Biogeosciences Discuss., 12, 11833–11861, 2015
www.biogeosciences-discuss.net/12/11833/2015/
doi:10.5194/bgd-12-11833-2015

Biogeosciences Discussions
Open Access

Modelling anomalies in the spring and autumn land surface phenology of the European forest

V. F. Rodriguez-Galiano[1,2], **M. Sanchez-Castillo**[3], **J. Dash**[2], **and P. M. Atkinson**[4,5,6,2]

[1]Physical Geography and Regional Geographic Analysis, University of Seville, Seville 41004, Spain
[2]Global Environmental Change and Earth Observation Research Group, Geography and Environment, University of Southampton, Southampton SO17 1BJ, UK
[3]Department of Haematology, Wellcome Trust and MRC Cambridge Stem Cell Institute and Cambridge Institute for Medical Research, University of Cambridge, Cambridge CB2 0XY, UK
[4]Faculty of Science and Technology, Engineering Building, Lancaster University, Lancaster LA1 4YR, UK
[5]Faculty of Geosciences, University of Utrecht, Heidelberglaan 2, 3584 CS Utrecht, the Netherlands
[6]School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, Belfast BT7 1NN, Northern Ireland, UK

Received: 9 July 2015 – Accepted: 13 July 2015 – Published: 30 July 2015

Correspondence to: V. F. Rodriguez-Galiano (vrgaliano@gmail.com)

Published by Copernicus Publications on behalf of the European Geosciences Union.

BGD

12, 11833–11861, 2015

Abstract

This research reveals new insights into the climatic drivers of anomalies in land surface phenology (LSP) across the entire European forest, while at the same time establishes a new conceptual framework for predictive modelling of LSP. Specifically, the Random Forest method, a multivariate, spatially non-stationary and non-linear machine learning approach, was introduced for phenological modelling across very large areas and across multiple years simultaneously: the typical case for satellite-observed LSP. The RF model was fitted to the relation between LSP anomalies and numerous climate predictor variables computed at biologically-relevant rather than human-imposed temporal scales. In addition, the legacy effect of an advanced or delayed spring on autumn phenology was explored. The RF models explained 81 and 62 % of the variance in the spring and autumn LSP anomalies, with relative errors of 10 and 20 %, respectively: a level of precision that has until now been unobtainable at the continental scale. Multivariate linear regression models explained only 36 and 25 %, respectively. It also allowed identification of the main drivers of the anomalies in LSP through its estimation of variable importance. This research, thus, shows clearly the inadequacy of the hitherto applied linear regression approaches for modelling LSP and paves the way for a new set of scientific investigations based on machine learning methods.

1 Introduction

Vegetation phenology has emerged as an important focus for scientific research in the last few decades. The interest in vegetation phenology is twofold: inter-annual recording of the timing of phenological events allows quantification of the impacts of climate change on vegetation; and a greater understanding of phenological responses enables meaningful projections of how ecosystems will respond to future changes in climate (Morisette et al., 2008; Menzel, 2002; Peñuelas and Filella, 2001; Peñuelas, 2009). Although different approaches have been devised for the study of vegetation phenology

BGD

12, 11833–11861, 2015

(Rafferty et al., 2013), the characterisation and modelling of vegetation phenology at global or regional scales has been undertaken mainly through the use of long-term time-series of satellite-sensor vegetation indices (termed land surface phenology, LSP, to reflect that satellite-observed phenology includes all land covers). Most studies of LSP analyse trends in phenological events across years (Delbart et al., 2008; Jeong et al., 2011; Karlsen et al., 2007; Myneni et al., 1997; Jeganathan et al., 2014), but more recent studies present process-based models to uncover cause–effect relationships between long-term trends in phenology and its key driving variables (Ivits et al., 2012; Maignan et al., 2008a, b; Stöckli et al., 2008, 2011; Zhou et al., 2001). This last group of studies focuses on changes in phenology produced by average changes in weather (mainly trends in warming). However, anomalies in LSP arising as a consequence of the inter-annual variability in weather are relatively unstudied, with model-based studies of this phenomenon being scarce (van Vliet, 2010).

A higher frequency in the occurrence of extreme climatic events has been observed in Europe, especially for summer temperatures (Luterbacher et al., 2004; Barriopedro et al., 2011). The summers of 2003 and 2010 in western and eastern Europe, respectively, were the warmest in the last 500 years (Barriopedro et al., 2011). Species and ecosystems respond more rapidly to these anomalies in weather than average climatic changes in most climatic scenarios (Zhao et al., 2013). Maignan et al. (2008b) and Rutishauser et al. (2008) reported that the LSP greening occurred 10 days earlier in 2007 than the average over the past three decades as a consequence of an exceptionally mild winter and spring. The study of the impacts of extreme inter-annual weather events on vegetation through the modelling of anomalies in spring and autumn phenologies can increase our knowledge about climate-driven changes in phenology, acting as natural experiments in climate change scenarios (Rafferty et al., 2013). On the other hand, the modelling of LSP has been less explored compared to the modelling of individual plant species, and there are many aspects that remain to be understood, which limits comprehensive understanding of LSP and, therefore, of phenology at regional or global scales. A more complete modelling of LSP considering the inter-annual variation across large areas would include the capacity to interpret observations and make meaningful projections in relation to disturbances and their subsequent impacts (Morisette et al., 2008).

Modelling efforts to characterize LSP have generally relied on simple functions (usually linear) of meteorological drivers, such as average temperature and precipitation (Ivits et al., 2012), growing degree days (GDD) (de Beurs and Henebry, 2005), light and temperature (Stöckli et al., 2011), minimum temperature, photoperiod, vapour pressure deficit (Jolly et al., 2005; Stöckli et al., 2008), or minimum relative humidity (Brown and de Beurs, 2008). However, there is lack of understanding on number of important aspects, such us the multivariate influence of meteorological variables (temperature, precipitation, solar radiation) driving phenology, or the effect of additional drivers in the modelling of autumnal phenophases (Morisette et al., 2008). For instance, Fu et al. (2014) found a "cause–effect relationship" between an earlier leaf senescence and an earlier spring flushing in leaves of warmed samples of *Fagus sylvatica* and *Quercus robur*. This legacy effect of spring phenology has been reported in recent studies using modified environments and plant species, but it has not been studied using LSP data. This latter aspect is particularly pertinent for studies that focus on inter-annual variation in phenology and could potentially contribute to increased knowledge of how climate change is affecting autumn phenology. On the other hand, many studies investigating the sensitivity of phenological events to climate variation use calendar seasonal or monthly mean climatic variables, which operate on human calendar scales (Maignan et al., 2008b), instead of using daily data to build climatic variables meaningful at biological scales (Pau et al., 2011). Pau et al. (2011) in a recent review highlighted the importance of using daily models related to vegetation circadian time scales and how climate change has influenced daily minima and maxima disproportionately. However, the modelling of LSP considering its potentially complicated relationship with climate in a multidimensional feature space (i.e. high number of multivariate climatic drivers at biological scales) might not be possible using traditional linear regression models (de Beurs and Henebry, 2005). In this sense, phenological modelling may benefit from machine learning techniques such as the Random Forest (RF) method (Breiman, 2001), reducing uncertainties and bias (Zhao et al., 2013). RFs have the potential to identify and model the complex non-linear relationships between phenology and climate, being able to handle a large number of predictors and determine their importance in explaining phenology. RFs has been applied with very promising results to other fields of ecology and biological sciences (Lawler et al., 2006; Archibald et al., 2009; Darling et al., 2012), as well as to the simulation of phenological shifts under different climatic change scenarios (Lebourgeois et al., 2010), but the potential for modelling climate-driven changes in phenology is still to be explored.

Understanding the effect of inter-annual climatic variation on LSP is an essential step to establish a plausible link between recent climate variability and vegetation phenological responses at global or regional scales, and importantly to make reliable forecasts about future vegetation responses to different future climatic scenarios. The aim of this study is, therefore, to provide an explanation of the observed anomalies in LSP of the entire European forest during the last decade, identifying the main climatic drivers of spring and autumnal LSP at the continental scale. Our research offers new insights into the study of LSP by modelling the climate-driven anomalies in phenology, rather than trends, and using innovative multivariate non-linear machine learning techniques to evaluate multiple climatic predictors at biological scales, and non-climatic predictors such as the legacy effect of the date of spring onset in leaf senescence. Climate predictors used range from monthly average values of temperature (max, min and avg), precipitation, short wave radiation and day length; trimestral cumulated values such as growing degree days or chilling requirements, among others; to the date of specific events such as the first freeze or the last freeze. Moreover, we considered flexible biological time scales in the analysis between climatic and phenological events rather than fixed calendar dates.

2 Materials and methods

2.1 Data

Three sources of data were used for this research: (i) satellite sensor derived temporal composites of MERIS Terrestrial Chlorophyll Index (MTCI), (ii) temperature and precipitation data from the European Climate Assessment and Data (ECA&D) project (http://www.ecad.eu) and (iii) surface radiation daylight (DAL) data and surface incoming shortwave (SIS) radiation data from the Climate Monitoring Satellite Application Facilities (CM SAF, http://www.cmsaf.eu).

We used weekly composites of MTCI data at 1 km spatial resolution from 2002 to 2012. This dataset was supplied by the European Space Agency and processed by Airbus Defence and Space. Daily temperature (mean, minimum and maximum) and daily precipitation data were derived from the European Climate Assessment & Dataset (ECA&D) time-series (version 10.0) with spatial resolution of $0.25^{\circ} \times 0.25^{\circ}$, covering the period from 2002 to 2011 (Haylock et al., 2008). The CM SAF DAL version CDR v001 (Müller and Trentmann, 2013) and SIS version CDR v002 (Posselt et al., 2011, 2012) were derived from Meteosat satellite sensors at a spatial resolution of $0.05^{\circ} \times 0.05^{\circ}$.

2.2 Phenology extraction and anomalies computation

The time-series of MERIS MTCI data was used to estimate both the onset of greenness (OG) and end of senescence (EOS). The yearly values of OG and EOS were estimated for each image pixel of the study area using the methodology described in Dash et al. (2010). These satellite derived LSP had a strong spatio-temporal correlation with ground observations (thousands of deciduous tree phenology records of the Pan European Phenology network, PEP725) (Rodriguez-Galiano et al., 2015a).

The anomalies in the LSP for a given year were defined as the difference from the long-term mean, normalized by the standard deviation across years. The value of the targeted year was excluded in the computation to enhance the inter-annual variation (Saleska et al., 2007). To match the spatial resolution of the ECA&D dataset, the LSP anomalies for each year were resampled to spatial resolution of $0.25^{\circ} \times 0.25^{\circ}$ by extracting the median of all the LSP anomaly values within this area after excluding the areas with less than 50 LSP estimates and the non-forest pixels according to the Globcover2005 and Globcover2009 land cover maps (http://due.esrin.esa.int/page_globcover.php). Globcover was selected for its greater consistency with MERIS MTCI time-series and its high geolocational accuracy (< 150 m) (Bicheron et al., 2011).

2.3 Computation of climatic predictors

A suite of climatic predictors were identified for each $0.25^{\circ} \times 0.25^{\circ}$ grid cell associated with the occurrence of positive or negative anomalies in LSP. Schwartz et al. (2006) found that most phenophases of plant observations in Europe correlated significantly with climatic predictors of the month of onset and the two preceding months. The different climatic measures were computed based on the 30 and 90 days previous to the Julian date (see previous section) of the phenological event at each pixel which experienced an anomaly in phenology in a given year (Fig. 1). The chilling requirements for spring modelling were an exception, as the period for its computation starts 90 days before the date of onset. Relative differences of the climatologies were computed to capture the inter-annual variability in climate at the pixel level for every predictor and to facilitate the modelling of climate-driven variation in phenology. The analysis included 31 and 27 predictor-related measures for spring and autumn modelling respectively. The predictor include temporal average values of temperature (T_{max}, T_{min} and T_{avg}), precipitation, surface radiation daylight (DAL) and surface incoming shortwave radiation (SIS); temporal cumulated predictors such as growing degree days, chilling, precipitation, SIS and DAL; and the date of specific events such as the onset of greenness (legacy effect for autumn phenology modelling) the first freeze or the last freeze, as well as the difference between both dates (freeze period) for the modelling of autumn only. Growing degree days were computed using temperature thresholds of 0 and 5°. Chilling requirements were computed as the sum of negative temperatures (temperatures

BGD
12, 11833–11861, 2015

below 0°). Freeze predictors were computed using temporal windows of 90 days only, and were defined as dates with minimum temperatures lower than −2 °C (Schwartz et al., 2006).

2.4 Modelling LSP anomalies

Conventional statistical models such as linear regression are inappropriate for investigating the drivers of anomalies in phenology because many of the relationships are likely to be non-linear. In this sense, machine learning methods have emerged as complementary alternatives to conventional statistical techniques. Within the branch of machine learning techniques, regression trees are particularly suitable when compared to global single predictive models, allowing for multiple regression models using recursive partitioning (Breiman et al., 1984). Assembling a single global model might not be representative of the phenomenon under study when there are many predictors which interact in complicated, non-linear ways and may vary spatially. An alternative approach is to sub-divide, or partition, the space into more homogeneous regions of similar characteristics. Regression trees use a sum of squares criterion to split the data into successively more homogeneous subsets. Therefore, different regression models can be fitted to different data subsets, which can represent different responses controlled by different drivers (Lawler et al., 2006; Archibald et al., 2009). For the purpose of this paper, this latter property makes regression trees particularly advantageous.

Different approaches have been proposed in the last few years to increase the predictive ability of regression tree models. Among all of them RFs are probably the most popular technique due to the simplicity of their applicability, interpretability of the models, and the robustness of predictions (Rodriguez-Galiano et al., 2015b). The RF method is an innovative machine learning approach that can perform multivariate non-linear regression, combining the performance of numerous regression tree algorithms to predict the anomalies in OG and EOS. The RF method receives a subset of (x) input vectors, made up of one phenology anomaly value and the values of the corresponding climatic predictors considered in the regression. RF builds a number K of regression trees (in-

BGD
12, 11833–11861, 2015

dividual regression models) and averages the results (Breiman, 2001). After K such trees $\{T(x)\}_1^K$ are grown, the RF regression predictor becomes:

$$\hat{f}_{\mathrm{rf}}^{K}(x) = \frac{1}{K}\sum_{k=1}^{K} T(x)$$

More details regarding the performance and the specific characteristics of a RF model can be seen in Rodriguez-Galiano et al. (2014, 2015b).

The locations with anomalies in LSP greater than 1 (positive and negative) were selected to build a RF predictive model on OG and EOS. Anomaly values of OG or EOS for each year were combined together with the different climatic predictors. The anomaly in OG was assessed as an extra predictor to evaluate the legacy effect of an advanced or delayed spring in the modelling of EOS. The values of these variables at the selected years and locations (spatiotemporal model) were combined into a set of input feature vectors (3900 feature vectors for the spring model and 3124 for autumn) as an input to the RF algorithm. These feature vectors were divided equally into two subsets, one for the training of the models and one as an additional test to the one internally computed by RF (oob) to evaluate performance. RF models composed of 2000 trees were grown using different subsets of predictors, varying the number of random predictors from 1 to 9. Random Forest method within the package implemented in the R statistical software was used to build the different models (Liaw and Wiener, 2002).

2.5 Selection of the most important predictors

The RF method can use the oob subset to estimate the relative importance of each predictor in the model. This property is especially useful for the present research, but also for other multivariate biological studies, where it is important to know the physical drivers of the phenomenon under investigation (Lawler et al., 2006; Archibald et al., 2009). However, the inclusion of different measures of climatic drivers may imply a large

BGD
12, 11833–11861, 2015

increase in the dimensionality of the datasets being used, as these variables are obtained by applying multiple functions or measures to the temperature, precipitation and radiation time-series. On the one hand, more information may be useful for the modelling process; on the other hand, an excessive number of correlated predictors or features can overwhelm the expected increase in accuracy and may introduce additional complexity limiting the ability of the method to point to possible cause–effect relationships between anomalies in phenology and their drivers, making interpretation challenging.

A feature selection approach, based on the ability of the RF to assess the relative importance of the predictors, was used to identify the minimum number of drivers which can better explain spring or autumn anomalies in phenology. To assess the importance of each climatic driver, the RF switches one of the input predictors while keeping the rest constant, and it re-evaluates the performance of the model measuring the decrease in node impurity (Breiman, 2001). The differences were averaged over all 2000 trees. In order to reduce the number of drivers the least important predictor was removed iteratively at different steps. Then, a 5-fold cross-validation was applied to obtain a stable estimate of the error of the model built after predictor deletions. Finally, the model with a better trade-off between number of predictors and error was chosen as the basis for interpreting the likely drivers of changes in phenology.

3 Results

Random Forest method, a multivariate, spatially non-stationary and non-linear machine learning approach, was applied to phenological modelling across very large areas and across multiple years simultaneously: the typical case for satellite-observed LSP. The RF model was fitted to the relation between LSP anomalies and numerous climate predictor variables computed at biologically-relevant rather than human-imposed temporal scales. We restricted our climate data choices to daily data (average, minimum and maximum temperatures, precipitation and radiation) to account for integrative forcing (that is, growing degree days, chilling requirements as well as cumulative precipitation and radiation), computed from the exact day of the phenological event backwards, rather than using the absolute fixed dates of the calendar months. Numerous models were built on the basis of different predictor combinations considering different temporal windows prior to the spring and autumn phenological events (see Sect. 2.3). The percentage of variation (pseudo-R^2) explained by different climatic-LSP models is shown in the Supplement (Tables S1, S2 and S3). No previous studies have investigated in depth the parametrization of GDD for LSP and climate inter-comparison, unlike for ground phenological studies (Snyder et al., 1999). Although, we did not carry out an exhaustive analysis of the optimum GDD parametrization, our results showed a systematic pattern in spring models, presenting slightly larger pseudo-R^2 for models which used 0 °C as a threshold for the computation of GDD (rather than 5 °C). Regarding, the length of the temporal windows for climatic function computation, spring models using 30 and 90 days for the computation of averaged and cumulative functions were more accurate, whereas for autumn models with 90 day-averaged predictors outperformed the rest.

The main drivers of anomalies in LSP were identified through the application of a feature selection procedure (see section "selection of the most important predictors"). Figure 2 shows the relative error in the prediction of different models after removing the least important predictor. Spring models were more accurate than autumn, with median relative error values of 10 to 27 % (12 to 1 predictor), vs. 26 to 60 % of autumn (14 to 1 predictor). Figure 3 shows the pseudo-R^2 of the models as well as the relative importance of each predictor. Spring models (explained a percentage of the variance up to 81 % (Fig. 3a), whereas autumn explained up to 61 % (Fig. 3b). Regarding the relative importance of the drivers, the same ranking in importance was observed within the different models of each phenophase, which reflected the stability in the RF importance estimation, and a high reliability of the results. To interpret the main climatic drivers of the anomalies in phenology, simplified models with reduced number of predictors were selected for spring and autumn, respectively. The spring model was composed of 6 predictors (pseudo-R^2 = 0.77 and median relative error of 10 %) and the autumn model of 5 predictors (pseudo-R^2 = 0.59 and median relative error of 28 %). Our results suggest that anomalies in the onset on greenness (LSP) of temperate forest species are mainly driven by daily temperature (but not necessarily the GDD), with the most important driver being the minimum temperature of the 30 days prior to onset. Photoperiod was also important, the most accurate empirical prediction was obtained by a combined temperature-radiation forcing, integrating the SIS of the previous 90 days. For senescence, temperature was suggested to be more important than photoperiod in controlling the senescence process (Yang et al., 2012; Vitasse et al., 2009; Jeong and Medvigy, 2014; Archetti et al., 2013), with the most important drivers being the date of the first freeze and the accumulation of chilling temperatures. However, we did not observe a legacy effect of a much earlier or later spring onset on the date of senescence. Autumn models that included the anomaly in the onset of greenness did not outperform the remaining models (see Tables S2 and S3) and the relative importance was low in comparison with other drivers.

4 Discussion

The selection and computation of the climatic predictors is an important step of phenological modelling. Most of studies on the sensitivity of phenological events to climate used human calendar scales, that is, seasonal or monthly calendar mean or cumulative climate predictors (Menzel et al., 2006; Schwartz et al., 2006; Maignan et al., 2008a, b), overlooking the importance of biological time-scales in phenology. However, with the increased availability of daily climatic datasets, current and future studies might benefit from the use of daily information to model the drivers of plants' circadian time-scales (Pau et al., 2011). Our study advanced in the modelling of vegetation phenology by improving the temporal matching between LSP anomalies and the preceding climatic conditions using daily data and biological scales. Regarding, the length of the temporal windows for climatic function computation, Menzel et al. (2006) showed that most

BGD
12, 11833–11861, 2015

phenological phases of plant species in Europe correlate significantly with mean temperatures of the month of onset and the two preceding months. However, in our study, when end of senescence was considered, a consistent divergent effect was observed between spring and autumn. From a computational point of view, considering larger temporal windows for calculating averages would induce a smoothing effect, degrading the information in the predictors, whereas cumulative functions such as GDD or chilling requirements would not be affected by this effect. However, we observed a divergent response between spring and autumn and consistent throughout the models of each phenophase suggests that a biological explanation for this phenomenon might be plausible. Autumn phenophases might be driven by longer-term changes in weather, while for spring the average conditions of the 30 days previous to the date of onset play a more important role (Tables S1–S3).

Understanding the drivers of anomalies in LSP amidst background inter-annual variation is a critical aspect of global change science (de Beurs and Henebry, 2005; Zhao et al., 2013). To this end, the RF method is particularly pertinent, as it allows the assessment of the importance of the predictors (Fig. 2). Our findings reveal that the accuracy of growing degree day-based models is overestimated using linear regression models and that non-linear multivariate relationships between temperature (especially minimum temperature) and radiation are needed to describe the relations between phenology and climatic drivers. This supports the findings of Stöckli et al. (2011) who explained temperate phenology using a combination of light and temperature. The highlighted importance of minimum temperatures might be related to the fact that minimum temperature is a better indicator of climatic changes than either the average or maximum temperature (Jolly et al., 2005; Duncan et al., 2014). Regarding GDD, although it has been applied extensively to predict vegetation phenophases, it is currently debated whether such models can detect when multiple environmental drivers are required to initiate a phenological event, or detect drivers that are relatively static across time, such as photoperiod (Stöckli et al., 2011). Our results support this hypothesis and also showed that the role of GDD alone in driving spring phenology might

BGD
12, 11833–11861, 2015

be overestimated when linear models were considered. GDD had the largest linear association with vegetation phenology changes, while the linear correlation between LSP and others drivers that were revealed as very important by the RF was small (see Tables 1 and 2). A simple linear analysis between GDD and phenology could ignore complex non-linear associations between phenology and predictors as well as synergies between climatic drivers. Regarding the senescence phase, the autumn models had a weaker predictive power compared the spring models. There is still lack of clear understanding of mechanism autumn senescence, however, temperature, and particularly the dates of freeze, has been suggested as major driver for autumn phenology.

The RF method provided an important alternative over simple, but less accurate analysis based on linear regression for the analysis of changes in spring and autumn phenology. A further comparison with a linear regression analysis suggested that there might be a non-linear relationship between the anomalies in LSP and the climatic drivers. Multivariate linear regression models were also fitted from the same combination of predictors selected as optimal by Random Forest. Multivariate linear models explained only 36 and 26 % of the variance in spring and autumn phenology anomalies across the continental scale. Additionally, a linear regression between predicted values from RF and observed anomalies in phenology produced R^2 values equal to 0.90 and 0.68 for spring and autumn LSP anomalies, respectively (Fig. 4a and b). On the other hand, the correlations between the predictions of linear regression models and observations were much weaker, with R^2 values of 0.39 and 0.25 (Fig. 4c and d). Linear models under-predicted the positive anomalies and over-predicted the negative.

A new approach to model changes in LSP was presented in this paper based on the application of the RF model to a set of climate predictors at biological scales. This new modelling technique has numerous advantages for the modelling of climate-driven changes in LSP. It is a non-parametric multivariate method which allows for non-linear relationships between (compared to traditional linear models) phenology and climate and can consider a large number of climatic predictors in the modelling process. This provides potential opportunity to capture the impact of all possible envi-

BGD
12, 11833–11861, 2015

ronmental/climatic drivers on vegetation phenology. The proposed method can recognize complex patterns between LSP and climate at multiple locations and times, integrating them into a unique overall model, rather than generating multiple models over a geographical area and for different years. Additionally it is data-driven, which means that there is no need to incorporate previous knowledge about the specific responses of vegetation to different predominant climatic controls (i.e. temperature, rainfall, and photoperiod), allowing climatic drivers to automatically shift both temporally and spatially. Therefore, it is highly generalizable, being applicable to different biogeographical regions where the phenology is controlled by different factors. This flexibility or generalization capacity of RF models to transition from one driver to another without the need for a model change also promotes its application to different climate change scenarios. We succeeded in modelling the anomalies in LSP phenology as observed from satellite-sensors in the European Forest, while using the same type of input data, the same model, and the same model parameters for the entire European continent.

The Supplement related to this article is available online at doi:10.5194/bgd-12-11833-2015-supplement.

Author contributions. V. F. Rodriguez-Galiano, J. Dash and P. M. Atkinson conceived and designed the experiments; V. F. Rodriguez-Galiano performed the experiments; V. F. Rodriguez-Galiano, M. Sanchez-Castillo and J. Dash contributed analysis tools; V. F. Rodriguez-Galiano drafted the paper. All authors contributed to the final paper.

Acknowledgements. The first author is a Marie Curie Grant holder (reference FP7-PEOPLE-2012-IEF-331667). The authors are grateful for the financial support given by the European Commission under the Seventh Framework Programme and the Spanish MINECO (project BIA2013-43462-P). P. M. Atkinson is grateful to the University of Utrecht for supporting him with The Belle van Zuylen Chair. We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu). Surface radiation data were obtained from EUMETSAT's Satellite Application Facility on Climate Monitoring (CM SAF).

BGD

12, 11833–11861, 2015

References

Archetti, M., Richardson, A. D., O'Keefe, J., and Delpierre, N.: Predicting climate change impacts on the amount and duration of autumn colors in a New England forest, PLoS ONE, 8, e57373, doi:10.1371/journal.pone.0057373, 2013.

Archibald, S., Roy, D. P., van Wilgen, B. W., and Scholes, R. J.: What limits fire? An examination of drivers of burnt area in southern Africa, Glob. Change Biol., 15, 613–630, doi:10.1111/j.1365-2486.2008.01754.x, 2009.

Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M., and García-Herrera, R.: The hot summer of 2010: redrawing the temperature record map of Europe, Science, 332, 220–224, doi:10.1126/science.1201224, 2011.

Bicheron, P., Amberg, V., Bourg, L., Petit, D., Huc, M., Miras, B., Brockmann, C., Hagolle, O., Delwart, S., Ranera, F., Leroy, M., and Arino, O.: Geolocation assessment of MERIS GlobCover orthorectified products, IEEE T. Geosci. Remote, 49, 2972–2982, doi:10.1109/tgrs.2011.2122337, 2011.

Breiman, L.: Random forests, Mach. Learn., 45, 5–32, 2001.

Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A.: Classification and Regression Trees, 1st edn., Chapman and Hall/CRC, Belmont, CA, 1984.

Brown, M. E. and de Beurs, K. M.: Evaluation of multi-sensor semi-arid crop season parameters based on NDVI and rainfall, Remote Sens. Environ., 112, 2261–2271, doi:10.1016/j.rse.2007.10.008, 2008.

Darling, E. S., Alvarez-Filip, L., Oliver, T. A., McClanahan, T. R., and Côté, I. M.: Evaluating life-history strategies of reef corals from species traits, Ecol. Lett., 15, 1378–1386, doi:10.1111/j.1461-0248.2012.01861.x, 2012.

Dash, J., Jeganathan, C., and Atkinson, P. M.: The use of MERIS Terrestrial Chlorophyll Index to study spatio-temporal variation in vegetation phenology over India, Remote Sens. Environ., 114, 1388–1402, 2010.

de Beurs, K. M. and Henebry, G. M.: Land surface phenology and temperature variation in the International Geosphere–Biosphere Program high-latitude transects, Glob. Change Biol., 11, 779–790, 2005.

Delbart, N., Picard, G., Le Toan, T., Kergoat, L., Quegan, S., Woodward, I., Dye, D., and Fedotova, V.: Spring phenology in boreal Eurasia over a nearly century time scale, Glob. Change Biol., 14, 603–614, 2008.

Duncan, J. M. A., Dash, J., and Atkinson, P. M.: Elucidating the impact of temperature variability and extremes on cereal croplands through remote sensing, Glob. Change Biol., 21, 1541–1551, doi:10.1111/gcb.12660, 2014.
Fu, Y. S. H., Campioli, M., Vitasse, Y., De Boeck, H. J., Van Den Berge, J., Abdelgawad, H., Asard, H., Piao, S., Deckmyn, G., and Janssens, I. A.: Variation in leaf flushing date influences autumnal senescence and next year's flushing date in two temperate tree species, P. Natl. Acad. Sci. USA, 111, 7355–7360, 2014.
Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, J. Geophys. Res., 113, D20119, doi:10.1029/2008JD010201, 2008.
Ivits, E., Cherlet, M., Tóth, G., Sommer, S., Mehl, W., Vogt, J., and Micale, F.: Combining satellite derived phenology with climate data for climate change impact assessment, Global Planet. Change, 88–89, 85–97, 2012.
Jeganathan, C., Dash, J., and Atkinson, P. M.: Remotely sensed trends in the phenology of northern high latitude terrestrial vegetation, controlling for land cover change and vegetation type, Remote Sens. Environ., 143, 154–170, doi:10.1016/j.rse.2013.11.020, 2014.
Jeong, S.-J. and Medvigy, D.: Macroscale prediction of autumn leaf coloration throughout the continental United States, Global Ecol. Biogeogr., 23, 1245–1254, doi:10.1111/geb.12206, 2014.
Jeong, S.-J., Ho, C.-H., Gim, H.-J., and Brown, M. E.: Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008, Glob. Change Biol., 17, 2385–2399, doi:10.1111/j.1365-2486.2011.02397.x, 2011.
Jolly, W. M., Nemani, R., and Running, S. W.: A generalized, bioclimatic index to predict foliar phenology in response to climate, Glob. Change Biol., 11, 619–632, doi:10.1111/j.1365-2486.2005.00930.x, 2005.
Karlsen, S. R., Solheim, I., Beck, P. S. A., Hogda, K. A., Wielgolaski, F. E., and Tommervik, H.: Variability of the start of the growing season in Fennoscandia, 1982–2002, Int. J. Biometeorol., 51, 513–524, doi:10.1007/s00484-007-0091-x, 2007.
Lawler, J. J., White, D., Neilson, R. P., and Blaustein, A. R.: Predicting climate-induced range shifts: model differences and model reliability, Glob. Change Biol., 12, 1568–1584, doi:10.1111/j.1365-2486.2006.01191.x, 2006.
Lebourgeois, F., Pierrat, J. C., Perez, V., Piedallu, C., Cecchini, S., and Ulrich, E.: Simulating phenological shifts in French temperate forests under two climatic change scenarios and four

driving global circulation models, Int. J. Biometeorol., 54, 563–581, doi:10.1007/s00484-010-0305-5, 2010.
Liaw, A. and Wiener, M.: Classification and regression by randomForest, R News, 2/3, 18–22, 2002.
Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., and Wanner, H.: European seasonal and annual temperature variability, trends, and extremes since 1500, Science, 303, 1499–1503, doi:10.1126/science.1093877, 2004.
Maignan, F., Bréon, F. M., Bacour, C., Demarty, J., and Poirson, A.: Interannual vegetation phenology estimates from global AVHRR measurements: comparison with in situ data and applications, Remote Sens. Environ., 112, 496–505, doi:10.1016/j.rse.2007.05.011, 2008a.
Maignan, F., Bréon, F. M., Vermote, E., Ciais, P., and Viovy, N.: Mild winter and spring 2007 over western Europe led to a widespread early vegetation onset, Geophys. Res. Lett., 35, L02404, doi:10.1029/2007GL032472, 2008b.
Menzel, A.: Phenology: its importance to the global change community, Climatic Change, 54, 379–385, doi:10.1023/a:1016125215496, 2002.
Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aaasa, A., Ahas, R., Alm-Kübler, K., Bissolli, P., Braslavská, O., Briede, A., Chmielewski, F. M., Crepinsek, Z., Curnel, Y., Dahl, Å., Defila, C., Donnelly, A., Filella, Y., Jatczak, K., Måge, F., Mestre, A., Nordli, Ø., Peñuelas, J., Pirinen, P., Remišová, V., Scheifinger, H., Striz, M., Susnik, A., Van Vliet, A. J. H., Wielgolaski, F. E., Zach, S., and Zust, A.: European phenological response to climate change matches the warming pattern, Glob. Change Biol., 12, 1969–1976, 2006.
Morisette, J. T., Richardson, A. D., Knapp, A. K., Fisher, J. I., Graham, E. A., Abatzoglou, J., Wilson, B. E., Breshears, D. D., Henebry, G. M., Hanes, J. M., and Liang, L.: Tracking the rhythm of the seasons in the face of global change: phenological research in the 21st century, Front. Ecol. Environ., 7, 253–260, doi:10.1890/070217, 2008.
Müller, R. and Trentmann, J.: CM SAF Meteosat Surface Radiation Daylight Data Set 1.0 – Monthly Means/Daily Means, Satellite Application Facility on Climate Monitoring, doi:10.5676/EUM_SAF_CM/DAL_MVIRI_SEVIRI/V001, 2013.
Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G., and Nemani, R. R.: Increased plant growth in the northern high latitudes from 1981 to 1991, Nature, 386, 698–702, doi:10.1038/386698a0, 1997.
Pau, S., Wolkovich, E. M., Cook, B. I., Davies, T. J., Kraft, N. J. B., Bolmgren, K., Betancourt, J. L., and Cleland, E. E.: Predicting phenology by integrating ecology, evolution and

BGD
12, 11833–11861, 2015

Modelling anomalies in the spring and autumn land surface phenology

V. F. Rodriguez-Galiano et al.

climate science, Glob. Change Biol., 17, 3633–3643, doi:10.1111/j.1365-2486.2011.02515.x, 2011.
Peñuelas, J.: Phenology feedbacks on climate change, Science, 324, 887–888, 2009.
Peñuelas, J. and Filella, I.: Phenology: responses to a warming world, Science, 294, 793–795, 2001.
Posselt, R., Müller, R., Stöckli, R., and Trentmann, J.: CM SAF Surface Radiation MVIRI Data Set 1.0 – Monthly Means/Daily Means/Hourly Means, Satellite Application Facility on Climate Monitoring, doi:10.5676/EUM_SAF_CM/RAD_MVIRI/V001, 2011.
Posselt, R., Mueller, R. W., Stöckli, R., and Trentmann, J.: Remote sensing of solar surface radiation for climate monitoring – the CM-SAF retrieval in international comparison, Remote Sens. Environ., 118, 186–198, doi:10.1016/j.rse.2011.11.016, 2012.
Rafferty, N. E., CaraDonna, P. J., Burkle, L. A., Iler, A. M., and Bronstein, J. L.: Phenological overlap of interacting species in a changing climate: an assessment of available approaches, Ecology and Evolution, 3, 3183–3193, doi:10.1002/ece3.668, 2013.
Rodriguez-Galiano, V. F., Chica-Olmo, M., and Chica-Rivas, M.: Predictive modelling of gold potential with the integration of multisource information based on random forest: a case study on the Rodalquilar area, southern Spain, Int. J. Geogr. Inf. Sci., 28, 1336–1354, doi:10.1080/13658816.2014.885527, 2014.
Rodriguez-Galiano, V., Dash, J., and Atkinson, P. M.: Inter-comparison of satellite sensor land surface phenology and ground phenology in Europe, Geophys. Res. Lett., 42, 2253–2260, doi:10.1002/2015GL063586, 2015a.
Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., and Chica-Rivas, M.: Machine learning predictive models for mineral prospectivity: an evaluation of neural networks, random forest, regression trees and support vector machines, Ore Geol. Rev., doi:10.1016/j.oregeorev.2015.01.001, in press, 2015b.
Rutishauser, T., Luterbacher, J., Defila, C., Frank, D., and Wanner, H.: Swiss spring plant phenology 2007: extremes, a multi-century perspective, and changes in temperature sensitivity, Geophys. Res. Lett., 35, L05703, 2008.
Saleska, S. R., Didan, K., Huete, A. R., and Da Rocha, H. R.: Amazon forests green-up during 2005 drought, Science, 318, 612, doi:10.1126/science.1146663, 2007.
Schwartz, M. D., Ahas, R., and Aasa, A.: Onset of spring starting earlier across the Northern Hemisphere, Glob. Change Biol., 12, 343–351, 2006.

Snyder, R. L., Spano, D., Cesaraccio, C., and Duce, P.: Determining degree-day thresholds from field observations, Int. J. Biometeorol., 42, 177–182, doi:10.1007/s004840050102, 1999.
Stöckli, R., Rutishauser, T., Dragoni, D., O'Keefe, J., Thornton, P. E., Jolly, M., Lu, L., and Denning, A. S.: Remote sensing data assimilation for a prognostic phenology model, J. Geophys. Res.-Biogeo., 113, G04021, doi:10.1029/2008JG000781, 2008.
Stöckli, R., Rutishauser, T., Baker, I., Liniger, M. A., and Denning, A. S.: A global reanalysis of vegetation phenology, J. Geophys. Res.-Biogeo., 116, G03020, doi:10.1029/2010JG001545, 2011.
van Vliet, A. H.: Societal adaptation options to changes in phenology, in: Phenological Research, edited by: Hudson, I. L. and Keatley, M. R., Springer, Netherlands, 75–98, 2010.
Vitasse, Y., Delzon, S., Dufrêne, E., Pontailler, J. Y., Louvet, J. M., Kremer, A., and Michalet, R.: Leaf phenology sensitivity to temperature in European trees: do within-species populations exhibit similar responses?, Agr. Forest Meteorol., 149, 735–744, doi:10.1016/j.agrformet.2008.10.019, 2009.
Yang, X., Mustard, J. F., Tang, J. W., and Xu, H.: Regional-scale phenology modeling based on meteorological records and remote sensing observations, J. Geophys. Res.-Biogeo., 117, G03029, doi:10.1029/2012jg001977, 2012.
Zhao, M. F., Peng, C. H., Xiang, W. H., Deng, X. W., Tian, D. L., Zhou, X. L., Yu, G. R., He, H. L., and Zhao, Z. H.: Plant phenological modeling and its application in global climate change research: overview and future challenges, Environ. Rev., 21, 1–14, doi:10.1139/er-2012-0036, 2013.
Zhou, L. M., Tucker, C. J., Kaufmann, R. K., Slayback, D., Shabanov, N. V., and Myneni, R. B.: Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999, J. Geophys. Res.-Atmos., 106, 20069–20083, doi:10.1029/2000jd000115, 2001.

BGD
12, 11833–11861, 2015

Modelling anomalies in the spring and autumn land surface phenology

V. F. Rodriguez-Galiano et al.

Table 1. Correlations between the predictors used in the modelling of spring anomalies. Significant correlations between the anomalies and the predictors are given in bold ($p < 0.05$).

		1	2	3	4	5	6	7	8	9	10	11	12	13
1	Anom.	**1.00**	**−0.40**	**−0.43**	**−0.11**	**−0.09**	**−0.12**	**−0.10**	**−0.11**	**−0.10**	**0.24**	−0.03	−0.03	−0.03
2	GDD090	**−0.40**	1.00	0.93	0.11	0.14	0.11	0.13	0.11	0.15	−0.64	0.00	−0.01	−0.01
3	GDD590	**−0.43**	0.93	1.00	0.11	0.10	0.11	0.10	0.11	0.11	−0.47	−0.01	−0.01	−0.01
4	MTG30	**−0.11**	0.11	0.11	1.00	0.99	1.00	0.99	1.00	0.98	−0.05	0.89	0.89	0.89
5	MTG90	**−0.09**	0.14	0.10	0.99	1.00	0.98	1.00	0.99	1.00	−0.13	0.88	0.88	0.88
6	MTX30	**−0.12**	0.11	0.11	1.00	0.98	1.00	0.99	0.99	0.98	−0.04	0.89	0.89	0.88
7	MTX90	**−0.10**	0.13	0.10	0.99	1.00	0.99	1.00	0.99	1.00	−0.11	0.89	0.89	0.89
8	MTN30	**−0.11**	0.11	0.11	1.00	0.99	0.99	0.99	1.00	0.98	−0.06	0.89	0.89	0.89
9	MTN90	**−0.10**	0.15	0.11	0.98	1.00	0.98	1.00	0.98	1.00	−0.15	0.88	0.88	0.88
10	FF	**0.24**	−0.64	−0.47	−0.05	−0.13	−0.04	−0.11	−0.06	−0.15	1.00	−0.01	0.00	0.00
11	FF	−0.03	0.00	−0.01	0.89	0.88	0.89	0.89	0.89	0.88	−0.01	1.00	1.00	1.00
12	LF	−0.03	−0.01	−0.01	0.89	0.88	0.89	0.89	0.89	0.88	0.00	1.00	1.00	1.00
13	PF	−0.03	−0.01	−0.01	0.89	0.88	0.88	0.89	0.89	0.88	0.00	1.00	1.00	1.00
14	CRR90	**−0.14**	0.23	0.16	0.20	0.25	0.19	0.23	0.21	0.26	−0.25	−0.01	−0.01	−0.02
15	MRR30	−0.04	0.01	0.01	0.97	0.96	0.96	0.96	0.96	0.96	0.00	0.88	0.88	0.88
16	MRR90	−0.04	0.01	0.01	0.96	0.96	0.96	0.96	0.96	0.96	0.00	0.88	0.88	0.88
17	CSIS90	**−0.33**	−0.12	0.03	0.02	−0.03	0.03	−0.03	0.02	−0.04	0.28	−0.04	−0.04	−0.04
18	MSIS30	**−0.16**	−0.06	0.04	0.00	−0.03	0.00	−0.03	0.01	−0.03	0.11	−0.05	−0.05	−0.05
19	MSIS90	**−0.16**	−0.06	0.04	0.00	−0.03	0.00	−0.03	0.01	−0.03	0.11	−0.05	−0.05	−0.05
20	CDAL90	−0.04	0.04	0.06	0.31	0.30	0.32	0.30	0.31	0.29	0.03	0.00	0.00	0.00
21	MDAL30	**−0.06**	−0.05	0.03	−0.01	−0.04	−0.01	−0.04	0.00	−0.03	0.06	−0.06	−0.06	−0.06
22	MDAL90	**−0.06**	−0.05	0.03	−0.01	−0.04	−0.01	−0.04	0.00	−0.03	0.06	−0.06	−0.06	−0.06
23	GDD030	**−0.45**	0.67	0.74	0.17	0.10	0.18	0.10	0.16	0.10	−0.24	0.00	−0.01	−0.01
24	GDD530	**−0.46**	0.64	0.75	0.15	0.09	0.16	0.09	0.14	0.09	−0.26	−0.01	−0.01	−0.01
25	CRR30	**−0.12**	0.18	0.16	0.28	0.29	0.27	0.28	0.29	0.30	−0.16	−0.01	−0.01	−0.01
26	CSIS30	**−0.31**	−0.11	0.03	0.07	0.02	0.08	0.02	0.06	0.02	0.26	−0.03	−0.03	−0.03
27	CDAL30	−0.03	0.05	0.06	0.31	0.31	0.32	0.31	0.31	0.30	0.01	0.00	0.00	0.00

BGD
12, 11833–11861, 2015

Table 1. Continued.

		14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	Anom.	**−0.14**	−0.04	−0.04	**−0.33**	**−0.16**	**−0.16**	−0.04	**−0.06**	**−0.06**	**−0.45**	**−0.46**	**−0.12**	**−0.31**	−0.03
2	GDD090	0.23	0.01	0.01	−0.12	−0.06	−0.06	0.04	−0.05	−0.05	0.67	0.64	0.18	−0.11	0.05
3	GDD590	0.16	0.01	0.01	0.03	0.04	0.04	0.06	0.03	0.03	0.74	0.75	0.16	0.03	0.06
4	MTG30	0.20	0.97	0.96	0.02	0.00	0.00	0.31	−0.01	−0.01	0.17	0.15	0.28	0.07	0.31
5	MTG90	0.25	0.96	0.96	−0.03	−0.03	−0.03	0.30	−0.04	−0.04	0.10	0.09	0.29	0.02	0.31
6	MTX30	0.19	0.96	0.96	0.03	0.00	0.00	0.32	−0.01	−0.01	0.18	0.16	0.27	0.08	0.32
7	MTX90	0.23	0.96	0.96	−0.03	−0.03	−0.03	0.30	−0.04	−0.04	0.10	0.09	0.28	0.02	0.31
8	MTN30	0.21	0.96	0.96	0.02	0.01	0.01	0.31	0.00	0.00	0.16	0.14	0.29	0.06	0.31
9	MTN90	0.26	0.96	0.96	−0.04	−0.03	−0.03	0.29	−0.03	−0.03	0.10	0.09	0.30	0.02	0.30
10	FF	−0.25	0.00	0.00	0.28	0.11	0.11	0.03	0.06	0.06	−0.24	−0.26	−0.16	0.26	0.01
11	FF	−0.01	0.88	0.88	−0.04	−0.05	−0.05	0.00	−0.06	−0.06	0.00	−0.01	−0.01	−0.03	0.00
12	LF	−0.01	0.88	0.88	−0.04	−0.05	−0.05	0.00	−0.06	−0.06	−0.01	−0.01	−0.01	−0.03	0.00
13	PF	−0.02	0.88	0.88	−0.04	−0.05	−0.05	0.00	−0.06	−0.06	−0.01	−0.01	−0.01	−0.03	0.00
14	CRR90	1.00	0.20	0.20	0.01	0.06	0.06	0.53	0.04	0.04	0.09	0.07	0.77	0.11	0.58
15	MRR30	0.20	1.00	1.00	0.00	−0.03	−0.03	0.31	−0.03	−0.03	0.03	0.03	0.26	0.05	0.31
16	MRR90	0.20	1.00	1.00	0.00	−0.03	−0.03	0.31	−0.03	−0.03	0.03	0.02	0.26	0.05	0.31
17	CSIS90	0.01	0.00	0.00	1.00	0.80	0.80	0.16	0.57	0.57	0.22	0.22	0.12	0.96	0.15
18	MSIS30	0.06	−0.03	−0.03	0.80	1.00	1.00	0.06	0.90	0.90	0.23	0.24	0.15	0.77	0.06
19	MSIS90	0.06	−0.03	−0.03	0.80	1.00	1.00	0.06	0.90	0.90	0.23	0.24	0.15	0.77	0.06
20	CDAL90	0.53	0.31	0.31	0.16	0.06	0.06	1.00	0.05	0.05	0.11	0.10	0.78	0.28	0.99
21	MDAL30	0.04	−0.03	−0.03	0.57	0.90	0.90	0.05	1.00	1.00	0.23	0.23	0.13	0.55	0.05
22	MDAL90	0.04	−0.03	−0.03	0.57	0.90	0.90	0.05	1.00	1.00	0.23	0.23	0.13	0.55	0.05
23	GDD030	0.09	0.03	0.03	0.22	0.23	0.23	0.11	0.23	0.23	1.00	0.97	0.16	0.23	0.11
24	GDD530	0.07	0.03	0.02	0.22	0.24	0.24	0.10	0.23	0.23	0.97	1.00	0.15	0.24	0.10
25	CRR30	0.77	0.26	0.26	0.12	0.15	0.15	0.78	0.13	0.13	0.16	0.15	1.00	0.18	0.79
26	CSIS30	0.11	0.05	0.05	0.96	0.77	0.77	0.28	0.55	0.55	0.23	0.24	0.18	1.00	0.28
27	CDAL30	0.58	0.31	0.31	0.15	0.06	0.06	0.99	0.05	0.05	0.11	0.10	0.79	0.28	1.00

BGD

12, 11833–11861, 2015

Modelling anomalies in the spring and autumn land surface phenology

V. F. Rodriguez-Galiano et al.

Table 2. Correlations between the predictors used in the modelling of autumn anomalies. Significant correlations between the anomalies and the predictors are given in bold ($p < 0.05$).

		1	2	3	4	5	6	7	8	9	10	11	12	13
1	Anom.	1	**0.10**	**0.31**	**0.34**	**0.33**	**0.36**	**0.28**	**0.30**	**0.28**	**0.27**	**0.26**	**0.34**	0.01
2	OGANO	0.10	1.00	0.06	0.08	0.14	0.16	0.05	0.15	0.02	0.07	0.05	0.19	−0.02
3	GDD030	0.31	0.06	1.00	0.97	0.54	0.58	0.94	0.53	0.88	0.42	0.87	0.62	−0.54
4	GDD530	0.34	0.08	0.97	1.00	0.53	0.60	0.86	0.49	0.80	0.37	0.80	0.59	−0.41
5	GDD090	0.33	0.14	0.54	0.53	1.00	0.98	0.49	0.95	0.54	0.90	0.36	0.85	−0.14
6	GDD590	0.36	0.16	0.58	0.60	0.98	1.00	0.49	0.92	0.54	0.85	0.37	0.84	−0.10
7	MTG30	0.28	0.05	0.94	0.86	0.49	0.49	1.00	0.56	0.93	0.44	0.94	0.63	−0.71
8	MTG90	0.30	0.15	0.53	0.49	0.95	0.92	0.56	1.00	0.61	0.93	0.43	0.89	−0.28
9	MTX30	0.28	0.02	0.88	0.80	0.54	0.54	0.93	0.61	1.00	0.58	0.78	0.60	−0.58
10	MTX90	0.27	0.07	0.42	0.37	0.90	0.85	0.44	0.93	0.58	1.00	0.28	0.73	−0.16
11	MTN30	0.26	0.05	0.87	0.80	0.36	0.37	0.94	0.43	0.78	0.28	1.00	0.61	−0.76
12	MTN90	0.34	0.19	0.62	0.59	0.85	0.84	0.63	0.89	0.60	0.73	0.61	1.00	−0.39
13	CHIL30	0.01	−0.02	−0.54	−0.41	−0.14	−0.10	−0.71	−0.28	−0.58	−0.16	−0.76	−0.39	1.00
14	CHIL90	−0.03	−0.04	−0.52	−0.40	−0.24	−0.20	−0.66	−0.36	−0.54	−0.24	−0.70	−0.48	0.91
15	FFN	**0.34**	0.01	0.25	0.24	0.12	0.14	0.24	0.12	0.20	0.09	0.26	0.19	−0.08
16	CRR30	**0.07**	0.02	0.09	0.11	0.05	0.07	0.04	−0.01	−0.09	−0.05	0.16	0.12	−0.05
17	MRR30	**0.07**	−0.05	0.10	0.11	0.13	0.13	0.10	0.13	0.12	0.13	0.08	0.13	0.00
18	MRR90	0.04	−0.07	0.11	0.10	0.09	0.09	0.09	0.09	0.07	0.05	0.09	0.12	0.01
19	CRR90	**−0.05**	0.06	0.03	0.07	−0.15	−0.11	−0.01	−0.18	−0.09	−0.31	0.08	0.04	−0.05
20	MSIS30	−0.05	−0.02	−0.09	−0.10	−0.07	−0.07	−0.02	0.02	0.03	0.02	−0.06	−0.02	−0.05
21	MSIS90	−0.05	−0.02	−0.09	−0.10	−0.07	−0.07	−0.02	0.02	0.03	0.02	−0.06	−0.02	−0.05
22	CSIS30	0.00	−0.10	−0.01	−0.03	0.04	0.02	0.02	0.07	0.23	0.17	−0.17	−0.07	0.09
23	CSIS90	−0.01	−0.11	0.01	−0.01	−0.05	−0.06	0.05	−0.01	0.14	0.07	−0.04	−0.12	−0.01
24	MDAL30	**−0.08**	0.01	−0.22	−0.23	−0.14	−0.14	−0.13	−0.03	−0.09	−0.03	−0.14	−0.06	−0.01
25	MDAL90	**−0.08**	0.01	−0.22	−0.23	−0.14	−0.14	−0.13	−0.03	−0.09	−0.03	−0.14	−0.06	−0.01
26	CDAL30	**−0.09**	−0.06	−0.11	−0.15	0.08	0.04	−0.09	0.09	0.17	0.23	−0.30	−0.08	0.17
27	CDAL90	**−0.15**	−0.10	−0.22	−0.25	−0.14	−0.19	−0.17	−0.11	−0.06	0.07	−0.24	−0.31	0.10

BGD

12, 11833–11861, 2015

Table 2. Continued.

		14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	Anom.	−0.03	**0.34**	**0.07**	**0.07**	0.04	**−0.05**	−0.05	−0.05	0.00	−0.01	**−0.08**	**−0.08**	**−0.09**	**−0.15**
2	OGANO	−0.04	0.01	0.02	−0.05	−0.07	0.06	−0.02	−0.02	−0.10	−0.11	0.01	0.01	−0.06	−0.10
3	GDD030	−0.52	0.25	0.09	0.10	0.11	0.03	−0.09	−0.09	−0.01	0.01	−0.22	−0.22	−0.11	−0.22
4	GDD530	−0.40	0.24	0.11	0.11	0.10	0.07	−0.10	−0.10	−0.03	−0.01	−0.23	−0.23	−0.15	−0.25
5	GDD090	−0.24	0.12	0.05	0.13	0.09	−0.15	−0.07	−0.07	0.04	−0.05	−0.14	−0.14	0.08	−0.14
6	GDD590	−0.20	0.14	0.07	0.13	0.09	−0.11	−0.07	−0.07	0.02	−0.06	−0.14	−0.14	0.04	−0.19
7	MTG30	−0.66	0.24	0.04	0.10	0.09	−0.01	−0.02	−0.02	0.02	0.05	−0.13	−0.13	−0.09	−0.17
8	MTG90	−0.36	0.12	−0.01	0.13	0.09	−0.18	0.02	0.02	0.07	−0.01	−0.03	−0.03	0.09	−0.11
9	MTX30	−0.54	0.20	−0.09	0.12	0.07	−0.09	0.03	0.03	0.23	0.14	−0.09	−0.09	0.17	−0.06
10	MTX90	−0.24	0.09	−0.05	0.13	0.05	−0.31	0.02	0.02	0.17	0.07	−0.03	−0.03	0.23	0.07
11	MTN30	−0.70	0.26	0.16	0.08	0.09	0.08	−0.06	−0.06	−0.17	−0.04	−0.14	−0.14	−0.30	−0.24
12	MTN90	−0.48	0.19	0.12	0.13	0.12	0.04	−0.02	−0.02	−0.07	−0.12	−0.06	−0.06	−0.08	−0.31
13	CHIL30	0.91	−0.08	−0.05	0.00	0.01	−0.05	−0.05	−0.05	0.09	−0.01	−0.01	−0.01	0.17	0.10
14	CHIL90	1.00	−0.09	−0.04	0.00	0.01	−0.05	−0.08	−0.08	0.08	0.01	−0.04	−0.04	0.16	0.15
15	FFN	−0.09	1.00	−0.10	0.05	0.04	−0.08	0.01	0.01	0.01	0.07	−0.05	−0.05	−0.08	−0.04
16	CRR30	−0.04	−0.10	1.00	0.12	0.04	0.51	−0.17	−0.17	−0.42	−0.25	−0.12	−0.12	−0.46	−0.25
17	MRR30	0.00	0.05	0.12	1.00	0.47	0.08	−0.03	−0.03	−0.02	−0.03	−0.03	−0.03	−0.02	−0.04
18	MRR90	0.01	0.04	0.04	0.47	1.00	0.06	−0.01	−0.01	−0.02	−0.04	−0.02	−0.02	−0.02	−0.08
19	CRR90	−0.05	−0.08	0.51	0.08	0.06	1.00	−0.04	−0.05	−0.14	−0.18	−0.05	−0.05	−0.20	−0.39
20	MSIS30	−0.08	0.01	−0.17	−0.03	−0.01	−0.04	1.00	1.00	0.56	0.66	0.88	0.88	0.05	−0.04
21	MSIS90	−0.08	0.01	−0.17	−0.03	−0.01	−0.05	1.00	1.00	0.55	0.66	0.88	0.88	0.05	−0.04
22	CSIS30	0.08	0.01	−0.42	−0.02	−0.02	−0.14	0.56	0.55	1.00	0.80	0.30	0.30	0.66	0.28
23	CSIS90	0.01	0.07	−0.25	−0.03	−0.04	−0.18	0.66	0.66	0.80	1.00	0.31	0.31	0.18	0.40
24	MDAL30	−0.04	−0.05	−0.12	−0.03	−0.02	−0.05	0.88	0.88	0.30	0.31	1.00	1.00	0.05	−0.05
25	MDAL90	−0.04	−0.05	−0.12	−0.03	−0.02	−0.05	0.88	0.88	0.30	0.31	1.00	1.00	0.05	−0.05
26	CDAL30	0.16	−0.08	−0.46	−0.02	−0.02	−0.20	0.05	0.05	0.66	0.18	0.05	0.05	1.00	0.41
27	CDAL90	0.15	−0.04	−0.25	−0.04	−0.08	−0.39	−0.04	−0.04	0.28	0.40	−0.05	−0.05	0.41	1.00

BGD
12, 11833–11861, 2015

V. F. Rodriguez-Galiano et al.

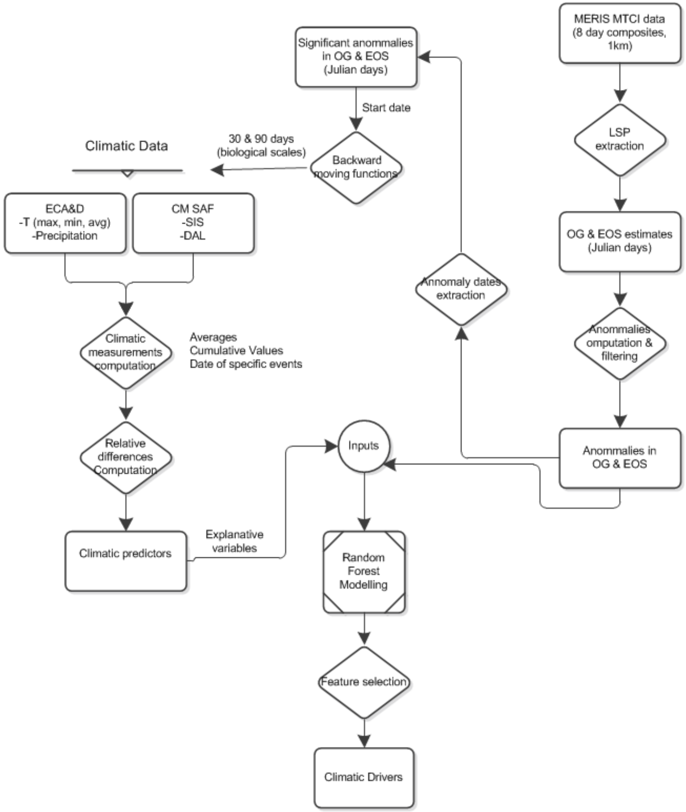

Figure 1. Flow-chart illustrating the methodology.

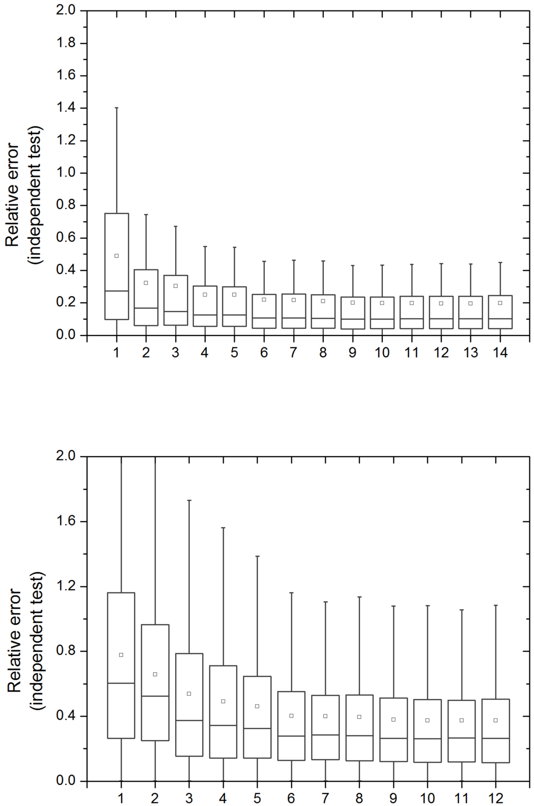

Figure 2. Relative error of the models fitted as a result of the feature selection approach. Median (interior horizontal line), mean (interior square), 1 and 99 % quantiles (edge of boxes), range (extremes). Relative errors were calculated for the prediction of 1974 and 1576 independent observations for spring **(a)** and autumn **(b)**, respectively.

BGD
12, 11833–11861, 2015

Modelling anomalies in the spring and autumn land surface phenology

V. F. Rodriguez-Galiano et al.

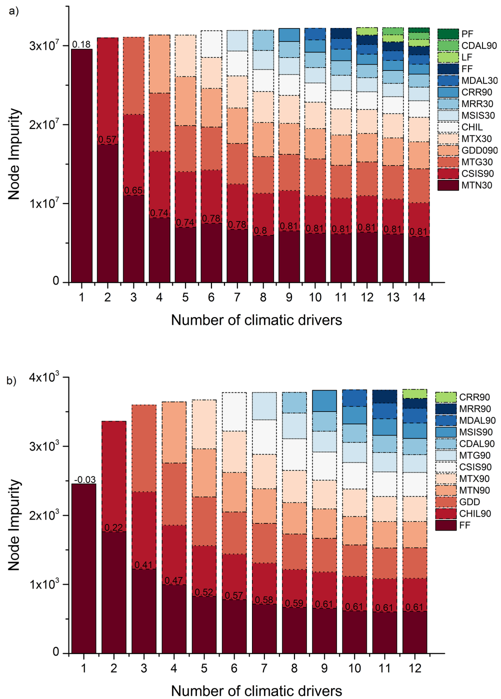

Figure 3. Relative importance of each independent variable in predicting phenology anomalies. Different models derived from the feature selection approach are represented in each column. Numbers given within each column represent the pseudo-square correlation coefficient of each model. Plots at the top and bottom represent the spring **(a)** and autumn anomalies **(b)**, respectively. The names of predictors follows the notation: prefix M and C represent the mean and cumulated functions; TX, TN and TG: maximum, minimum and average temperature, respectively; PP: precipitation; SIS: surface incoming shortwave radiation; DAL: surface radiation daylight; GDD: growing degree days; CHIL: chilling requirements; FF, LF and PF: first, last and period of freeze, respectively.

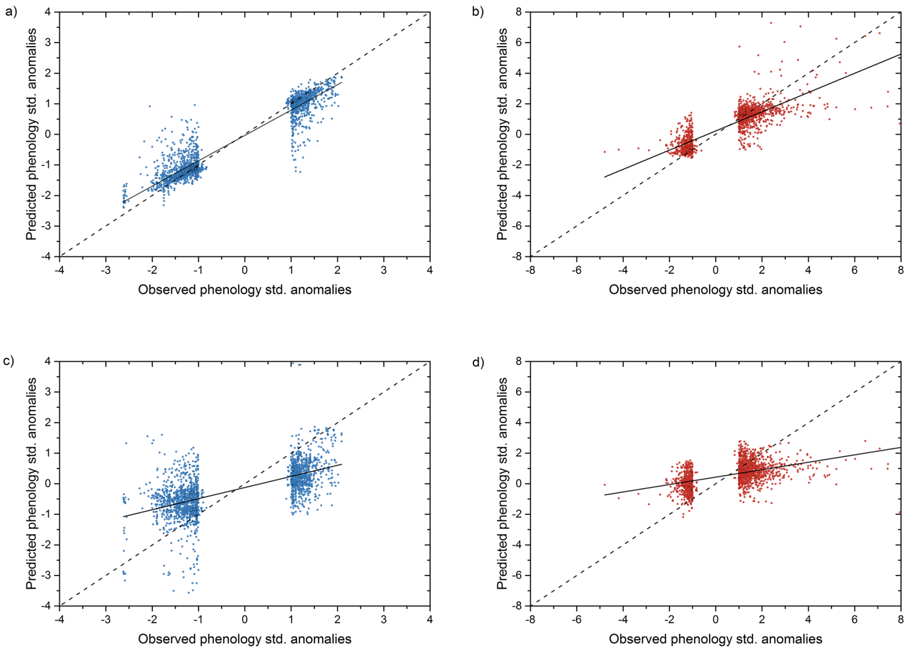

Figure 4. Scatterplots between observed anomalies in LSP and the predictions calculated using a selection of climatic predictors (see Figs. 2 and 3). Plots for spring phenology are shown on the left panel (blue; **a**, **c**) and autumn on the right (red; **b**, **d**). Random Forest predictions are given in the upper panel **(a, b)** and those of the linear regression in the bottom **(c, d)** panel. The dashed lines represent an exact 1 : 1 relationship (expected fitting), the solid lines show a linear regression of these data. The explained variances (percentage R^2 of regression line) are 90 % (spring Random Forest model), 68 % (autumn Random Forest model), 39 % (spring Linear model) and 25 % (autumn linear model).

BGD

12, 11833–11861, 2015