Phenology Network (de Natuurkalender). The identified inconsistent observations were reported to the Natuurkalender for further analysis.

The provided consistency information helps scientists and decision makers to address the impact of climate change on vegetation seasonality. Inconsistent-free VPOs are crucial because they can be used to calibrate models, estimating the time of phenophases and its rate of change and dispersion. These estimates not only provide information about vegetation seasonality but they are used as inputs to analyse geographic phenomena in other domain (e.g., agriculture and food security).

Table 6.1 Comparison of the two consistency-check workflows

	The workflow designed in Chapter 2	**The workflow designed in Chapter 3**
Methods	- t-SNE dimensionality reduction - Expectation Maximization - Model-based clustering and intra-cluster boxplot	- Triangulation - Regression modelling - Constraint satisfaction
Environmental variables	- Daily Maximum temperature - Daily minimum temperature - Daily average temperature - Daily precipitation - Daily water vapor pressure - Daily solar radiation - Daily day length - Daily snow water equivalent	Daily average temperature
Role of geographical distance	Implicitly takes the geographical distance between VPOs into account	Explicitly takes the geographical distance between VPOs into account
Assumptions	Environmental contextual variables are very strong that clustering method can group the context conditions of VPOs	- The Average daily temperature drives the synchrony of VPOs - Within 100-km of each VPO, temperature does not change substantially and there is at least one other VPO
Strengths	Use as many environmental contextual variables as available	Allows using domain information (i.e., phenological synchrony driven by temperature).
Limitations	- variables that do not, or less, drive the phenophase may result in meaningless clusters - Weak environmental gradients might result in a small number of clusters	- The optimal distance to connect observations sites cannot be known a priori - The low density of VPOs might also result in low-level connectivity

6.2 Reflection on types and inputs effects of SPP models

SPP models use various (mathematical) formulations and can be executed with multiple gridded datasets and at different spatial resolutions. This results in an array of SPP patterns and trends that deserves further attention. Chapters 4 and 5 present geocomputational workflows designed to analyse the effect of model type and model inputs. Both workflows analyse the results at continental scales and for multiples decades. This provides a comprehensive evaluation of the main SPP models and their inputs to study vegetation seasonality over space and time.

The workflow presented in chapter 4 validates and compares various types of SPP models as well as their patterns and trends. It uses the same data and approach so that we can compare models outputs, and consequently the patterns and trends that can be extracted from these outputs. This workflow uses gridded weather data to validate models calibrated by point weather data. The calibration data (1961-1994) is older than the validation one (2000-2014). This allows evaluating if the models are reliable when their inputs come from different sources and periods.

The workflow described in chapter 5 evaluates the effect of gridded inputs of models, from different sources and at various spatial resolutions. This workflow builds on the one presented in chapter 4. The workflow in chapter 5 uses the most accurate model from chapter 4 (Extended Spring Indices or SI-x). The selected model and gridded input sources (Daymet and gridMET) of the workflow are widely used for environmental studies in the CONUS. Studies that validate SPP models and their inputs often use root-mean-square error (RMSE) at observational sites to evaluate the quality of the model. However, our workflows show the importance of analysing patterns and trends of the models (i.e., spatial and pattern oriented validation). Both workflows use VPOs checked for consistency using the workflow described in chapter 2. Inconsistent-free VPOs provide a more realistic estimation of model performance.

The workflows presented in chapters 4 and 5 provide automatic access to historical and contemporary VPOs and gridded inputs. The spatial aggregation of model inputs, the generation of model outputs and the analysis of their patterns and trends are implemented in open-source programming languages and cloud-based processing environments. These environments support the analysis of 9 million 1-km grid cells and make the workflows reproducible in a few hours; something that would take more than a year using a local machine with one CPU core.

In addition to chapters 4 and 5, which describe the workflows in detail, the workflows are documented in a way that is easy to understand via comments added to the scripts. These comments help to make the workflows more reusable.

6.3 Answers to research questions

RQ1: How to use environmental contextual information to check the consistency of volunteered phenological information?

As shown in chapter 2, contextual environmental information can be used to group VPO sites. These groups can consequently be used to detect outliers. The selected dimensionality reduction method, t-SNE, facilitates the parametrization of model-based clustering methods. In chapter 2 we also show that an expectation maximization (EM) method can be used to cluster data automatically. EM finds the maximum likelihood parameters of mixture models and specifies the mixture component to which each data point belongs. This helps to capture the complexity of the context where VPOs were collected. Unlike global models to check VPO consistency such as boxplots, the use of contextual information allows analysing the observations at a higher level of abstraction (i.e., per clusters). In our case, we found 12 clusters of environmental conditions across CONUS. The proposed workflow works well if relevant contextual information is used (e.g., climate drives phenological events). Besides, our workflow is flexible and allows multiple layers of contextual geoinformation to be integrated into consistency checks.

RQ2: How to integrate domain information (i.e., phenological synchrony) with contextual information to check the consistency of volunteered phenological observations?

Chapter 3 shows that phenological synchrony information can be combined with contextual information to define, propagate and optimize consistency constraints. Our workflow starts by using the Delaunay triangulation to spatially group VPOs. Then, we use a linear regression to model the difference in date of VPOs as a function of accumulated daily temperatures. This regression helps to define consistency constraints that are set by introducing domain knowledge (i.e. the fact that synchrony is mostly driven by temperature). Our workflow, which was used to identify inconsistent VPOs in the Netherlands, was benchmarked against the traditional boxplot filter method. This latter method flagged many VPOs as outliers whereas their DOYs could be explained by differences in environmental conditions (i.e. temperature). Important "by-products" of our workflow are that we could confirm that colder days in late winter and early spring lead to more synchronous spring onset across the Netherlands than warmer ones. Moreover,

our workflow gave us an idea of phenological diversity for plants growing under similar temperature regimes and helped to check the strength of the relationship between synchrony and temperature for the plant species being studied in chapter 3.

RQ3: How to analyse the impact of the type of phenological model on the patterns and trends that can be derived from it?

As demonstrated in chapter 4, VPOs and simulated annealing can be used to (re)calibrate various kinds of SPP models. The quality of these models can then be assessed using, for instance, the consistent contemporary VPOs produced in chapter 2. A cloud-based platform can subsequently be used to efficiently upscale the generation of model outputs so that phenological patterns and trends can be analysed in detail. Cloud computing helps to explore these patterns and trends at various scales, from continental to global. The significance of the temporal trend of the various models show areas with clear phenological changes. In particular, we explored the impact of modelling spring plant phenology using the SI-x, Spring Warming (SW), UNIFORC, Photothermal SW and Photothermal UNIFORC models across CONUS. From a computation point of view, the workflow is highly efficient because we could processes daily temperature data for about 9 million grid cells in a few hours. The presented workflow does require daily gridded inputs, but these are becoming more and more accessible and available at high spatial resolution.

RQ4: How to analyse the effect of using various gridded model inputs and of their spatial resolution on the patterns and trends derived from a phenological model?

As shown in chapter 5, cloud computing also offers an efficient mechanism to spatially aggregate gridded models inputs. This is crucial for large-scale analysis of the impact of input data sources and their spatial resolution. Further, the long-term average of the model's annual output can be used to characterize the phenological spatial patterns derived from the various input data sources and scales. In chapter 5 we also show that a regression analyses can be used to quantify trends and the significance of their difference from different model inputs and scales. More precisely, the workflow created for chapter 5 facilitates the aggregation of gridded model inputs to study and explore multi-scale effects. Our workflow, illustrated using the SI-x models for CONUS, shows that phenological information is more affected by the source of input data than by its spatial resolution.

6.4   Main contributions

Previous consistency checks for VPOs assume that most observations are consistent and hence they only consider their distribution. Here, environmental contextual information is used for checking VPOs consistency. Our checks can use relevant contextual information to identify and filter out inconsistent VPOs automatically.

We also designed an alternative approach to check consistency based on domain knowledge. In particular, on the fact that plant phenophases show a high level of synchrony and that this is mostly driven by temperature. In other words: under the same environmental conditions (temperature) plants of the same species should leaf or bloom at approximately the same time (even in dispersed geographic locations). Synchrony can thus be used for checking the consistency checks of VPOs. This second approach to consistency check relies on a network graph and the definition of constraints to automatically find inconsistent VPOs. This approach thus goes beyond traditional consistency checks. Moreover, this consistency check approach confirmed that colder days in late winter and early spring lead to more synchronous spring onset than warmer ones.

Both consistency checks show that inconsistent observations can significantly bias spatial and temporal trends in spring onset. These checks also show that the most common method to filter outliers in VPOs (boxplot) can lead to misleading synchrony results. From an operational perspective, consistency checks were applied to data from the USA-NPN and the Natuurkalender datasets, helping to curate the data and make it better suited for further analyses.

In this PhD thesis, we go beyond common model validation statistics such as RMSE. Our work considers the variation in space (spatial pattern) and the variation in space and time (spatio-temporal trends). In the past, such an evaluation of patterns and trends in space and time was not easy because of the lack of continuous input data and/or the need for high geocomputational power. However, gridded (weather) inputs and cloud-based geoprocessing tools are nowadays available to support a spatio-temporal validation of (SPP) models.

The evaluation of SPP models confirms that it is not realistic to assume that these models provide similar outputs when using different model inputs. Especially when we use large-scale (e.g., continental scale) data, variations in model outputs and in their patterns and trends are substantial. Evaluating and mapping such variations via the workflows presented in chapters 4 and 5 support the choice of input datasets more suited for phenological modelling

and has the potential to help managers as well as researchers and policy makers to design climate change adaptation and mitigation actions or policies.

Studying the impact of climate change on vegetation seasonality requires the production of large-scale phenological information that is fully transparent and reproducible. Such phenological information is key to support other climate change studies, management and policy making. In this research, an attempt has been made to develop workflows that can run over large areas at a fine spatial resolution.

6.5   Future research avenues

In line with the performed work, several future research avenues are recommended in the following sub-sections:

Combination and improvement of the consistency checks

Although the workflows presented in chapters 2 and 3 are applicable to different situations (see section 6.1), there is a value in combining them. Future work can be directed to replacing the intra-cluster outlier detection step of the workflow presented in chapter 2 with the workflow presented in chapter 3. This replacement helps to explicitly consider the role of geographic distance and might result in more meaningful clusters of contextual conditions. Further, the clustering and graphing steps of these workflows are similar in that both yield homogeneous groups of VPOs. The application of the workflow described in chapter 3 to areas with sparse VPOs might lead to disconnected graphs (Figure *6.2*). In such cases, intra-graph outlier detection may help to identify inconsistent VPOs.

The consistency constraint defined in chapter 3 uses only one contextual variable. In situations where more than one variable drives synchrony, the constraint could be redefined. For example, it is known that autumn plant phenology is mainly influenced by temperature and precipitation (Forkner, 2014). In such cases, multiple linear regression could be used to define consistency constraints. Phenological sequential information such as "flowering must happen earlier than leafing" can also be used to define consistency constraints. For example, Ettinger et al. (2018) recently studied how early spring phenophases define those that follow.
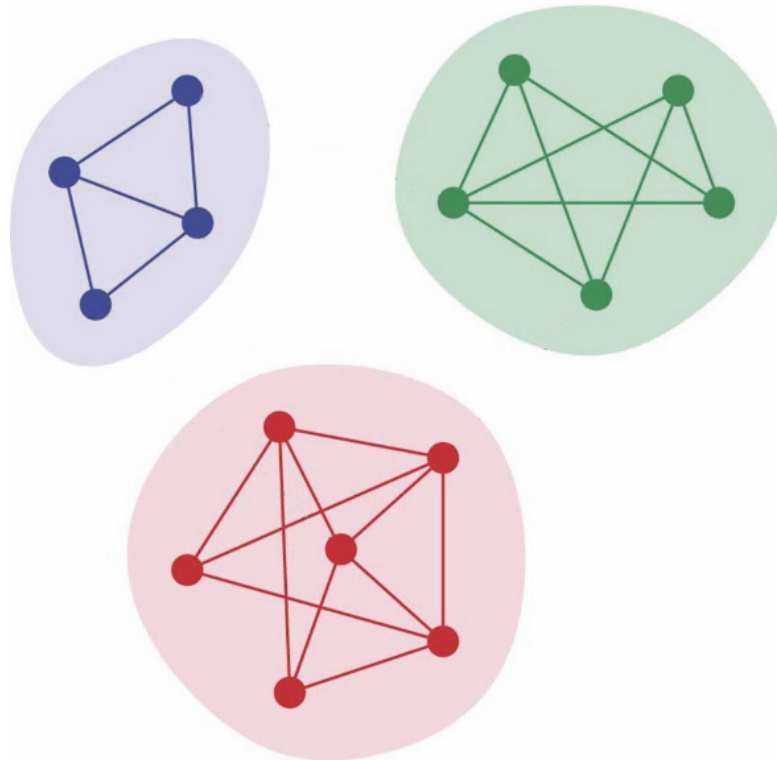
Figure 6.2 Disconnected graph of VPOs as groups that have similar environmental conditions.

Development of new SPP models

The workflows presented in chapters 4 and 5 use existing SPP models. Although these models allow estimating patterns and trends in spring phenology using gridded inputs, they might not be the best models to explain phenological variability in space and time. Developing new models using data-driven approaches and exploiting the increasing availability of gridded environmental data is a worthy line of research. Moreover, existing SPP models often assume that temperature is the key driver of SPP, however, climate is continuously changing, and the importance of phenological drivers might vary. Hence, we recommend analysing the effect of stationarity of SPP model parameters. Such analysis may highlight the need for period-specific SPP models. For example, several studies indicate that a substantial change in the temperature regime of the Northern Hemisphere took place after 1980.

Development and virtualisation of workflows

The workflows described in this thesis were implemented in various programming environments such as R and GEE. This choice was made because

not all methods are functional in GEE (e.g., Delaunay triangulation). Future work is directed towards virtualizing the workflows implementations in a standard unit of software to ensure that scripts and all their dependencies (e.g., libraries) runs quickly and reliably from one computing environment. For example, Docker containers provide such an environment (Boettiger and Carl, 2015). Moreover, open-source cluster-computing frameworks such as Apache Spark are used to enhance the process of large ecological studies. Recently, the Netherlands eScience Center and the University of Twente provided a Spark-based infrastructure for phenological studies (Zurita-Milla et al., 2017). We suggest improving and adding our workflows to this infrastructure so that they are truly scalable.

In this PhD thesis, we focused on reproducible geocomputational workflow for analysing SPP in space and time. There is an urgent need to design workflows that address other principals of phenological geo-information. For example, findability, accessibility, interoperability, and reusability (FAIR; Wilkinson et al., 2016) of this information has not been addressed yet.

Application for other types of VPO and VGI

The increasing interest in volunteered geographic information (VGI) has led to the collection of various national VGI databases. The consistency, or more generally quality, of these databases is often unclear. The workflows presented in chapters 2 and 3 could spark similar work in other disciplines and lead to qualified VGI databases. A natural follow up of our work could be to analyse the increasing number of VPOs on autumn seasonality. Compared to SPP, autumn plant phenology might be derived from other environmental variables than temperature (e.g., precipitation, photoperiod).

The workflow proposed in chapter 2 is generic and could be applied to other types of VGI as long as relevant contextual information is available. In addition, there are various types of VGI (e.g., geotagged tweets) about geographic phenomena (e.g., air pollution or disease in the urban environment) with synchrony. For instance, Morris et al. (2018) have shown strong synchrony in seasonal influenza epidemics across Norway, Sweden, and Denmark. Or, Garcia-Marti et al. (2018) have studied the impact of environmental conditions on the time window when ticks become active in spring. The workflows designed in this PhD thesis could thus be adapted to check the consistency of these observations as well as to evaluate the impact of using various models and/or model inputs to estimate the occurrence of these geographic phenomena.

Appendix A: supplementary material for chapter 3
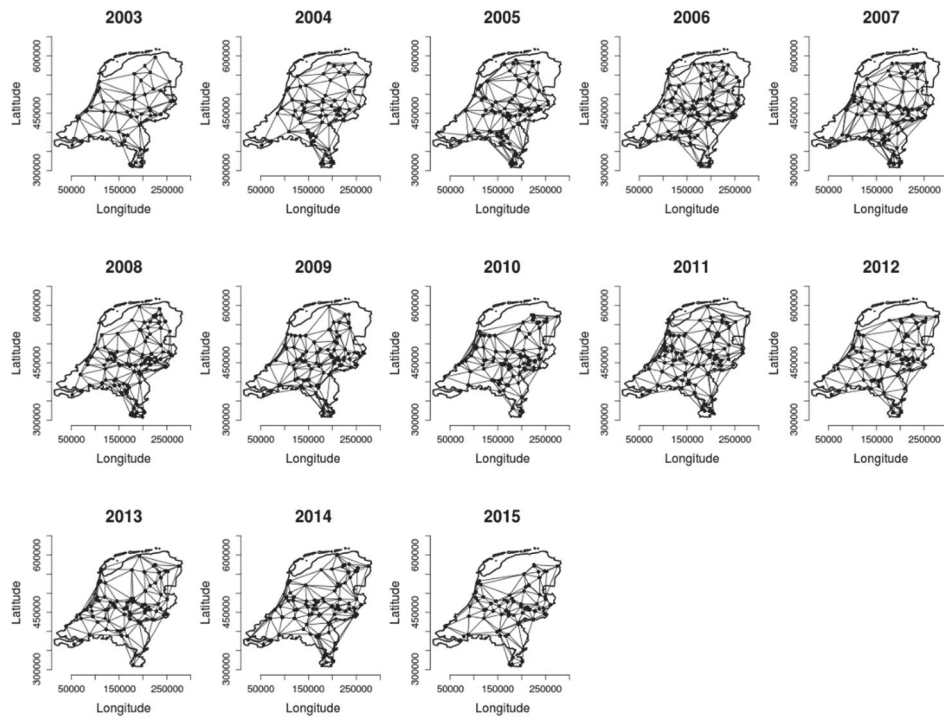
A.1 Annual graphs for wood anemone flowering onset observations. The graphs were made using a Delaunay triangulation and edges longer than 100-km were pruned.
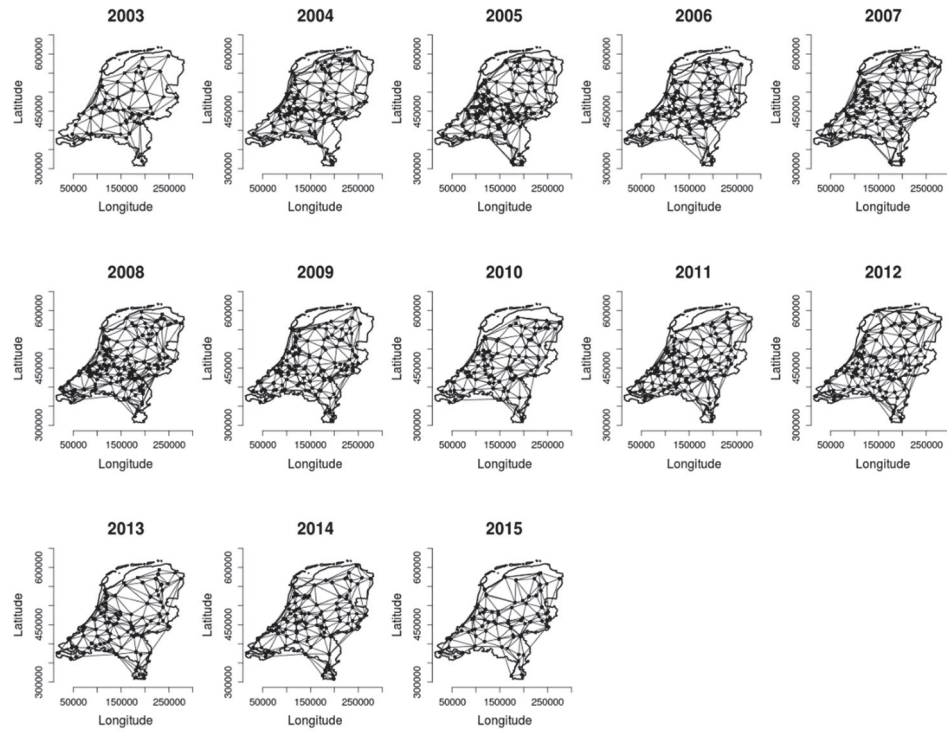
A.2 Annual graphs for cow parsley flowering onset observations. The graphs were made using a Delaunay triangulation and edges longer than 100-km were pruned.
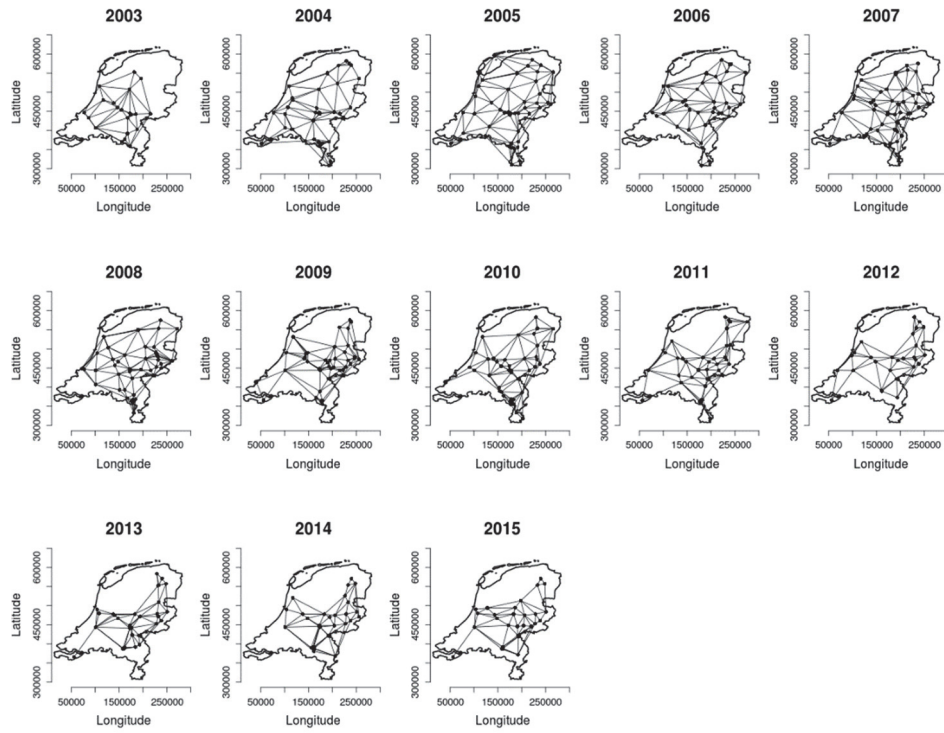
A.3 Annual graphs for cow parsley flowering onset observations. The graphs were made using a Delaunay triangulation and edges longer than 100-km were pruned.
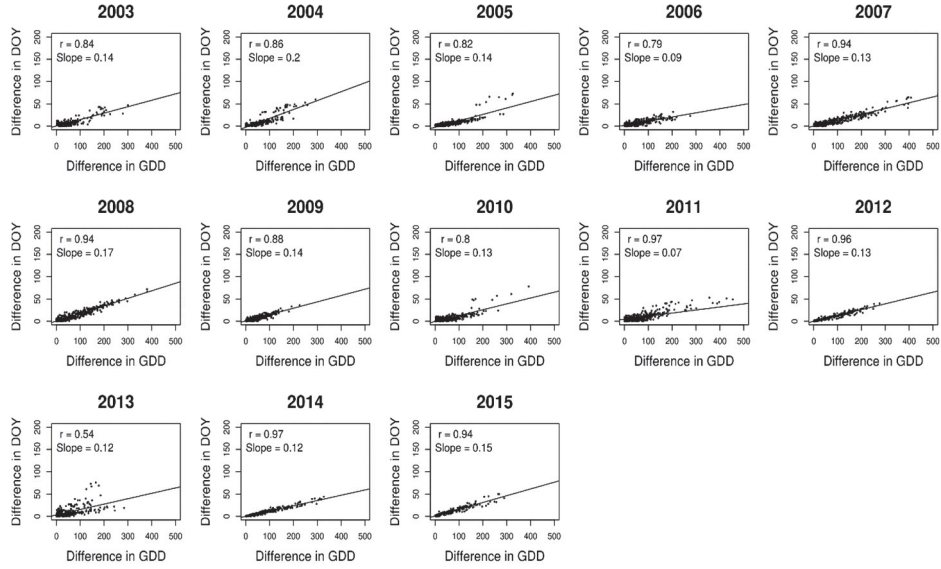
A.4 Linear regression fit between the difference in observed flowering DOY (number of days) and the difference in modelled GDD (number of degree days) for pairs of spatially connected observations of wood anemone flowering onset (see Figure A.1). Correlation coefficient and rate of spatial change in DOY per unit of GDDs (i.e., slope of regression line) are given for each observation year.
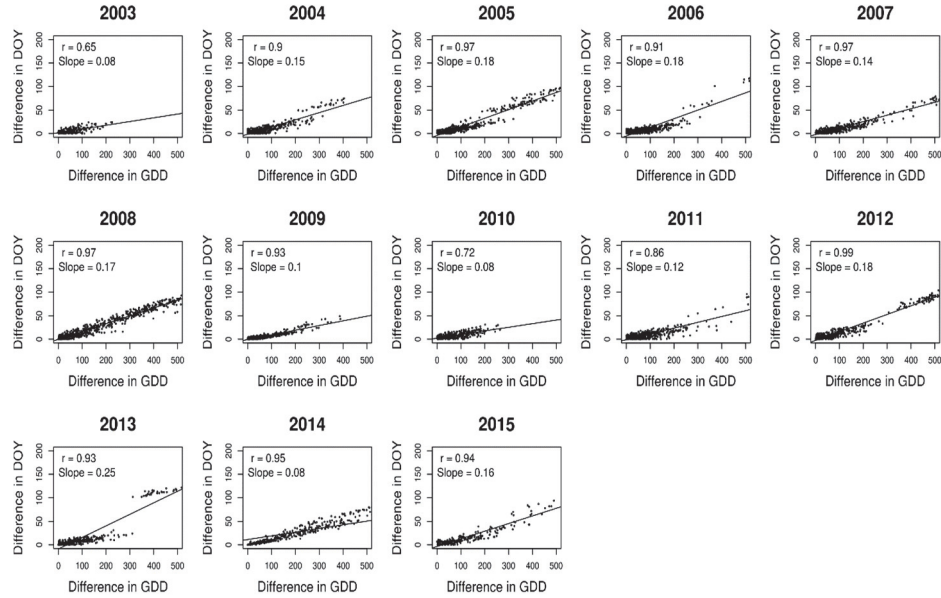
A.5 Linear regression fit between the difference in observed flowering DOY (number of days) and the difference in modelled GDD (number of degree days) for pairs of spatially connected observations of cow parsley flowering onset (see Figure A.2). Correlation coefficient and rate of spatial change in DOY per unit of GDDs (i.e., slope of regression line) are given for each observation year.
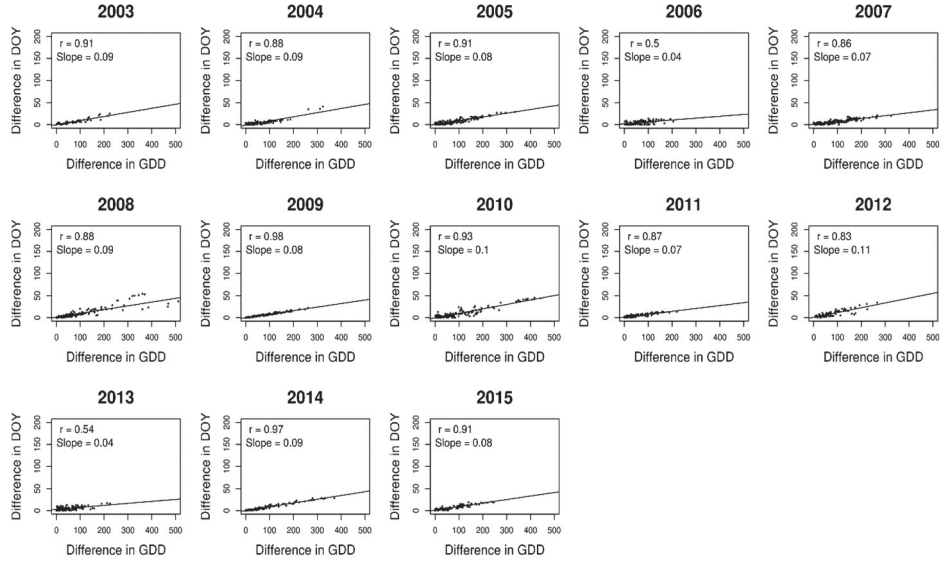
A.6 Linear regression fit between the difference in observed flowering DOY (number of days) and the difference in modelled GDD (number of degree days) for pairs of spatially connected observations of pedunculate oak leafing onset (see Figure A.3). Correlation coefficient and rate of spatial change in DOY per unit of GDDs (i.e., slope of regression line) are given for each observation year.
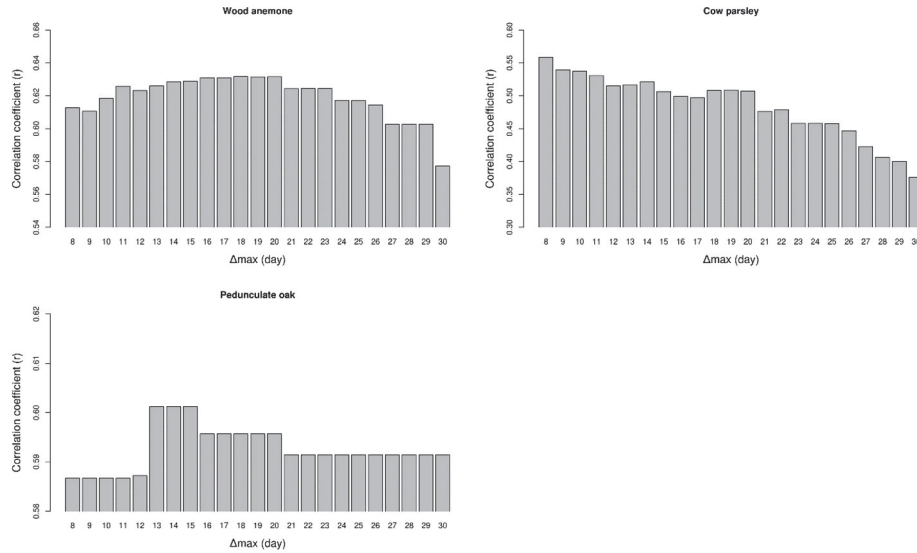
A.7 Correlation coefficient (r) between the standard deviation of the reported DOYs for the flowering onset of wood anemone, cow parseley and pedunculate oak and the average GDD for ΔMax values varying from one week to 30 days.
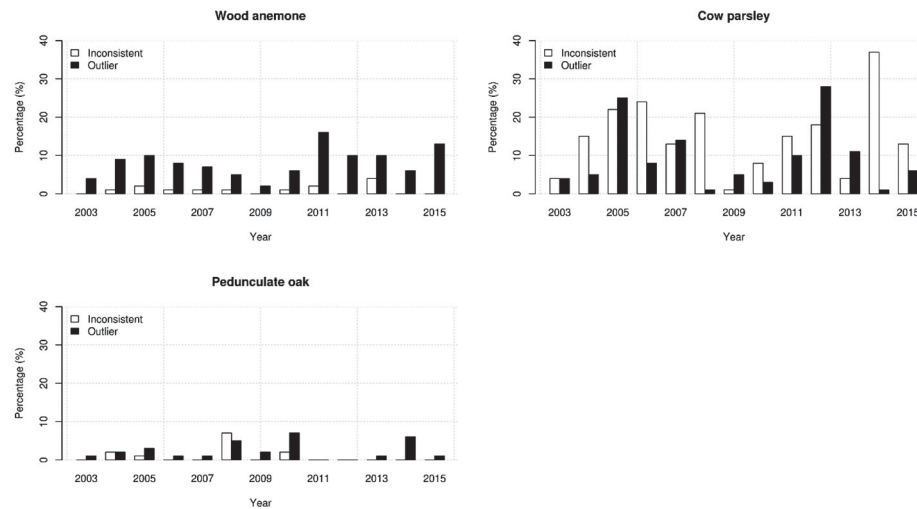
A.8 Annual percentages of inconsistent observations and of boxplot outliers in volunteered observations of wood anemone, cow parseley and pedunculate oak. Inconsistent observations are unusually early or late DOYs with respect to the regional temperature regime of the observation sites, while the outliers are only very early or late DOY.
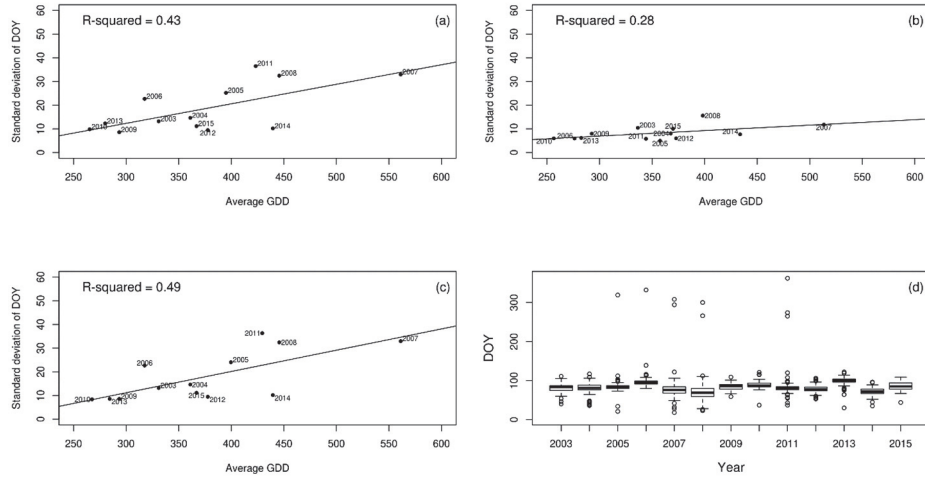
A.9 Wood anemone flowering onset synchrony models for (a) original, (b) outlier-free and (c) consistent observations. Panel (d) shows annual boxplots of the reported DOYs for the original observations.
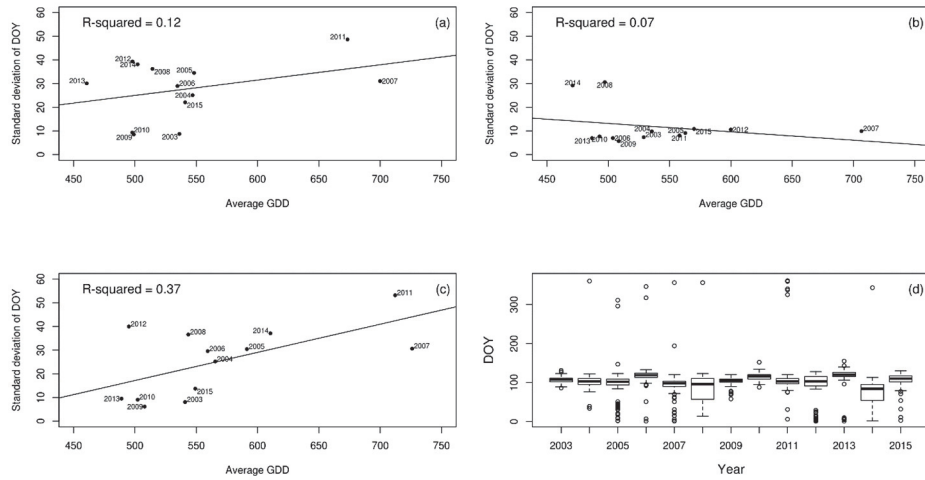
A.10 Cow parsley flowering onset synchrony models for (a) original, (b) outlier-free and (c) consistent observations. Panel (d) shows annual boxplots of the reported DOYs for the original observations.
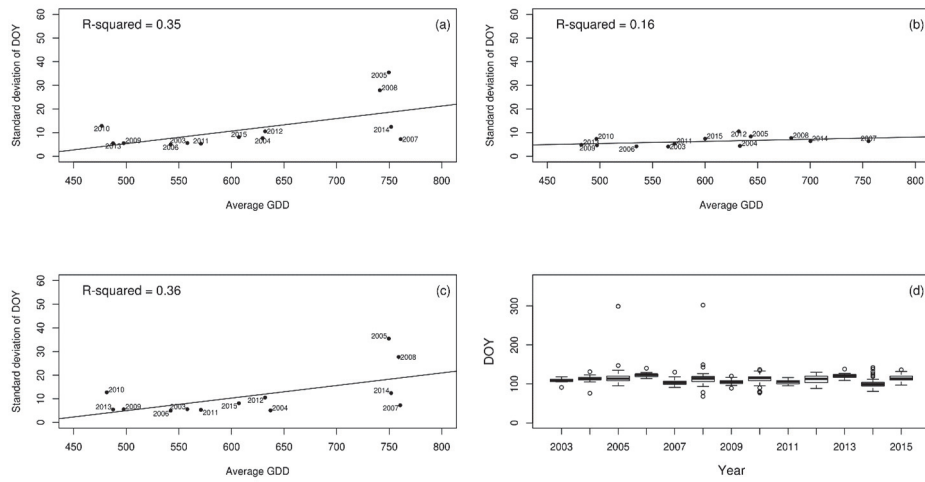
A.11 Pedunculate oak leafing onset synchrony models for (a) original, (b) outlier-free and (c) consistent observations. Panel (d) shows annual boxplots of the reported DOYs for the original observations.

Appendix B: supplementary material for chapter 5
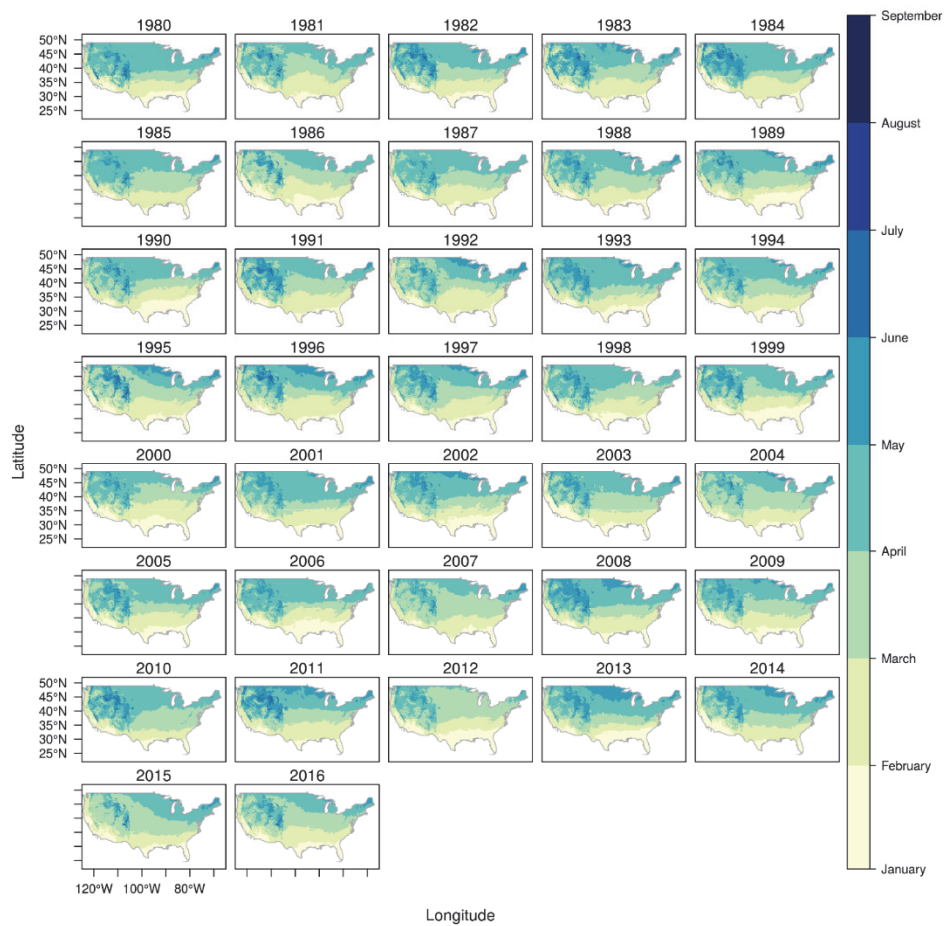
B.1 Annual SI-x FL from 1-km Daymet (1980-2016).
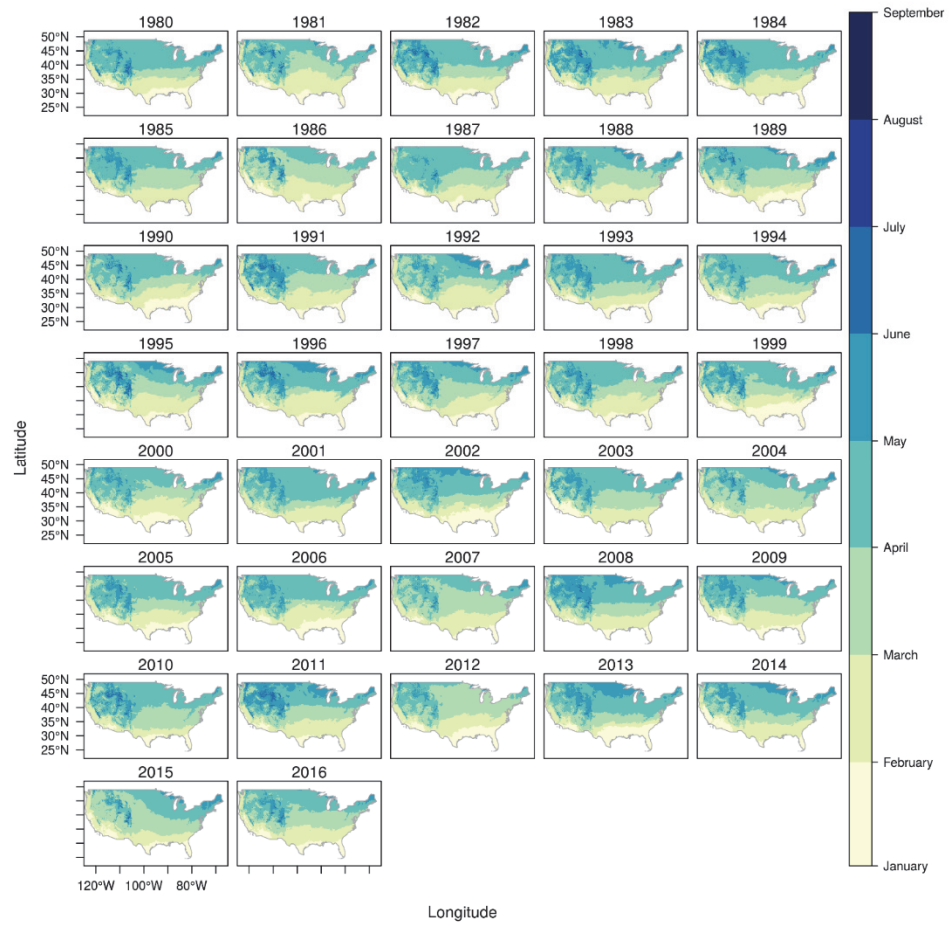
B.2 Annual SI-x FL from 4-km Daymet (1980-2016).
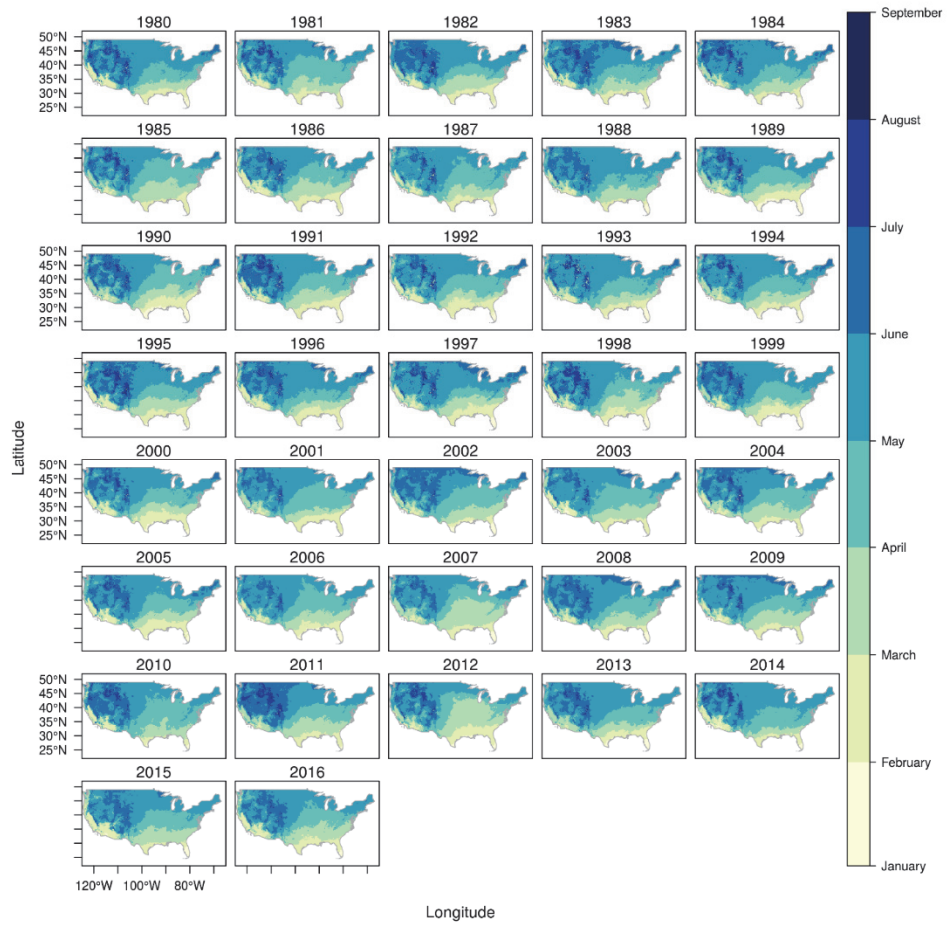
B.3 Annual SI-x FL from 4-km gridMET (1980-2016).
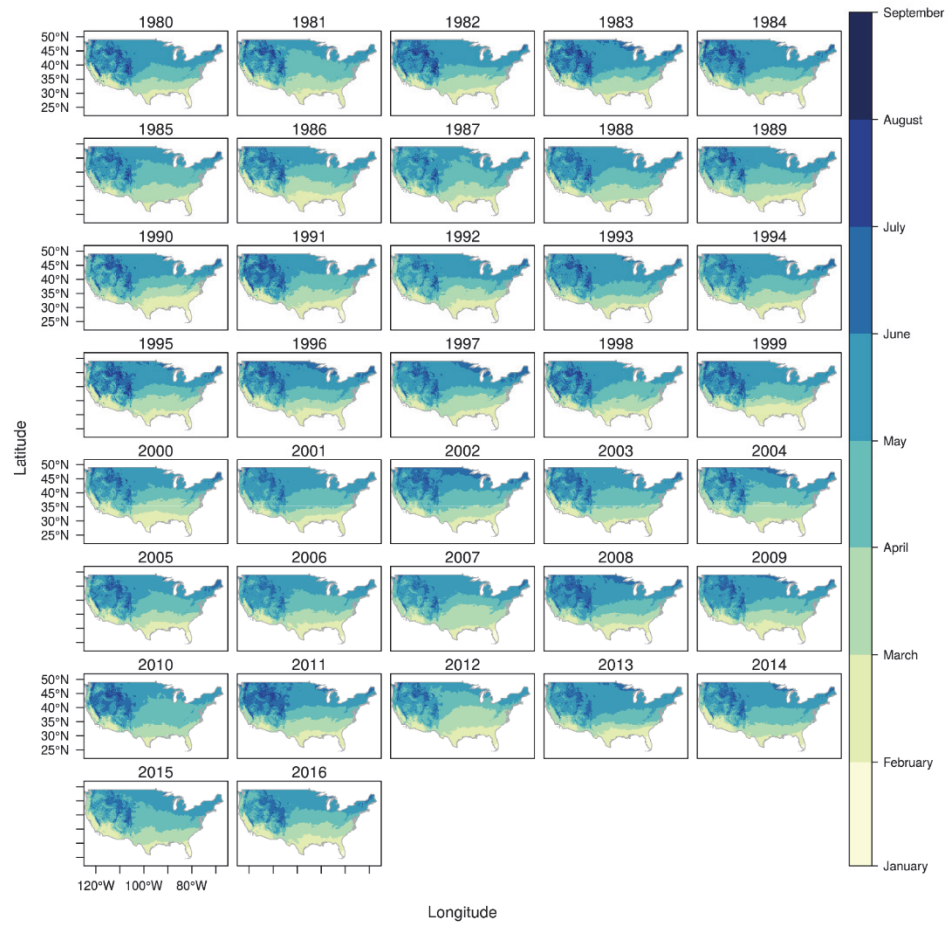
B.4 Annual SI-x FB from 1-km Daymet (1980-2016).
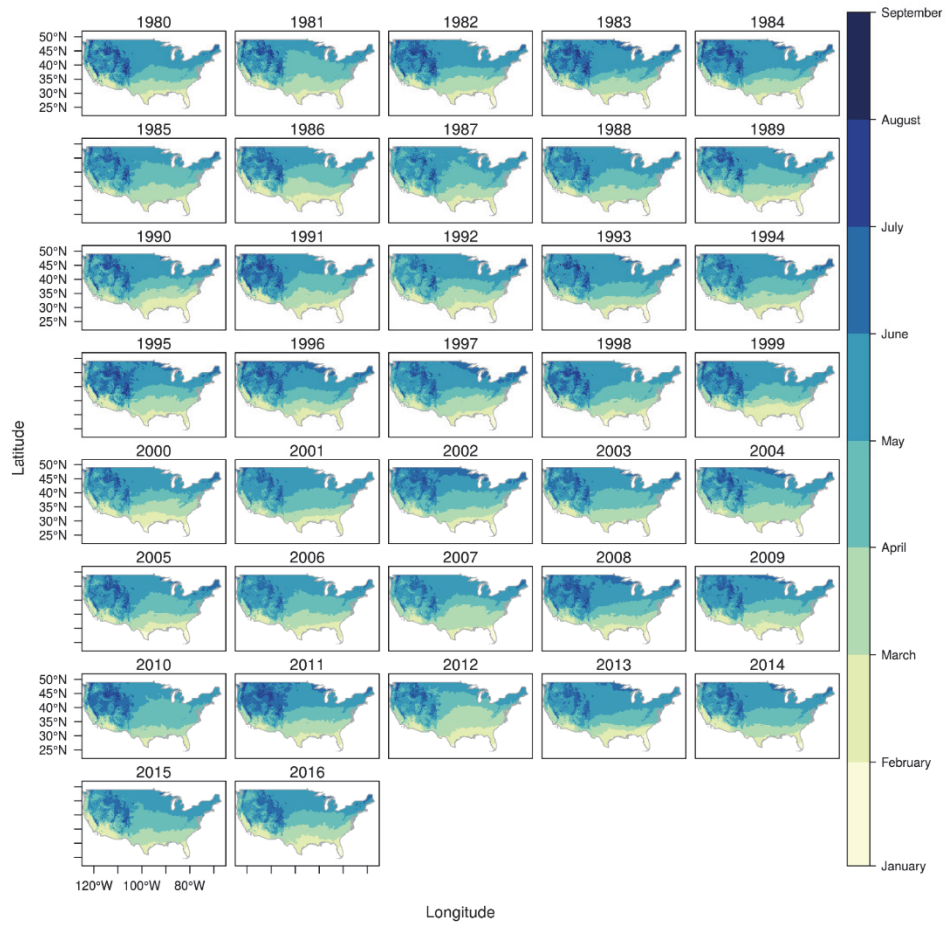
B.5 Annual SI-x FB from 4-km Daymet (1980-2016).
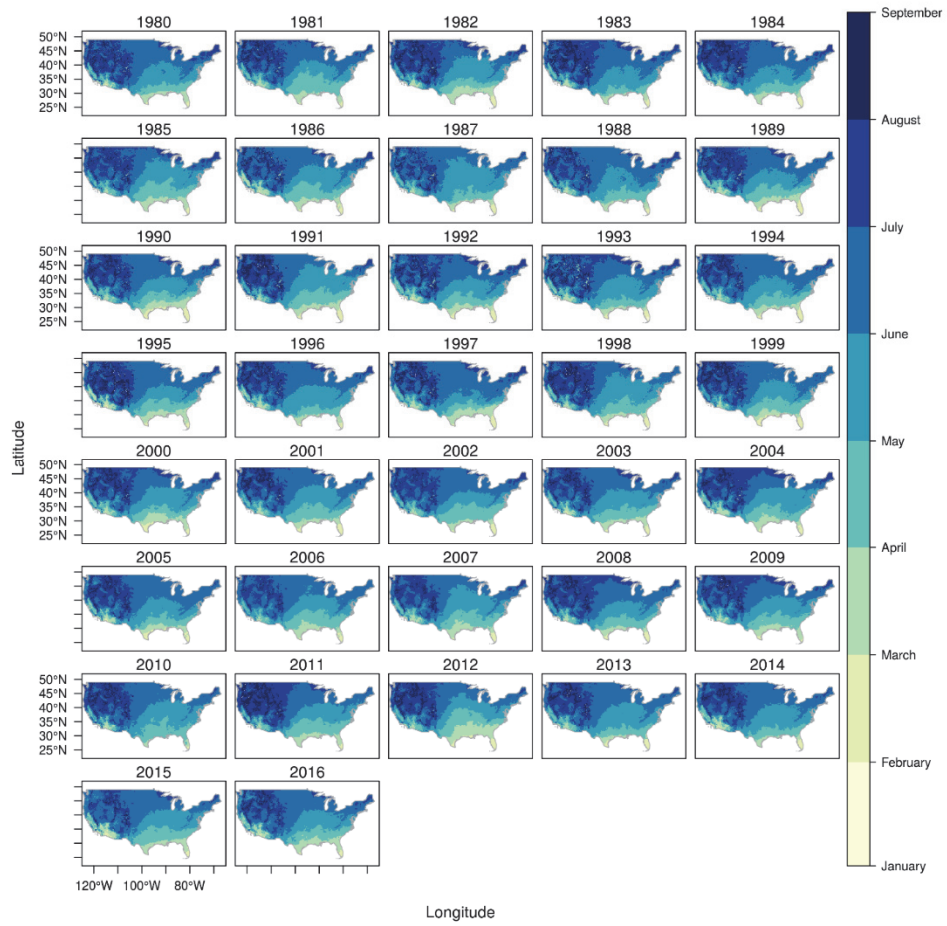
B.6 Annual SI-x FB from 4-km gridMET (1980-2016).
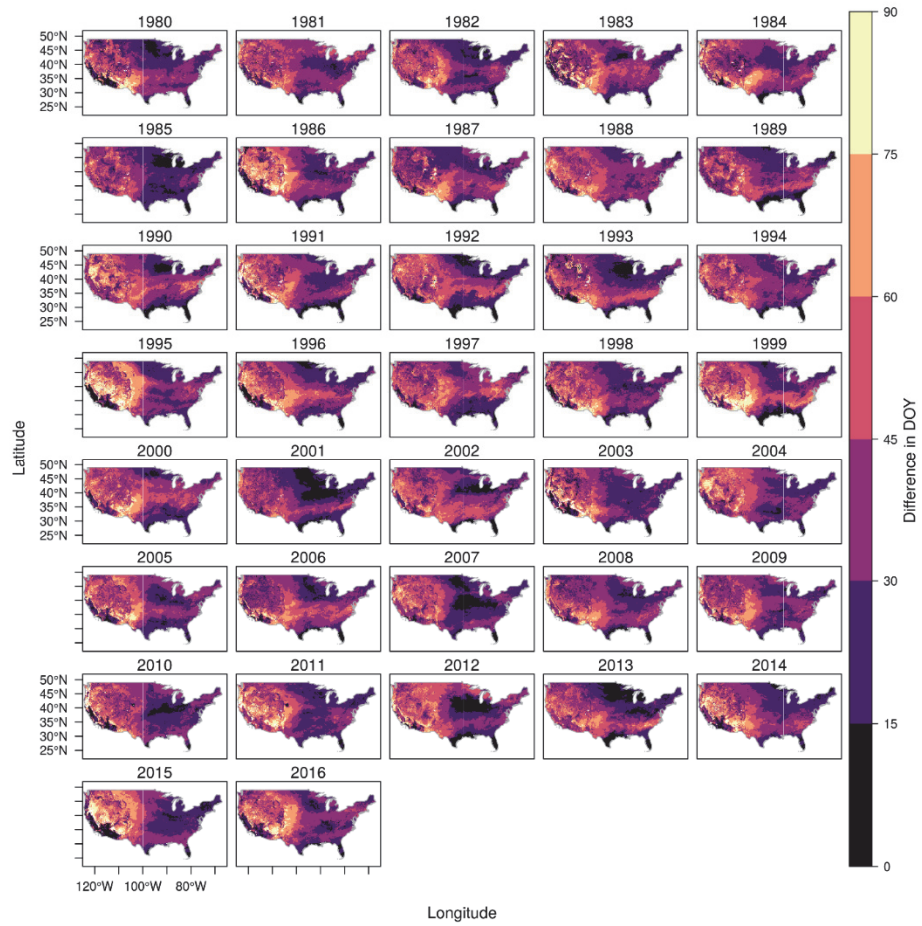
B.7 Maps of the difference between SI-x FL from 1-km Daymet and 4-km gridMET (1980-2016).
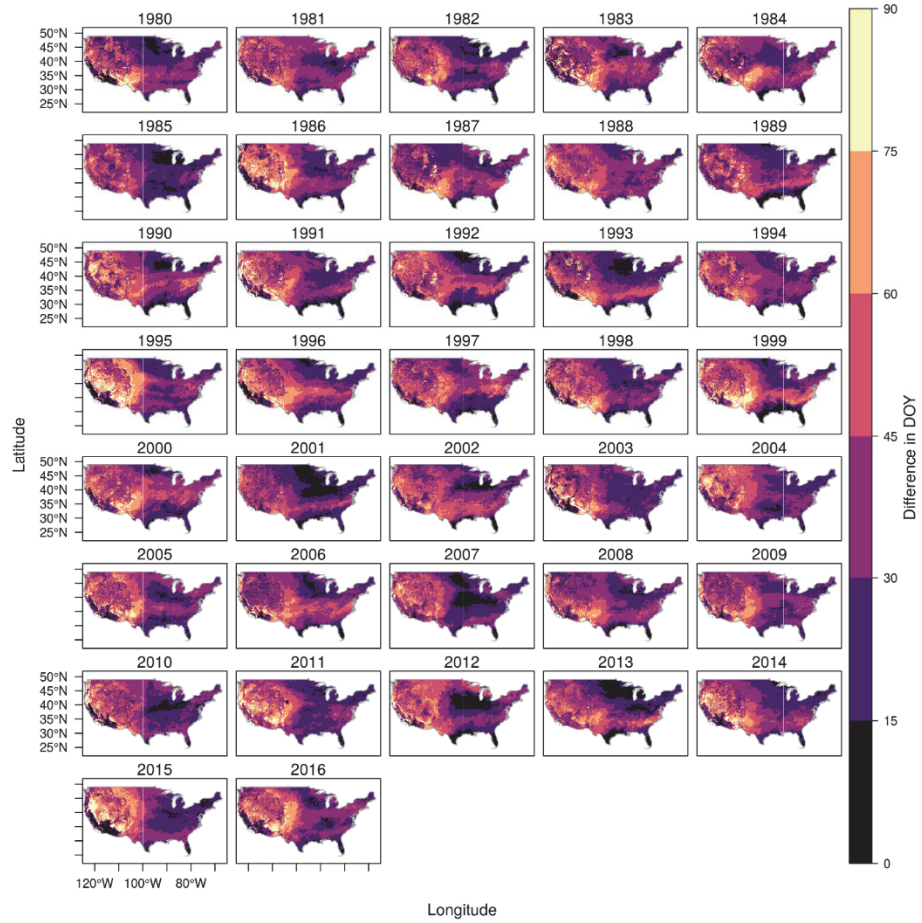
B.8 Maps of the difference between SI-x FL from 4-km Daymet and 4-km gridMET (1980-2016).
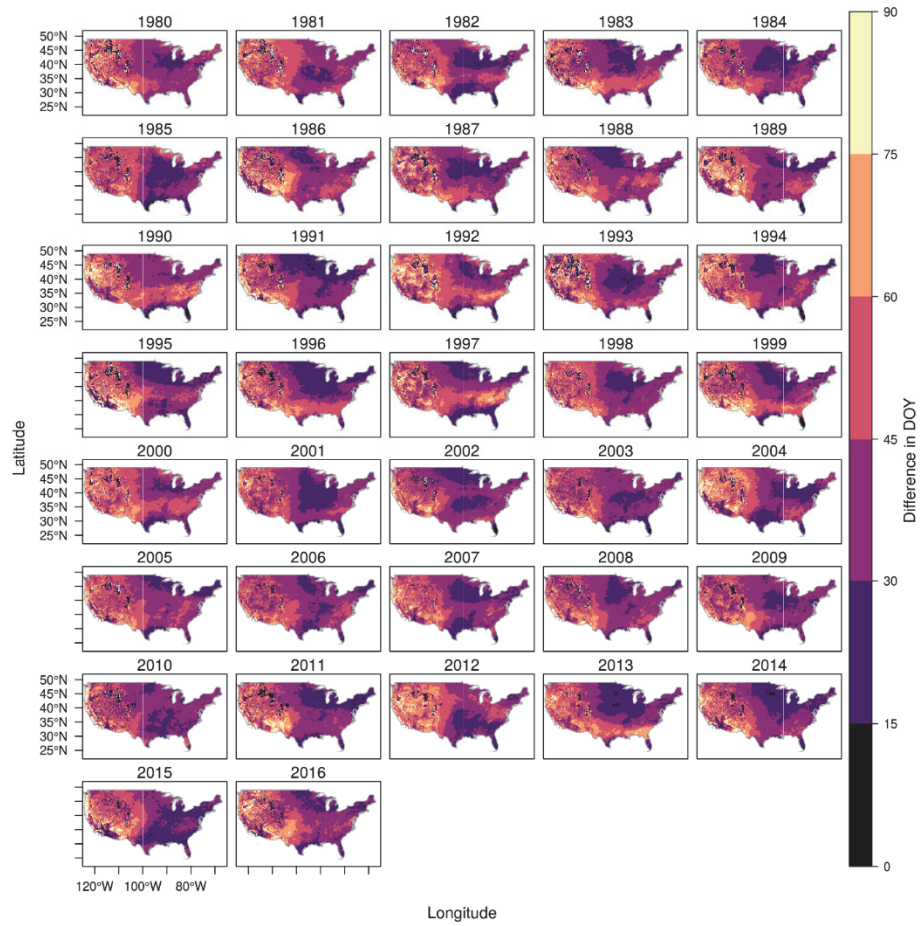
B.9 Maps of the difference between SI-x FB from 1-km Daymet and 4-km gridMET (1980-2016).
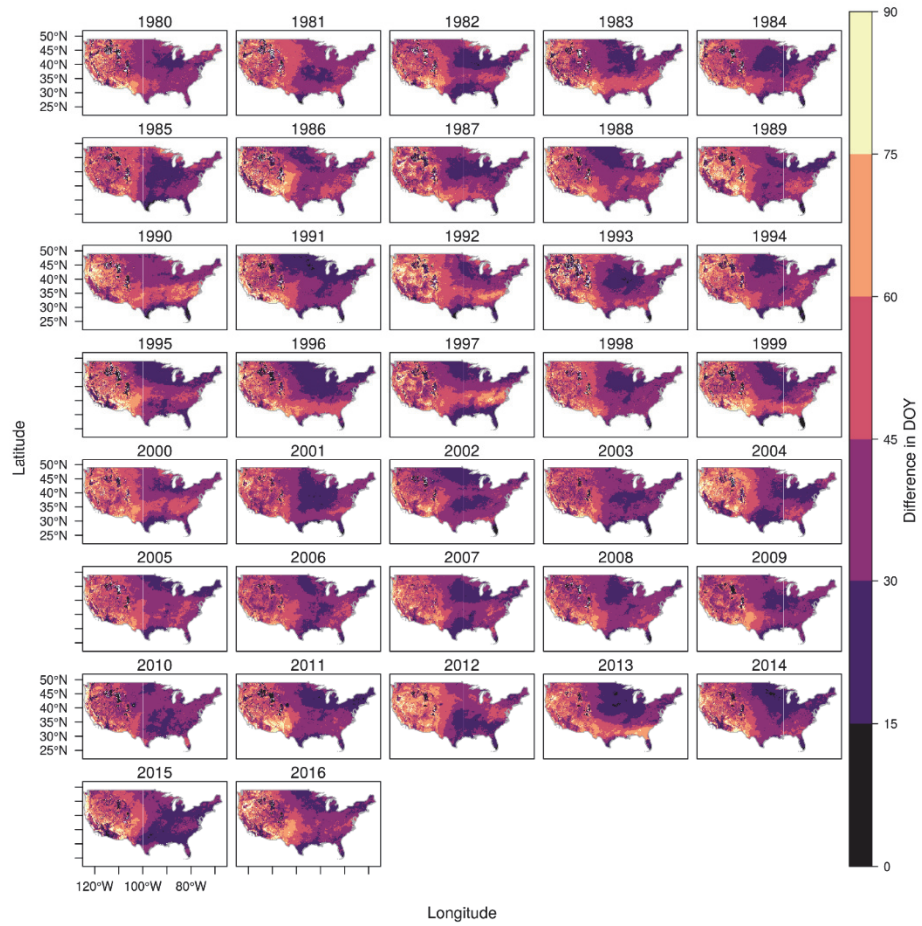
B.10 Maps of the difference between SI-x FB from 4-km Daymet and 4-km gridMET (1980-2016).

B.11 Histograms of the difference between SI-x FL from 4-km SI-x Daymet and 4-km SI-x gridMET (1980-2016).

B.12 Histograms of the difference between SI-x FB from 4-km SI-x Daymet and 4-km SI-x gridMET (1980-2016).
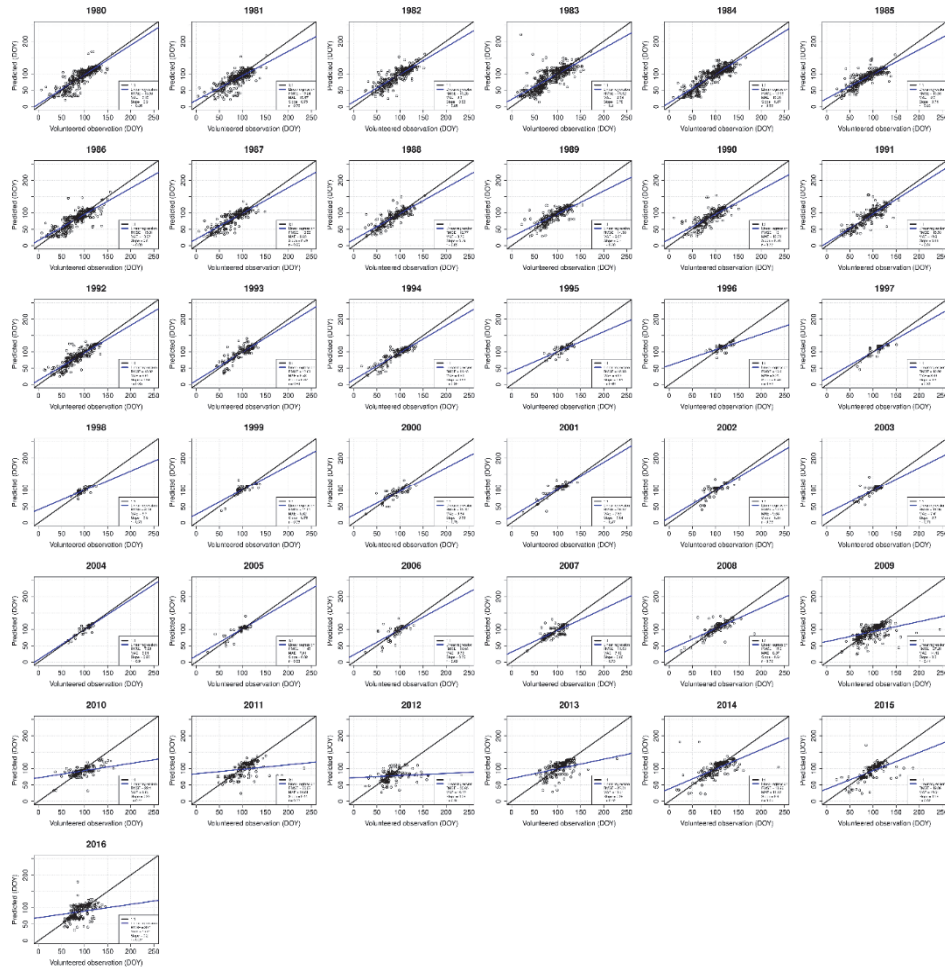
B.13 Scatter plots of volunteered observations versus modelled FL index from 1-km Daymet. The RMSE, MAE, correlation, slope of regression line (blue) fit between observed and estimated and identity line (black) are also included in Scatter plots (1980-2016).
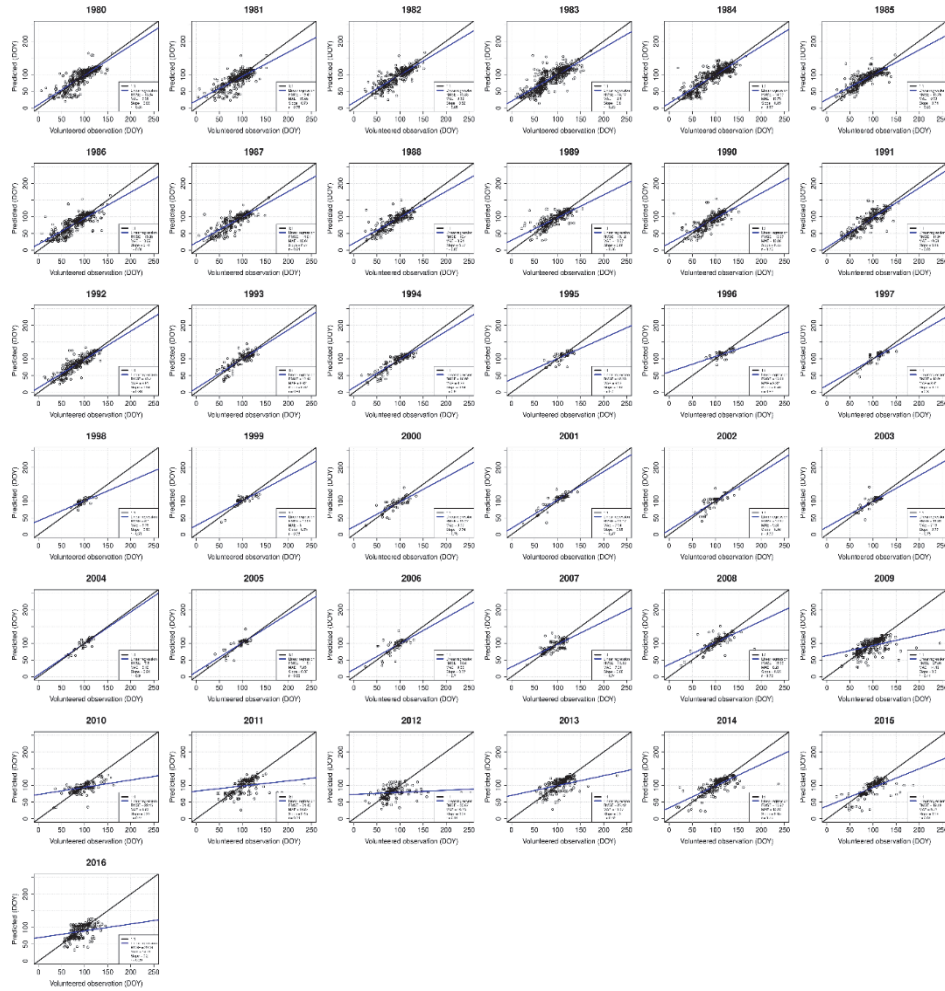
B.14 Scatter plots of volunteered observations versus modelled FL index from 4-km Daymet. The RMSE, MAE, correlation, slope of regression line (blue) fit between observed and estimated and identity line (black) are also included in Scatter plots (1980-2016).
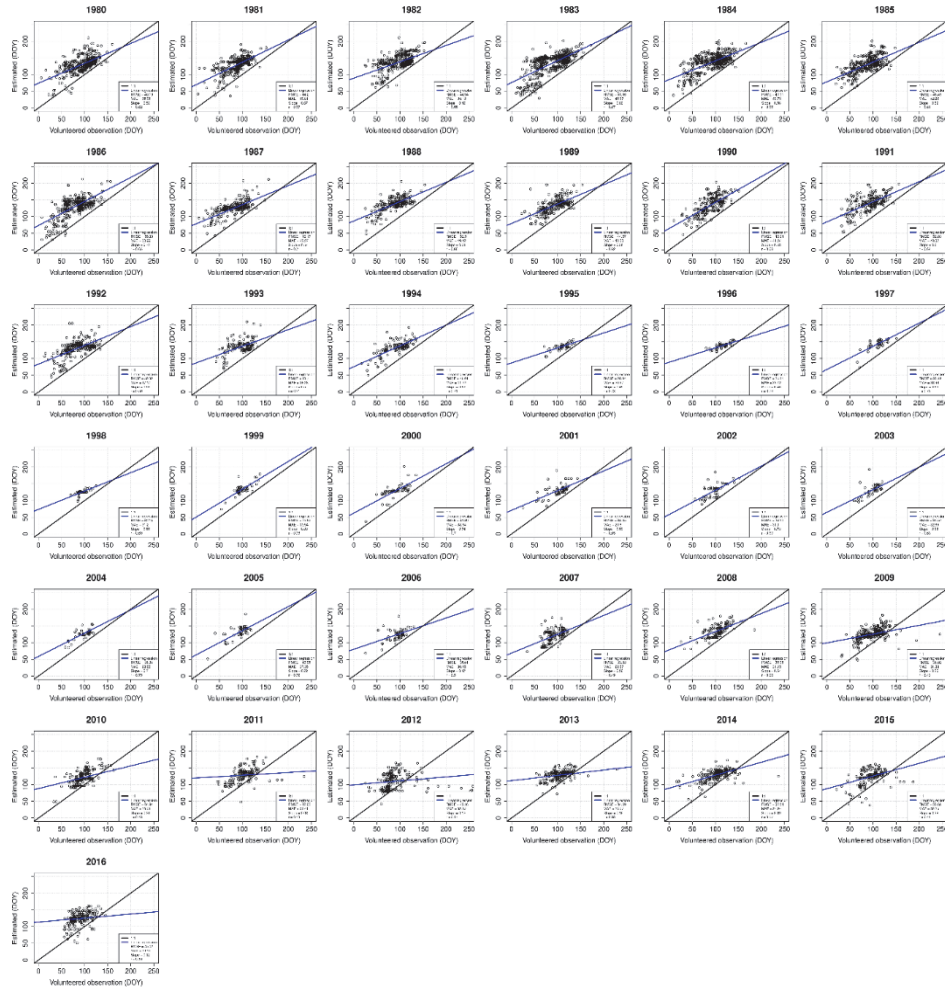
B.15 Scatter plots of volunteered observations versus modelled FL index from 4-km gridMET. The RMSE, MAE, correlation, slope of regression line (blue) fit between observed and estimated and identity line (black) are also included in Scatter plots (1980-2016).
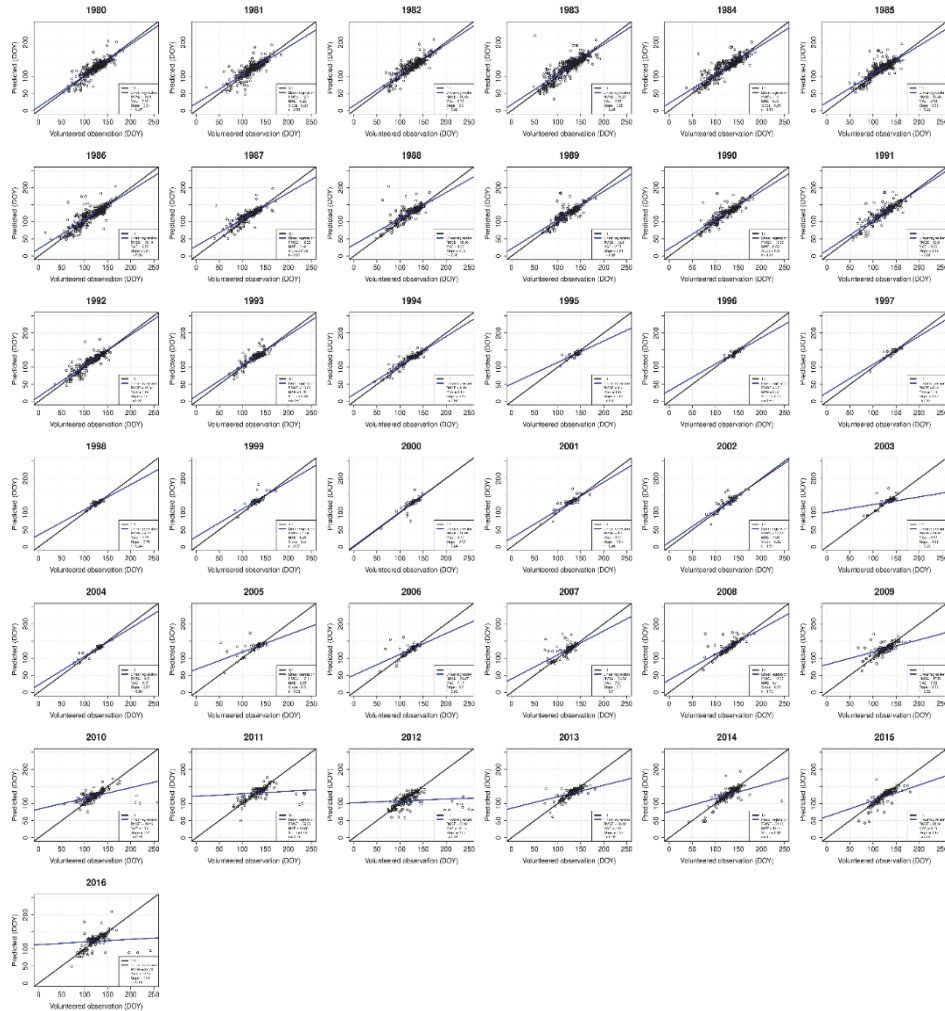
B.16 Scatter plots of volunteered observations versus modelled FB index from 1-km Daymet. The RMSE, MAE, correlation, slope of regression line (blue) fit between observed and estimated and identity line (black) are also included in Scatter plots (1980-2016).
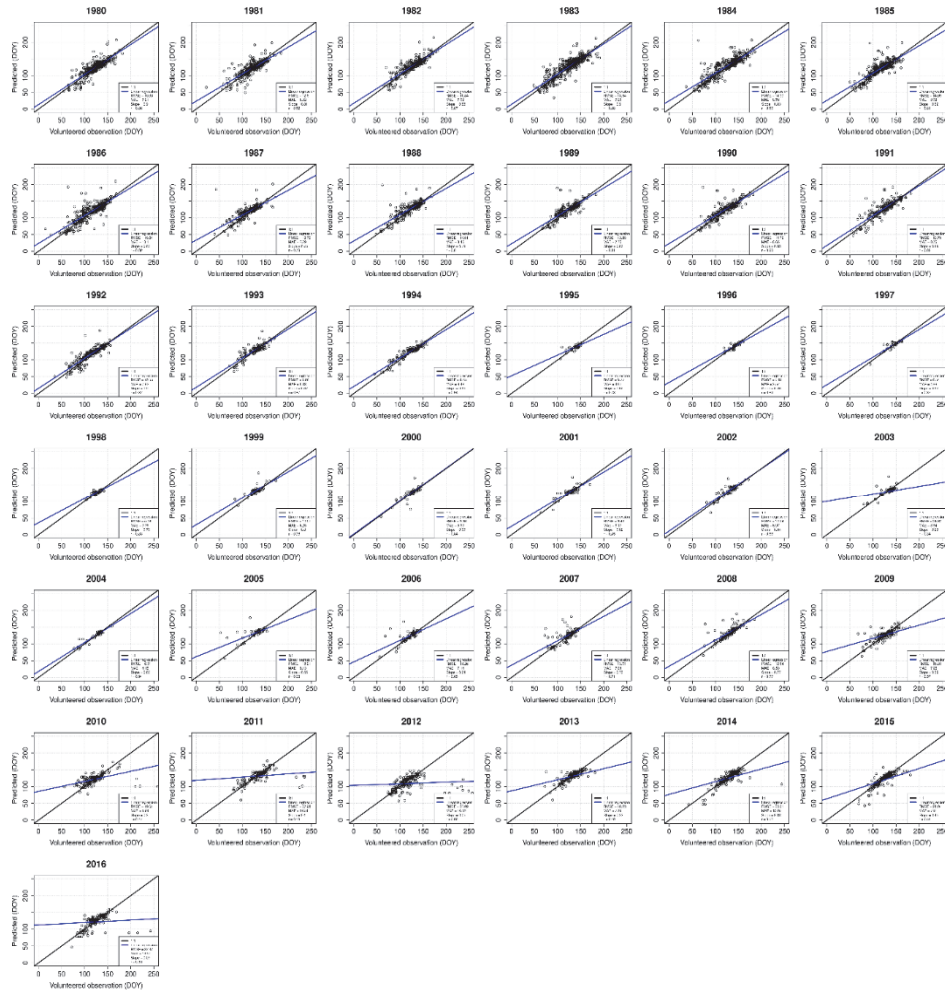
B.17 Scatter plots of volunteered observations versus modelled FB index from 4-km Daymet. The RMSE, MAE, correlation, slope of regression line (blue) fit between observed and estimated and identity line (black) are also included in Scatter plots (1980-2016).
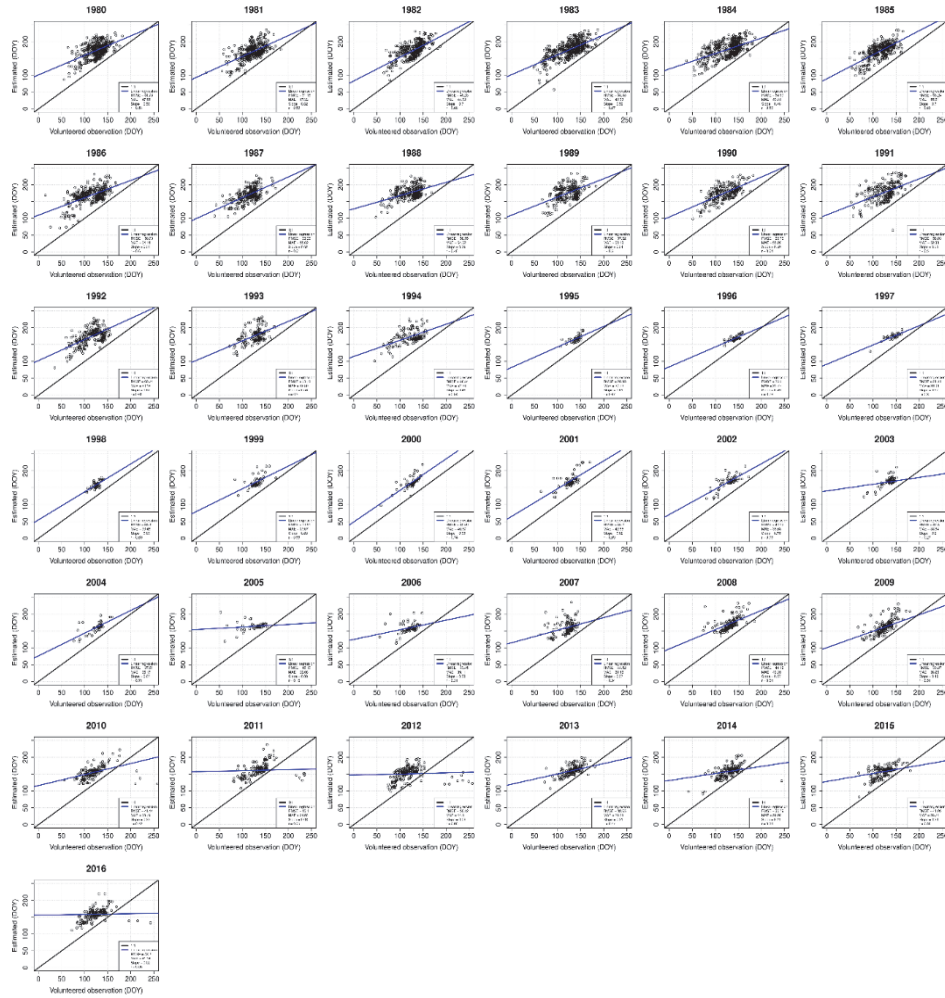
B.18 Scatter plots of volunteered observations versus modelled FB index from 4-km gridMET. The RMSE, MAE, correlation, slope of regression line (blue) fit between observed and estimated and identity line (black) are also included in Scatter plots (1980-2016).

Bibliography

Aarts, E., & Korst, J. (1988). Simulated annealing and boltzmann machines. New York, NY; John Wiley and Sons Inc.

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, *33*(1), 121–131.

Abatzoglou, J. T., & Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(42), 11770–11775.

Ali, A. L., & Schmid, F. (2014). Data Quality Assurance for Volunteered Geographic Information. In *Geographic Information Science* (pp. 126–141). Springer.

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. *Environmental Research Letters*, *10*(10), 104008.

Anyamba, A., & Tucker, C. J. (2005). Analysis of Sahelian vegetation dynamics using NOAA-AVHRR NDVI data from 1981–2003. *Journal of Arid Environments*, *63*(3), 596–614.

Arthur, D., & Vassilvitskii, S. (2007). K-means++: The Advantages of Careful Seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (pp. 1027–1035). Philadelphia, PA, USA: Society for Industrial and Applied Mathematics.

Askeyev, O. V., Tischin, D., Sparks, T. H., & Askeyev, I. V. (2005). The effect of climate on the phenology, acorn crop and radial increment of pedunculate oak (Quercus robur) in the middle Volga region, Tatarstan, Russia. *International Journal of Biometeorology*, *49*(4), 262–266.

Atkinson, M., Gesing, S., Montagnat, J., & Taylor, I. (2017). Scientific workflows: Past, present and future. *Future Generation Computer Systems*, *75*, 216–227.

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, Transactions American Geophysical Union*, *94*(20), 181–182.

Ault, T. R., Macalady, A. K., Pederson, G. T., Betancourt, J. L., & Schwartz, M. D. (2011). Northern Hemisphere modes of variability and the timing of spring in western North America. *Journal of Climate*, *24*(15), 4003–4014.

Ault, T. R., Schwartz, M. D., Zurita-Milla, R., Weltzin, J. F., & Betancourt, J. L. (2015). Trends and natural variability of spring onset in the coterminous United States as evaluated by a new gridded dataset of spring indices : abstract.

Ault, T. R., Zurita-Milla, R., & Schwartz, M. D. (2015). A Matlab© toolbox for

calculating spring indices from daily meteorological data. *Computers & Geosciences*, *83*, 46–53.

Badeck, F.-W., Bondeau, A., Bottcher, K., Doktor, D., Lucht, W., Schaber, J., & Sitch, S. (2004). Responses of spring phenology to climate change. *New Phytologist*, *162*(2), 295–309.

Baeten, L. (2010). Recruitment and performance of forest understorey plants in post-agricultural forests Lander Baeten, 169.

Bakkenes, M., Alkemade, J. R. M., Ihle, F., Leemans, R., & Latour, J. B. (2002). Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. *Global Change Biology*, *8*(4), 390–407.

Ballatore, A., & Zipf, A. (2015). A Conceptual Quality Framework for Volunteered Geographic Information (pp. 89–107). Springer, Cham.

Banfield, J. D., & Raftery, A. E. (1993). Model-Based Gaussian and Non-Gaussian Clustering. *Biometrics*, *49*(3), 803–821.

Barr, A., Black, T. A., & McCaughey, H. (2009). Climatic and Phenological Controls of the Carbon and Energy Balances of Three Contrasting Boreal Forest Ecosystems in Western Canada. In *Phenology of Ecosystem Processes* (pp. 3–34). Springer New York.

Basler, D. (2016). Evaluating phenological models for the prediction of leaf-out dates in six temperate tree species across central Europe. *Agricultural and Forest Meteorology*, *217*, 10–21.

Basler, D., & Körner, C. (2012). Photoperiod sensitivity of bud burst in 14 temperate forest tree species. *Agricultural and Forest Meteorology*, *165*, 73–81.

Batty, M. (2017). Geocomputation. *Environment and Planning B: Urban Analytics and City Science*.

Beaubien, E. G., & Hamann, A. (2011). Plant phenology networks of citizen scientists: recommendations from two decades of experience in Canada. *International Journal of Biometeorology*, *55*(6), 833–841.

Belmecheri, S., Babst, F., Hudson, A. R., Betancourt, J., Trouet, V., Belmecheri, S., … Trouet, V. (2017). Northern Hemisphere Jet Stream Position Indices as Diagnostic Tools for Climate and Ecosystem Dynamics. *Earth Interactions*, *21*(8), 1–23.

Biernacki, C., Celeux, G., & Govaert, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *Pattern Analysis and Machine Intelligence, IEEE Transactions On*, *22*(7), 719–725.

Bimonte, S., Boucelma, O., Machabert, O., & Sellami, S. (2014). From Volunteered Geographic Information to Volunteered Geographic OLAP: A VGI Data Quality-Based Approach. In *Computational Science and Its Applications* (Vol. 8582, pp. 69–80). Springer International Publishing.

Bock, C. E., & Root, T. L. (1981). The Christmas bird count and avian ecology.

Boettiger, C., & Carl. (2015). An introduction to Docker for reproducible research. *ACM SIGOPS Operating Systems Review*, *49*(1), 71–79.

Bolmgren, K., & Eriksson, O. (2015). Are mismatches the norm? Timing of flowering, fruiting, dispersal and germination and their fitness effects in Frangula alnus (Rhamnaceae). *Oikos*, *124*(5), 639–648.

Bonsal, B. R., Zhang, X., Vincent, L. A., Hogg, W. D., Bonsal, B. R., Zhang, X., Vincent, L. A., & Hogg, W. D. (2001). Characteristics of Daily and Extreme Temperatures over Canada. *Journal of Climate*, *14*(9), 1959–1976.

Both, C., Van Asch, M., Bijlsma, R. G., Van Den Burg, A. B., & Visser, M. E. (2009). Climate change and unequal phenological changes across four trophic levels: constraints or adaptations? *Journal of Animal Ecology*, *78*(1), 73–83.

Both, C., & Visser, M. E. (2001). Adjustment to climate change is constrained by arrival date in a long-distance migrant bird. *Nature*, *411*(6835), 296–298.

Boyd, P. W., & Doney, S. C. (2002). Modelling regional responses by marine pelagic ecosystems to global climate change. *Geophysical Research Letters*, *29*(16), 53-1-53–54.

Brabham, D. C. (2009). Crowdsourcing the public participation process for planning projects. *Planning Theory*, *8*(3), 242–262.

Briske, D. D., Joyce, L. A., Polley, H. W., Brown, J. R., Wolter, K., Morgan, J. A., McCarl, B. A., & Bailey, D. W. (2015). Climate-change adaptation on rangelands: linking regional exposure with diverse adaptive capacity. *Frontiers in Ecology and the Environment*, *13*(5), 249–256.

Broussard, M. A., Mas, F., Howlett, B., Pattemore, D., & Tylianakis, J. M. (2017). Possible mechanisms of pollination failure in hybrid carrot seed and implications for industry in a changing climate. *PLOS ONE*, *12*(6), e0180215.

Brunsdon, C., & Comber, A. (2012). Assessing the changing flowering date of the common lilac in North America: a random coefficient model approach. *GeoInformatica*, *16*(4), 675–690.

Burghardt, L. T., Metcalf, C. J. E., Wilczek, A. M., Schmitt, J., & Donohue, K. (2015). Modeling the Influence of Genetic and Environmental Variation on the Expression of Plant Life Cycles across Landscapes. *The American Naturalist*, *185*(2), 212–227.

Butcher, G. S., & Niven, D. K. (2007). Combining data from the Christmas Bird Count and the Breeding Bird Survey to determine the continental status and trends of North America birds. *National Audubon Society*.

Calinger, K. M., Queenborough, S., & Curtis, P. S. (2013). Herbarium specimens reveal the footprint of climate change on flowering trends across north-central North America. *Ecology Letters*, *16*(8), 1037–1044.

Cannell, M. G. R., & Smith, R. I. (1983). Thermal Time, Chill Days and Prediction of Budburst in Picea sitchensis. *The Journal of Applied Ecology*, *20*(3), 951.

Capiro, J. (1993). Western Regional Phenological Summary of Information on Honeysuckle and Lilac First Bloom Phase Covering the Period 1956-1991.

Caprio, J. M. (1974). The Solar Thermal Unit Concept in Problems Related to Plant Development and Potential Evapotranspiration (pp. 353–364). Springer, Berlin, Heidelberg.

Cayan, D. R., Dettinger, M. D., Kammerdiener, S. A., Caprio, J. M., & Peterson, D. H. (2001). Changes in the Onset of Spring in the Western United States. *Bulletin of the American Meteorological Society*, *82*(3), 399–415.

Chen, X. (2017). Process-Based Simulation and Prediction of Plant Phenology Spatiotemporal Variations (pp. 45–66). Springer Berlin Heidelberg.

Chew, P. L. (1989). Constrained delaunay triangulations. *Algorithmica*, *4*(1), 97–108.

Chmielewski, F. M. (2013). Phenology in Agriculture and Horticulture. In *Phenology: An Integrative Environmental Science* (pp. 539–561). Springer.

Chmielewski, F. M., Götz, K.-P., Rawel, H. M., & Homann, T. (2014). Towards new approaches in phenological modelling. *EGU General Assembly 2014, Held 27 April - 2 May, 2014 in Vienna, Austria, Id.2205*, *16*.

Chmielewski, F. M., Müller, A., & Bruns, E. (2004). Climate changes and trends in phenology of fruit trees and field crops in Germany, 1961–2000. *Agricultural and Forest Meteorology*, *121*(1–2), 69–78.

Chuine, I., Bonhomme, M., Legave, J.-M., García de Cortázar-Atauri, I., Charrier, G., Lacointe, A., & Améglio, T. (2016). Can phenological models predict tree phenology accurately in the future? The unrevealed hurdle of endodormancy break. *Global Change Biology*, *22*(10), 3444–3460.

Chuine, I., Cambon, G., & Comtois, P. (2000). Scaling phenology from the local to the regional level: advances from species-specific phenological models. *Global Change Biology*, *6*(8), 943–952.

Chuine, I., de Cortazar-Atauri, I. G., Kramer, K., & Hänninen, H. (2013). Plant Development Models. In *Phenology: An Integrative Environmental Science* (pp. 275–293). Dordrecht: Springer Netherlands.

Chuine, I., Yiou, P., Viovy, N., Seguin, B., Daux, V., & Ladurie, E. L. R. (2004). Grape ripening as a past climate indicator. *Nature*, *432*(7015), 289–290.

Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A., & Schwartz, M. D. (2007). Shifting plant phenology in response to global change. *Trends in Ecology & Evolution*, *22*(7), 357–365.

Cohen-Boulakia, S., Belhajjame, K., Collin, O., Chopard, J., Froidevaux, C., Gaignard, A., … Blanchet, C. (2017). Scientific workflows for computational reproducibility in the life sciences: Status, challenges and opportunities. *Future Generation Computer Systems*, *75*, 284–298.

Coleman, D. J., Georgiadou, Y., & Labonte, J. (2009). Volunteered geographic information: The nature and motivation of produsers. *International Journal of Spatial Data Infrastructures Research*.

Comber, A., See, L., Fritz, S., Van der Velde, M., Perger, C., & Foody, G. (2013). Using control data to determine the reliability of volunteered geographic information about land cover. *International Journal of Applied*

Earth Observation and Geoinformation, *23*, 37–48.

Cook, B. I., Wolkovich, E. M., Davies, T. J., Ault, T. R., Betancourt, J. L., Allen, J. M., … Travers, S. E. (2012). Sensitivity of Spring Phenology to Warming Across Temporal and Spatial Climate Gradients in Two Independent Databases. *Ecosystems*, *15*(8), 1283–1294.

Council, N. R. (1998). Future of the National Weather Service Cooperative Observer Network. Washington, DC: The National Academies Press.

Črepinšek, Z., Kajfež-Bogataj, L., & Bergant, K. (2006). Modelling of weather variability effect on fitophenology. *Ecological Modelling*, *194*(1–3), 256–265.

Crimmins, A., Kolian, M., Bacanskas, L., & Rosseel, K. (2016). *Climate Change Indicators in the United States 2016 (fourth edition)*.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., & Pasteris, P. P. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, *28*(15), 2031–2064.

Daly, C., Smith, J. W., Smith, J. I., & McKane, R. B. (2007). High-Resolution Spatial Modeling of Daily Weather Elements for a Catchment in the Oregon Cascade Mountains, United States. *Journal of Applied Meteorology and Climatology*, *46*(10), 1565–1586.

De Frenne, P., Kolb, A., Verheyen, K., Brunet, J., Chabrerie, O., Decocq, G., … Graae, B. J. (2009). Unravelling the effects of temperature, latitude and local environment on the reproduction of forest herbs. *Global Ecology and Biogeography*, *18*(6), 641–651.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B*, *39*(1), 1–38.

Denny, E., Gerst, K., Miller-Rushing, A., Tierney, G., Crimmins, T., Enquist, C. F., … Weltzin, J. (2014). Standardized phenology monitoring methods to track plant and animal activity for science and resource management applications. *International Journal of Biometeorology*, *58*(4), 591–601.

Devictor, V., Whittaker, R. J., & Beltrame, C. (2010). Beyond scarcity: citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions*, *16*(3), 354–362.

Dhar, M., Sharma, M., & Bhat, A. (2017). Functional genomics of apocarotenoids in saffron: insights from chemistry, molecular biology and therapeutic applications. *Briefings in Functional*.

Dickinson, J. L., Zuckerberg, B., & Bonter, D. N. (2010). Citizen science as an ecological research tool: challenges and benefits. *Annual Review of Ecology, Evolution, and Systematics*, *41*, 149–172.

Doi, H., & Katano, I. (2008). Phenological timings of leaf budburst with climate change in Japan. *Agricultural and Forest Meteorology*, *148*(3), 512–516.

Donaldson, J. (2010). Package 'tsne.'

Donnelly, A., Crowe, O., Regan, E., Begley, S., & Caffarra, A. (2014). The role of citizen science in monitoring biodiversity in Ireland. *International Journal of Biometeorology*, *58*(6), 1237–1249.

Ehlen, J., Caldwell, D. R., & Harding, S. (2002). GeoComputation: what is it? *Computers, Environment and Urban Systems*, *26*(4), 257–265.

Ellwood, E. R., Temple, S. A., Primack, R. B., Bradley, N. L., & Davis, C. C. (2013). Record-Breaking Early Flowering in the Eastern United States. *PLoS ONE*, *8*(1), e53788.

Elwood, S. (2008). Volunteered geographic information: key questions, concepts and methods to guide emerging research and practice. *Geojournal*, *72*(3), 133–135.

Elwood, S., Goodchild, M. F., & Sui, D. (2013). Prospects for VGI Research and the Emerging Fourth Paradigm. In *Crowdsourcing Geographic Knowledge* (pp. 361–375). Springer Netherlands.

English-Loeb, G. M., & Karban, R. (1992). Consequences of variation in flowering phenology for seed head herbivory and reproductive success in Erigeron glaucus (Compositae). *Oecologia*, *89*(4), 588–595.

Enquist, C. A. F., Kellermann, J. L., Gerst, K. L., & Miller-Rushing, A. J. (2014). Phenology research for natural resource management in the United States. *International Journal of Biometeorology*, *58*(4), 579–589.

Ettinger, A. K., Gee, S., & Wolkovich, E. M. (2018). Phenological sequences: how early-season events define those that follow. *American Journal of Botany*, *105*(10), 1771–1780.

Feick, R., & Roche, S. (2013). Understanding the Value of VGI. In *Crowdsourcing Geographic Knowledge* (pp. 15–29). Springer Netherlands.

Ferster, C. J., & Coops, N. C. (2013). A review of earth observation using mobile personal communication devices. *Computers & Geosciences*, *51*, 339–349.

Fitchett, J. M., Grab, S. W., & Thompson, D. I. (2015). Plant phenology and climate change. *Progress in Physical Geography*, *39*(4), 460–482.

Flanagin, A. J., & Metzger, M. J. (2008). The credibility of volunteered geographic information. *Geojournal*, *72*(3–4), 137–148.

Forkner, R. E. (2014). Simulated herbivory advances autumn phenology in Acer rubrum. *International Journal of Biometeorology*, *58*(4), 499–507.

Fowler, H. J., Blenkinsop, S., & Tebaldi, C. (2007). Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, *27*(12), 1547–1578.

Fraga, H., García de Cortázar Atauri, I., Malheiro, A. C., & Santos, J. A. (2016). Modelling climate change impacts on viticultural yield, phenology and stress conditions in Europe. *Global Change Biology*, *22*(11), 3774–3788.

Fraley, C., & Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical*

Association, *97*(458), 611–631.

Fraley, C., Raftery, A. E., & Scrucca, L. (2012). Normal mixture modeling for model-based clustering, classification, and density estimation. *Department of Statistics, University of Washington*.

Frigge, M., Hoaglin, D. C., & Iglewicz, B. (1989). Some Implementations of the Boxplot. *The American Statistician*, *43*(1), 50–54.

Fuccillo, K. K., Crimmins, T. M., de Rivera, C. E., & Elder, T. S. (2015). Assessing accuracy in citizen science-based plant phenology monitoring. *International Journal of Biometeorology*, *59*(7), 917–926.

Galindo, A., Díaz, P. L., & Huerta, J. (2011). A quality approach to volunteer geographic information. *Proceedings of International Symposium on Spatial Data Quality*.

Garcia-Marti, I., Zurita-Milla, R., Harms, M. G., & Swart, A. (2018). Using volunteered observations to map human exposure to ticks. *Scientific Reports*, *8*(1), 15435.

García de Cortázar-Atauri, I., Brisson, N., & Gaudillere, J. P. (2009). Performance of several models for predicting budburst date of grapevine (Vitis vinifera L.). *International Journal of Biometeorology*, *53*(4), 317–326.

Gerst, K. L., Kellermann, J. L., Enquist, C. A. F., Rosemartin, A. H., & Denny, E. G. (2016). Estimating the onset of spring from a complex phenology database: trade-offs across geographic scales. *International Journal of Biometeorology*, *60*(3), 391–400.

Gerstmann, H., Doktor, D., Gläßer, C., & Möller, M. (2016). PHASE: A geostatistical model for the Kriging-based spatial prediction of crop phenology using public phenological and climatological observations. *Computers and Electronics in Agriculture*, *127*, 726–738.

Gil, Y., Deelman, E., Ellisman, M., Fahringer, T., Fox, G., Gannon, D., … Myers, J. (2007). Examining the Challenges of Scientific Workflows. *Computer*, *40*(12), 24–32.

Giorgi, F., & Mearns, L. O. (1991). Approaches to the simulation of regional climate change: A review. *Reviews of Geophysics*, *29*(2), 191.

Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, *69*(4), 211–221.

Goodchild, M. F. (2009). NeoGeography and the nature of geographic expertise. *Journal of Location Based Services*, *3*(2), 82–96.

Goodchild, M. F., & Glennon, J. A. (2010). Crowdsourcing geographic information for disaster response: a research frontier. *International Journal of Digital Earth*, *3*(3), 231–241.

Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics*, *1*, 110–120.

Gordo, O., & Sanz, J. J. (2010). Impact of climate change on plant phenology in Mediterranean ecosystems. *Global Change Biology*, *16*(3), 1082–1106.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R.

(2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27.

Gornish, E. S., & Tylianakis, J. M. (2013). Community shifts under climate change: Mechanisms at multiple scales. *American Journal of Botany*, *100*(7), 1422–1434.

Gouveia, C., & Fonseca, A. (2008). New approaches to environmental monitoring: the use of ICT to explore volunteered geographic information. *GeoJournal*, *72*(3–4), 185–197.

Gu, L., Hanson, P. J., Post, W. Mac, Kaiser, D. P., Yang, B., Nemani, R., Pallardy, S. G., & Meyers, T. (2008). The 2007 Eastern US Spring Freeze: Increased Cold Damage in a Warming World. *BioScience*, *58*(3), 253.

Guo, W., Gong, J., Jiang, W., Liu, Y., & She, B. (2010). OpenRS-Cloud: A remote sensing image processing platform based on cloud computing environment. *Science China Technological Sciences*, *53*(S1), 221–230.

Hamunyela, E., Verbesselt, J., Roerink, G., & Herold, M. (2013). Trends in Spring Phenology of Western European Deciduous Forests. *Remote Sensing*, *5*(12), 6159–6179.

Hasenauer, H., Merganicova, K., Petritsch, R., Pietsch, S. A., & Thornton, P. E. (2003). Validating daily climate interpolations over complex terrain in Austria. *Agricultural and Forest Meteorology*, *119*(1–2), 87–107.

Haynes, D., Jokela, A., & Manson, S. (2018). IPUMS-Terra: integrated big heterogeneous spatiotemporal data analysis system. *Journal of Geographical Systems*, *20*(4), 343–361.

Henderson, A., Fischer, B., Scariot, A., Whitaker Pacheco, M. A., & Pardini, R. (2000). Flowering phenology of a palm community in a central Amazon forest. *Brittonia*, *52*(2), 149–159.

Heppenstall, A., & Harland, K. (2014). Evolutionary Algorithms. In *GeoComputation, Second Edition* (pp. 143–168). CRC Press.

Hochachka, W. M., & Fink, D. (2012). Broad-scale citizen science data from checklists: prospects and challenges for macroecology.

Huang, Q., Li, J., & Li, Z. (2018). A geospatial hybrid cloud platform based on multi-sourced computing and model resources for geosciences. *International Journal of Digital Earth*, *11*(12), 1184–1204.

Hufkens, K., Basler, D., Milliman, T., Melaas, E. K., & Richardson, A. D. (2018). An integrated phenology modelling framework in r. *Methods in Ecology and Evolution*, *9*(5), 1276–1285.

Hunter, A. F., & Lechowicz, M. J. (1992). Predicting the Timing of Budburst in Temperate Trees. *The Journal of Applied Ecology*, *29*(3), 597.

Ims, R. A. (1990). The ecology and evolution of reproductive synchrony. *Trends in Ecology & Evolution*, *5*(5), 135–140.

Inouye, D. W. (2008). Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. *Ecology*, *89*(2), 353–362.

Izquierdo-Verdiguier, E., Zurita-Milla, R., Ault, T. R., & Schwartz, M. D. (2018). Development and analysis of spring plant phenology products: 36 years

of 1-km grids over the conterminous US. *Agricultural and Forest Meteorology*, *262*, 34–41.

Janssen, P. H. M., & Heuberger, P. S. C. (1995). Calibration of process-oriented models. *Ecological Modelling*, *83*(1–2), 55–66.

Jie, C., Bao-hua, G., Ying, G. N., & Yuk-sing Gilbert, C. (2004). Comparative studies on phenotypic plasticity of two herbs Changium smyrnioides and Anthriscus sylvestris. *J Zhejiang Univ SC1 Journal of Zhejiang University SCIENCE ISSN*, *5*(6), 656–662.

Jochner, S., Caffarra, A., & Menzel, A. (2014). Phenological modelling: Can spatial data substitute temporal data? *EGU General Assembly 2014, Held 27 April - 2 May, 2014 in Vienna, Austria, Id.7547*, *16*.

Jolly, W. M., Nemani, R., & Running, S. W. (2005). A generalized, bioclimatic index to predict foliar phenology in response to climate. *Global Change Biology*, *11*(4), 619–632.

Karimi, H. A., Roongpiboonsopit, D., & Wang, H. (2011). Exploring Real-Time Geoprocessing in Cloud Computing: Navigation Services Case Study. *Transactions in GIS*, *15*(5), 613–633.

Keatley, M., & Hudson, I. (2010). Introduction and Overview. In *Phenological Research* (pp. 1–22). Springer Netherlands.

Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., … Richardson, A. D. (2014). Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Climate Change*, *4*(7), 598–604.

Kelling, S., Gerbracht, J., Fink, D., Lagoze, C., Wong, W.-K., Yu, J., Damoulas, T., & Gomes, C. P. (2012). eBird: A Human/Computer Learning Network for Biodiversity Conservation and Research. In *IAAI*.

Kelling, S., Yu, J., Gerbracht, J., & Wong, W. K. (2011). Emergent Filters: Automated Data Verification in a Large-scale Citizen Science Project. In *e-Science Workshops (eScienceW), 2011 IEEE Seventh International Conference on* (pp. 20–27). IEEE.

Koch, E. (2010). Global Framework for Data Collection – Data Bases, Data Availability, Future Networks, Online Databases. In *Phenological Research* (pp. 23–61). Springer Netherlands.

Kramer, K. (1996). *Phenology and growth of European trees in relation to climate change.*

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., & Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, *19*(1).

Kullback, S., & Leibler, R. A. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics*, *22*(1), 79–86.

Labe, Z., Ault, T. R., & Zurita-Milla, R. (2017). Identifying anomalously early spring onsets in the CESM large ensemble project. *Climate Dynamics*, *48*(11–12), 3949–3966.

Lappalainen, H., & Heikinheimo, M. (1994). The effect of temperature on the phenology of perennial plant species. *The Finnish Research Programme on Climate*.

Laube, J., Sparks, T. H., Estrella, N., H?fler, J., Ankerst, D. P., & Menzel, A. (2014). Chilling outweighs photoperiod in preventing precocious spring development. *Global Change Biology*, *20*(1), 170–182.

Leinonen, I., & Kramer, K. (2002). Applications of Phenological Models to Predict the Future Carbon Sequestration Potential of Boreal Forests. *Climatic Change*, *55*(1/2), 99–113.

Lieth, H. (1974). *Phenology and seasonality modeling*. Springer.

Linkosalo, T., Lappalainen, H. K., & Hari, P. (2008). A comparison of phenological models of leaf bud burst and flowering of boreal trees using independent observations. *Tree Physiology*, *28*(12), 1873–1882.

Liu, J., Xue, Y., Palmer-Brown, D., Chen, Z., & He, X. (2015). High-Throughput Geocomputational Workflows in a Grid Environment. *Computer*, *48*(11), 70–80.

Liu, Y., Padmanabhan, A., & Wang, S. (2015). CyberGIS Gateway for enabling data-rich geospatial research and education. *Concurrency and Computation: Practice and Experience*, *27*(2), 395–407.

Logan, M. (2010). Analysis of Covariance (ANCOVA). In *Biostatistical Design and Analysis Using R* (pp. 448–465). Wiley-Blackwell.

Luber, G., & Lemery, J. (2014). *Global Climate Change and Human Health : From Science to Practice.* Wiley.

Macbean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F.-M., & Ciais, P. (2015). Using satellite data to improve the leaf phenology of a global terrestrial biosphere model. *Biogeosciences*, *12*, 7185–7208.

MacDonald, A. J. (2018). Abiotic and habitat drivers of tick vector abundance, diversity, phenology and human encounter risk in southern California. *PLOS ONE*, *13*(7), e0201665.

Marra, P. P., Francis, C. M., Mulvihill, R. S., & Moore, F. R. (2005). The influence of climate on the timing and rate of spring bird migration. *Oecologia*, *142*(2), 307–315.

Masle, J., Doussinault, G., Farquhar, G. D., & Sun, B. (1989). Foliar stage in wheat correlates better to photothermal time than to thermal time. *Plant, Cell and Environment*, *12*(3), 235–247.

Matyas, S., Kiefer, P., Schlieder, C., & Kleyer, S. (2011). Wisdom about the crowd: assuring geospatial data quality collected in location-based games. In *Entertainment Computing–ICEC 2011* (pp. 331–336). Springer.

Mayer, A. (2010). Phenology and Citizen Science. *BioScience*, *60*(3), 172–175.

McMaster, G. S., & Wilhelm, W. W. (1997). Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology*, *87*(4), 291–300.

Meehl, G. A., Arblaster, J. M., Branstator, G., Meehl, G. A., Arblaster, J. M., &

Branstator, G. (2012). Mechanisms Contributing to the Warming Hole and the Consequent U.S. East–West Differential of Heat Extremes. *Journal of Climate*, *25*(18), 6394–6408.

Mehdipoor, H., Izquierdo-Verdiguier, E., & Zurita-Milla, R. (2017). Continental-Scale Monitoring and Mapping of False Spring : A Cloud Computing Solution + powerpoint.

Mehdipoor, H., Zurita-Milla, R., Augustijn, E., & van Vliet, A. (2016). Analyzing phenological synchronicity using volunteered geographic information. Association of Geographic Information Laboratories for Europe (AGILE).

Mehdipoor, H., Zurita-Milla, R., Izquierdo-Verdiguier, E., & Betancourt, J. L. (2018). Influence of source and scale of gridded temperature data on modelled spring onset patterns in the conterminous United States. *International Journal of Climatology*.

Mehdipoor, H., Zurita-Milla, R., Rosemartin, A., Gerst, K. L., & Weltzin, J. F. (2015). Developing a Workflow to Identify Inconsistencies in Volunteered Geographic Information: A Phenological Case Study. *PLOS ONE*, *10*(10), e0140811.

Melaas, E. K., Friedl, M. A., & Richardson, A. D. (2016). Multiscale modeling of spring phenology across Deciduous Forests in the Eastern United States. *Global Change Biology*, *22*(2), 792–805.

Mendoza, I., Peres, C. A., & Morellato, L. P. C. (2017). Continental-scale patterns and climatic drivers of fruiting phenology: A quantitative Neotropical review. *Global and Planetary Change*, *148*, 227–241.

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., & Houston, T. G. (2012). An Overview of the Global Historical Climatology Network-Daily Database. *Journal of Atmospheric and Oceanic Technology*, *29*(7), 897–910.

Menzel, A. (2005). A 500 year pheno-climatological view on the 2003 heatwave in Europe assessed by grape harvest dates. *Meteorologische Zeitschrift*, *14*(1), 75–77.

Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., … Zust, A. (2006). European phenological response to climate change matches the warming pattern. *Global Change Biology*, *12*(10), 1969–1976.

Mihorski, M., Sparks, T. H., & Tryjanowski, P. (2012). The weekend bias in recording rare birds: mechanisms and consequencess. *Acta Ornithologica*, *47*(1), 87–94.

Miller-Rushing, A. J., Høye, T. T., Inouye, D. W., & Post, E. (2010). The effects of phenological mismatches on demography. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*(1555), 3177–3186.

Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., … Bailey, A. A. (2004). The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research*, *109*(D7), D07S90.

Morisette, J. T., Jarnevich, C. S., Holcombe, T. R., Talbert, C. B., Ignizio, D., Talbert, M. K., … Young, N. E. (2013). VisTrails SAHM: visualization and workflow management for species habitat modeling. *Ecography*, *36*(2), 129–135.

Morisette, J. T., Richardson, A. D., Knapp, A. K., Fisher, J. I., Graham, E. A., Abatzoglou, J., … Liang, L. (2009). Tracking the rhythm of the seasons in the face of global change: phenological research in the 21st century. *Frontiers in Ecology and the Environment*, *7*(5), 253–260.

Morris, S. E., Freiesleben de Blasio, B., Viboud, C., Wesolowski, A., Bjørnstad, O. N., & Grenfell, B. T. (2018). Analysis of multi-level spatial data reveals strong synchrony in seasonal influenza epidemics across Norway, Sweden, and Denmark. *PLOS ONE*, *13*(5), e0197519.

Munson, S. M., & Sher, A. A. (2015). Long-term shifts in the phenology of rare and endemic Rocky Mountain plants. *American Journal of Botany*, *102*(8), 1268–1276.

Newell, D. A., Pembroke, M. M., & Boyd, W. E. (2012). Crowd sourcing for conservation: Web 2.0 a powerful tool for biologists. *Future Internet*, *4*(2), 551–562.

Nissanka, S. P., Karunaratne, A. S., Weerakoon, W. M. W., Thorburn, P. J., Boote, K. J., & Jones, J. W. (2017). Accounting for both parameter and model structure uncertainty in crop model predictions of phenology: A case study on rice. *European Journal of Agronomy*, *88*, 53–62.

Park, S. J., Hwang, C. S., & Vlek, P. L. G. (2005). Comparison of adaptive techniques to predict crop yield response under varying soil and land management conditions. *Agricultural Systems*, *85*(1), 59–81.

Parker, C. J. (2014). The Rise of Volunteered Information. *The Fundamentals of Human Factors Design for Volunteered Geographic Information*.

Parker, L. E., & Abatzoglou, J. T. (2018). Shifts in the thermal niche of almond under climate change. *Climatic Change*, *147*(1–2), 211–224.

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, *421*(6918), 37–42.

Penuelas, J., Filella, I., & Comas, P. (2002). Changed plant and animal life cycles from 1952 to 2000 in the Mediterranean region. *Global Change Biology*, *8*(6), 531–544.

Piekielek, N. B., Hansen, A. J., & Chang, T. (2015). Using custom scientific workflow software and GIS to inform protected area climate adaptation planning in the Greater Yellowstone Ecosystem. *Ecological Informatics*, *30*, 40–48.

Polgar, C. A., & Primack, R. B. (2011). Leaf-out phenology of temperate woody plants: from trees to ecosystems. *New Phytologist*, *191*(4), 926–941.

Post, E., Steinman, B. A., & Mann, M. E. (2018). Acceleration of phenological advance and warming with latitude over the past century. *Scientific Reports*, *8*(1), 3927.

Puppi, G. (2007). Origin and development of phenology as a science. *Ital J*

Agrometeorol, *3*, 24–29.

Ranta, E., Lindström, J., Kaitala, V., Crone, E., Lundberg, P., Hokkanen, T., & Kubin, E. (2010). Life History Mediated Responses to Weather, Phenology and Large-Scale Population Patterns. In *Phenological Research* (pp. 321–338). Dordrecht: Springer Netherlands.

Rasmussen, C. E. (1999). The infinite Gaussian mixture model. *NIPS*.

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. *Agricultural and Forest Meteorology*, *169*, 156–173.

Robeson, S. M. (2004). Trends in time-varying percentiles of daily minimum and maximum temperature over North America. *Geophysical Research Letters*, *31*(4), L04203.

Rodriguez-Galiano, V. F., Sanchez-Castillo, M., Dash, J., Atkinson, P. M., & Ojeda-Zujar, J. (2016). Modelling interannual variation in the spring and autumn land surface phenology of the European forest. *Biogeosciences*, *13*(11), 3305–3317.

Root, T. L., Price, J. T., Hall, K. R., Schneider, S. H., Rosenzweig, C., & Pounds, J. A. (2003). Fingerprints of global warming on wild animals and plants. *Nature*, *421*(6918), 57–60.

Rosemartin, A. H., Denny, E. G., Weltzin, J. F., Lee Marsh, R., Wilson, B. E., Mehdipoor, H., Zurita-Milla, R., & Schwartz, M. D. (2015). Lilac and honeysuckle phenology data 1956-2014. *Scientific Data*, *2*, 150038.

Santer, B. (1985). The use of general circulation models in climate impact analysis: A preliminary study of the impacts of a CO2- induced climatic change on West European agriculture. *Climatic Change*, *7*(1), 71–93.

Schaber, J., Badeck, F., Doktor, D., & von Bloh, W. (2010). Combining Messy Phenological Time Series BT - Phenological Research: Methods for Environmental and Climate Change Analysis. In I. L. Hudson & M. R. Keatley (Eds.) (pp. 147–158). Dordrecht: Springer Netherlands.

Schlieder, C., & Yanenko, O. (2010). Spatio-temporal proximity and social distance: a confirmation framework for social reporting. *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*. San Jose, California: ACM.

Schwartz, M. D. (1999). Advancing to full bloom: planning phenological research for the 21st century. *International Journal of Biometeorology*, *42*(3), 113–118.

Schwartz, M. D. (1993). Assessing the onset of spring: a climatological perspective. *Physical Geography*, *14*(6), 536–550.

Schwartz, M. D. (1990). Detecting the onset of spring: a possible application of phenological models. *Climate Research*.

Schwartz, M. D. (1998). Green-wave phenology. *Nature*, *394*(6696), 839–840.

Schwartz, M. D. (1994). Monitoring global change with phenology: The case of the spring green wave. *International Journal of Biometeorology*, *38*(1),

18–22.

Schwartz, M. D. (2003). *Phenology: an integrative environmental science* (Vol. 39). Springer.

Schwartz, M. D. (2013). Phenology: An Integrative Environmental Science. In *Phenological Data, Networks, and Research*. Springer.

Schwartz, M. D. (1997). Spring index models: an approach to connecting satellite and surface phenology. *Phenology in Seasonal Climates I*, 23–38.

Schwartz, M. D., Ahas, R., & Aasa, A. (2006). Onset of spring starting earlier across the Northern Hemisphere. *Global Change Biology*, *12*(2), 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. *International Journal of Climatology*, *33*(13), 2917–2922.

Schwartz, M. D., Betancourt, J. L., & Weltzin, J. F. (2012). From Caprio's lilacs to the USA National Phenology Network. *Frontiers in Ecology and the Environment*, *10*(6), 324–327.

Schwartz, M. D., & Chen, X. (2002). Examining the onset of spring in China. *Climate Research*, *21*(2), 157–164.

Schwartz, M. D., & Hanes, J. M. (2010). Intercomparing multiple measures of the onset of spring in eastern North America. *International Journal of Climatology*, *30*(11), 1614–1626.

Schwartz, M. D., & Marotz, G. A. (1988). Synoptic Events and Spring Phenology. *Physical Geography*, *9*(2), 151–161.

Schwartz, M. D., Reed, B. C., & White, M. A. (2002). Assessing satellite-derived start-of-season measures in the conterminous USA. *International Journal of Climatology*, *22*(14), 1793–1805.

Schwartz, M. D., & Reiter, B. E. (2000). Changes in north American spring. *International Journal of Climatology*, *20*(8), 929–932.

Scully, R. (2010). Intercomparison of PRISM and Daymet Temperature Interpolation from 1980 to 2003. *All Graduate Theses and Dissertations*.

See, L., Comber, A., Salk, C., Fritz, S., van der Velde, M., Perger, C., … Obersteiner, M. (2013). Comparing the quality of crowdsourced data contributed by expert and non-experts. *PLoS ONE*, *8*(7), e69958.

Senaratne, H., Mobasheri, A., Ali, A. L., Capineri, C., & Haklay, M. (Muki). (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, *31*(1), 139–167.

Shen, M., Cong, N., & Cao, R. (2015). Temperature sensitivity as an explanation of the latitudinal pattern of green-up date trend in Northern Hemisphere vegetation during 1982-2008. *International Journal of Climatology*, *35*(12), 3707–3712.

Shi, Y., Deng, M., Yang, X., & Liu, Q. (2016). Adaptive detection of spatial point

event outliers using multilevel constrained Delaunay triangulation. *Computers, Environment and Urban Systems*, *59*, 164–183.

Sogaard, G., Johnsen, O., Nilsen, J., & Junttila, O. (2008). Climatic control of bud burst in young seedlings of nine provenances of Norway spruce. *Tree Physiology*, *28*(2), 311–320.

Song, C. X., Zhao, L., Biehl, L. L., Merwade, V., & Villoria, N. (2016). GABBs: Cyberinfrastructure for Self-Service Geospatial Data Exploration, Computation, and Sharing. *American Geophysical Union, Fall Meeting 2016, Abstract #IN31C-1775*.

Soroye, P., Ahmed, N., & Kerr, J. T. (2018). Opportunistic citizen science data transform understanding of species distributions, phenology, and diversity gradients for global change research. *Global Change Biology*.

Sparks, T. H., Huber, K., & Tryjanowski, P. (2008). Something for the weekend? Examining the bias in avian phenological recording. *International Journal of Biometeorology*, *52*(6), 505–510.

Stodden, V. (2010). The Scientific Method in Practice: Reproducibility in the Computational Sciences. *SSRN Electronic Journal*.

Studer, S., Appenzeller, C., & Defila, C. (2005). Inter-Annual Variability and Decadal Trends in Alpine Spring Phenology: A Multivariate Analysis Approach. *Climatic Change*, *73*(3), 395–414.

Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D., & Kelling, S. (2009). eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, *142*(10), 2282–2292.

Talbert, D. A., Tinker, P., Crowther, T., & Walker, D. (2017). *Using Machine Learning to Understand Top-Down Effects in an Ecosystem: Opportunities, Challenges, and Lessons Learned*.

Thackeray, S. J., Henrys, P. A., Hemming, D., Bell, J. R., Botham, M. S., Burthe, S., … Wanless, S. (2016). Phenological sensitivity to climate across taxa and trophic levels. *Nature*.

Thill, J.-C., & Dragicevic, S. (2018). GeoComputational Research on Regional Systems (pp. 3–6). Springer, Cham.

Thornton, P. E., Running, S. W., & White, M. A. (1997). Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal Journal of Hydrology*, *190*, 214–251.

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

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, *46*, 234.

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE - suplement. *Journal of Machine Learning Research*, *9*(2579–2605), 85.

van Laarhoven, P. J. M., & Aarts, E. H. L. (1987). Simulated annealing. In *Simulated Annealing: Theory and Applications* (pp. 7–15). Dordrecht: Springer Netherlands.

van Rooijen, N. M., de Keersmaecker, W., Ozinga, W. A., Coppin, P., Hennekens, S. M., Schaminée, J. H. J., Somers, B., & Honnay, O. (2015). Plant Species Diversity Mediates Ecosystem Stability of Natural Dune Grasslands in Response to Drought. *Ecosystems*, *18*(8), 1383–1394.

van Vliet, A. J. H., Bron, W. A., & Mulder, S. (2014). The how and why of societal publications for citizen science projects and scientists. *International Journal of Biometeorology*, *58*(4), 565–577.

van Vliet, A. J. H., de Groot, R. S., Bellens, Y., Braun, P., Bruegger, R., Bruns, E., … Sparks, T. (2003). The European Phenology Network. *International Journal of Biometeorology*, *47*(4), 202–212.

Vintrou, E., Soumaré, M., Bernard, S., Bégué, A., Baron, C., & Lo Seen, D. (2012). Mapping Fragmented Agricultural Systems in the Sudano-Sahelian Environments of Africa Using Random Forest and Ensemble Metrics of Coarse Resolution MODIS Imagery. *Photogrammetric Engineering & Remote Sensing*, *78*(8), 839–848.

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J.-M., Hoegh-Guldberg, O., & Bairlein, F. (2002). Ecological responses to recent climate change. *Nature*, *416*(6879), 389–395.

Wang, C., Tang, Y., & Chen, J. (2016). Plant phenological synchrony increases under rapid within-spring warming. *Scientific Reports*, *6*(1), 25460.

Wang, H., Ge, Q., Rutishauser, T., Dai, Y., & Dai, J. (2015). Parameterization of temperature sensitivity of spring phenology and its application in explaining diverse phenological responses to temperature change. *Scientific Reports*, *5*(1), 8833.

Weis, A., & Kossler, T. (2004). Genetic variation in flowering time induces phenological assortative mating: quantitative genetic methods applied to Brassica rapa. *American Journal of Botany*.

Wiersma, Y. F. (2010). Birding 2.0: Citizen Science and Effective Monitoring in the Web 2.0 World. *Avian Conservation and Ecology*, *5*(2), art13.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018.

Wolfe, D. W., Schwartz, M. D., Lakso, A. N., Otsuki, Y., Pool, R. M., & Shaulis, N. J. (2005). Climate change and shifts in spring phenology of three horticultural woody perennials in northeastern USA. *International Journal of Biometeorology*, *49*(5), 303–309.

Worner, S. P. (1992). Performance of Phenological Models Under Variable Temperature Regimes: Consequences of the Kaufmann or Rate Summation Effect. *Environmental Entomology*, *21*(4), 689–699.

Wu, X., Zurita-Milla, R., & Kraak, M.-J. (2016). A novel analysis of spring phenological patterns over Europe based on co-clustering. *Journal of Geophysical Research: Biogeosciences*, *121*(6), 1434–1448.

Yanenko, O., & Schlieder, C. (2012). Enhancing the Quality of Volunteered

Geographic Information: A Constraint-Based Approach. In J. Gensel, D. Josselin, & D. Vandenbroucke (Eds.), *Bridging the Geographic Information Sciences: International AGILE'2012 Conference, Avignon (France), April, 24-27, 2012* (pp. 429–446). Berlin, Heidelberg: Springer Berlin Heidelberg.

Yenni, G. M., Christensen, E. M., Bledsoe, E. K., Supp, S. R., Diaz, R. M., White, E. P., & Ernest, S. K. M. (2018). Developing a modern data workflow for living data. *BioRxiv*, 344804.

Yin, D., Liu, Y., Padmanabhan, A., Terstriep, J., Rush, J., & Wang, S. (2017). A CyberGIS-Jupyter Framework for Geospatial Analytics at Scale. In *Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact - PEARC17* (pp. 1–8). New York, New York, USA: ACM Press.

Yuan, W., Piao, S., Qin, D., Dong, W., Xia, J., Lin, H., & Chen, M. (2018). Influence of Vegetation Growth on the Enhanced Seasonality of Atmospheric CO_2. *Global Biogeochemical Cycles*, *32*(1), 32–41.

Zavalloni, C., Andresen, J. A., & Flore, J. A. (2006). Phenological Models of Flower Bud Stages and Fruit Growth of `Montmorency' Sour Cherry Based on Growing Degree-day Accumulation. *Journal of the American Society for Horticultural Science*, *131*(5), 601–607.

Zhu, L., Meng, J., Li, F., & You, N. (2017). Predicting the patterns of change in spring onset and false springs in China during the twenty-first century. *International Journal of Biometeorology*, 1–16.

Zurita-Milla, R., Goncalves, R., Izquierdo-Verdiguier, E., & Ostermann, F. O. (2017). Exploring Vegetation Phenology At Continental Scales : Linking Temperature-Based Indices And Land Surface Phenological Metrics.

Zurita-Milla, R., Izquierdo-Verdiguier, E., & Mehdipoor, H. (2015). Phenological modelling using volunteered observations and machine learning methods. In *Phenology 2015: third international conference on phenology, 5-8 October 2015, Kusadasi, Turkey*.

Zurita-Milla, R., van Gijsel, J. A. E., Hamm, N. A. S., Augustijn, P. W. M., & Vrieling, A. (2013). Exploring Spatiotemporal Phenological Patterns and Trajectories Using Self-Organizing Maps. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(4), 1914–1921.

Summary

Among the various research questions raised by climate change, the question: "how does climate change affect vegetation seasonality?" is crucial. This is because changes in vegetation seasonality have both global and substantial implications for our planet. In particular, phenology is the science that studies seasonal plant and animal life cycle events (phenophases) and how annual and inter-annual variations in weather and environmental conditions affect them. In this respect, volunteered phenological observations (VPOs) and phenological model outputs are key information to study phenology in space and time. This phenological information helps to better understand the impact of climate change on vegetation seasonality. In this research, we designed novel geocomputational workflows to explore spring plant phenology (SPP) at large scale and over long periods using VPOs and temperature-driven phenological models. The workflows focus on 1) checking the consistency of VPOs using contextual geo-information and domain knowledge and 2) analyzing the impact of the type of phenological model as well as of its input data sources and their spatial resolution on the patterns and trends derived from the model.

The first workflow uses the statistical distribution of the VPOs as well as environmental contextual information to provide robust consistency checks. It relies on all the available environmental information for the VPO sites (in our case weather variables) and it is suited for areas with strong environmental gradients. The workflow was used to identify inconsistent (i.e., anomalously early or late in relation to their associated environmental conditions) VPOs from the USA National Phenology Network (USA-NPN). In particular, it was applied to flowering observations of common and cloned lilac plants (Syringa vulgaris and Syringa x chinensis) in the coterminous United States for the period 1980 to 2013. About 97% of the VPOs were flagged as consistent, indicating that volunteers provided reliable information. Relative to the original dataset, the exclusion of inconsistent observations changed the apparent rate of change in lilac bloom dates by two days per decade, indicating the importance of inconsistency checking as a key step to analyze volunteered geographic information.

The second workflow uses variables that are related to the phenology domain. In this case, temperature as it is an important phenological driver in temperate climates and it strongly influences phenological synchrony. This knowledge forms the basis for defining, searching and optimizing consistency constraints. This workflow was tested using VPOs collected in the Netherlands during the period 2003–2015. The average percentage of inconsistent observations was low to moderate (ranging from 1% for wood anemone (Anemone nemorosa L.) and pedunculate oak (Quercus robur L.) to 15% for cow parsley species (Anthriscus sylvestris (L.) Hoffm)). This again indicates that volunteers provide

reliable phenological information. We also found a significant correlation between the standard deviation of DOY of the observed events and the accumulation of daily temperature (with correlation coefficients ranging from 0.78 for lesser celandine (Ficaria verna Huds), and 0.60 for pedunculate oak). This confirms that colder days in late winter and early spring lead to synchronous flowering and leafing onsets. Our results highlight the potential of combining environmental and domain information and knowledge to check the consistency of (phenological) volunteered geographic information.

The third workflow was designed to analyse and compare patterns and trends from a suite of temperature-based phenological models, namely the Extended Spring Indices (SI-x), Thermal Time and Photothermal Time models. These models were first calibrated using historical lilac leaf onset observations for the period 1961-1994. Then, contemporary VPOs and daily gridded temperature data were used to validate the models. Results show that the root-mean-square errors (RMSEs) of the SI-x and Thermal Time models are similar, and about two days lower than those provided by the other models. Yet, the dates of leaf out provided by each of the models are up to 11 days different, and the trends are up to a week per decade different. These results also show that the statistical significance of phenological trends strongly depends on the type of model. Therefore, current approaches for validating phenological models based on global statistics such as RMSE do not provide information about the variability of patterns and trends in different regions. Studies using phenological models and gridded input data to study climate change impact on plant seasonality should check both the spatial and temporal variability of the model outputs at large-scale. Using a model that is found less valid than another one (i.e., with a "worse" RMSE for a given set of validation observations) may still provide more realistic patterns and trends when compared with large-scale phenological data and/or information.

The fourth workflow was used to evaluate how the source and spatial resolution (i.e., scale) of the input data might affect phenological models and indices that track variations and shifts of vegetation seasonality at continental scales. The workflow, based on cloud computing and volunteered phenological observations, focuses on the SI-x, which estimate the day of year (DOY) for first leaf (FL) and first bloom (FB) in plants sensitive to accumulation of warmth in early to mid-spring. The SI-x products obtained using Daymet (at 1, 4, 35, and 100 km spatial resolution) and gridMET (at 4, 35, and 100 km) data, from 1980 to 2016. These products and their resulting patterns and trends across the coterminous United States, were affected more by the source of input data than by their spatial resolution. The SI-x estimates DOY of FL (FB) are about 3 (4) weeks more accurate using Daymet than gridMET. We noted significant differences, and even contradictory rates of change in DOY, when the models were driven by Daymet or gridMET temperatures. The SI-x products generated

from gridMET poorly estimate the timing of spring onset, whereas Daymet SI-x products and actual volunteered observations are moderately correlated (r = 0.7). Daymet better captures temperature regimes, particularly in the western United States, and hence it is more appropriate for generating high-resolution SI-x products at continental scale.

The main conclusion of this PhD thesis is that a careful analysis of phenological data and models is necessary to avoid misleading results. Geocomputational principles and approaches provide an ideal paradigm to integrate and scale up geographical and time contexts in such an analysis. The illustrated workflows reveal that variations in the consistency and source of phenological model inputs as well as the choice of phenological model significantly alter the estimation of long-term spring onset patterns and trends. Future studies that leverage volunteered observations and gridded (weather) datasets can adapt the proposed workflows to improve the robustness of their scientific analyses.

Samenvatting

Van de verschillende onderzoeksvragen die klimaatverandering oproept, is de vraag: "Hoe beïnvloedt klimaatverandering de seizoen gebondenheid van vegetatie?" cruciaal. Dit komt omdat veranderingen in seizoen gebondenheid van vegetaties wereldwijde en substantiële implicaties heeft voor onze planeet. Fenologie is de wetenschap die periodieke gebeurtenissen in de levenscyclus van planten en dieren (fenofasen) onderzoekt en bestudeert hoe jaarlijkse en wisselende variaties in weers- en milieuomstandigheden op hen van invloed zijn. In dit opzicht leveren vrijwillige fenologische observaties (VFO's) en uitkomsten van fenologische modellen belangrijke informatie om fenologie in ruimte en tijd te bestuderen. Deze fenologische informatie kan helpen om de impact van klimaatverandering op de seizoen gebondenheid van de vegetatie beter te begrijpen. Dit onderzoek ontwerpt nieuwe geocomputationele werkprocessen om voorjaars planten fenologie (VPF) op grote schaal en gedurende lange perioden te onderzoeken met behulp van VFO's en temperatuur gestuurde fenologische modellen. De workflows zijn gericht op 1) het controleren van de consistentie van VFO's met behulp van contextuele geo-informatie en domeinkennis en 2) het analyseren van de impact van het type fenologisch model, de inputgegevensbronnen en hun ruimtelijke resolutie op de patronen en trends afgeleid van het model.

De eerste workflow maakt gebruik van de statistische distributie van de VFO's en gebruikt contextuele informatie over de omgeving voor robuuste consistentiecontroles. Het gebruikt alle weervariabelen die beschikbaar zijn voor de VFO-locaties (in dit geval weersvariablen) en is geschikt voor gebieden met een sterke milieuverloop. De werkstroom werd gebruikt om inconsistente VFO's (d.w.z. abnormaal vroeg of laat in relatie tot hun omgevingscondities) van het USA National Phenology Network (USA-NPN) te identificeren. De workflow werd toegepast op VFO's van bloeiobservaties van gewone en gekloonde seringen (Syringa vulgaris en Syringa x chinensis) op het vasteland van de Verenigde Staten voor de periode 1980-2013. Ongeveer 97% van de VFO's werden aangemerkt als consistent, wat aangeeft dat vrijwilligers betrouwbare informatie hebben verstrekt voor deze studie. Ten opzichte van de oorspronkelijke dataset veranderde de uitsluiting van inconsistente waarnemingen de schijnbare veranderingssnelheid in sering-bloeiperiodes met twee dagen per decennium, wat het belang van inconsistentiecontrole aangeeft als een belangrijke stap in de beoordeling van de gegevenskwaliteit voor vrijwillige geografische informatie.

De tweede workflow gebruikt de sleutelvariabele van de VFO's. In dit geval, temperatuur, omdat dit een belangrijke fenologische indicator is in gematigde klimaatzones en grote invloed heeft op de synchronie van de fenologische waarnemingen. Dit vormt de basis voor het definiëren, doorzoeken en

optimaliseren van consistentiebeperkingen. Deze workflow is getest met VFO's die in de periode 2003-2015 in Nederland zijn verzameld. Het gemiddelde percentage inconsistente waarnemingen was laag tot matig (variërend van 1% voor bosanemoon (Anemone nemorosa L.) en zomereik (Quercus robur L.) tot 15% voor fluitenkruid (Anthriscus sylvestris (L.) Hoffm)). Dit toont nogmaals aan dat vrijwilligers betrouwbare fenologische informatie verstrekken. We vonden ook een significante correlatie tussen de standaardafwijking van de dag van het jaar (DJ) van de waargenomen gebeurtenissen en de accumulatie van de dagelijkse temperatuur (met correlatiecoëfficiënten variërend van 0,78 voor speenkruid (Ficaria verna Huds) en 0,60 voor zomereik). Dit bevestigde dat koudere dagen in de late winter en het vroege voorjaar tot synchrone bloei en begin van de bladontwikkeling leiden. Onze resultaten benadrukten het potentieel van synchronisatie-informatie en geografische context voor het controleren van de consistentie van fenologische vrijwillige geografische informatie.

De derde workflow analyseert en vergelijkt patronen en trends van een reeks temperatuur gebaseerde fenologische modellen: Extended Spring Indices (SI-x), Thermal Time en Photothermal Time-modellen. De modellen werden eerst gekalibreerd met behulp van historische sering bladaanzet waarnemingen voor de periode 1961-1994. Vervolgens werden de modellen gebruikt om moderne VFO's en dagelijkse temperatuurgegevens te valideren. De resultaten van de workflow tonen aan dat root-mean-square errors (RMSE's) van SI-x en Thermal Time-modellen vergelijkbaar zijn en ongeveer twee dagen lager dan die van de andere modellen. Toch zijn de data van bladontwikkeling die door elk van de modellen wordt verstrekt tot 11 dagen verschillend, en de trends variëren tot een week per decennium. De workflow resultaten laten ook zien dat de statistische significantie van deze trends sterk afhankelijk is van het type model. Daarom kunnen huidige benaderingen voor het valideren van fenologische modellen op basis van globale statistieken zoals RMSE niet worden gebruikt om informatie te krijgen over de variabiliteit van patronen en trends in verschillende regio's. Onderzoeken gebaseerd op fenologische modellen en gerasterde inputgegevens om de invloed van klimaatverandering op seizoensinvloeden te bestuderen, moeten zowel de ruimtelijke als de temporele variabiliteit op grote schaal controleren. Het gebruik van een minder valide model (d.w.z. "grotere" RMSE voor een gegeven set observaties) kan nog steeds meer realistische patronen en trends opleveren in vergelijking tot grootschalige fenologische gegevens en / of informatie.

De vierde werkstroom evalueert hoe de specifieke bron en ruimtelijke resolutie (d.w.z. schaal) van de invoergegevens van invloed kan zijn op fenologische modellen en indices die variaties en verschuivingen op continentale schaal volgen. De workflow, gebaseerd op cloud computing en vrijwillige fenologische waarnemingen, richt zich op SI-x die de dag van het jaar (DJ) schat voor de

opkomst van het eerste blad en de eerste bloei in planten die bijzonder gevoelig zijn voor warmte-accumulatie in het vroege voorjaar tot het midden van de lente. De SI-x-producten zijn verkregen met Daymet (op 1, 4, 35 en 100 km ruimtelijke resolutie) en gridMET (op 4, 35 en 100 km), van 1980 tot 2016. De resulterende patronen en trends in de Verenigde Staten, worden meer beïnvloed door de bron van de invoergegevens dan door hun ruimtelijke resolutie. The SI-x-schattingen DJ van bloei en bladontwikkeling zijn respectievelijk ongeveer 3 en 4 weken nauwkeuriger met Daymet dan gridMET. Dit leidt tot aanzienlijke verschillen, en zelfs tegenstrijdige veranderingssnelheden in DJ, aangedreven door Daymet versus gridMET-temperaturen. SI-x-producten gegenereerd uit gridMET schatten de timing van het begin van de lente slecht, terwijl de Daymet SI-x-producten en de werkelijke vrijwillige observaties matig gecorreleerd zijn (r = 0,7). Daymet legt temperatuurregimes beter vast, met name in het westen van de Verenigde Staten, en is meer geschikt voor het genereren van hoge resolutie SI-x-indices op continentale schaal.

De belangrijkste conclusie van dit proefschrift is dat een zorgvuldige analyse van fenologische gegevens en modellen noodzakelijk is om misleidende resultaten te voorkomen. Geocomputationele principes en methodes vormen een ideaal paradigma om geografische en tijdscontexten in een dergelijke analyse te integreren en op te schalen. De geïllustreerde workflows onthulden dat variaties in de consistentie en inputs van het fenologische modellen, evenals de keuze van het fenologische modelparadigma, de voorspelling van patronen en trends op de lange termijn aanzienlijk veranderen. Toekomstige studies die gebruikmaken van vrijwillige waarnemingen en gerasterde (weer) gegevenssets kunnen de voorgestelde workflows aanpassen om de robuustheid van hun wetenschappelijke analyses te verbeteren.

Biography

First Name: Hamed

Last Name: Mehdi Poor

Date of birth: 15-January-1986

Contact: mr.h.mehdipoor@gmail.com mr.h.mehdipoor@gmail.com

2004 - 2005	National Organization for Development of Exceptional Talents, Tehran, Iran. PUC diploma in mathematics and physics.
2005 - 2009	Khajeh Nasir Toosi University of Technology, Tehran, Iran. BSc. in Geomatics. Specialization: surveying, geographic information system (GIS) and remote sensing (RS).
2011 - 2013	Department of geo-information processing, ITC faculty, University of Twente, Enschede, The Netherlands. MSc. in Geoinformatics. Graduated with Distinction. MSc Thesis: Identification of Inconsistency and Bias in volunteered geographic information Supervision: Prof.dr. R. Zurita-Milla and Dr.ir. P.W.M. Augustijn-Beckers
2013 - 2019	Department of geo-information processing, ITC faculty, University of Twente, Enschede, The Netherlands. PhD Thesis: Geocomputational workflows for analysing spring plant phenology in space and time Supervision: Prof. dr. R. Zurita-Milla and Dr. ir. P.W.M. Augustijn-Beckers

List of publications:

Peer reviewed journals:

Mehdipoor, H., Zurita-Milla, R., Augustijn, E.W. and van Vliet, A., 2018. Checking the Consistency of Volunteered Phenological Observations While Analysing Their Synchrony. ISPRS International Journal of Geo-Information, 7(12), p.487.

Mehdipoor, H., Zurita-Milla, R., Izquierdo-Verdiguier, E. and Betancourt, J.L., 2018. Influence of source and scale of gridded temperature data on modelled spring onset patterns in the conterminous United States. International journal of climatology, 38(14), pp.5430-5440.

Mehdipoor, H., Zurita-Milla, R., Augustijn, E.W. and Izquierdo-Verdiguier, E., 2018. Exploring differences in spatial patterns and temporal trends of phenological models at continental scale using gridded temperature time-series. International Journal of Biometeorology, in review.

Allen, M.J., Vanos, J., Hondula, D.M., Vecellio, D.J., Knight, D., **Mehdipoor, H.**, Lucas, R., Fuhrmann, C., Lokys, H., Lees, A. and Nascimento, S.T., 2017. Supporting sustainability initiatives through biometeorology education and training. International journal of biometeorology, 61(1), pp.93-106.

Perkins IV, D.R., Vanos, J., Fuhrmann, C., Allen, M., Knight, D., Lee, C.C., Lees, A., Leung, A., Lucas, R., **Mehdipoor, H** and Tavares Nascimento, S., 2017. Enhancing the teaching and learning of biometeorology in higher education. Bulletin of the American Meteorological Society, 98(9), pp.ES239-ES242.

Mehdipoor, H., Vanos, J.K., Zurita-Milla, R. and Cao, G., 2017. emerging technologies for biometeorology. International journal of biometeorology, 61(1), pp.81-88.

Rashidi, P., Wang, T., Skidmore, A., **Mehdipoor, H.**, Darvishzadeh, R., Ngene, S., Vrieling, A. and Toxopeus, A.G., 2016. Elephant poaching risk assessed using spatial and non-spatial Bayesian models. Ecological modelling, 338, pp.60-68.

Mehdipoor, H., Zurita-Milla, R., Rosemartin, A., Gerst, K.L. and Weltzin, J.F., 2015. Developing a workflow to identify inconsistencies in volunteered geographic information: a phenological case study. PloS one, 10(10), p.e0140811.

Rosemartin, A.H., Denny, E.G., Weltzin, J.F., Marsh, R.L., Wilson, B.E., **Mehdipoor, H.**, Zurita-Milla, R. and Schwartz, M.D., 2015. Lilac and honeysuckle phenology data 1956–2014. Scientific data, 2, p.150038.

Other scientific contributions:

Mehdipoor, H., Izquierdo-Verdiguier, E. and Zurita-Milla, R., 2017. Continental-Scale Monitoring and Mapping of False Spring: A Cloud Computing Solution+ powerpoint. In 2017 International Conference on GeoComputation, GeoComputation 2017: Celebrating 21 Years of GeoComputation

Mehdipoor, H., Izquierdo-Verdiguier, E. and Zurita-Milla, R., 2016, April. Revisiting the extended spring indices using gridded weather data and machine learning. In EGU General Assembly Conference Abstracts (Vol. 18, p. 16698).

Mehdipoor, H., Zurita-Milla, R., Augustijn-Beckers, P. and van Vliet, A., 2016. Analyzing phenological synchronicity using volunteered geographic information. In Geospatial data in a changing world: proceedings of the 19th AGILE conference on geographic information science, 14-17 June 2016, Helsinki, Finland.

Mehdipoor, H., Zurita-Milla, R. and Kraak, M.J., 2015, May. Developing a geocomputational workflow to check the consistency of volunteered geographic information. In Proceedings of the 13th international conference on GeoComputation geospatial information sciences, 20-23 May 2015, Richardson TX, United States of America.

Zurita-Milla, R., Izquierdo-Verdiguier, E. and **Mehdipoor, H.**, 2015. Phenological modelling using volunteered observations and machine

learning methods. In Phenology 2015: third international conference on phenology, 5-8 October 2015, Kusadasi, Turkey.

Mehdipoor, H. and Zurita-Milla, R., 2015. Checking for inconsistent volunteered phenological observations. In Phenology 2015: third international conference on phenology.

Mehdipoor, H., 2015. Exploring climate change and its impact on agriculture using volunteered geographic information. In AGILE PhD School.

Mehdipoor, H., 2013. Identification of Inconsistency and Bias in Volunteerd Phenological Observations. University of Twente Faculty of Geo-Information and Earth Observation (ITC).