

1 **Modelling interannual variation in the spring and autumn**
2 **land surface phenology of the European forest**

3
4 **V.F. Rodriguez-Galiano^{1,2}, M. Sanchez-Castillo³, J. Dash², P.M. Atkinson^{4,5,6,2} and**
5 **J. Ojeda-Zujar¹**

6 [1] Physical Geography and Regional Geographic Analysis, University of Seville, Seville
7 41004, Spain

8 [2] Global Environmental Change and Earth Observation Research Group, Geography and
9 Environment, University of Southampton, Southampton SO17 1BJ, United Kingdom

10 [3] Department of Haematology, Wellcome Trust and MRC Cambridge Stem Cell Institute and
11 Cambridge Institute for Medical Research, University of Cambridge, Cambridge CB2 0XY,
12 United Kingdom

13 [4] Faculty of Science and Technology, Engineering Building, Lancaster University, Lancaster
14 LA1 4YR, United Kingdom

15 [5] Faculty of Geosciences, University of Utrecht, Heidelberglaan 2, 3584 CS Utrecht, The
16 Netherlands

17 [6] School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, Belfast
18 BT7 1NN, Northern Ireland, UK

19 Correspondence to: V.F. Rodriguez-Galiano (vrgaliano@us.es)

20

21 **1. Abstract**

22 This research reveals new insights into the weather drivers of interannual variation in land
23 surface phenology (LSP) across the entire European forest, while at the same time establishes
24 a new conceptual framework for predictive modelling of LSP. Specifically, the Random Forest
25 method, a multivariate, spatially non-stationary and non-linear machine learning approach, was

26 introduced for phenological modelling across very large areas and across multiple years
27 simultaneously: the typical case for satellite-observed LSP. The RF model was fitted to the
28 relation between LSP interannual variation and numerous climate predictor variables computed
29 at biologically-relevant rather than human-imposed temporal scales. In addition, the legacy
30 effect of an advanced or delayed spring on autumn phenology was explored. The RF models
31 explained 81% and 62% of the variance in the spring and autumn LSP interannual variation,
32 with relative errors of 10% and 20%, respectively: a level of precision that has until now been
33 unobtainable at the continental scale. Multivariate linear regression models explained only 36%
34 and 25%, respectively. It also allowed identification of the main drivers of the interannual
35 variation in LSP through its estimation of variable importance. This research, thus, shows an
36 alternative to the hitherto applied linear regression approaches for modelling LSP and paves
37 the way for further scientific investigation based on machine learning methods.

38 **2. Introduction**

39 Vegetation phenology has emerged as an important focus for scientific research in the last few
40 decades. The interest in vegetation phenology is twofold: inter-annual recording of the timing
41 of phenological events allows quantification of the impacts of climate change on vegetation;
42 and a greater understanding of phenological responses enables meaningful projections of how
43 ecosystems will respond to future changes in climate (Menzel, 2002; Morisette et al., 2008;
44 Peñuelas, 2009; Peñuelas and Filella, 2001). Although different approaches have been devised
45 for the study of vegetation phenology (Rafferty et al., 2013), the characterisation and modelling
46 of vegetation phenology at global or regional scales has been undertaken mainly through the
47 use of long-term time-series of satellite-sensor vegetation indices (termed land surface
48 phenology, LSP, to reflect that satellite-observed phenology includes all land covers). Most
49 studies of LSP analyse trends in phenological events across years (Delbart et al., 2008;
50 Jeganathan et al., 2014; Jeong et al., 2011; Karlsen et al., 2007; Myneni et al., 1997), but more
51 recent studies present process-based models to uncover cause-effect relationships between
52 long-term trends in phenology and its key driving variables (Ivits et al., 2012; Maignan et al.,
53 2008a; Maignan et al., 2008b; Stöckli et al., 2011; Stöckli et al., 2008; Yu et al., 2015; Zhou et
54 al., 2001). This last group of studies focuses on trends in phenology produced by trends in
55 weather (mainly warming). However, interannual variation in LSP arising as a consequence of
56 the inter-annual variability in weather are less studied (Cook et al., 2005; De Beurs and

57 Henebry, 2008; Menzel et al., 2005; Post and Stenseth, 1999; Zhang et al., 2004), with model-
58 based studies of this phenomenon being scarce (van Vliet, 2010).

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

76 Modelling efforts to characterize LSP have generally relied on functions (usually linear) of
77 meteorological drivers, such as average temperature and precipitation (Ivits et al., 2012),
78 growing degree days (GDD) (de Beurs and Henebry, 2005), light and temperature (Stöckli et
79 al., 2011), minimum temperature, photoperiod, vapour pressure deficit (Jolly et al., 2005;
80 Stöckli et al., 2008), or minimum relative humidity (Brown and de Beurs, 2008). However,
81 there is lack of understanding on number of important aspects, such as the multivariate
82 influence of meteorological variables (temperature, precipitation, solar radiation) driving
83 phenology, or the effect of additional drivers in the modelling of autumnal phenophases
84 (Morisette et al., 2008). For instance, Fu et al. (2014) found a “cause-effect relationship”
85 between an earlier leaf senescence and an earlier spring flushing in leaves of warmed samples
86 of *Fagus sylvatica* and *Quercus robur*. This legacy effect of spring phenology has been
87 reported in recent studies using modified environments and plant species, but it has not been
88 studied using LSP data. This latter aspect is particularly pertinent for studies that focus on inter-
89 annual variation in phenology and could potentially contribute to increased knowledge of how

90 climate change is affecting autumn phenology. On the other hand, many studies investigating
91 the sensitivity of phenological events to climate variation use calendar seasonal or monthly
92 mean climatic variables, which operate on fixed human calendar scales with a start date of 1st
93 of January (Maignan et al., 2008b), instead of using biological scales, for example, time relative
94 to the growing phase of plants (Pau et al., 2011). However, the modelling of interannual
95 variation in LSP considering its potentially complicated relationship with climate in a
96 multidimensional feature space (i.e. high number of multivariate weather drivers) might not be
97 possible using traditional linear regression models (de Beurs and Henebry, 2005). In this sense,
98 phenological modelling may benefit from machine learning techniques such as the Random
99 Forest (RF) method (Breiman, 2001), reducing uncertainties and bias (Zhao et al., 2013). RFs
100 have the potential to identify and model the complex non-linear relationships between
101 phenology and climate, being able to handle a large number of predictors and determine their
102 importance in explaining phenology. RFs has been applied with very promising results to other
103 fields of ecology and biological sciences (Archibald et al., 2009; Darling et al., 2012; Lawler
104 et al., 2006), as well as to the simulation of phenological shifts under different climatic change
105 scenarios (Lebourgeois et al., 2010), but the potential for modelling climate-driven interannual
106 variation in phenology is still to be explored.

107 Understanding the effect of inter-annual weather variation on LSP is an essential step to
108 establish a plausible link between recent climate variability and vegetation phenological
109 responses at global or regional scales, and importantly to make reliable forecasts about future
110 vegetation responses to different future climatic scenarios. The aim of this study is, therefore,
111 to provide an explanation of the observed interannual variation in LSP of the entire European
112 forest during the last decade, identifying the main weather drivers for spring and autumn at the
113 continental scale. Our research offers new insights into the study of LSP by modelling the
114 climate-driven past interannual variation in phenology, rather than trends, and using innovative
115 multivariate non-linear machine learning techniques to evaluate multiple weather predictors at
116 biological scales, and non-weather predictors such as the legacy effect of the date of spring
117 onset in leaf senescence. Climate predictors used range from 30 days average values of
118 temperature variables (max, min and avg) such as precipitation, short wave radiation and day
119 length; trimestral cumulated values such as growing degree days or chilling requirements,
120 among others; to the date of specific events such as the first freeze or the last freeze. Moreover,
121 we considered flexible biological time scales in the analysis between weather and phenological
122 events rather than calendar months.

123

124 **3. Materials and Methods**

125 **3.1 Data**

126 Three sources of data were used for this research: i) Satellite sensor derived temporal
127 composites of MERIS Terrestrial Chlorophyll Index (MTCI), ii) temperature and precipitation
128 data from the European Climate Assessment and Data project (<http://www.ecad.eu>) and iii)
129 surface radiation daylight (DAL; w/m^2) data and surface incoming shortwave (SIS; w/m^2)
130 radiation data from the Climate Monitoring Satellite Application Facilities
131 (<http://www.cmsaf.eu>). We used weekly composites of MTCI data at 1 km spatial resolution
132 from 2002 to 2012. This dataset was supplied by the European Space Agency and processed
133 by Airbus Defence and Space. Daily temperature (mean, minimum and maximum) and daily
134 precipitation data were derived from the European Climate Assessment & Dataset time-series
135 (version 10.0) with spatial resolution of $0.25^\circ \times 0.25^\circ$, covering the period from 2002 to 2011
136 (Haylock et al., 2008). Both radiation datasets, DAL (Müller and Trentmann, 2013) and SIS
137 (Posselt et al., 2012; Posselt et al., 2011) were derived from Meteosat satellite sensors at a
138 spatial resolution of $0.05^\circ \times 0.05^\circ$ covering the same period as temperature and precipitation
139 datasets.

140 **3.2 Phenology extraction and interannual variation in LSP computation**

141 The time-series of MERIS MTCI data was used to estimate both the onset of greenness
142 (OG) and end of senescence (EOS) from 2003 to 2011. Data for every estimation year
143 considered 1.5 years of data (from October in the previous year to July in the next year)
144 because the annual pattern of vegetation growth in some parts of Europe spans across
145 calendar years and, hence, insufficient information about LSP is captured using a single
146 year of data. The yearly values of OG and EOS were estimated for each image pixel of the
147 study area using the methodology described in Dash et al. (2010). This methodology
148 consists of two major procedures: data smoothing and LSP estimation (Figure 2a).
149 Smoothed MTCI time-series data were obtained using a discrete Fourier transform because
150 of its advantage of requiring fewer user-defined parameters compared to other methods
151 (Atkinson et al., 2012). The peak in the annual profile was defined as a point on the
152 phenological curve where the first derivative changes sign from positive to negative. Next,
153 the derived data were searched backward and forward departing from the maximum annual

154 peak to estimate the OG and EOS, respectively. OG was defined as a valley at the
155 beginning of the growing season point (a change in derivative value from positive to
156 negative) and EOS was defined as a valley point occurring at the decaying end of a
157 phenology cycle (a change in derivative value from negative to positive). These satellite-
158 derived LSP estimates were compared to ground observations of the thousands of
159 deciduous tree phenology records of the Pan European Phenology network (PEP725)
160 (Rodriguez-Galiano et al., 2015a). This comparison resulted in a large spatio-temporal
161 correlation of the phenology estimates with the spring phenophase (OG vs leaf unfolding;
162 pseudo- $R^2=0.70$) and autumn phenophase (EOS vs autumnal colouring; pseudo- $R^2=0.71$).

163 Z-score values during the study period were used as a proxy to measure interannual
164 variation in the LSP parameters. The z-score values for a given year were defined as the
165 difference from the multi-year mean, normalized by the standard deviation across years.
166 The value of the targeted year was excluded in the computation of multiyear mean to
167 enhance the inter-annual variation (Saleska et al., 2007). The spatio-temporal distribution
168 of spring and autumn LSP z-score values is shown in Figures S1 and S2 of the supporting
169 information, respectively.

170 To match the spatial resolution of the weather predictors, the LSP z-score values for each year
171 were resampled to a spatial resolution of $0.25^\circ \times 0.25^\circ$ by calculating the median of all the LSP
172 z-score values within this area after excluding the areas with fewer than 50 LSP estimates and
173 the non-forest pixels according to the Globcover2005 and Globcover2009 land cover maps
174 (<http://due.esrin.esa.int/globcover/>) (Figure 1). Only LSP estimates of Globcover forest
175 categories with complete temporal coverage (2003–2011) were included in the analysis to
176 reduce the likelihood of natural and human disturbances (Potter et al., 2003) and to minimize
177 the effects of human management (i.e. irrigation in croplands). Globcover was selected for its
178 greater consistency with the MERIS MTCI time-series and its high geolocational accuracy
179 (<150 m) (Bicheron et al., 2011).

180 **3.3 Computation of weather predictors**

181 A suite of weather predictors were computed for each $0.25 \times 0.25^\circ$ grid cell associated with the
182 occurrence of positive or negative z-score values in LSP (see Table 1). The predictors include
183 temporal average values of temperature variables (Tmax, Tmin and Tavg), precipitation, DAL
184 and SIS; temporal cumulated predictors such as growing degree days, chilling, precipitation,
185 SIS and DAL; and the date of specific events such as the onset of greenness (legacy effect for

186 autumn phenology modelling) the first freeze or the last freeze, as well as the difference
187 between both dates (freeze period) for the modelling of autumn only. Growing degree days
188 were computed using temperature thresholds of 0° and 5°. Chilling requirements were
189 computed as the sum of negative temperatures (temperatures below 0°). Freeze was defined as
190 dates with minimum temperatures lower than -2° (Schwartz et al., 2006).

191 The different weather predictors were computed based on the 30 and 90 days previous to the
192 day of the year (DOY) of the z-score values in OG and EOS (Figure 2b) following Schwartz
193 et al. (2006) and Menzel et al. (2006), who found that most phenophases of plant observations
194 in Europe correlated significantly with weather predictors representing the month of onset and
195 the two preceding months. The chilling requirements for spring modelling and freeze predictors
196 were an exception, as the period for its computation starts 90 days prior to the OG. Relative
197 differences between each predictor and its multi-year average for the same period were
198 computed to capture the inter-annual variability in climate variables at the pixel level for every
199 predictor and to facilitate the modelling of climate-driven variation in phenology (Table 1).

200 **3.4 Modelling interannual variation in LSP**

201 Conventional statistical models such as linear regression might be inappropriate for
202 investigating the drivers of interannual variation in phenology because many of the
203 relationships are likely to be non-linear (De Beurs and Henebry, 2008). In this sense, machine
204 learning methods have emerged as complementary alternatives to conventional statistical
205 techniques. Within the branch of machine learning techniques, regression trees are particularly
206 suitable when compared to global single predictive models, allowing for multiple regression
207 models using recursive partitioning (Breiman, 1984). Assembling a single global model might
208 not be representative of LSP of the entire European continent, when there are many climatic
209 drivers which interact in complicated, non-linear ways and may vary spatially and temporally.
210 For the purpose of this paper, an alternative approach is to sub-divide, or partition, the data
211 space into more homogeneous regions of similar climates and ecological factors.

212 Regression trees use a sum of squares criterion to split the data into successively more
213 homogeneous subsets contained at many different structural units called nodes. Each of the
214 terminal nodes, has attached to it a simple regression which applies in that node only. Therefore,
215 different regressions can be fitted to different data subsets within one single regression tree,
216 which can represent different responses controlled by different drivers (Archibald et al., 2009;
217 Lawler et al., 2006). Additionally, the performance of multiple regression trees can be

218 combined to increase the predictive ability of a single regression tree model, following the
219 Random Forest technique (Figure 3). The RF method is an innovative machine learning
220 approach that can perform multivariate non-linear regression, combining the performance of
221 numerous regression tree algorithms to predict the interannual variation in OG and EOS. More
222 details regarding the performance and the specific characteristics of a RF model can be seen in
223 Rodriguez-Galiano et al. (2015b); Rodriguez-Galiano et al. (2014), and Figure 3.

224 The Random Forest method was applied to phenological modelling across very large areas and
225 across multiple years simultaneously: the typical case for satellite-observed LSP. The RF
226 model was fitted to the relation between LSP interannual variation and numerous climate
227 predictor variables computed at biologically-relevant rather than human-imposed temporal
228 scales. We restricted our climate data choices to daily data (average, minimum and maximum
229 temperatures, precipitation and radiation) to account for integrative forcing (that is, growing
230 degree days, chilling requirements as well as cumulative precipitation and radiation), computed
231 from the exact day of the phenological event backwards, rather than using the calendar months.
232 The locations with z-score in LSP greater than 1 (positive and negative) were selected to build
233 a RF predictive model on OG and EOS. Z-score values of OG or EOS for each year were
234 combined together with the different weather predictors. The z-score values in OG were
235 assessed as an extra predictor to evaluate the legacy effect of an advanced or delayed spring in
236 the modelling of EOS. The values of these variables at the selected years and locations
237 (spatiotemporal model) were combined into a set of input feature vectors (3900 feature vectors
238 for the spring model and 3124 for autumn) as an input to the RF algorithm. These feature
239 vectors were divided equally into two subsets, one for the training of the models (inbag) and
240 one as an additional test to the one internally computed by RF (out of bag; oob) to evaluate
241 performance. RF models composed of 2000 trees were grown using different subsets of
242 predictors, varying the number of random predictors from 1 to 9. The Random Forest method
243 within the package implemented in the R statistical software was used to build the different
244 models (Liaw and Wiener, 2002).

245 **3.5 Selection of the most important predictors**

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

250 inclusion of different measures of weather predictors may imply a large increase in the
251 dimensionality of the datasets being used, as these variables are obtained by applying multiple
252 functions or measures to the temperature, precipitation and radiation time-series. On the one
253 hand, more information may be useful for the modelling process; on the other hand, an
254 excessive number of correlated predictors or features can overwhelm the expected increase in
255 accuracy and may introduce additional complexity limiting the ability of the method to point
256 to possible cause-effect relationships between interannual variation in phenology and their
257 drivers, making interpretation challenging.

258 A feature selection approach, based on the ability of the RF to assess the relative importance
259 of the predictors, was used to identify the minimum number of drivers which can better explain
260 spring or autumn interannual variation in phenology. To assess the importance of each weather
261 predictor, the RF switches one of the input predictors while keeping the rest constant, and it re-
262 evaluates the performance of the model measuring the decrease in node impurity (Breiman,
263 2001). The differences were averaged over all 2000 trees to compute the general drivers for the
264 interannual variation in Europe. However, different subsets of variables could be used to
265 characterize different climates and ecological factors at every single regression tree model or
266 node (see previous section). In order to reduce the number of drivers the least important
267 predictor was removed iteratively at different steps. Then, a 5-fold cross-validation was applied
268 to obtain a stable estimate of the error of the model built after predictor deletions. Finally, the
269 model with a better trade-off between number of predictors and error was chosen as the basis
270 for interpreting the likely drivers of interannual variation in phenology.

271 **4. Results**

272 Numerous models were built on the basis of different predictor combinations considering
273 different temporal windows prior to the spring and autumn phenological events (see section
274 “computation of weather predictors”). The percentage of variation (pseudo- R^2) explained by
275 different weather-LSP models is shown in the supplementary information (Table S1, S2 and
276 S3). No previous studies have investigated in depth the parametrization of GDD for LSP and
277 climate inter-comparison, unlike for ground phenological studies (Snyder et al., 1999).
278 Although, we did not carry out an exhaustive analysis of the optimum GDD parametrization,
279 our results showed a systematic pattern in spring models, presenting slightly larger pseudo- R^2
280 for models which used 0° C as a threshold for the computation of GDD (rather than 5° C).
281 Regarding, the length of the temporal windows for weather function computation, spring

282 models using 30 and 90 days for the computation of averaged and cumulative functions were
283 more accurate, whereas for autumn models with 90 day-averaged predictors outperformed the
284 rest.

285 The main drivers of interannual variation in LSP were identified through the application of a
286 feature selection procedure (see section “selection of the most important predictors”). Spring
287 models were more accurate than autumn, with median relative error values of 10% to 27% (12
288 to 1 predictor), versus 26% to 60% of autumn (14 to 1 predictor). Figure 4 shows the pseudo-
289 R^2 of the models as well as the relative importance of each predictor. Spring models (explained
290 a percentage of the variance up to 81% (Figure 4a), whereas autumn explained up to 61%
291 (Figure 4b). Cook et al. (2005), using a model based on GDD only, explained 63% on the
292 variance of onset date for mixed and boreal forest. Figure 5 shows the relative error in the
293 prediction of different models after removing the least important predictor. Regarding the
294 relative importance of the drivers, the same ranking in importance was observed within the
295 different models of each phenophase, which reflected the stability in the RF importance
296 estimation, and a high reliability of the results (Figure 4). To interpret the main weather drivers
297 of the interannual variation in phenology, simplified models with reduced number of predictors
298 were selected for spring and autumn (see section 3.5), respectively. The spring model was
299 composed of 6 predictors (pseudo- $R^2=0.77$ and median relative error of 10%) and the autumn
300 model of 5 predictors (pseudo- $R^2=0.59$ and median relative error of 28%) (Figure 6). Our
301 results suggest that interannual variation in the onset on greenness (LSP) of temperate forest
302 species are driven mainly by the daily temperature of the 30 days prior to onset (but not
303 necessarily the GDD), with the most important driver being the minimum temperature.
304 Photoperiod was also important, the most accurate empirical prediction was obtained by a
305 combined temperature-radiation forcing, integrating the SIS of the previous 90 days. For
306 senescence, temperature was suggested to be more important than photoperiod in controlling
307 the senescence process (Archetti et al., 2013; Jeong and Medvigy, 2014; Vitasse et al., 2009;
308 Yang et al., 2012), with the most important drivers being the date of the first freeze and the
309 accumulation of chilling temperatures. However, we did not observe a legacy effect of a much
310 earlier or later spring onset on the date of senescence. Autumn models that included the
311 interannual variation (z-score values) in the onset of greenness did not outperform the
312 remaining models (see Table S2 and S3 in supplementary information) and the relative
313 importance was low in comparison with other drivers.

314 5. Discussion

315 The selection and computation of the weather predictors is an important step of phenological
316 modelling. Most of studies on the sensitivity of phenological events to climate used human
317 calendar scales, that is, seasonal or monthly calendar mean or cumulative climate predictors
318 (Maignan et al., 2008a; Maignan et al., 2008b; Menzel et al., 2006; Schwartz et al., 2006),
319 overlooking the importance of biological time-scales in phenology. However, with the
320 increased availability of daily weather datasets, current and future studies might benefit from
321 the use of daily information to model the drivers of plants' circadian time-scales (Pau et al.,
322 2011). Our study advanced the modelling of vegetation phenology by improving the temporal
323 matching between LSP interannual variation and the preceding weather conditions by
324 analysing daily data at biological scales. Regarding, the length of the temporal windows for
325 weather function computation, Menzel et al. (2006) showed that most phenological phases of
326 plant species in Europe correlate significantly with mean temperatures of the month of onset
327 and the two preceding months. However, in our study, when end of senescence was considered,
328 a consistent divergent effect was observed between spring and autumn. Autumn phenophases
329 might be driven by longer-term changes in weather, while for spring the average conditions of
330 the 30 days previous to the date of onset play a more important role (Table S1, S2 and S3 in
331 supplementary information). From a computational point of view, considering larger temporal
332 windows for calculating averages would induce a smoothing effect, degrading the information
333 in the predictors, whereas cumulative functions such as GDD or chilling requirements would
334 not be affected by this effect. However, we observed a divergent response between spring and
335 autumn and consistent throughout the models of each phenophase suggests that a biological
336 explanation for this phenomenon might be plausible.

337 Understanding the drivers of interannual variation in LSP amidst background inter-annual
338 variation is a critical aspect of global change science (de Beurs and Henebry, 2005; Zhao et al.,
339 2013). To this end, the RF method is particularly pertinent, as it allows the assessment of the
340 importance of the predictors (Figure 4). Our findings reveal that the accuracy of growing degree
341 day-based models might be overestimated using linear regression models and that non-linear
342 multivariate relationships between temperature (especially minimum temperature) and
343 radiation are needed to describe the relations between phenology and weather drivers. This
344 supports the findings of Stöckli et al. (2011) who explained temperate phenology using a
345 combination of light and temperature. The highlighted importance of minimum temperatures
346 might be related to the fact that minimum temperature is a better indicator of weather changes

347 than either the average or maximum temperature (Duncan et al., 2014; Jolly et al., 2005).
348 Regarding GDD, although it has been applied extensively to predict vegetation phenophases ,
349 it is currently debated whether such models can detect when multiple environmental drivers
350 are required to initiate a phenological event, or detect drivers that are relatively static across
351 time, such as photoperiod (Stöckli et al. 2011). Our results reveal that multiple environmental
352 drivers are required to initiate phenological events of Europe and also showed that the role of
353 GDD alone in driving spring phenology might be overestimated due to an over-reliance on
354 linear models. GDD had the largest linear association with vegetation phenology interannual
355 variation, while the linear correlation between LSP and others drivers that were revealed as
356 very important by the RF was small (see Tables 1 and 2). A simple linear analysis between
357 GDD and phenology could ignore complex non-linear associations between phenology and
358 predictors as well as synergies between weather drivers. Regarding the senescence phase, the
359 autumn models had a weaker predictive power compared the spring models. There is still lack
360 of clear understanding of mechanism autumn senescence, however, temperature, and
361 particularly the dates of freeze, has been suggested as major driver for autumn phenology.

362 The RF method provided an important alternative over simple, but less accurate analysis based
363 on linear regression for the analysis of interannual variation in spring and autumn phenology.
364 A further comparison with a linear regression analysis suggested that there might be a non-
365 linear relationship between the interannual variation in LSP and the weather drivers.
366 Multivariate linear regression models were also fitted from the same combination of predictors
367 selected as optimal by Random Forest. Multivariate linear models explained only 36% and 26%
368 of the variance in spring and autumn phenology interannual variation across the continental
369 scale. Additionally, a linear regression between predicted values from RF and observed
370 interannual variation in phenology produced R^2 values equal to 0.90 and 0.68 for spring and
371 autumn LSP interannual variation, respectively (Figure 6a and 6b). On the other hand, the
372 correlations between the predictions of linear regression models and observations were much
373 weaker, with R^2 values of 0.39 and 0.25 (Figure 6c and 6d). Linear models under-predicted a
374 delay in the phenophases (positive z-score values) and over-predicted the advances (negative
375 z-score values). The spatial distribution of the relative errors for RF and multivariate linear
376 regression is shown in Figures S3 to S6 of the supporting information. The relative errors of
377 the latter were significantly higher. Additionally, the residuals seemed not to be homoscedastic
378 suggesting that linear models might not be able to deal with the complex patterns between LSP

379 and climate patterns at multiple locations and times, integrating them into a unique overall
380 model.

381 A new approach to model interannual variation in LSP was presented in this paper based on
382 the application of the RF model to a set of climate predictors at biological scales. This new
383 modelling technique has numerous advantages for the modelling of climate-driven interannual
384 variation in LSP. It is a non-parametric multivariate method which allows for non-linear
385 relationships between (compared to traditional linear models) phenology and climate and can
386 consider a large number of weather predictors in the modelling process. This provides potential
387 opportunity to capture the impact of all possible environmental/weather drivers on vegetation
388 phenology. The proposed method can recognize complex patterns between LSP and climate at
389 multiple locations and times, integrating them into a unique overall model, rather than
390 generating multiple models over a geographical area and for different years. Additionally it is
391 data-driven, which means that there is no need to incorporate previous knowledge about the
392 specific responses of vegetation to different predominant weather controls (i.e. temperature,
393 rainfall, and photoperiod), allowing weather drivers to automatically shift both temporally and
394 spatially. Therefore, it is highly generalizable, being applicable to different biogeographical
395 regions where the phenology is controlled by different factors. This flexibility or generalization
396 capacity of RF models to transition from one driver to another without the need for a model
397 change also promotes its application to different climate change scenarios. We succeeded in
398 modelling the interannual variation in LSP phenology as observed from satellite-sensors in the
399 European Forest, while using the same type of input data, the same model, and the same model
400 parameters for the entire European continent.

401 **Author Contributions**

402 V.F.R.G., J.D. and P.M.A. conceived and designed the experiments; V.F.R.G. performed the
403 experiments; V.F.R.G., M.S.C. and J.D. contributed analysis tools; V.F.R.G. drafted the paper.
404 All authors contributed to the final paper.

405 **Acknowledgements**

406 The first author is a Marie Curie Grant holder (reference FP7-PEOPLE-2012-IEF-331667).
407 The authors are grateful for the financial support given by the European Commission under the
408 Seventh Framework Programme and the Spanish MINECO (project BIA2013-43462-P). PMA
409 is grateful to the University of Utrecht for supporting him with The Belle van Zuylen Chair.

410 We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES
411 (<http://ensembles-eu.metoffice.com>) and the data providers in the ECA&D project
412 (<http://www.ecad.eu>". Surface radiation data were obtained from EUMETSAT's Satellite
413 Application Facility on Climate Monitoring (CM SAF).

414 **References**

- 415 Archetti, M., Richardson, A. D., O'Keefe, J., and Delpierre, N.: Predicting Climate Change
416 Impacts on the Amount and Duration of Autumn Colors in a New England Forest, PLoS ONE,
417 8, 2013.
- 418 Archibald, S., Roy, D. P., van Wilgen, B. W., and Scholes, R. J.: What limits fire? An
419 examination of drivers of burnt area in Southern Africa, *Glob. Change Biol.*, 15, 613-630, 2009.
- 420 Atkinson, P. M., Jeganathan, C., Dash, J., and Atzberger, C.: Inter-comparison of four models
421 for smoothing satellite sensor time-series data to estimate vegetation phenology, *Remote Sens.*
422 *Environ.*, 123, 400-417, 2012.
- 423 Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M., and García-Herrera, R.: The Hot
424 Summer of 2010: Redrawing the Temperature Record Map of Europe, *Science*, 332, 220-224,
425 2011.
- 426 Bicheron, P., Amberg, V., Bourg, L., Petit, D., Huc, M., Miras, B., Brockmann, C., Hagolle,
427 O., Delwart, S., Ranera, F., Leroy, M., and Arino, O.: Geolocation Assessment of MERIS
428 GlobCover Orthorectified Products, *IEEE Trans. Geosci. Remote Sensing*, 49, 2972-2982,
429 2011.
- 430 Breiman, L.: Classification and regression trees, Chapman & Hall/CRC, 1984.
- 431 Breiman, L.: Random forests, *Machine Learning*, 45, 5-32, 2001.
- 432 Brown, M. E. and de Beurs, K. M.: Evaluation of multi-sensor semi-arid crop season
433 parameters based on NDVI and rainfall, *Remote Sens. Environ.*, 112, 2261-2271, 2008.
- 434 Cook, B. I., Smith, T. M., and Mann, M. E.: The North Atlantic Oscillation and regional
435 phenology prediction over Europe, *Glob. Change Biol.*, 11, 919-926, 2005.
- 436 Darling, E. S., Alvarez-Filip, L., Oliver, T. A., McClanahan, T. R., and Côté, I. M.: Evaluating
437 life-history strategies of reef corals from species traits, *Ecology Letters*, 15, 1378-1386, 2012.
- 438 Dash, J., Jeganathan, C., and Atkinson, P. M.: The use of MERIS Terrestrial Chlorophyll Index
439 to study spatio-temporal variation in vegetation phenology over India, *Remote Sens. Environ.*,
440 114, 1388-1402, 2010.
- 441 de Beurs, K. M. and Henebry, G. M.: Land surface phenology and temperature variation in the
442 International Geosphere-Biosphere Program high-latitude transects, *Glob. Change Biol.*, 11,
443 779-790, 2005.
- 444 De Beurs, K. M. and Henebry, G. M.: Northern annular mode effects on the land surface
445 phenologies of northern Eurasia, *Journal of Climate*, 21, 4257-4279, 2008.
- 446 Delbart, N., Picard, G., Le Toan, T., Kergoat, L., Quegan, S., Woodward, I., Dye, D., and
447 Fedotova, V.: Spring phenology in boreal Eurasia over a nearly century time scale, *Glob.*
448 *Change Biol.*, 14, 603-614, 2008.
- 449 Duncan, J. M. A., Dash, J., and Atkinson, P. M.: Elucidating the impact of temperature
450 variability and extremes on cereal croplands through remote sensing, *Glob. Change Biol.*, 21,
451 1541-1551, 2014.
- 452 Fu, Y. S. H., Campioli, M., Vitasse, Y., De Boeck, H. J., Van Den Berge, J., Abdelgawad, H.,
453 Asard, H., Piao, S., Deckmyn, G., and Janssens, I. A.: Variation in leaf flushing date influences
454 autumnal senescence and next year's flushing date in two temperate tree species, *Proceedings*
455 *of the National Academy of Sciences of the United States of America*, 111, 7355-7360, 2014.

456 Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A
457 European daily high-resolution gridded data set of surface temperature and precipitation for
458 1950-2006, *J. Geophys. Res.*, 113, 2008.

459 Ivits, E., Cherlet, M., Tóth, G., Sommer, S., Mehl, W., Vogt, J., and Micale, F.: Combining
460 satellite derived phenology with climate data for climate change impact assessment, *Global*
461 *and Planetary Change*, 88-89, 85-97, 012.

462 Ivits, E., Cherlet, M., Tóth, G., Sommer, S., Mehl, W., Vogt, J., and Micale, F.: Combining
463 satellite derived phenology with climate data for climate change impact assessment, *Glob.*
464 *Planet. Change*, 88-89, 85-97, 2012.

465 Jeganathan, C., Dash, J., and Atkinson, P. M.: Remotely sensed trends in the phenology of
466 northern high latitude terrestrial vegetation, controlling for land cover change and vegetation
467 type, *Remote Sens. Environ.*, 143, 154-170, 2014.

468 Jeong, S.-J., Ho, C.-H., Gim, H.-J., and Brown, M. E.: Phenology shifts at start vs. end of
469 growing season in temperate vegetation over the Northern Hemisphere for the period 1982–
470 2008, *Glob. Change Biol.*, 17, 2385-2399, 2011.

471 Jeong, S.-J. and Medvigy, D.: Macroscale prediction of autumn leaf coloration throughout the
472 continental United States, *Global Ecology and Biogeography*, 23, 1245-1254, 2014.

473 Jolly, W. M., Nemani, R., and Running, S. W.: A generalized, bioclimatic index to predict
474 foliar phenology in response to climate, *Glob. Change Biol.*, 11, 619-632, 2005.

475 Karlsen, S. R., Solheim, I., Beck, P. S. A., Hogda, K. A., Wielgolaski, F. E., and Tommervik,
476 H.: Variability of the start of the growing season in Fennoscandia, 1982-2002, *Int. J.*
477 *Biometeorol.*, 51, 513-524, 2007.

478 Lawler, J. J., White, D., Neilson, R. P., and Blaustein, A. R.: Predicting climate-induced range
479 shifts: Model differences and model reliability, *Glob. Change Biol.*, 12, 1568-1584, 2006.

480 Lebourgeois, F., Pierrat, J. C., Perez, V., Piedallu, C., Cecchini, S., and Ulrich, E.: Simulating
481 phenological shifts in French temperate forests under two climatic change scenarios and four
482 driving global circulation models, *Int. J. Biometeorol.*, 54, 563-581, 2010.

483 Liaw, A. and Wiener, M.: Classification and Regression by randomForest, *R News*, 2/3, 18-22,
484 2002.

485 Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., and Wanner, H.: European Seasonal
486 and Annual Temperature Variability, Trends, and Extremes Since 1500, *Science*, 303, 1499-
487 1503, 2004.

488 Maignan, F., Bréon, F. M., Bacour, C., Demarty, J., and Poirson, A.: Interannual vegetation
489 phenology estimates from global AVHRR measurements: Comparison with in situ data and
490 applications, *Remote Sens. Environ.*, 112, 496-505, 2008a.

491 Maignan, F., Bréon, F. M., Vermote, E., Ciais, P., and Viovy, N.: Mild winter and spring 2007
492 over western Europe led to a widespread early vegetation onset, *Geophys. Res. Lett.*, 35,
493 L02404, 2008b.

494 Menzel, A.: Phenology: Its Importance to the Global Change Community, *Clim. Change*, 54,
495 379-385, 2002.

496 Menzel, A., Sparks, T. H., Estrella, N., and Eckhardt, S.: 'SSW to NNE' - North Atlantic
497 Oscillation affects the progress of seasons across Europe, *Glob. Change Biol.*, 11, 909-918,
498 2005.

499 Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aaasa, A., Ahas, R., Alm-Kübler, K., Bissolli,
500 P., Braslavská, O., Briede, A., Chmielewski, F. M., Crepinsek, Z., Curnel, Y., Dahl, Å., Defila,
501 C., Donnelly, A., Filella, Y., Jatczak, K., Måge, F., Mestre, A., Nordli, Ø., Peñuelas, J., Pirinen,
502 P., Remišová, V., Scheifinger, H., Striz, M., Susnik, A., Van Vliet, A. J. H., Wielgolaski, F. E.,
503 Zach, S., and Züst, A.: European phenological response to climate change matches the warming
504 pattern, *Glob. Change Biol.*, 12, 1969-1976, 2006.

505 Morisette, J. T., Richardson, A. D., Knapp, A. K., Fisher, J. I., Graham, E. A., Abatzoglou, J.,
506 Wilson, B. E., Breshears, D. D., Henebry, G. M., Hanes, J. M., and Liang, L.: Tracking the
507 rhythm of the seasons in the face of global change: phenological research in the 21st century,
508 *Frontiers in Ecology and the Environment*, 7, 253-260, 2008.

509 Müller, R. and Trentmann, J.: CM SAF Meteosat Surface Radiation Daylight Data Set 1.0 -
510 Monthly Means / Daily Means . Satellite Application Facility on Climate Monitoring, doi:
511 10.5676/EUM_SAF_CM/DAL_MVIRI_SEVIRI/V001, 2013. 2013.

512 Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G., and Nemani, R. R.: Increased plant
513 growth in the northern high latitudes from 1981 to 1991, *Nature*, 386, 698-702, 1997.

514 Pau, S., Wolkovich, E. M., Cook, B. I., Davies, T. J., Kraft, N. J. B., Bolmgren, K., Betancourt,
515 J. L., and Cleland, E. E.: Predicting phenology by integrating ecology, evolution and climate
516 science, *Glob. Change Biol.*, 17, 3633-3643, 2011.

517 Peñuelas, J.: Phenology feedbacks on climate change, *Science*, 324, 887-888, 2009.

518 Peñuelas, J. and Filella, I.: Phenology: Responses to a warming world, *Science*, 294, 793-795,
519 2001.

520 Posselt, R., Mueller, R. W., Stöckli, R., and Trentmann, J.: Remote sensing of solar surface
521 radiation for climate monitoring — the CM-SAF retrieval in international comparison, *Remote
522 Sens. Environ.*, 118, 186-198, 2012.

523 Posselt, R., Müller, R., Stöckli, R., and Trentmann, J.: CM SAF Surface Radiation MVIRI Data
524 Set 1.0 - Monthly Means / Daily Means / Hourly Means, Satellite Application Facility on
525 Climate Monitoring, doi: 10.5676/EUM_SAF_CM/RAD_MVIRI/V001, 2011. 2011.

526 Post, E. and Stenseth, N. C.: Climatic variability, plant phenology, and northern ungulates,
527 *Ecology*, 80, 1322-1339, 1999.

528 Potter, C., Tan, P. N., Steinbach, M., Klooster, S., Kumar, V., Myneni, R., and Genovese, V.:
529 Major disturbance events in terrestrial ecosystems detected using global satellite data sets, *Glob.
530 Change Biol.*, 9, 1005-1021, 2003.

531 Rafferty, N. E., CaraDonna, P. J., Burkle, L. A., Iler, A. M., and Bronstein, J. L.: Phenological
532 overlap of interacting species in a changing climate: an assessment of available approaches,
533 *Ecology and Evolution*, 3, 3183-3193, 2013.

534 Rodriguez-Galiano, V., Dash, J., and Atkinson, P. M.: Inter-comparison of satellite sensor land
535 surface phenology and ground phenology in Europe, *Geophys. Res. Lett.*, 42, 2253-2260,
536 2015a.

537 Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., and Chica-Rivas, M.: Machine
538 learning predictive models for mineral prospectivity: An evaluation of neural networks,
539 random forest, regression trees and support vector machines, *Ore Geology Reviews*, 71, 804-
540 818, 2015b.

541 Rodriguez-Galiano, V. F., Chica-Olmo, M., and Chica-Rivas, M.: Predictive modelling of gold
542 potential with the integration of multisource information based on random forest: a case study
543 on the Rodalquilar area, Southern Spain, *International Journal of Geographical Information
544 Science*, 28, 1336-1354, 2014.

545 Rutishauser, T., Luterbacher, J., Defila, C., Frank, D., and Wanner, H.: Swiss spring plant
546 phenology 2007: Extremes, a multi-century perspective, and changes in temperature sensitivity,
547 *Geophys. Res. Lett.*, 35, 2008.

548 Saleska, S. R., Didan, K., Huete, A. R., and Da Rocha, H. R.: Amazon forests green-up during
549 2005 drought, *Science*, 318, 612, 2007.

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

552 Snyder, R. L., Spano, D., Cesaraccio, C., and Duce, P.: Determining degree-day thresholds
553 from field observations, *Int. J. Biometeorol.*, 42, 177-182, 1999.

554 Stöckli, R., Rutishauser, T., Baker, I., Liniger, M. A., and Denning, A. S.: A global reanalysis
555 of vegetation phenology, *Journal of Geophysical Research: Biogeosciences*, 116, G03020,
556 2011.

557 Stöckli, R., Rutishauser, T., Dragoni, D., O'Keefe, J., Thornton, P. E., Jolly, M., Lu, L., and
558 Denning, A. S.: Remote sensing data assimilation for a prognostic phenology model, *Journal*
559 *of Geophysical Research: Biogeosciences*, 113, G04021, 2008.

560 van Vliet, A. H.: Societal adaptation Options to Changes in Phenology. In: *Phenological*
561 *Research*, Hudson, I. L. and Keatley, M. R. (Eds.), Springer Netherlands, Netherlands, 2010.

562 Vitasse, Y., Delzon, S., Dufrêne, E., Pontailleur, J. Y., Louvet, J. M., Kremer, A., and Michalet,
563 R.: Leaf phenology sensitivity to temperature in European trees: Do within-species populations
564 exhibit similar responses?, *Agric. For. Meteorol.*, 149, 735-744, 2009.

565 Yang, X., Mustard, J. F., Tang, J. W., and Xu, H.: Regional-scale phenology modeling based
566 on meteorological records and remote sensing observations, *Journal of Geophysical Research-*
567 *Biogeosciences*, 117, 2012.

568 Yu, R., Schwartz, M. D., Donnelly, A., and Liang, L.: An observation-based progression
569 modeling approach to spring and autumn deciduous tree phenology, *Int. J. Biometeorol.*, doi:
570 10.1007/s00484-015-1031-9, 2015. 2015.

571 Zhang, X., Friedl, M. A., Schaaf, C. B., and Strahler, A. H.: Climate controls on vegetation
572 phenological patterns in northern mid- and high latitudes inferred from MODIS data, *Glob.*
573 *Change Biol.*, 10, 1133-1145, 2004.

574 Zhao, M. F., Peng, C. H., Xiang, W. H., Deng, X. W., Tian, D. L., Zhou, X. L., Yu, G. R., He,
575 H. L., and Zhao, Z. H.: Plant phenological modeling and its application in global climate
576 change research: overview and future challenges, *Environmental Reviews*, 21, 1-14, 2013.

577 Zhou, L. M., Tucker, C. J., Kaufmann, R. K., Slayback, D., Shabanov, N. V., and Myneni, R.
578 B.: Variations in northern vegetation activity inferred from satellite data of vegetation index
579 during 1981 to 1999, *J. Geophys. Res.-Atmos.*, 106, 20069-20083, 2001.

580

581

582 Table 1. Predictors used in the modelling of the interannual variation in LSP. * predicted over
 583 a period of 90 days. ** predicted over a period of the 30 and 90 days previous to the date of
 584 the z-score value.

OG anomalies	EOS anomalies
Averages (M):	
Maximum temperature (TX)**	Maximum temperature (TX)**
Minimum temperature (TN)**	Minimum temperature (TN)**
Average temperature (TG)**	Average temperature (TG)**
Precipitation (PP)**	Precipitation (PP)**
Surface incoming shortwave radiation (SIS)**	Surface incoming shortwave radiation (SIS)**
Surface radiation daylight (DAL)**	Surface radiation daylight (DAL)**
Cumulates (C)	
Growing Degree Days (0° C threshold) (GDD)**	Growing Degree Days (0° C threshold) (GDD)**
Growing Degree Days (5° C threshold) (GDD)**	Growing Degree Days (5° C threshold) (GDD)**
Chilling requirements (CHIL)*	Chilling requirements (CHIL)**
Precipitation (PP)**	Precipitation (PP)**
Surface incoming shortwave radiation (SIS)**	Surface incoming shortwave radiation (SIS)**
Surface radiation daylight (DAL)**	Surface radiation daylight (DAL)**
Date of specific events	
First freeze (FF)*	First freeze (FF)*
Last freeze (LF)*	OG z-score value (OGA) (legacy effect of an advanced or delayed spring)
Period of freeze (PF)*	

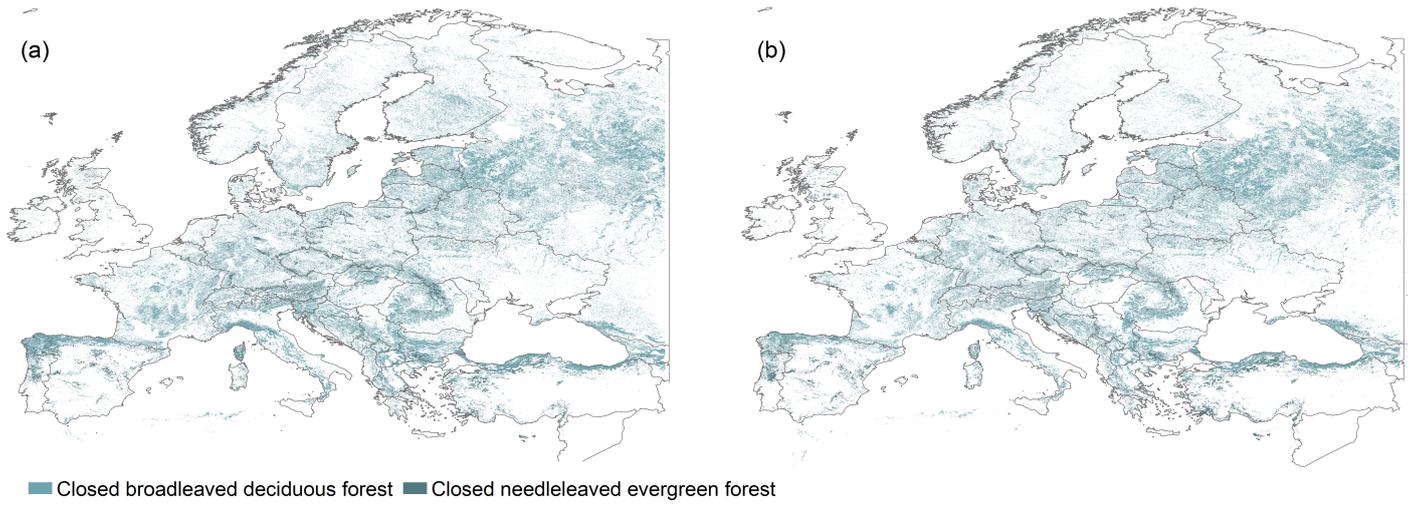
585

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

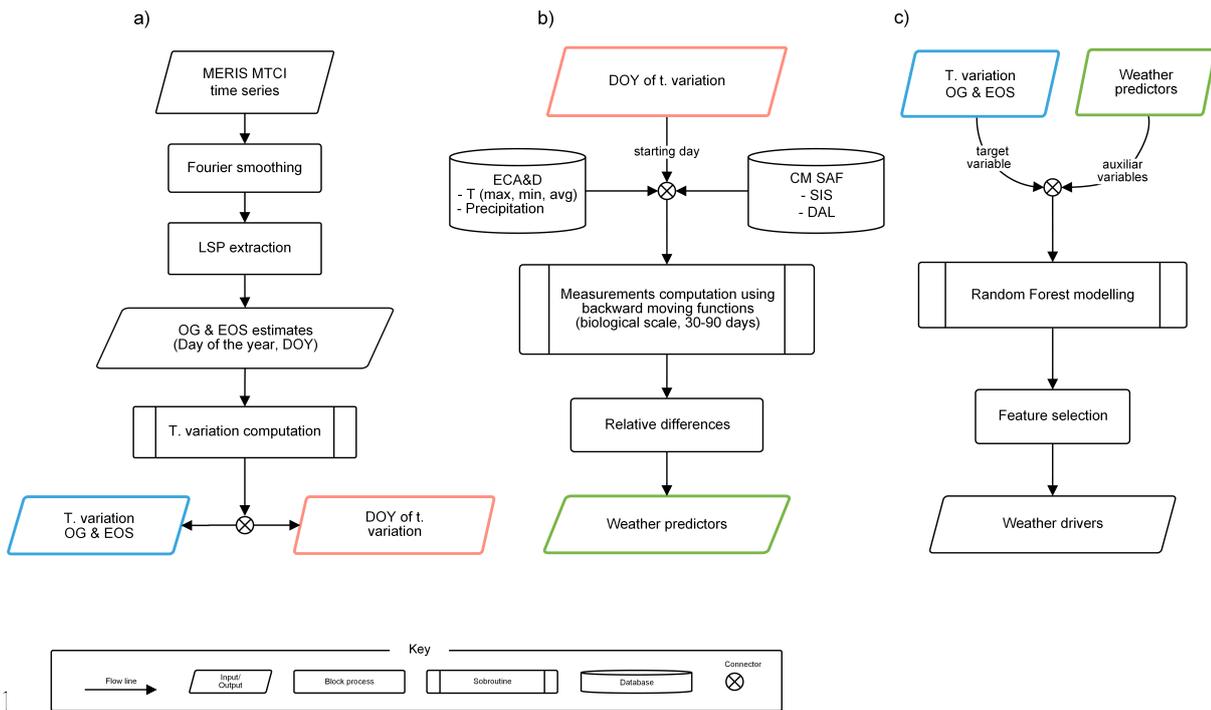
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1 Anom.	1.00	-0.40	-0.43	-0.11	-0.09	-0.12	-0.10	-0.11	-0.10	0.24	-0.03	-0.03	-0.03	-0.14	-0.04	-0.04	-0.33	-0.16	-0.16	-0.04	-0.06	-0.06	-0.45	-0.46	-0.12	-0.31	-0.03
2 GDD090	-0.40	1.00	0.93	0.11	0.14	0.11	0.13	0.11	0.15	-0.64	0.00	-0.01	-0.01	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	-0.43	0.93	1.00	0.11	0.10	0.11	0.10	0.11	0.11	-0.47	-0.01	-0.01	-0.01	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	-0.11	0.11	0.11	1.00	0.99	1.00	0.99	1.00	0.98	-0.05	0.89	0.89	0.89	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	-0.09	0.14	0.10	0.99	1.00	0.98	1.00	0.99	1.00	-0.13	0.88	0.88	0.88	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	-0.12	0.11	0.11	1.00	0.98	1.00	0.99	0.99	0.98	-0.04	0.89	0.89	0.88	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	-0.10	0.13	0.10	0.99	1.00	0.99	1.00	0.99	1.00	-0.11	0.89	0.89	0.89	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	-0.11	0.11	0.11	1.00	0.99	0.99	0.99	1.00	0.98	-0.06	0.89	0.89	0.89	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	-0.10	0.15	0.11	0.98	1.00	0.98	1.00	0.98	1.00	-0.15	0.88	0.88	0.88	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 CHIL	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	-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	-0.03	0.00	-0.01	0.89	0.88	0.89	0.89	0.89	0.88	-0.01	1.00	1.00	1.00	-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	-0.03	-0.01	-0.01	0.89	0.88	0.89	0.89	0.89	0.88	0.00	1.00	1.00	1.00	-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	-0.03	-0.01	-0.01	0.89	0.88	0.88	0.89	0.89	0.88	0.00	1.00	1.00	1.00	-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	-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	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	-0.04	0.01	0.01	0.97	0.96	0.96	0.96	0.96	0.96	0.00	0.88	0.88	0.88	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	-0.04	0.01	0.01	0.96	0.96	0.96	0.96	0.96	0.96	0.00	0.88	0.88	0.88	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	-0.33	-0.12	0.03	0.02	-0.03	0.03	-0.03	0.02	-0.04	0.28	-0.04	-0.04	-0.04	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	-0.16	-0.06	0.04	0.00	-0.03	0.00	-0.03	0.01	-0.03	0.11	-0.05	-0.05	-0.05	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	-0.16	-0.06	0.04	0.00	-0.03	0.00	-0.03	0.01	-0.03	0.11	-0.05	-0.05	-0.05	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	-0.04	0.04	0.06	0.31	0.30	0.32	0.30	0.31	0.29	0.03	0.00	0.00	0.00	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	-0.06	-0.05	0.03	-0.01	-0.04	-0.01	-0.04	0.00	-0.03	0.06	-0.06	-0.06	-0.06	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	-0.06	-0.05	0.03	-0.01	-0.04	-0.01	-0.04	0.00	-0.03	0.06	-0.06	-0.06	-0.06	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	-0.45	0.67	0.74	0.17	0.10	0.18	0.10	0.16	0.10	-0.24	0.00	-0.01	-0.01	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	-0.46	0.64	0.75	0.15	0.09	0.16	0.09	0.14	0.09	-0.26	-0.01	-0.01	-0.01	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	-0.12	0.18	0.16	0.28	0.29	0.27	0.28	0.29	0.30	-0.16	-0.01	-0.01	-0.01	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	-0.31	-0.11	0.03	0.07	0.02	0.08	0.02	0.06	0.02	0.26	-0.03	-0.03	-0.03	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	-0.03	0.05	0.06	0.31	0.31	0.32	0.31	0.31	0.30	0.01	0.00	0.00	0.00	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

1 Table 3. Correlations between the predictors used in the modelling of autumn interannual variation in LSP. Significant correlations between the
 2 anomalies and the predictors are given in bold ($p < 0.05$).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
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	-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 OGA	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	-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	0.31	0.06	1.00	0.97	0.54	0.58	0.94	0.53	0.88	0.42	0.87	0.62	-0.54	-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	0.34	0.08	0.97	1.00	0.53	0.60	0.86	0.49	0.80	0.37	0.80	0.59	-0.41	-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	0.33	0.14	0.54	0.53	1.00	0.98	0.49	0.95	0.54	0.90	0.36	0.85	-0.14	-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	0.36	0.16	0.58	0.60	0.98	1.00	0.49	0.92	0.54	0.85	0.37	0.84	-0.10	-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	0.28	0.05	0.94	0.86	0.49	0.49	1.00	0.56	0.93	0.44	0.94	0.63	-0.71	-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	0.30	0.15	0.53	0.49	0.95	0.92	0.56	1.00	0.61	0.93	0.43	0.89	-0.28	-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	0.28	0.02	0.88	0.80	0.54	0.54	0.93	0.61	1.00	0.58	0.78	0.60	-0.58	-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	0.27	0.07	0.42	0.37	0.90	0.85	0.44	0.93	0.58	1.00	0.28	0.73	-0.16	-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	0.26	0.05	0.87	0.80	0.36	0.37	0.94	0.43	0.78	0.28	1.00	0.61	-0.76	-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	0.34	0.19	0.62	0.59	0.85	0.84	0.63	0.89	0.60	0.73	0.61	1.00	-0.39	-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	0.01	-0.02	-0.54	-0.41	-0.14	-0.10	-0.71	-0.28	-0.58	-0.16	-0.76	-0.39	1.00	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	-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	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 FF	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	-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	0.07	0.02	0.09	0.11	0.05	0.07	0.04	-0.01	-0.09	-0.05	0.16	0.12	-0.05	-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	0.07	-0.05	0.10	0.11	0.13	0.13	0.10	0.13	0.12	0.13	0.08	0.13	0.00	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	0.04	-0.07	0.11	0.10	0.09	0.09	0.09	0.09	0.07	0.05	0.09	0.12	0.01	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	0.06	0.03	0.07	-0.15	-0.11	-0.01	-0.18	-0.09	-0.31	0.08	0.04	-0.05	-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	-0.05	-0.02	-0.09	-0.10	-0.07	-0.07	-0.02	0.02	0.03	0.02	-0.06	-0.02	-0.05	-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	-0.05	-0.02	-0.09	-0.10	-0.07	-0.07	-0.02	0.02	0.03	0.02	-0.06	-0.02	-0.05	-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	0.00	-0.10	-0.01	-0.03	0.04	0.02	0.02	0.07	0.23	0.17	-0.17	-0.07	0.09	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	-0.11	0.01	-0.01	-0.05	-0.06	0.05	-0.01	0.14	0.07	-0.04	-0.12	-0.01	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	-0.08	0.01	-0.22	-0.23	-0.14	-0.14	-0.13	-0.03	-0.09	-0.03	-0.14	-0.06	-0.01	-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	-0.08	0.01	-0.22	-0.23	-0.14	-0.14	-0.13	-0.03	-0.09	-0.03	-0.14	-0.06	-0.01	-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	-0.09	-0.06	-0.11	-0.15	0.08	0.04	-0.09	0.09	0.17	0.23	-0.30	-0.08	0.17	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	-0.10	-0.22	-0.25	-0.14	-0.19	-0.17	-0.11	-0.06	0.07	-0.24	-0.31	0.10	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

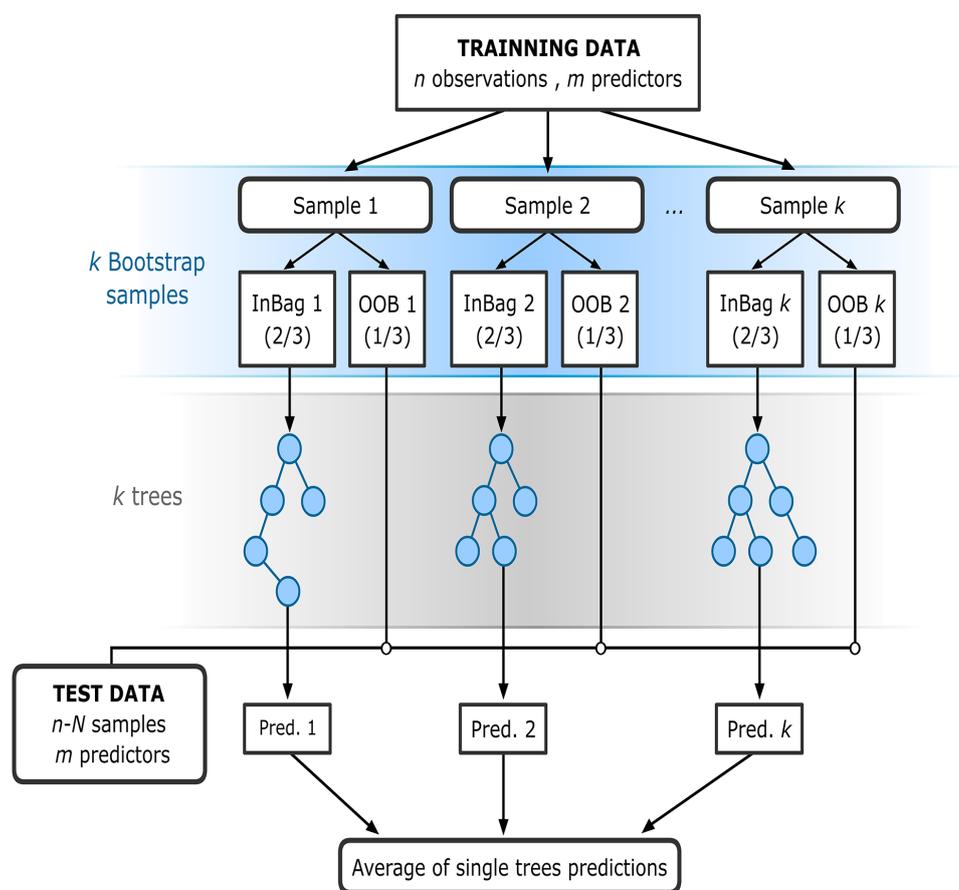


2 Figure 1. Spatial distribution of Globcover broadleaved deciduous forest and needleleaved
3 evergreen forest in 2005 (a) and 2009 (b).



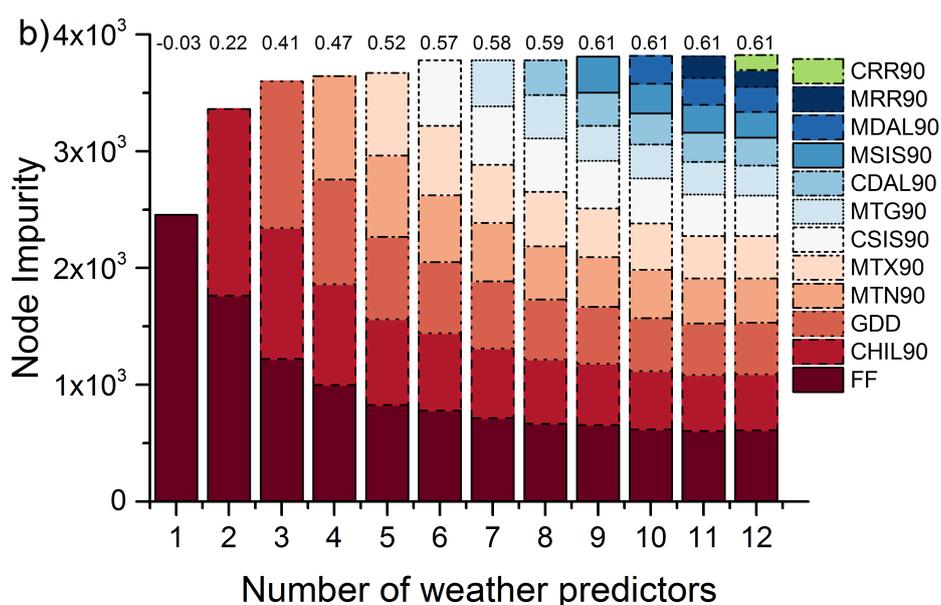
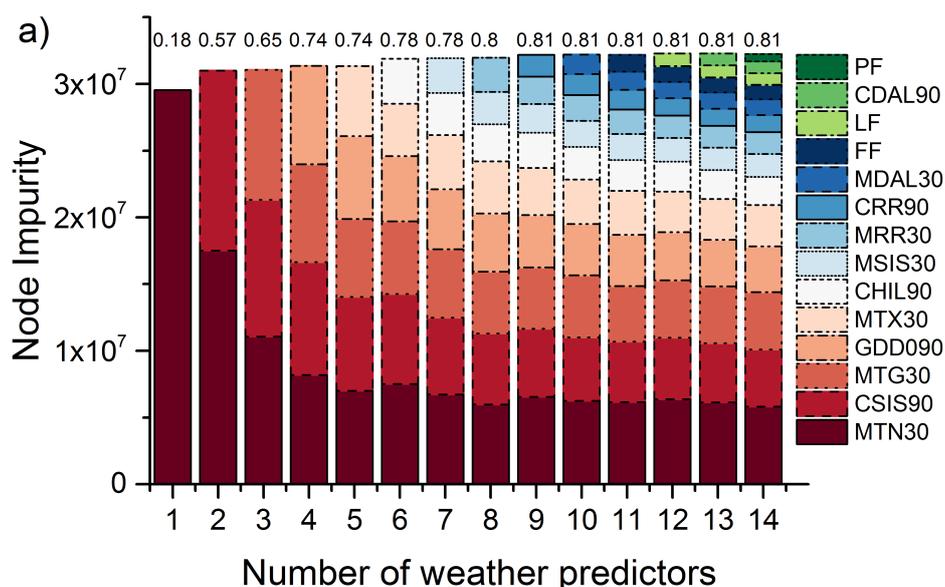
1
 2 Figure 2. Flow-chart illustrating the methodology. A) Phenology extraction and interannual
 3 variation in LSP computation. B) Computation of weather predictors. C) Modelling of
 4 interannual variation in phenology.

5

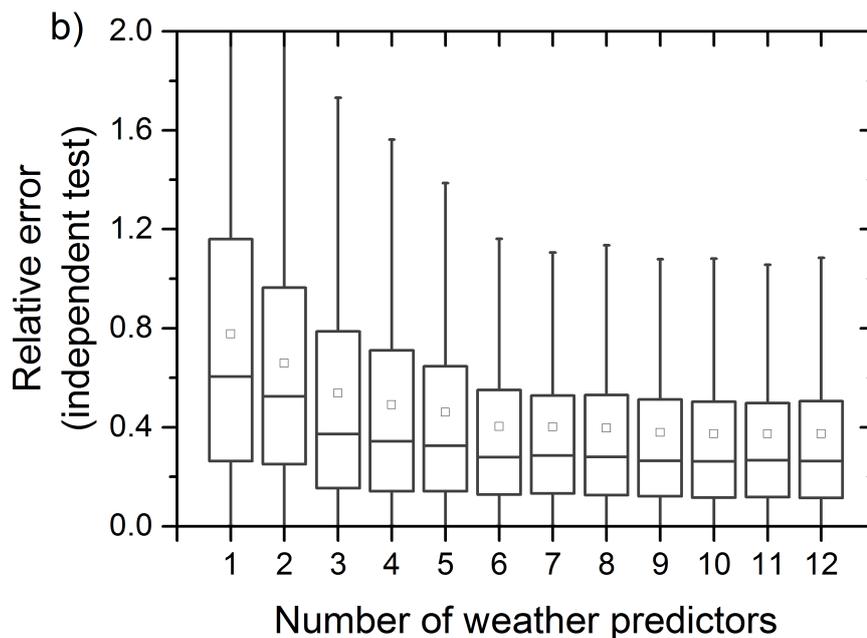
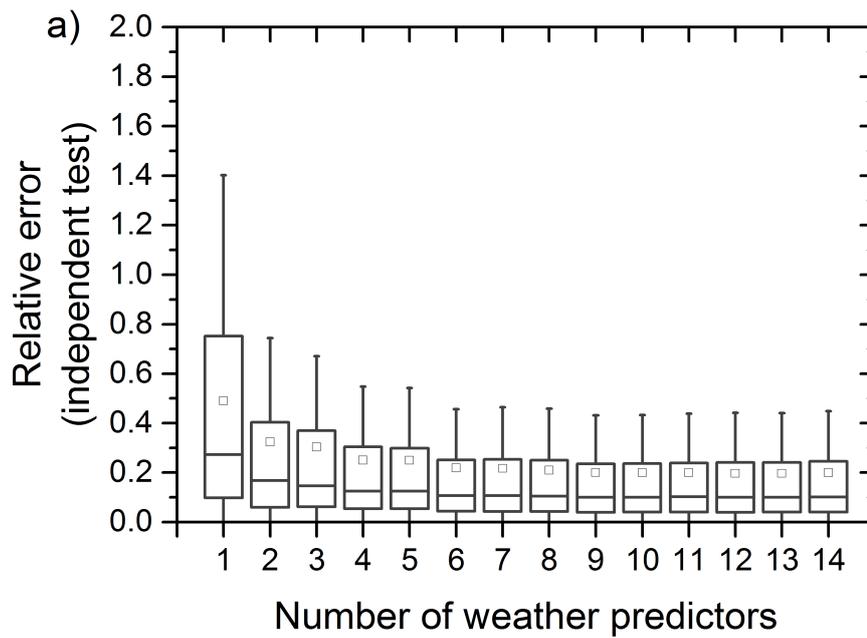


1

2 Figure 3. The flowchart of Random Forest for regression (adapted from Rodriguez-Galiano et
 3 al. 2015b). The RF method receives a subset of input vectors (n), made up of one phenology z-
 4 score value and the values of the corresponding weather predictors for a given location and
 5 year. RF builds a number K of regression trees making them grow from different training data
 6 subsets, resampling randomly the original dataset with replacement. Hence, most data will be
 7 used multiple times in different models. On the other hand, when the RF makes a tree grow, it
 8 uses the best predictor within a subset of predictors (m) which has been selected randomly from
 9 the overall set of input predictors. These especial characteristics of RF confer a greater
 10 prediction stability and accuracy and, at the same time, avoid the correlation of the different
 11 RTs, increasing the diversity of patterns that can be learnt from data. The multiple predictions
 12 of all k RTs for a given vector used as training are then averaged to obtain a unique estimation
 13 of the phenology z-score value.



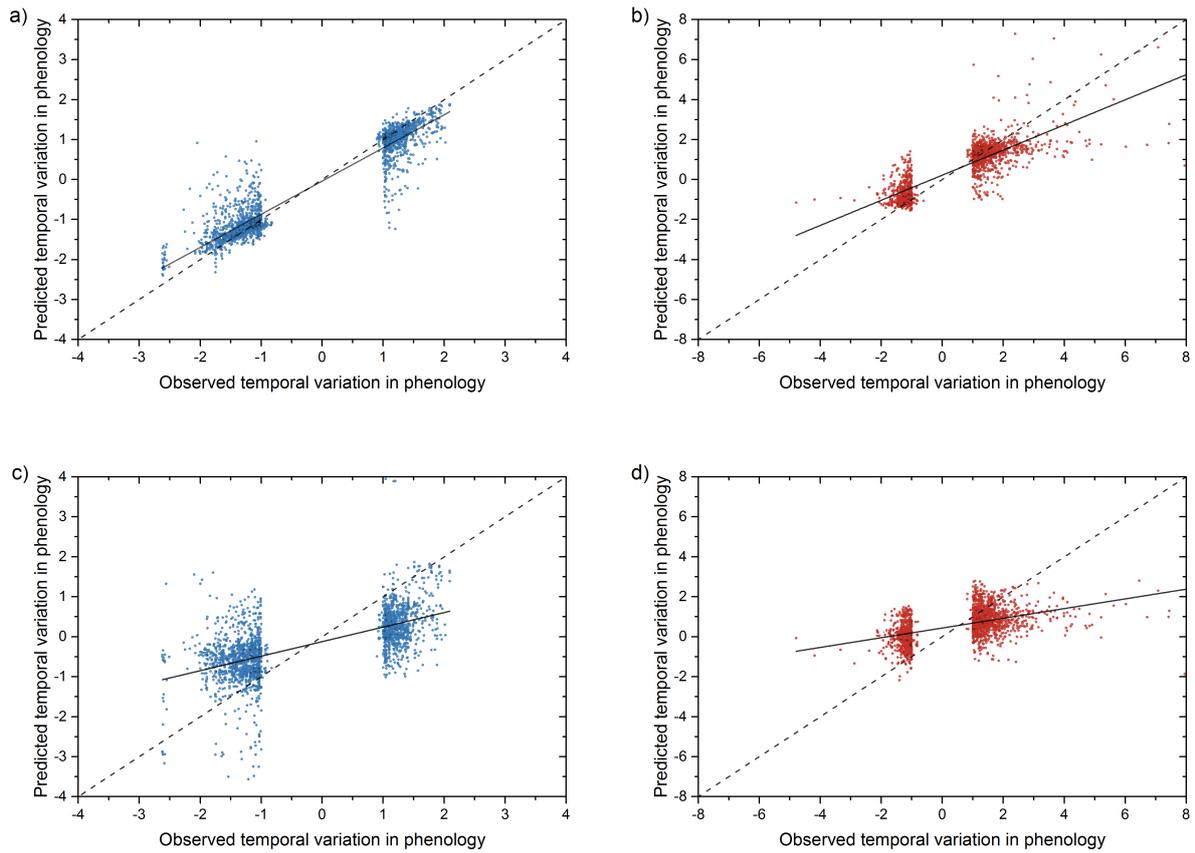
1
 2 Figure 4. Relative importance of each independent variable in predicting phenology interannual
 3 variation in Europe. Different models derived from the feature selection approach are
 4 represented in each column. Numbers given over each column represent the coefficient
 5 determination of each model. Plots at the top and bottom represent the spring (a) and autumn
 6 interannual variation in LSP (b), respectively. The names of predictors follows the notation:
 7 Prefix M and C represent the mean and cumulated functions; TX, TN and TG: maximum,
 8 minimum and average temperature, respectively; PP: precipitation; SIS: surface incoming
 9 shortwave radiation; DAL: surface radiation daylight; GDD: growing degree days; CHIL:
 10 chilling requirements; FF, LF and PF: first, last and period of freeze, respectively.



1

2 Figure 5. Relative error of the models fitted as a result of the feature selection approach. Median
 3 (interior horizontal line), mean (interior square), 1% and 99% quantiles (edge of boxes), range
 4 (extremes). Relative errors were calculated for the prediction of 1,974 and 1,576 independent
 5 observations for spring (a) and autumn (b), respectively. See previous figure for the weather
 6 predictor variables in the models, as shown in the x-axis.

7



1

2 Figure 6. Scatterplots between observed anomalies in LSP and the predictions calculated using
 3 a selection of weather predictors (see Figure 2 and Figure 3). Plots for spring phenology are
 4 shown on the left panel (blue; a, c) and autumn on the right (red; b, d). Random Forest
 5 predictions are given in the upper panel (a, b) and those of the linear regression in the bottom
 6 (c, d) panel. The dashed lines represent an exact 1:1 relationship (expected fitting), the solid
 7 lines show a linear regression of these data. The explained variances (percentage R^2) and RMSE
 8 values are 90% and 0.43 (spring Random Forest model), 68% and 0.92 (autumn Random Forest
 9 model), 39% and 1.04 (spring Linear model) and 25% 1.40 (autumn linear model).