Timing leaf senescence: a generalized additive models for location, scale and shape approach.

Bertold Mariën[1,*], Dimitri Papadimitriou[2], Titta Kotilainen[3], Paolo Zuccarini[4], Inge Dox[1], Melanie Verlinden[1], Thilo Heinecke[1], Joachim Mariën[5], Patrick Willems[6], Mieke Decoster[1], Aina Gascó[1], Holger Lange[7], Josep Peñuelas[5,8], and Matteo Campioli[1]

[1]PLECO (Plants and Ecosystems), Department of Biology, University of Antwerp, 2610 Wilrijk, Belgium
[2]IDLab (Internet Data Lab), Department of Computer Science, University of Antwerp, 2000 Antwerp, Belgium
[3]Luke (Natural Resources Institute Finland), 20520 Turku, Finland
[4]Centre for Research on Ecology and Forestry Applications (CREAF), 08290 Cerdanyola del Vallès, Spain
[5]EVECO (Evolutionary Ecology Group), Department of Biology, University of Antwerp, 2610 Wilrijk, Belgium
[6]Hydraulics and Geotechnics Section, Department of Civil Engineering, KU Leuven, Kasteelpark Arenberg 40, 3001, Leuven, Belgium
[7]Division of Environment and Natural Resources, Department of Biogeochemistry and Soil Quality, Norwegian Institute of Bioeconomy Research, 1433 Ås, Norway
[8]Global Ecology Unit CREAF-CSIC-UAB, CSIC, 08193 Bellaterra, Spain

*Author for correspondence:
Bertold Mariën
Tel: 032659333
Email: bertold.marien@uantwerpen.be

Total word count (excluding summary, author contributions, references and legends:	8424	No. of figures:	4
Summary:	300	No. of Tables	3
Introduction:	1760	No. of Supporting Information files:	2
Materials and Methods:	3507		
Results	1403		
Discussion:	1527		
Conclusion:	171		
Acknowledgements:	46		
Author Contributions:	85		
Funding:	21		

Summary

Accurate estimations of phenophases in deciduous trees are important to understanding forest ecosystems and their feedback on the climate. In particular, the timing of leaf senescence is of fundamental importance to trees' nutrient stoichiometry and drought tolerance and therefore to their vigor and fecundity. Nevertheless, there is no integrated view on the significance, and direction, of seasonal trends in leaf senescence, especially for years characterized by extreme weather events. Difficulties in the acquisition and analyses of hierarchical data can account for this.

Our study advances science by collecting four years of chlorophyll content index (CCI) measurements in thirty-eight individuals of four deciduous tree species (*Betula pendula*, *Fagus sylvatica*, *Populus tremula* and *Quercus robur*) in Belgium, Norway and Spain, and analyzing these data using generalized additive models for location, scale and shape. As a result, (I) the phenological strategy and seasonal trend of leaf senescence in these tree species could be clarified for exceptionally dry and warm years, and (III) the average temperature, global radiation, and vapor pressure deficit could be established as main drivers behind the variation in the timing of the senescence transition date.

Our results show that the onset of the re-organization phase in the leaf senescence, which we approximated and defined as local minima in the second derivative of a CCI graph, was in all species mainly negatively affected by the average temperature, global radiation and vapor pressure deficit. All together the variables explained 89% to 98% of the variability in the leaf senescence timing. An additional finding is that the generalized beta type 2 and generalized gamma distributions are well suited to model the chlorophyll content index, while the senescence transition date can be modeled using the normal-exponential-student-*t*, generalized gamma and zero-inflated Box-Cox Cole and Green distributions for beech, oak and birch, and poplar, respectively.

Keywords

Betula pendula; *Fagus sylvatica*; Generalized additive models for location, scale and shape; Leaf senescence; Phenology; *Populus tremula*; *Quercus robur*

1. Introduction

1.1. What is leaf senescence?

To survive changing environmental conditions and avoid potential trade-off effects, trees constantly have to optimize their nutrient stoichiometry and water balance (Keskitalo et al., 2005; Munne-Bosch and Alegre, 2004). For example, during tree growth, trees might be unable to compensate fluctuations in the soil's nutrients using their regular mechanisms (i.e. increasing the expression of root transporter genes, root growth and branching, or root organic compound exudation) and need to remobilize nutrients from reserves (Amtmann and Armengaud, 2009; Dakora and Phillips, 2002; Gruber et al., 2013; Maillard et al., 2015).

A significant part of the tree's nutrients is present in the leaves and can be remobilized during the process of leaf senescence, a highly coordinated developmental stage of the cells with leaf death as a consequence (Hörtensteiner and Feller, 2002; Keskitalo et al., 2005; Medawar, 1957). It is evident that changes in the tree crown through leaf senescence will affect the tree's water balance and capacity to uptake carbon. Alternatively, leaf senescence can be triggered to correct the water balance, if the tree fails to compensate the water deficit via stomata regulation, with nutrient remobilization as co-occurring effect (Matos et al., 2020; Munne-Bosch and Alegre, 2004). Under drought, premature leaf senescence might occur in certain species (i.e. 'drought-deciduous' species; e.g. *Betula pendula* or *Quercus robur*) to avoid nutrient losses through an unanticipated abscission of green leaves, as it occurs in other species (i.e. 'drought-evergreen' species; e.g. *Fagus sylvatica*) (Estiarte and Penuelas, 2015; González, 2012; Harvey and Driessche, 2011; Manzoni et al., 2015; Marchin et al., 2010; Matos et al., 2020; Sohrt et al., 2018; Vitasse et al., 2011; Wendler and Millard, 1996).

The main drivers of leaf senescence are the photoperiod (through phytochrome), the light intensity and spectral quality (through photo-oxidative stress), cold temperatures (through an increasing electrolyte leakage and decreasing antioxidant metabolism), warm temperatures (through a reduction in the carboxylation and an increased respiration) and drought stress (through a reduction in the water potential of the roots cells and increase in the hormones abscisic acid and ethylene) (Estrella and Menzel, 2006; Feller and Fischer, 1994; Lang et al., 2019; Matos et al., 2020). In all cases, the senescence process is guided by increased concentrations of reactive oxygen species (ROS) controlled by antioxidant levels (Jajic et al., 2015; Juvany et al., 2013; Munne-Bosch and Alegre, 2004). During premature senescence, immature leaves will initiate senescence later than mature leaves, as the former have higher concentrations of cytokines and auxins (i.e. senescence delaying hormones) than the latter, and are less sensitive to ethylene (Diamantoglou and Kull, 1988; Marchin et al., 2010; Matos et al., 2020).

The variety in drivers that affect the timing and rate of senescence is reflected in different leaf senescence strategies of deciduous species (Manzoni et al., 2015). For example, 'drought-evergreen' species will only initiate leaf senescence in autumn, typically as a result of photoperiodic changes (Vitasse et al., 2009). During a drought, these species can lose carbon and nutrients through respiration or hydraulic failure with the abscission of green leaves as a consequence. Therefore, they might become outcompeted when droughts occur more frequently and persist longer (Crabbe et al., 2016; IPCC, 2014; Manzoni et al., 2015). Alternatively, 'drought-deciduous' species will shed their leaves gradually, or rapidly, when the carbon balance in the leaves becomes negative and the leaf carbon reserves are depleted. When the circumstances improve, 'drought-deciduous' species might develop a new leaf flush to rebalance the trade-off between the carbon uptake and the nutrient cost of making new leaves, and as long as the

photosynthesis surpasses the respiration (Manzoni et al., 2015). Some plant genera (e.g. *Alnus*), associated with nitrogen-fixing bacteria, might simply desiccate green leaves without initiating senescence (Gill et al., 2015; Keskitalo et al., 2005). Furthermore, the interaction among different environmental stressors might also affect the timing of the leaf senescence and explain its variance within one or multiple individuals (Archetti et al., 2013; Delpierre et al., 2009; Gressler et al., 2015; Liu et al., 2020; Panchen et al., 2015).

The onset of leaf senescence in 'drought-deciduous' species can occur at any time, while it is conservatively timed in drought-evergreen species. In addition, the rate and duration of premature and autumn leaf senescence can also differ among species and years, and leaves can abscise without initiating or completing leaf senescence (Primka Iv and Smith, 2019). Therefore, the timing of leaf senescence, as opposed to the timing of the leaf abscission, is described best on a species-specific basis and in regard to the actually occurring nutrient remobilization (Gill et al., 2015; Munné-Bosch, 2015; Panchen et al., 2015; Xie et al., 2018a). For example, studies found that *Quercus robur* can efficiently remobilize nutrients (e.g. N, P) and microelements (e.g. Ni, B), can undergo premature senescence, and has strong nutrient sinks (e.g. reproductive tissue and sinks that are mostly root-based, instead of bark-based) (Maillard et al., 2015; Villar-Salvador et al., 2015). Studies also noted that a significant decline in chlorophyll is a good indicator of the onset of leaf senescence because the detoxification of chlorophyll always immediately precedes the dismantling of nutrient-rich macromolecules (e.g. rubisco, RNA, et cet.) (Aerts, 1996; Estiarte and Penuelas, 2015; Etienne et al., 2018; Feller and Fischer, 1994; Giraldo et al., 2013; Hörtensteiner and Feller, 2002; Maillard et al., 2015; Marchin et al., 2010; Matile, 2000).

1.2. Leaf senescence on the ecological scale?

Arguably the timing of the onset of leaf senescence is the moment when the senescence-associated genes (SAGs) are expressed (Andersson et al., 2004; Gepstein et al., 2003; Munne-Bosch and Alegre, 2004; Schippers et al., 2015). Unfortunately, observing the expression of SAGs on an ecological scale is unattainable. Since leaf senescence consists of multiple steps and requires an integrative and multi-scale analysis, assessing the best proxy for detecting temporal trends in leaf senescence therefore becomes of crucial importance (Bresson et al., 2017; Gill et al., 2015; Keskitalo et al., 2005). Examples in the literature, of such proxies include visual coloration assessments, chlorophyll measurements, remote sensing observations of standard indices and coloration analysis through phenocams or drones (Maleki et al., 2020; Piao et al., 2019). Portillo-Estrada et al. (2020) even showed that the onset of senescence corresponded with an outburst in oxygenated volatile organic compounds (VOCs), suggesting the use of VOCs to assess the onset of senescence on a regional scale. Because studies use different definitions of leaf senescence, different proxies, different trend estimation or filtering methods, and even different indicators for phenological transition dates, the reproducibility, comparability and interpretation of results on leaf senescence is also hampered (Gallinat et al., 2015; Gill et al., 2015; Gu et al., 2009; Panchen et al., 2015).

The use of different trend estimation methods and different indictors for phenological transition dates should be based on theoretical grounds and good statistical practices (i.e. where the model assumptions are met and ideally the start, rate and end of the leaf senescence process are described separately) (Gallinat et al., 2015; Houlahan et al., 2017; Hudson and Keatley, 2010). Especially because the temporal dynamics of ecological processes are inherently complex (Ryo et al., 2019). For example, the trend estimation methods in the literature range from the fitting of double-logistic functions to complex

network-based modelling (Alberton et al., 2019; Bush et al., 2017; Diao, 2019; Menzel et al., 2008; Vander Mijnsbrugge et al., 2016; Xu et al., 2014; Zhang and Goldberg, 2011; Zhang et al., 2003; Zhao et al., 2019). Likewise, phenological transition dates are studied using a variety of methods (e.g. threshold values, spectral signature changes, etc.) (Diao, 2019; Gill et al., 2015; Hudson and Keatley, 2010; Keenan and Richardson, 2015; Lim et al., 2018; Verbesselt et al., 2010; Wingate et al., 2015; Xie et al., 2018b; Xie and Wilson, 2020; Zhang et al., 2003).

Zhao et al. (2019) deduced that the detection of phenological transition dates is inherently linked to the accuracy of the model trend and seasonality, and that many methods have additional pitfalls (e.g. the inadequate handling of non-linear trends, too restrictive model assumptions, model misspecification, data noise and usability limitations). Their approach (i.e. ensemble learning) is especially useful for inference but less suited for exploring the drivers of a process and predicting. We therefore use here generalized additive models for location, scale and shape (GAMLSS) to assess the temporal trend, while we use local minima in the second derivative of partial effect functions as phenological transition dates. These should be interpreted as a proxy for the onset of the 're-organization phase' of senescence, a phase defined by large changes inside the leaf cells (e.g. chlorophyll degradation, loss of cellular integrity, decrease in photosynthetic activity, et cet.) (Munne-Bosch and Alegre, 2004).

1.3. Research questions and hypotheses

This study aims (I) to clarify the phenological strategy and seasonal trend of leaf senescence in four common deciduous tree species (*Fagus sylvatica*, *Quercus robur*, *Betula pendula* and *Populus tremula*) during exceptionally warm and dry years, and (II) to determine the environmental drivers of the leaf senescence timing. The link between the estimation of trends, seasonality and phenological transition dates was addressed, while using GAMLSS to assess autumn phenology (Akanztiliotou et al., 2002; Rigby and Stasinopoulos, 2001; Rigby and Stasinopoulos, 2005; Zhao et al., 2019).

We expected no inter-annual trend in the transition dates in *Fagus sylvatica* because it is assumed to be a 'drought-evergreen species' of which the timing of its leaf senescence is determined by the photoperiod and cold temperatures (Matos et al., 2020; Vitasse et al., 2013; Vitasse et al., 2009). Likewise, no trend was expected in the transition dates in *Populus tremula* because studies suggest that only the photoperiod and light-derived factors can explain its stable onset of senescence (Brelsford et al., 2019a; Brelsford et al., 2019b; Fracheboud et al., 2009; Keskitalo et al., 2005; Michelson et al., 2018). In contrast, as the study period (2017 – 2020) was warm and extremely dry, we expected advanced transition dates in *Betula pendula* and *Quercus robur* because they are 'drought-deciduous' species that will initiate premature senescence when exposed to drought stress (Estrella and Menzel, 2006; Maillard et al., 2015; Wendler and Millard, 1996). Additionally, the growth pattern between *Fagus sylvatica* (closed canopy with one leaf flush and senescence starting in the sun-exposed leaves), *Populus tremula* (open canopy with one leaf flush at a mature age), *Quercus robur* (closed canopy with the potential for multiple leaf flushes and senescence starting in the sun-exposed leaves) and *Betula pendula* (open canopy with continuous leaf flushing and senescence starting in the inner parts of the canopy), might be reflected in the timing of the transition dates (Koike, 1990).

2. Methods

2.1. Description of the sites

From mid-July to late November, for each year within the study period, the chlorophyll content index (CCI) was measured for leaves of trees in Belgium (from 2017 to 2020), Norway (from 2017 to 2019) and Spain (from 2018 to 2020). We chose these countries because Belgium, Norway and Spain approximately cover the center, and Northern and Southern edge of our study species' spatial distribution in Europe. In Belgium, we measured ten leaves of twenty mature trees in three forests; the Klein Schietveld (KS), Park of Brasschaat (PB) and Fortress of Borsbeek (BB). Likewise, in Spain, we measured ten leaves of nine trees in three forests; Fogars de Montclús (FM), Sant Joan de les Abadesses (SJA) and El Puig (EP). In Norway, we only measured six leaves of nine trees in a Norwegian forest on the farm of Hoxmark (NO).

Our five Belgian stands (i.e. beech KS, birch KS, beech PB, oak PB and poplar BB) cover an area of ca 0.5 to 3 ha and are all monospecific and homogeneous. The beech and oak trees (ca 60 – 70 and 60 – 120 years old, respectively) were planted but experienced little human management. The birch and poplar trees (ca 50 – 60 and 20 - 50 years old, respectively) were naturally established. The soil in the BB (i.e. technozol; sandy loam) is disturbed and wet, while the soil in the KS and PB (i.e. podzol; sandy) is dry to moderately wet, less nutrient rich, and undisturbed (Mariën et al., 2019). The beech trees in Norway were part of a small stand (ca 0.05 ha) planted for educational activities ca 15 years before sampling at the Hoxmark Experimental Farm in Ås. The birch and oak trees (ca. 10 y old) have been spontaneously growing in the vicinity of the small beech stand. The soil has a thick humus top layer over a layer with a large amount of clay. The Spanish stands of beech FM and poplar EP are part of a large forested area, mainly covered by beech, within the Natural park of the Montseny Massif. The poplar trees at EP have been growing spontaneously on old agricultural land partly recolonized by forest ca 100 years ago. The soils in EP (leptozol; sandy with granite and granodiorite) and FM are similar (leptozol; sandy with leucogranite), although the soil in FM contains less organic matter. The Betula trees of SJA are in the park of Sant Joan de les Abadesses (leptozol; sandy). They were scattered individuals on a pasture, remaining after a clear-cut removing a large birch stand ca. 50 years ago. In Spain, the poplars are estimated to be ca 20 – 40 years old, while the beech and birch trees are estimated to be ca 50 – 100 years old. All individual trees were selected for dominance and vitality.

Our study included four poplars (*Populus tremula* L.) in the BB, four beeches (*Fagus sylvatica* L.) and four birches (*Betula pendula* Roth) in the KS, four oaks (*Quercus robur* L.) and four beeches in the PB, three beeches, three birches and three oaks in NO, four beeches in FM, three birches in SJA and two poplars in EP. Every two weeks from 2017 to 2019, and every week in 2020, we (or tree-climbers in Belgium) collected five sun-leaves and five shade-leaves from each tree in Belgium and Spain. In Norway, only three sun-leaves and three shade-leaves were collected from each tree. Subsequently, the CCI of these leaves was measured immediately using a chlorophyll content meter (CCM-200 plus, Opti-Sciences Inc., Hudson, NH, USA) measuring the ratio of the transmission of radiation from the red (653 nm) and near-infrared (931 nm) wavelengths emitted by a diode (Parry et al., 2014). The leaves from each tree were measured approximately on the same moment of the day; and using the same side of the leaf. Due to its curvilinear relationship, the CCI acts as a proxy for chlorophyll concentrations and senescence (Bresson et al., 2017; Michelson et al., 2018).

All analyses were done using R v.3.6.3. (R Core Team, 2020). R/dplyr was used for data handling, while R/ggplot2, R/grid, R/viridis and R/cowplot were used for visualization (Garnier, 2018; Wickham, 2009; Wickham et al., 2018; Wilke, 2019).

2.2. GAMLSS

2.2.1. Why use GAMLSS

The assumptions of Generalized additive mixed models (GAMMs) were violated (see Text S1). For example, the residuals of the GAMMs were non-independent and identically distributed (i.i.d.), heteroscedastic and non-normal, while the data were likely modeled better using different family distributions, showed over-dispersion, were skewed with heavy tails, and showed bi-modality. A framework capable of addressing these issues is GAMLSS (Rigby and Stasinopoulos, 2005).

GAMLSS were introduced to model data where the distribution of the response variable does not necessarily follows an exponential family distribution (e.g. data that is discrete, censored, heterogeneous, truncated, skewed or kurtotic, etc.) (Akanztiliotou et al., 2002; Rigby and Stasinopoulos, 2001; Rigby and Stasinopoulos, 2005). Unlike GAMMs, the GAMLSS inferential framework models not only the distribution parameter μ, but also the distribution parameters σ, v and τ. These four distribution parameters correspond to the location, scale and shape of the response variable's distribution and can generally be interpreted using the distribution's moments (i.e. the mean, variance, skewness and kurtosis, respectively) (Stasinopoulos and Rigby, 2007). In practice, the semi-parametric GAMLSS framework can use many distribution families to model the response variable whilst providing not only information on changes in the mean but also on the variance, skewness and kurtosis (Rigby et al., 2019).

Another advantage of the 'complete distribution' approach of GAMLSS, unlike quantile regressions, is that it offers tools for both rigorous testing of the parametric model assumptions and model selection (Voudouris et al., 2013). However, the effectiveness of GAMLSS depends largely on choices made by the user. For example, in GAMLSS, one has to decide the distribution of the response variable, the link functions for each parameter, the explanatory terms for each parameter and the amount of smoothing (Voudouris et al., 2013).

2.2.2. The GAMLSS model

GAMLSS can be written as:

$$\mathbf{Y}_i \sim \mathbf{D}(\boldsymbol{\mu}_i, \boldsymbol{\sigma}_i, \boldsymbol{v}_i, \boldsymbol{\tau}_i)$$
$$\mathbf{g}_1(\boldsymbol{\mu}_i) = \boldsymbol{\eta}_1$$
$$\mathbf{g}_2(\boldsymbol{\sigma}_i) = \boldsymbol{\eta}_2$$
$$\mathbf{g}_3(\boldsymbol{v}_i) = \boldsymbol{\eta}_3$$
$$\mathbf{g}_4(\boldsymbol{\tau}_i) = \boldsymbol{\eta}_4$$

with Y_i as the independent response variable observations for i = 1,…, n, observations, D as the distribution of the response variable and g being the monotonic link function relating the predictor η to the distribution parameters (μ_i, σ_i, v_i, τ_i) (Rigby et al., 2019; Stasinopoulos and Rigby, 2007; Stasinopoulos et al., 2017; Stasinopoulos et al., 2018).

To model the CCI as a function of its covariates, we used the *gamlss* and *refit* functions in R/gamlss (Rigby and Stasinopoulos, 2005). First, we used the lowest Akaike Information Criterion (AIC) values returned by the *fitDist* function in R/gamlss to select the best potential distributions. Then, we built several models with

different distributions and checked their parametric assumptions using the *GAIC*, *plot* and *wp* functions in R/gamlss. The *GAIC* function returns the generalized AIC values of the models. The *plot* function returns a summary comprising the mean, variance, coefficient of skewness, coefficient of kurtosis and Filliben correlation coefficient (Filliben, 1975). It also outputs four diagnostic plots of the normalized quantile residuals, which allow to test for homogeneity of variance and deviations of normality (Dunn and Smyth, 1996; Stasinopoulos et al., 2018). In addition, the *plot* function gives the ACF and pACF plots to test for autocorrelation of the regression residuals. The *wp* function gives a wormplot (a detrended quantile-quantile plot) and is used to test for skewness and kurtosis, or to ensure normality in the residuals whenever the plot of the residuals is within the two elliptic 95% pointwise confidence intervals (Buuren and Fredriks, 2001). We selected the most suited distributions based on the generalized R^2 of Nagelkerke (given by the *Rsq* function in R/gamlss), the normalized root mean square error (RMSE; given by the *performance_RMSE* function in R/performance) the model assumptions and the results on the *fitDist* function (Lüdecke et al., 2021; Nagelkerke, 1991). We chose the generalized beta type 2 (GB2) distribution with default logarithmic link functions for our beech and poplar data, while we considered the generalized gamma (GG) distribution with default logarithmic and identity link functions for our birch and oak data (Harter, 1967; Lopatatzidis and Green, 2000; McDonald, 1984; McDonald, 1996; McDonald and Xu, 1995; Stasinopoulos et al., 2018). Both distributions are continuous distributions between 0 to $+\infty$ and can be found in R/gamlss.dist (Stasinopoulos and Rigby, 2020). Observe that the GB2 and GG distributions have five and three parameters, respectively (Rigby et al., 2019).

The hierarchical character of our data determined our choice for the 'most suited' model. For example, when identical values are observed in an additive term within a categorical independent variable (e.g. the same amount of precipitation per *site*), the additive term behaves as a constant term. The choice to characterize the dependency among observations of the same point predictors (*tree individual* or *site*) is not trivial. Therefore, together with constrains imposed by the link functions, autocorrelation and multimodality in the data, we decided to run a GAMLSS model for each species, year and site (Pregibon, 1980). This method yields twenty-six simpler models in which the fixed covariates of the CCI were the *leaf type* (categorical with two levels), *tree individual* (categorical with four levels) and *day of the year* (continuous). In another six models, modeling the CCI of birch and oak in NO, the *leaf type* was not included as a covariate. A consequence of the separate models was that we sacrificed predictive power in favor of yearly trend estimation accuracy. To maximize the penalized log-likelihood and reach convergence, we used the default RS algorithm for fitting mean and dispersion additive models (Rigby and Stasinopoulos, 1996a; Rigby and Stasinopoulos, 1996b; Stasinopoulos et al., 2017). To reuse the smoothers of R/mgcv, we implemented the *ga* function from R/gamlss.add as argument of the *gamlss* function (Stasinopoulos and Rigby, 2020; Wood, 2017; Wood, 2004; Wood, 2011; Zuur et al., 2007). The *ga* function hampered stepwise model selection and the extraction of results (the *getSmo* function was required to extract smoother information) but its usage reduced and increased the AIC and R^2 of the models, respectively. Within the ga argument, the fixed covariates were the *leaf type* (categorical with two levels) and *day of the year* (continuous). The latter covariate was smoothed using P-splines, while the dependency among observations of the same *tree individual* was incorporated outside the ga argument by using *tree individual* as random intercept. We chose P-splines because they are low rank smoothers that ease the selection of the optimal number and position of the knots, and because they are suitable for nonparametric and mixed modeling (Eilers et al., 2015; Eilers and Marx, 1996). We specified the sigma.fo, nu.fo and tau.fo arguments like we modeled the covariates for the mean and specified the usage of the restricted maximum likelihood (REML) in the ga argument. The REML argument was chosen as

smoothness selection method because it is preferred by literature and less prone to overfitting (Models S1 - S2) (Reiss and Ogden, 2009; Wood, 2011).

The parametric assumptions of the models were ensured using R/gamlss's *summary*, *getSmo*, *plot*, *Rsq*, and *wp* functions. The graphs were obtained using R/gamlss's *plot* and *term.plot* functions. To evaluate the risk that a leaf has to present CCI values below a given threshold, we constructed partial quantile plots. These were constructed using R/gamlss's *getQuantile* function and R/graphics *curve* function, and show the partial effect of the explanatory variable *Day of the year* on the quantiles.

2.2.3. Smoother dynamics in GAMLSS

2.2.3.1. The rate of change in the curvature

To asses the rate of change in the leaf senescence progression, Zhang et al. (2003) suggested the use of the curvature (κ) in order to characterize the phenological transition date. This transition date corresponds to the time at which the curvature value decelerates most rapidly. Its decrease can be observed from the function's graph as the time where its curvature changed the most. Zhang et al. (2003) suggested to consider a local minimum in the first derivative (i.e. rate of change) of the curvature κ'(t). To track the rate of variation in the curvature over time, one needs to approximate the signed curvature κ in the graph of a function y using the parametrized version of equation 1.

Eq. 1

$$\mathbf{x} = \boldsymbol{t}$$
$$\mathbf{y} = \boldsymbol{f}(\boldsymbol{x})$$
$$\boldsymbol{\kappa} = \frac{\boldsymbol{y}''(\boldsymbol{x})}{(\mathbf{1} + (\boldsymbol{y}'(\boldsymbol{x}))^{\mathbf{2}})^{\frac{\mathbf{3}}{\mathbf{2}}}}$$

With x as value for the explanatory variable, y as the function obtained by the penalized smoother in a gamlss model and κ as its curvature.

The signed curvature can be approximated using the second derivative, if the first derivative (i.e. slope) of the function is small (Eq. 2). In this case, when y is sufficiently smooth and y' is not bigger than y".

Eq. 2

$$\boldsymbol{\kappa} = \boldsymbol{y}''(\mathbf{1} + \boldsymbol{O}(\boldsymbol{y}'^{\mathbf{2}}))$$

with κ as the curvature of the graph of the function y obtained by the penalized smoother in the gamlss model, and O is the Bachmann-Landau symbol.

2.2.3.2. The second derivative of a function

To find the transition dates of interest using the functions in R/gamlss, we first ran the *getPEF* function which returned the partial effect that the explanatory variable *day of the year* had on the predictor. Subsequently, we used the *pef* function to calculate the elasticity (i.e. derivatives) of the partial effect function. Then, we could calculate our transition date of interest: a local minima in the second derivative (i.e. the moment when the CCI decline accelerated most rapidly; in other words, when the curvature of the graph changed the most). To assess the confidence on this date, we incorporated its calculation in a resampling procedure iterated 200 times. The number of basis functions (k) in the P-splines was set to five

or six. The results were plotted as density plots using the geom_density_ridges2 argument in R/ggridges (Wilke, 2020).

2.2.4. Drivers of the variation in the transition date

We aimed to determine which environmental variables drove the variation in the transition dates. The selected candidates were the daily average temperature, global radiation, precipitation, vapor pressure deficit and year (i.e. year represents also all other environmental factors that change throughout time; e.g. N deposition). Accounting for the nature of the phytochrome system, we opted to include the global radiation, rather than the day length, as predictor for the transition dates (Legris et al., 2019). The relationship between the sum of the global radiation and the day length is strong, although not the same for each year (Fig. S1 – S2). For Belgium, the meteorological data were derived from half-hourly measurements done at the meteorological station in Brasschaat (21 m.a.s.l.; courtesy from INBO and ICOS; Fig. S3; see Mariën et al. (2021)). When necessary, the data were gap-filled using data from the Dutch meteorological station in Woensdrecht (14 m.a.s.l). Meteorological data from Norway and Spain were measured at the meteorological stations in Ås (92 m.a.s.l.; Fig. S4), and Sant Pau de Segúries (852 m.a.s.l; Fig. S5) and Viladrau (953 m.a.s.l; Fig. S6), respectively. All meteorological data from the Netherlands, Norway and Spain were taken from the Dutch (KNMI, 2021), Norwegian (MET Norway, 2021) and Catalan (MeteoCat, 2021) Meteorological Institutes.

The vapor pressure deficit was calculated following the equations in Buck (1981) (Eq. 3).

Eq. 3

$$\boldsymbol{e_0 = 613.75 \times e^{\left(\frac{(17.502 \times T)}{(240.97 + T)}\right)}}$$

$$\boldsymbol{e = \left(\frac{RH}{100}\right) \times e_0}$$

$$\boldsymbol{VPD = e_0 - e}$$

with e_0 as the saturation vapor pressure (in Pa), T as the temperature (in °C), e as the actual vapor pressure deficit (in Pa), RH as the relative humidity (in %) and VPD as the vapor pressure deficit (in Pa).

To indicate the drought stress for the hydrological years (i.e. from the 1st of April to the 31st of March) 2017 to 2021, we computed the daily rainfall deficit using solar radiation, wind speed, temperature, relative humidity and precipitation data from the meteorological station in Ukkel (Fig. S7). We first computed the potential evapotranspiration using the Bultot et al. (1983) method, which is similar to Penman (1948)'s method but has parameters calibrated specifically for the local conditions (Baguis et al., 2010). Subsequently, we derived the daily rainfall deficit (I) per hydrological year and (II) using continuous computation by accumulating the daily potential evapotranspiration minus the daily sum of the precipitation. Unlike for the rainfall deficit starting from a zero deficit at the start of the hydrological year (i.e. the first of April), we accounted in the calculation of the continuously computed rainfall deficit for the hydrological fraction in wet periods that does not contribute to building up ground water reserves. As such, we can account for potential effects of droughts in successive years. Long-term (i.e. since 1898) daily precipitation and potential evapotranspiration data are available at the station of Ukkel. The data for the period 1901 - 2000 was considered here as the reference period for the long-term rainfall deficit statistics (Mariën et al., 2021).

To combine the half-hourly Belgian meteorological data with our transition dates, we averaged the temperature and vapor pressure deficit, summed the precipitation and global radiation, and used only

vapor pressure deficit and global radiation data between 7 a.m. and 7 p.m. (to avoid negative values) for each day.

To assess the relevance of the potential drivers behind the variation in the transition dates in Belgium, we used three methods for each species. First, we made correlation matrices using the *rcorr* and *corrplot* functions in R/Hmisc and R/corrplot, respectively (Harrel Jr, 2020; Wei and Simko, 2017). These provide Pearson's correlation coefficient and indicate which explanatory variables will not influence the variation in the transition dates (Friendly, 2002; Murdoch and Chow, 1996). Then, we ran Random Forest models to determine which explanatory variable explains most variation (i.e. R^2) in the transition dates. The Random Forests, an ensemble learning method that classifies multiple regression trees and returns the average regression, were built by first splitting the data in a training and a validation dataset using R's *floor* and *sample* functions. Subsequently, the percentage increase in the mean square error and the node purity of the predictions was calculated using the *randomForest* function in R/randomForest (Breiman, 2001; Liaw and Wiener, 2002). Finally, to automatically select the 'significant' drivers behind the transition dates, we used multivariate adaptive regression splines (MARS) to perform enhanced adaptive (non-parametric) regressions through hinges (EARTH) (Friedman, 1991; Stasinopoulos et al., 2017). EARTH uses a form of piecewise linear regressions, which capture the non-linear relationships in the data, to select the most significant explanatory variables in a similar manner to step functions. Hinge functions first create kinks or points were linear regressions intersect. These are then continuously assessed using a generalized cross-validation procedure and pruned until an optimal number is established based on changes in the R^2 that are less than 0.001 (Boehmke and Greenwell, 2020). To implement the MARS algorithm using the *earth* function in R/earth as an argument in the *gamlss* function, we required R/gamlss.add2's interface in the GAMLSS framework (Milborrow, 2020; Stasinopoulos et al., 2017; Stasinopoulos, 2020). Note that R/gamlss.add2 is not openly available and its usage is courtesy of the author.

To model the transition dates as a function of its covariates, we constructed a GAMLSS model for each species. We determined the most suitable distribution for the response variable of each species using the lowest AIC value returned by the *FitDist* function and the lowest Kullback-Leibler divergence. The latter performs as a proxy for the information loss that occurs when one describes a real dataset using a theoretical distribution and is returned by the *KLD* function in R/LaplacesDemon (Statisticat, 2020). However, the lack of data hampered the use of these functions, forcing us to fall back on diagnostic plots to assess the most suitable distribution for the response variable in a trial-and-error fashion. We selected the normal-exponential-student-*t* (NET) distribution with default identity and logarithmic link functions for beech, the GG distribution with default logarithmic and identity link functions for oak and birch, and the zero-inflated Box-Cox Cole and Green (BCCGo) distribution with default logarithmic and identity link functions for poplar (Cole and Green, 1992; Harter, 1967; Lopatatzidis and Green, 2000; Rigby and Stasinopoulos, 1994). The NET and BCCGo distributions are both continuous distributions where the former can have values between $-\infty$ to $+\infty$ and the latter between 0 to $+\infty$. Both distributions are found in R/gamlss.dist (Stasinopoulos and Rigby, 2020). The NET distribution has four parameters with fixed shape parameters, while the BCCGo distribution has three parameters (Rigby et al., 2019). For simplicity and to exploit the automatic variable selection, we only specified the MARS algorithm in the μ parameter of the model to allow the fitting of smooth non-linear functions based on the continuous explanatory variables *year*, *average temperature*, *average vapor pressure deficit*, *global radiation* and *precipitation*. No interactions were considered, while the default RS algorithm was used (Model S3 – S5) (Rigby and Stasinopoulos, 1996a; Rigby and Stasinopoulos, 1996b; Stasinopoulos et al., 2017). Plots of the results were retrieved using R/gamlss's *getSmo* function and R/earth's *plotmo* function. For comparison, we ran each model again using the resulting significant explanatory variables as linear predictor variables which we also specified in the sigma.fo and nu.fo arguments when possible (Models S6 – S8).

3. Results

3.1. Modeling the chlorophyll content index trend

The diagnostic plots for beech and poplar (Model S1), and oak and birch (Model S2) showed that the normalized quantile residuals of the GAMLSS models did not violate any prior assumptions, suggesting that the CCI was adequately modeled using the GB2 and GG distributions (Fig. S8 – S24). However, the CCI was better modeled for beech (using the GB2 distribution) than oak, birch and poplar (using the GG and GB2 distributions) for two reasons. First, the width of the confidence intervals and the quantile functions in the regression terms and partial quantile plots, respectively, of the GAMLSS models for oak, birch and poplar was larger in some years than those of the GAMLSS models for beech (Fig. S25 – S36). Second, the R^2 of Nagelkerke of the GAMLSS models for beech ranged generally higher (from 0.66 to 0.92) than the R^2 of Nagelkerke of the GAMLSS models for oak and birch (from 0.29 to 0.89; Table S1). In general, the fitted GAMLSS models for beech also had a lower global deviance and AIC value, and more degrees of freedom.

Note that the GB2 and GG distributions were chosen here as the best quantitative distributions with a bias towards particular years. They were not necessarily the most suitable distributions to model the CCI at a particular year, but they were the most suitable distributions for modeling the CCI of a given species during all years.

3.2. Characterizing the transition date

The CCI of the trees was high in summer and rapidly declined in autumn (Fig. 2). Its decline started first in oak and was followed by poplar, birch and beech, respectively. The standard errors on the average CCI were consistently low.

Despite the simple resampling approach, the histograms and density plots of the transition dates reflect each species phenological strategy of leaf senescence (Fig. 3; Table 1). For example, the beech trees in Belgium, Norway and Spain all show a stable timing in their transition dates centered around the third, first and second week of October, respectively. There are exceptions, for example: the beech trees of the KS in 2018 and the beech trees of NO in 2017. The former shows a much earlier and very small peak in the transition dates around the end of July, while the latter show transition dates centered around the fourth week of September. Likewise, in 2018 and 2019, there are earlier and very small peaks in the transition dates around the end of July in the beech trees in FM. In 2019 and 2020, in the beech trees in FM, one can also observe slightly bigger peaks in the transition dates around one to half a month earlier than the major peak in mid-October.

The oak trees in the PB show two peaks, with a first small peak (except in 2019 when the peak is large) in the distribution of the transition dates centered around the fourth week of July (2017 - 2019) or late September (2020), and a second much bigger peak around the fourth week of October. The oak trees in NO also show two peaks in 2017 and 2018. However, unlike for the oak trees in the PB, the oak trees in NO show a first bigger peak around the second (2018) and third (2017) week of August and a second smaller peak around the second (2017) and fourth (2018) week of September. In 2019, the oak trees in NO shows only one big peak in the transition dates centered around the second week of September.

The transition dates of the poplars in the BB occurred largely in the third week of October in 2018 and 2019, and ca one week earlier in 2020. A similar pattern can be observed in the poplars in EP where the transition dates peaked in the third week of October in 2018, in the first week of November in 2019 and

the fourth week of September in 2020. However, the long tail in the histograms and density plots of the poplar data suggest that poplars are susceptible to an advance in their timing of the transition dates.

For birch trees, in most stands, spread out histograms and density plots of the transition dates suggest a high probability of premature senescence events rather than a clearly defined peak distribution in autumn. However, birch shows a variable pattern in transition dates among years. For example, in 2018, the birch trees of the KS, NO and SJA show only one peak around the fourth and first week of October, and third week of September, respectively. The timing of birch seems similar to that of beech in the regard that birch trees in Belgium have transition dates that generally occurred later than those of Spain. The birch trees of NO show the latest transition dates centered around the first week of September in 2017 and the fourth week of August in 2019.

3.3. Determining the drivers of the leaf senescence timing

The results of the correlation matrices, the Random Forests and the EARTH procedure in the fitted GAMLSS models with MARS (Models S3 – S5) were similar (Fig. 4; Fig. S37 – S39). For example, all methods show the concurring result that the transition date in all four species was - with a different degree of importance - negatively related to the average temperature, the vapor pressure deficit and the global radiation. In addition, most methods suggest that the precipitation did not much affect the timing of the transition dates in any species, while the effect of the year was likely only substantial in birch (Table 2). Note here that the poor man's partial dependence plots showing the precipitation effect on the transition dates seems horizontal for all species, suggesting that the precipitation did not have a substantial effect. Likewise, the poor man's partial dependence plots do not show a clear yearly trend in the transition dates in any of the species, despite the correlation matrices suggesting a small positive and negative correlation in beech and birch, respectively (Fig. 4; panel C).

The diagnostic plots of the GAMLSS models with MARS show violations of the assumptions on the normalized quantile residuals (Fig. S40 – S43; Table 3). First, the residuals in the wormplots suggest leptokurtosis, indicating that the kurtosis is modeled too light (Stasinopoulos et al., 2017). Likely, because MARS is a local nonparametric algorithm sensitive to insufficient data at key knots. Although the combination of GAMLSS (semi-parametric) and MARS avoids the use of more complex global nonparametric algorithms (e.g. neural networks) and the problems related to parametric methods (e.g. error distribution knowledge), the explanatory variables could only be specified in the μ parameter. The consequential assumption that the process is driven by the mean shows the limits to modeling the response variable using one distribution parameter. Second, the ACF and pACF show serial autocorrelation. One reason might be that insufficient data hampered efficient use of specialized functions to determine the most suitable distribution or specialized smoothers to cope with non-linear effects. The interactions between the explanatory variables were also not considered in favor of computability. Third, the Durbin-Wu-Hausman, and Breusch-Pagan's LM and Pesaran's CD tests showed the presence of endogenous variables (i.e. variables correlated with the error term) and cross-sectional dependence, respectively (Breusch and Pagan, 1980; Durbin, 1954; Hausman, 1978; Pesaran, 2004; Pesaran, 2014; Wu, 1973). The former and latter tests were executed using the *ivreg*, and *plm* and *pcdtest* functions in R/ivreg and R/plm, respectively (Croissant and Millo, 2008; Fox et al., 2020; Millo, 2017). The consequence of endogeneity is that the likelihood of reporting significant, but biased, coefficient estimates increased, and that there is always a bias-variance tradeoff in the current experimental set-up. The cross-sectional dependence shows that the explanatory variables influenced the response variable at

different velocities and the cofounding effects that arise when integrating different datasets. While the model output is valid in its local context, future improvements might include probabilistic approaches or lagged/dynamic variable modeling.

The above reasons also explain why the diagnostic plots of the GAMLSS models with linear predictors indicate violations on the assumptions of the residuals, and why the direction and magnitude of the effect shown in some of the regression terms in the alternative GAMLSS models might not concur with the results found in the GAMLSS models with MARS (e.g. the effect of the global radiation on the transition dates in beech; Fig. S44 – S51, Tables S2 – S3).

4. Discussion

4.1. Timing the transition date

Our results show that the GB2 and GG distributions are well-suited to infer trends in the CCI of beech, oak, birch and poplar. Henceforward, their probability density functions (see the dGB2 and dGG functions in R/gamlss) can be used to predict the probability that a leaf reaches a certain CCI value at a given date. Alternatively, their inverse cumulative distribution functions (see qGB2 and qGG functions in R/gamlss) can now be used to predict the quantile value of the CCI at any given probability (i.e. the value of the CCI for which the probability is at or below a given quantile). One could test whether more specialized members of the GB distribution family are more suited than the GB2 and GG distributions to infer trends from more frequently sampled CCI measurements, and whether the GB2 and GG distributions can model the CCI of other deciduous tree species.

The resampling approach used to extract the transition dates is less sensitive to outliers and large deviations than parametric methods. However, to deal with heteroscedasticity and spatio-temporal correlation in the data, a case could be made for future implementation of wild bootstrap methods or block non-uniform resampling approaches. Nevertheless, although the variance in the phenological transition dates refers here only to the variance among species and sites, the histograms and density plots of the transition dates reflect our hypotheses well.

We did not expect a trend in the transition dates in beech and poplar, while we expected an advance in the transition dates in oak and birch in response to the dry and warm weather. The distributions of the transition dates in beech (and poplar in Belgium) overlapped each year, indicating that the onset of senescence in beech is rather conservative with no clear trend. However, the kurtosis in the distribution of the transition dates in poplars did differ between years. In addition, the distribution parameters for beech and poplar changed slightly each year suggesting that although the bulk of the leaves initiated the re-organization phase of senescence at one given moment, the leaves were not completely insensitive to the effect of environmental variables. This is certainly true for poplar in Spain, where the distribution of the transition dates in 2020 peaks around one month earlier than in 2017. Nevertheless, the onset of senescence in beech can be considered a singular event in correspondence with a singular leaf flush. Premature shedding of beech leaves in summer, as occurred during our study (personal observation) and the study of Bréda et al. (2006), was barely reflected by our estimations of the transition dates or CCI measurements. Therefore, the amount of nutrient remobilization during these events is likely negligible, meaning that these events, in beech, can hardly be considered true senescence (Bréda et al., 2006).

The situation is different for oak where the transition date can perhaps be characterized as two events, one low probability event in summer (i.e. marked by a small peak in the distribution of the transition date) and one high probability event in autumn (i.e. marked by a big peak in the distribution of the transition date) in Belgium, and vice-versa in Norway. This result corresponds with the literature where oak is found to show premature senescence and rare polycyclic behavior with a large intraspecific variability (Bobinac et al., 2012; Maillard et al., 2015). Finally, our result concur with our hypothesis that the onset of senescence in birch can occur in several events, concurring with birch's nondeterministic growth pattern. As a consequence, it is sometimes hard to pinpoint the onset of senescence in birch to one specific moment.

It remains difficult to forecast the direction of the transition date in our deciduous trees (an advance or delay) because the years 2017 to 2020 were, in Belgium, all exceptionally warm and dry, offering little background on the 'normal' timing for the onset of senescence (Mariën et al., 2019). The extreme drought of 2018, and the heat stress with increased aridity in 2019 were nevertheless not associated with a substantially larger probability for an advanced transition date.

We do notice that beech, birch and oak trees in Norway generally have earlier transition dates than trees of the same species in Belgium. Likewise, beech and birch trees in Spain have earlier and later transition dates than trees of the same species in Belgium and Norway, respectively. In 2020, the poplar trees also showed earlier transition dates in both Belgium and Spain. However, drawing generalized conclusion from this result is difficult due to the age differences between the trees. The meteorological circumstances in Belgium, Norway and Spain were also profoundly different, although 2018 was extremely dry and warm in both Belgium and Norway.

4.2. What drives the variation in the transition date?

Our results show that the NET, GG and BCCGo distributions can be used to model the transition dates for beech, oak, birch and poplar (see the dNET, qNET, dGG, qGG, dBCCGo, and qBCCGo functions in R/gamlss). However, the AIC and Kullback-Leibler divergence values, and the deviations in the diagnostic plots of our models indicate that more data, more suitable or specialized distributions (e.g. finite mixture distributions), and more suitable link functions might be required to improve (I) our estimations of the timing of the transition dates and (II) our estimations of the direction and magnitude of the effects of environmental parameters on the transition dates (Aitkin et al., 2009; Everitt, 2014; Leisch, 2004; Stasinopoulos et al., 2017).

Nevertheless, the magnitude in which the timing of senescence in deciduous trees is affected by environmental parameters is clearly species-specific (Xie et al., 2018b). In addition, our results showing that the average temperature, vapor pressure deficit and global radiation significantly and negatively affected the transition dates in all four species supports the idea that the balance between the ROS (especially H_2O_2) and antioxidants drives the timing of senescence (Andersson et al., 2004; Jajic et al., 2015; Juvany et al., 2013; Khanna-Chopra et al., 2013; Munne-Bosch and Alegre, 2004). Senescence would then be initiated by a variety of stress factors, whether or not the leaves first reached or required a 'point of no return' established by the effect of the photoperiod on the phytochrome system (Keskitalo et al., 2005; Olsen et al., 1997). All species would still maintain different mechanisms and thresholds to control their timing of senescence. Note here that the transition date remains a proxy for the actual onset of senescence (i.e. the 'initiation phase' or moment the SAG genes are expressed); a proxy whose value

depends on the length between the initiation phase and the moment a significant change occurs in the curvature of the CCI graph (Munne-Bosch and Alegre, 2004).

In contrast to literature, our data does not suggest that the photoperiod is solely responsible for the actual onset of senescence in beech and poplar (Keskitalo et al., 2005; Michelson et al., 2018). In fact, the global radiation explained only around 17% and 25% of the variation in the transition dates in beech and poplar, respectively. On the other hand, our results indicate that the temperature and vapor pressure both explain more than 19% to 25%, and 19% to 24% of the variation in the transition dates in beech, poplar, oak and birch. Although the transition date is just a proxy, it relates strongly to the onset of relevant ecological processes (i.e. maintenance of the tree's nutrient stoichiometry and response to drought stress) defining the onset of senescence (i.e. or at least its re-organization phase). As expected, and shown by the small variation that is found in the transition dates of beech and poplar, the transition date in oak and birch is more sensitive to environmental variables. In future studies, changes in the photon ratios of blue to green, and blue to red could be could be examined as potential cues related to the diurnal cycle and the timing of senescence (Brelsford et al., 2019a; Brelsford et al., 2019b; Chiang et al., 2019; Kotilainen et al., 2020). In addition, the water vapor column thickness and the total ozone column thickness affect the red to far-red photon ratio. The annual pattern of these atmospheric factors could therefore also have a role as a seasonal cue for changes in the timing of senescence as, for example, studies predicts an increase in future atmospheric water vapor levels (IPCC, 2012; Kotilainen et al., 2020; Solomon et al., 2009).

We highlight three additional results. First, the variation in the data explained by the EARTH GAMLSS models (R^2 ranges from 0.71 to 0.93) and the Random Forest models (the percentage of variance explained ranges from 89% to 98%) is rather high. Second, the precipitation did not have large immediate effects on the transition date in our species, certainly not in beech. Third, the year (interpreted here as a measure for the unexplained variance, rather than its temporal meaning) explained only around 12% to 23% of the variation of the transition dates. Any legacy effect or changes in the timing of the leaf out are not expected to explain more than 23% of the variation in the transition dates (Chen et al., 2019; Fu et al., 2014; Keenan and Richardson, 2015).

5. Conclusion

The GB2 distribution is well suited to model the CCI in beech and poplar leaves, while the GG distribution yields satisfactory results to model the CCI in oak and birch. The distributions on the transition dates, (NET, BCCGo and GG), give a good indication of the uncertainty that surrounds the onset of senescence in beech, oak, birch and poplar, and can be used as prior information for Bayesian analyses. They also allow the inference of environmental variables, mainly the average temperature, vapor pressure deficit and global radiation, that negatively affected the yearly and species-specific variation in the transition dates. The effect of the precipitation and year on the transition dates in beech (and poplar in Belgium) was small. However, the effect of the year on the scale and shape of the transition date distribution was substantial in oak and birch, two deciduous tree species known for their premature senescence. To conclude, we highlight the link that has been made between the estimation of trends, seasonality and phenological transition dates.

Acknowledgements

We express our gratitude to prof. dr. Mikis Stasinopoulos, dr. Craig Brelsford, the Flemish Institute for Nature and Forest (INBO), the Integrated Carbon Observation System (ICOS), the Flemish Agency for Forest and Nature (ANB), the Belgian Armed Forces, the Municipality of Brasschaat and Bergen Boomverzorging.

Author contributions

MC designed the sampling methodology; the research team of MC, which includes ID, MV, TH, BM, PZ, MD, AG, HL and others, provided the CCI data; BM and DP conceived the ideas for the statistical approach and subsequently analyzed the data; JM provided initial code for resampling; TK investigated the effect of light spectrum indicators in Belgium; PW collected and analyzed the data to calculate the rainfall deficit; BM wrote the manuscript. All authors read the final draft and gave approval for publication.

Funding

BM and MC acknowledge the DOCPRO4 fellowship (University of Antwerp) and ERC Starting Grant LEAF-FALL (714916) (European Research Council), respectively.

References

- Aerts, R., 1996. Nutrient Resorption from Senescing Leaves of Perennials: Are there General Patterns? J. Ecol., 84(4): 597-608.
- Aitkin, M., Francis, B., Hinde, J. and Darnell, R., 2009. Statistical Modelling in R.
- Akanztiliotou, K., Rigby, R. and Stasinopoulos, D., 2002. The R implementation of Generalized Additive Models for Location, Scale and Shape. Statistical modelling in Society: Proceedings of the 17th International Workshop on statistical modelling, 54.
- Alberton, B. et al., 2019. Leafing Patterns and Drivers across Seasonally Dry Tropical Communities. Remote Sensing, 11(19): 2267.
- Amtmann, A. and Armengaud, P., 2009. Effects of N, P, K and S on metabolism: New knowledge gained from multi-level analysis. Curr. Opin. Plant Biol., 12: 275-83.
- Andersson, A. et al., 2004. A transcriptional timetable of autumn senescence. Genome Biol, 5(4): R24.
- Archetti, M., Richardson, A.D., O'Keefe, J. and Delpierre, N., 2013. Predicting Climate Change Impacts on the Amount and Duration of Autumn Colors in a New England Forest. PLoS ONE, 8(3): e57373.
- Baguis, P., Roulin, E., Willems, P. and Ntegeka, V., 2010. Climate change scenarios for precipitation and potential evapotranspiration over central Belgium. Theoretical and Applied Climatology, 99: 273-286.
- Bobinac, M., Batos, B., Miljković, D. and Radulovic, S., 2012. Polycyclism and Phenological Variability in the Common Oak (Quercus robur L.). Arch. Biol. Sci., 64: 97-105.
- Boehmke, B. and Greenwell, B.M., 2020. Hands-On Machine Learning with R. Chapman and Hall/CRC, 488 pp.
- Bréda, N., Huc, R., Granier, A. and Dreyer, E., 2006. Temperate forest trees and stands under severe drought: a review of ecophysiological responses, adaptation processes and long-term consequences. Ann For Sci, 63(6): 625-644.
- Breiman, L., 2001. Random Forests. Machine Learning, 45(1): 5-32.
- Brelsford, C.C., Nybakken, L., Kotilainen, T.K. and Robson, T.M., 2019a. The influence of spectral composition on spring and autumn phenology in trees. Tree Physiol, 39(6): 925-950.
- Brelsford, C.C., Trasser, M., Paris, T., Hartikainen, S.M. and Robson, T.M., 2019b. Understory light quality affects leaf pigments and leaf phenology in different plant functional types. bioRxiv: 829036.
- Bresson, J., Bieker, S., Riester, L., Doll, J. and Zentgraf, U., 2017. A guideline for leaf senescence analyses: From quantification to physiological and molecular investigations. J. Exp. Bot., 69.
- Breusch, T.S. and Pagan, A., 1980. The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics. Rev. Econ. Stud., 47(1): 239-253.
- Buck, A.L., 1981. New Equations for Computing Vapor Pressure and Enhancement Factor. Journal of Applied Meteorology, 20(12): 1527-1532.
- Bultot, F., Coppens, A. and Dupriez, G.L., 1983. Estimation de l'évapotranspiration potentielle en Belgique : (procédure révisée). Bruxelles : Institut royal météorologique de Belgique.
- Bush, E.R. et al., 2017. Fourier analysis to detect phenological cycles using long-term tropical field data and simulations. Methods Ecol. Evol., 8(5): 530-540.
- Buuren, S. and Fredriks, M., 2001. Worm plot: A simple diagnostic device for modelling growth reference curves. Stat. Med., 20: 1259-77.
- Chen, L. et al., 2019. Long-term changes in the impacts of global warming on leaf phenology of four temperate tree species. Global Change Biol., 25(3): 997-1004.

- Chiang, C., Olsen, J.E., Basler, D., Bankestad, D. and Hoch, G., 2019. Latitude and Weather Influences on Sun Light Quality and the Relationship to Tree Growth. Forests, 10(8): 610.
- Cole, T.J. and Green, P.J., 1992. Smoothing reference centile curves: The lms method and penalized likelihood. Stat. Med., 11(10): 1305-1319.
- Crabbe, R.A. et al., 2016. Extreme warm temperatures alter forest phenology and productivity in Europe. Sci Total Environ, 563-564: 486-95.
- Croissant, Y. and Millo, G., 2008. Panel Data Econometrics in R: The plm Package. Journal of Statistical Software; Vol 1, Issue 2 (2008).
- Dakora, F. and Phillips, D., 2002. Root exudates as mediators of mineral acquisition in low-nutrient environments. Plant Soil, 245: 35-47.
- Delpierre, N. et al., 2009. Modelling interannual and spatial variability of leaf senescence for three deciduous tree species in France. Agricultural and Forest Meteorology, 149(6-7): 938-948.
- Diamantoglou, S. and Kull, U., 1988. Der Stickstoffhaushalt immergrüner mediterraner Hartlaubblätter. Flora, 180(5): 377-390.
- Diao, C., 2019. Complex network-based time series remote sensing model in monitoring the fall foliage transition date for peak coloration. Remote Sens. Environ., 229: 179-192.
- Dunn, P.K. and Smyth, G.K., 1996. Randomized Quantile Residuals. Journal of Computational and Graphical Statistics, 5(3): 236-244.
- Durbin, J., 1954. Errors in Variables. Revue de l'Institut International de Statistique / Review of the International Statistical Institute, 22(1/3): 23-32.
- Eilers, P., Marx, B. and Durbán, M., 2015. Twenty years of P-splines. SORT (Statistics and Operations Research Transactions), 39: 149-186.
- Eilers, P.H.C. and Marx, B.D., 1996. Flexible smoothing with B -splines and penalties. Statist. Sci., 11(2): 89-121.
- Estiarte, M. and Penuelas, J., 2015. Alteration of the phenology of leaf senescence and fall in winter deciduous species by climate change: effects on nutrient proficiency. Glob. Chang. Biol., 21(3): 1005-17.
- Estrella, N. and Menzel, A., 2006. Responses of leaf colouring in four deciduous tree species to climate and weather in Germany. Clim. Res., 32(3): 253-267.
- Etienne, P. et al., 2018. Macro and Micronutrient Storage in Plants and Their Remobilization When Facing Scarcity: The Case of Drought. Agriculture, 8: 14.
- Everitt, B.S., 2014. Finite Mixture Distributions, Wiley StatsRef: Statistics Reference Online.
- Feller, U. and Fischer, A., 1994. Nitrogen Metabolism in Senescing Leaves. Crit. Rev. Plant Sci., 13(3): 241-273.
- Filliben, J.J., 1975. The Probability Plot Correlation Coefficient Test for Normality. Technometrics, 17(1): 111-117.
- Fox, J., Kleiber, C. and Zeileis, A., 2020. ivreg: Two-Stage Least-Squares Regression with Diagnostics.
- Fracheboud, Y. et al., 2009. The control of autumn senescence in European aspen. Plant Physiol., 149(4): 1982-91.
- Friedman, J.H., 1991. Multivariate Adaptive Regression Splines. The Annals of Statistics, 19(1): 1-67.
- Friendly, M., 2002. Corrgrams. The American Statistician, 56(4): 316-324.
- Fu, Y.S. et al., 2014. Variation in leaf flushing date influences autumnal senescence and next year's flushing date in two temperate tree species. Proc Natl Acad Sci U S A, 111(20): 7355-60.
- Gallinat, A.S., Primack, R.B. and Wagner, D.L., 2015. Autumn, the neglected season in climate change research. Trends Ecol Evol, 30(3): 169-76.

- Garnier, S., 2018. viridis: Default Color Maps from 'matplotlib'.
- Gepstein, S. et al., 2003. Large-scale identification of leaf senescence-associated genes. The Plant Journal, 36(5): 629-642.
- Gill, A.L. et al., 2015. Changes in autumn senescence in northern hemisphere deciduous trees: a meta-analysis of autumn phenology studies. Ann. Bot., 116(6): 875-88.
- Giraldo, J.P., Wheeler, J.K., Huggett, B.A. and Holbrook, N.M., 2013. The role of leaf hydraulic conductance dynamics on the timing of leaf senescence. Funct. Plant Biol., 41(1): 37-47.
- González, E., 2012. Seasonal patterns of litterfall in the floodplain forest of a large Mediterranean river. Limnetica, 31.
- Gressler, E., Jochner, S., Capdevielle-Vargas, R.M., Morellato, L.P.C. and Menzel, A., 2015. Vertical variation in autumn leaf phenology of Fagus sylvatica L. in southern Germany. Agricultural and Forest Meteorology, 201: 176-186.
- Gruber, B.D., Giehl, R.F.H., Friedel, S. and von Wirén, N., 2013. Plasticity of the Arabidopsis Root System under Nutrient Deficiencies. Plant Physiol., 163(1): 161.
- Gu, L. et al., 2009. Characterizing the Seasonal Dynamics of Plant Community Photosynthesis Across a Range of Vegetation Types.
- Harrel Jr, F.E., 2020. Harrell Miscellaneous.
- Harter, H.L., 1967. Maximum-Likelihood Estimation of the Parameters of a Four-Parameter Generalized Gamma Population from Complete and Censored Samples. Technometrics, 9(1): 159-165.
- Harvey, H. and Driessche, R., 2011. Poplar nutrient resorption in fall or drought: influence of nutrient status and clone. Can J For Res, 29: 1916-1925.
- Hausman, J.A., 1978. Specification Tests in Econometrics. Econometrica, 46(6): 1251-1271.
- Hohberg, M., Pütz, P. and Kneib, T., 2020. Treatment effects beyond the mean using distributional regression: Methods and guidance. PLoS ONE, 15(2): e0226514.
- Hörtensteiner, S. and Feller, U., 2002. Nitrogen metabolism and remobilization during senescence. J. Exp. Bot., 53(370): 927-937.
- Houlahan, J.E., McKinney, S.T., Anderson, T.M. and McGill, B.J., 2017. The priority of prediction in ecological understanding. Oikos, 126(1): 1-7.
- Hudson, I.L. and Keatley, M.R., 2010. Phenological Research. Springer.
- IPCC, 2012. Near-term Climate Change: Projections and Predictability. Notes 953–1028.
- IPCC, 2014. Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change. In: R.K.P.a.L.A.M.e. Core Writing Team (Editor). IPCC, Geneva, Switzerland, pp. 151.
- Jajic, I., Sarna, T. and Strzalka, K., 2015. Senescence, Stress, and Reactive Oxygen Species. Plants (Basel, Switzerland), 4(3): 393-411.
- Juvany, M., Muller, M. and Munne-Bosch, S., 2013. Photo-oxidative stress in emerging and senescing leaves: a mirror image? J Exp Bot, 64(11): 3087-98.
- Keenan, T.F. and Richardson, A.D., 2015. The timing of autumn senescence is affected by the timing of spring phenology: implications for predictive models. Glob. Chang. Biol., 21(7): 2634-2641.
- Keskitalo, J., Bergquist, G., Gardestrom, P. and Jansson, S., 2005. A cellular timetable of autumn senescence. Plant Physiol., 139(4): 1635-48.
- Khanna-Chopra, R., Nutan, K. and Pareek, A., 2013. Regulation of Leaf Senescence: Role of Reactive Oxygen Species, pp. 393-416.
- KNMI, 2021. *Dagwaarnemingen*. Retrieved 06/07/2021 from https://daggegevens.knmi.nl/

- Koike, T., 1990. Autumn coloring, photosynthetic performance and leaf development of deciduous broad-leaved trees in relation to forest succession. Tree Physiology, 7(1_2_3_4): 21-32.
- Kotilainen, T. et al., 2020. Patterns in the spectral composition of sunlight and biologically meaningful spectral photon ratios as affected by atmospheric factors. Agricultural and Forest Meteorology, 291: 108041.
- Lang, W., Chen, X., Qian, S., Liu, G. and Piao, S., 2019. A new process-based model for predicting autumn phenology: How is leaf senescence controlled by photoperiod and temperature coupling? Agricultural and Forest Meteorology, 268: 124-135.
- Legris, M., Ince, Y.C. and Fankhauser, C., 2019. Molecular mechanisms underlying phytochrome-controlled morphogenesis in plants. Nat Commun, 10(1): 5219.
- Leisch, F., 2004. FlexMix: A General Framework for Finite Mixture Models and Latent Class Regression in R. 2004, 11(8): 18.
- Liaw, A. and Wiener, M., 2002. Classification and Regression by randomForest. R news, 2(3): 18 - 22.
- Lim, C. et al., 2018. Ecological consideration for several methodologies to diagnose vegetation phenology. Ecol. Res., 33.
- Liu, Q. et al., 2020. Modeling leaf senescence of deciduous tree species in Europe. Global Change Biol., 26(7): 4104-4118.
- Lopatatzidis, A. and Green, P., 2000. Nonparametric quantile regression using the gamma distribution. submitted for publication.
- Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P. and Makowski, D., 2021. performance: An R Package for Assessment, Comparison and Testing of Statistical Models. Journal of Open Source Software, 6(60).
- Maillard, A. et al., 2015. Leaf mineral nutrient remobilization during leaf senescence and modulation by nutrient deficiency. Frontiers in Plant Science, 6(317).
- Maleki, M. et al., 2020. Estimation of Gross Primary Productivity (GPP) Phenology of a Short-Rotation Plantation Using Remotely Sensed Indices Derived from Sentinel-2 Images. Remote Sensing, 12: 2104.
- Manzoni, S., Vico, G., Thompson, S., Beyer, F. and Weih, M., 2015. Contrasting leaf phenological strategies optimize carbon gain under droughts of different duration. Advances in Water Resources, 84: 37-51.
- Marchin, R., Zeng, H. and Hoffmann, W., 2010. Drought-deciduous behavior reduces nutrient losses from temperate deciduous trees under severe drought. Oecologia, 163: 845-54.
- Mariën, B. et al., 2019. Detecting the onset of autumn leaf senescence in deciduous forest trees of the temperate zone. New Phytol, 224(1): 166-176.
- Mariën, B. et al., 2021. Does drought advance the onset of autumn leaf senescence in temperate deciduous forest trees? Biogeosciences, 18(11): 3309-3330.
- Matile, P., 2000. Biochemistry of Indian summer: physiology of autumnal leaf coloration. Exp Gerontol, 35(2): 145-158.
- Matos, F.S., Borges, L.P. and Muller, C., 2020. Ecophysiology Of Leaf Senescence. Agronomy & Agricultural Science, 3(1): 1-6.
- McDonald, J., 1984. Some Generalized Functions for the Size Distribution of Income. Econometrica, 52(3): 647-63.
- McDonald, J.B., 1996. 14 Probability distributions for financial models, Handbook of Statistics. Elsevier, pp. 427-461.

- McDonald, J.B. and Xu, Y.J., 1995. A generalization of the beta distribution with applications. J. Econometrics, 66(1): 133-152.
- Medawar, P.B., 1957. The Uniqueness of the individual, by P.B. Medawar. Methuen, London.
- Menzel, A. et al., 2008. Bayesian analysis of the species-specific lengthening of the growing season in two European countries and the influence of an insect pest. Int J Biometeorol, 52(3): 209-218.
- MET Norway, 2021. *Norsk Klimaservicesenter*. Retrieved 06/07/2021 from https://seklima.met.no/observations/
- MeteoCat, 2021. *RuralCat*. Retrieved 06/07/2021 from https://ruralcat.gencat.cat/agrometeo.estacions
- Michelson, I.H. et al., 2018. Autumn senescence in aspen is not triggered by day length. Physiol Plant, 162(1): 123-134.
- Milborrow, S., 2020. Derived from mda:mars by T. Hastie and R. Tibshirani. earth: Multivariate Adaptive Regression Splines.
- Millo, G., 2017. Robust Standard Error Estimators for Panel Models: A Unifying Approach. Journal of statistical software, 82.
- Munné-Bosch, S., 2015. Senescence: Is It Universal or Not? Trends Plant Sci., 20(11): 713-720.
- Munne-Bosch, S. and Alegre, L., 2004. Die and let live: leaf senescence contributes to plant survival under drought stress. Funct. Plant Biol., 31(3): 203-216.
- Murdoch, D.J. and Chow, E.D., 1996. A Graphical Display of Large Correlation Matrices. The American Statistician, 50(2): 178-180.
- Nagelkerke, N.J.D., 1991. A note on a general definition of the coefficient of determination. Biometrika, 78(3): 691-692.
- Olsen, J.E. et al., 1997. Ectopic expression of oat phytochrome A in hybrid aspen changes critical daylength for growth and prevents cold acclimatization. The Plant Journal, 12(6): 1339-1350.
- Panchen, Z.A. et al., 2015. Substantial variation in leaf senescence times among 1360 temperate woody plant species: implications for phenology and ecosystem processes. Ann. Bot., 116(6): 865-73.
- Parry, C., Blonquist Jr, J. and Bugbee, B., 2014. In situ measurement of leaf chlorophyll concentration: Analysis of the optical/absolute relationship. Plant, Cell Environ., 37.
- Penman, H.L., 1948. Natural evaporation from open water, hare soil and grass. Proc R Soc Lond A Math Phys Sci, 193(1032): 120-45.
- Pesaran, M.H., 2004. General Diagnostic Tests for Cross Section Dependence in Panels, Cambridge Working Papers in Economics. Faculty of Economics, University of Cambridge, Cambridge.
- Pesaran, M.H., 2014. Testing Weak Cross-Sectional Dependence in Large Panels. Econometric Rev., 34(6-10): 1089-1117.
- Piao, S. et al., 2019. Plant phenology and global climate change: Current progresses and challenges. Global Change Biol., 25(6): 1922-1940.
- Portillo-Estrada, M., Ariza-Carricondo, C. and Ceulemans, R., 2020. Outburst of senescence-related VOC emissions from a bioenergy poplar plantation. Plant Physiol. Biochem., 148: 324-332.
- Pregibon, D., 1980. Goodness of Link Tests for Generalized Linear Models. Journal of the Royal Statistical Society, Series C, 29(1): 15-24.
- Primka Iv, E. and Smith, W., 2019. Synchrony in fall leaf drop: chlorophyll degradation, color change, and abscission layer formation in three temperate deciduous tree species. Am. J. Bot., 106.

- R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Reiss, P. and Ogden, R., 2009. Smoothing parameter selection for a class of semiparametric linear models. Journal of the Royal Statistical Society Series B, 71: 505-523.
- Rigby, R. and Stasinopoulos, D., 2001. The GAMLSS project: a flexible approach to statistical modelling.
- Rigby, R., Stasinopoulos, D., Heller, G. and De Bastiani, F., 2019. Distributions for Modeling Location, Scale, and Shape: Using GAMLSS in R.
- Rigby, R.A. and Stasinopoulos, D.M., 1996a. A semi-parametric additive model for variance heterogeneity. Statistics and Computing, 6(1): 57-65.
- Rigby, R.A. and Stasinopoulos, D.M., 2005. Generalized additive models for location, scale and shape. Journal of the Royal Statistical Society Series C-Applied Statistics, 54(3): 507-544.
- Rigby, R.A. and Stasinopoulos, M.D., 1994. Robust Fitting of an Additive Model for Variance Heterogeneity. Compstat. Physica-Verlag HD, Heidelberg, pp. 263-268.
- Rigby, R.A. and Stasinopoulos, M.D., 1996b. Mean and Dispersion Additive Models. In: W. Härdle and M.G. Schimek (Editors), Statistical Theory and Computational Aspects of Smoothing. Physica-Verlag HD, Heidelberg, pp. 215-230.
- Ryo, M., Aguilar-Trigueros, C.A., Pinek, L., Muller, L.A.H. and Rillig, M.C., 2019. Basic Principles of Temporal Dynamics. Trends Ecol. Evol., 34(8): 723-733.
- Schippers, J., Schmidt, R., Wagstaff, C. and Jing, H.-C., 2015. Living to Die and Dying to Live: The Survival Strategy behind Leaf Senescence. Plant Physiol., 169.
- Sohrt, J., Herschbach, C. and Weiler, M., 2018. Foliar P- but not N resorption efficiency depends on the P-concentration and the N:P ratio in trees of temperate forests. Trees.
- Solomon, S., Plattner, G.-K., Knutti, R. and Friedlingstein, P., 2009. Irreversible climate change due to carbon dioxide emissions. Proceedings of the National Academy of Sciences, 106(6): 1704.
- Stasinopoulos, D. and Rigby, R., 2007. Generalized additive models for Location Scale and Shape (GAMLSS) in R. Journal of Statistical Software, 23.
- Stasinopoulos, D., Rigby, R., Heller, G., Voudouris, V. and De Bastiani, F., 2017. Flexible regression and smoothing: Using GAMLSS in R.
- Stasinopoulos, D.M. and Rigby, B., 2020. gamlss.dist: Distributions for Generalized Additive Models for Location Scale and Shape.
- Stasinopoulos, M.D., 2020. gamlss.add2: More Extra Additive Terms for GAMLSS Models.
- Stasinopoulos, M.D., Rigby, R.A. and Bastiani, F.D., 2018. GAMLSS: A distributional regression approach. Statistical Modelling, 18(3-4): 248-273.
- Statisticat, L., 2020. LaplacesDemon: Complete Environment for Bayesian Inference. Bayesian-Inference.com.
- Vander Mijnsbrugge, K. et al., 2016. Repeated Summer Drought and Re-watering during the First Growing Year of Oak (Quercus petraea) Delay Autumn Senescence and Bud Burst in the Following Spring. Front Plant Sci, 7(419): 419.
- Verbesselt, J., Hyndman, R., Zeileis, A. and Culvenor, D., 2010. Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. Remote Sens. Environ., 114: 2970-2980.
- Villar-Salvador, P., Uscola, M. and Jacobs, D., 2015. The role of stored carbohydrates and nitrogen in the growth and stress tolerance of planted forest trees. New Forests, 46: 813-839.
- Vitasse, Y. et al., 2011. Assessing the effects of climate change on the phenology of European temperate trees. Agricultural and Forest Meteorology, 151(7): 969-980.

- Vitasse, Y. et al., 2013. Elevational adaptation and plasticity in seedling phenology of temperate deciduous tree species. Oecologia, 171(3): 663-678.
- Vitasse, Y., Porte, A.J., Kremer, A., Michalet, R. and Delzon, S., 2009. Responses of canopy duration to temperature changes in four temperate tree species: relative contributions of spring and autumn leaf phenology. Oecologia, 161(1): 187-98.
- Voudouris, V., Rigby, R. and Stasinopoulos, D., 2013. Discussion: A comparison of GAMLSS with quantile regression. Statistical Modelling, 13: 335-348.
- Wei, T. and Simko, V., 2017. R package "corrplot": Visualization of a Correlation Matrix.
- Wendler, R. and Millard, P., 1996. Impacts of water and nitrogen supplies on the physiology, leaf demography and nitrogen dynamics of Betula pendula. Tree physiology, 16: 153-159.
- Wickham, H., 2009. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.
- Wickham, H., Francois, R., Henry, L. and Müller, K., 2018. dplyr: A Grammar of Data Manipulation.
- Wilke, C.O., 2019. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'.
- Wilke, O.C., 2020. ggridges: Ridgeline Plots in 'ggplot2'.
- Wingate, L. et al., 2015. Interpreting canopy development and physiology using a European phenology camera network at flux sites. Biogeosciences, 12(20): 5995-6015.
- Wood, S., 2017. Generalized additive Models. Chapman and Hall/CRC, New York.
- Wood, S.N., 2004. Stable and Efficient Multiple Smoothing Parameter Estimation for Generalized Additive Models. J. Amer. Statistical Assoc., 99(467): 673-686.
- Wood, S.N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J. Roy. Stat. Soc. Ser. B. (Stat. Method.), 73(1): 3-36.
- Wu, D.-M., 1973. Alternative Tests of Independence between Stochastic Regressors and Disturbances. Econometrica, 41(4): 733-750.
- Xie, Q. et al., 2018a. Vegetation Indices Combining the Red and Red-Edge Spectral Information for Leaf Area Index Retrieval. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(5): 1482-1493.
- Xie, Y., Wang, X., Wilson, A.M. and Silander, J.A., 2018b. Predicting autumn phenology: How deciduous tree species respond to weather stressors. Agricultural and Forest Meteorology, 250-251: 127-137.
- Xie, Y. and Wilson, A.M., 2020. Change point estimation of deciduous forest land surface phenology. Remote Sens. Environ., 240: 111698.
- Xu, H., Twine, T. and Yang, X., 2014. Evaluating Remotely Sensed Phenological Metrics in a Dynamic Ecosystem Model, 6, 4660-4686 pp.
- Zhang, X. and Goldberg, M.D., 2011. Monitoring fall foliage coloration dynamics using time-series satellite data. Remote Sens. Environ., 115(2): 382-391.
- Zhang, X.Y. et al., 2003. Monitoring vegetation phenology using MODIS. Remote Sens. Environ., 84(3): 471-475.
- Zhao, K. et al., 2019. Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm. Remote Sens. Environ., 232: 111181.
- Zuur, A., Ieno, E. and Smith, G., 2007. Analysing Ecological Data. Statistics for Biology and Health, 75. Springer, New York, XXVI, 672 pp.

Figures

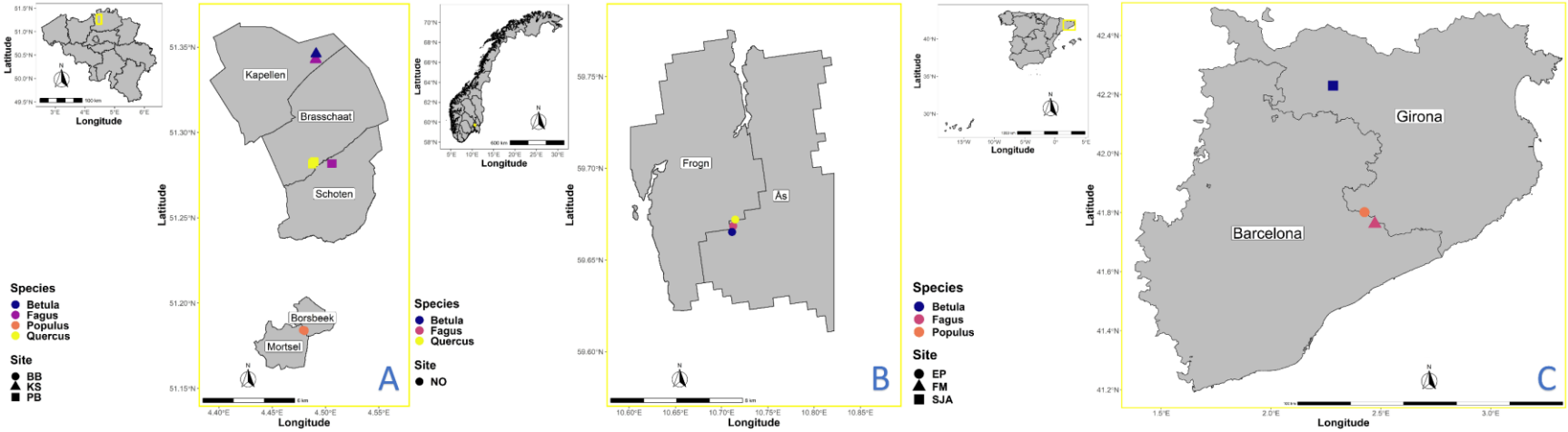
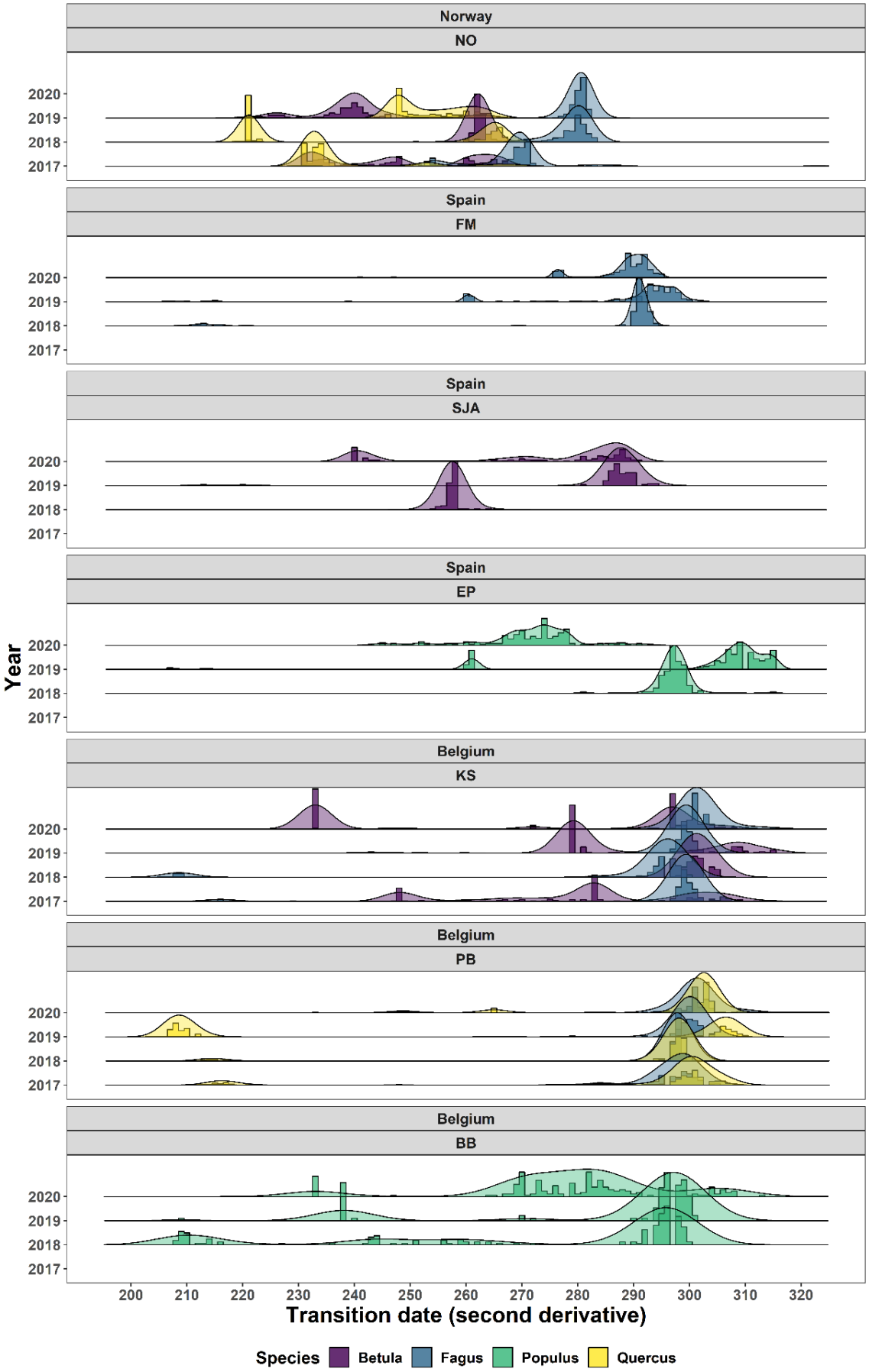


Fig. 1: The location of each tree within the different municipalities in the province of Antwerp (Belgium; A), the county of Viken (Norway; B) and the provinces of Barcelona and Girona (Spain; C; yellow rectangles). The colored dots represent the location of each individual *Betula pendula* (blue), *Fagus sylvatica* (purple), *Populus tremula* (red) and *Quercus robur* (yellow) tree. The shape of each dot shows the site where each tree is located. In Belgium, BB stands for the Fortress of Borsbeek (circle), KS stands for the Klein Schietveld (triangle) and PB stands for the Park of Brasschaat (square). In Norway, NO stands for the farm of Hoxmark (circle). In Spain, EP stands for El Puig (circle), SJA stands for Sant Joan de les Abadesses (square) and FM stands for Fogars de Montclús (triangle). Municipality, county or province names are printed in Bold.

Fig. 2: The chlorophyll content index (CCI) of the mature *Fagus sylvatica* (n_{KS} = 4, n_{PB}= 4, n_{FM} = 4 and n_{NO} = 3), *Quercus robur* (n_{PB} = 4 and n_{NO} = 3), *Betula pendula* (n_{KS} = 4, n_{SJA} = 3 and n_{NO} = 3) and *Populus tremula* (n_{BB} = 4 and n_{EP} = 2) trees at the sampling sites of the Klein Schietveld (KS), Park of Brasschaat (PB) and Fortress of Borsbeek (BB) in Belgium, El Puig (EP), Sant Joan de les Abadesses (SJA) and Fogars de Montclús (FM) in Spain, and the farm of Hoxmark (NO) in Norway. Dots and error bars represent the mean CCI with standard errors in the years 2017 (blue), 2018 (purple), 2019 (green) and 2020 (yellow).


Norway
NO
Spain
FM
Spain
SJA
Spain
EP
Belgium
KS
Belgium
PB
Belgium
BB
Year
2020
2019
2018
2017
200
210
220
230
240
250
260
270
280
290
300
310
320
Transition date (second derivative)
Species
Betula
Fagus
Populus
Quercus

998

Fig. 3: Histograms and density plots of the resampled values showing the distribution of the second derivative's transition dates for the mature *Fagus sylvatica* (n_{KS} = 4, n_{PB}= 4, n_{FM} = 4 and n_{NO} = 3; teal), *Quercus robur* (n_{PB} = 4 and n_{NO} = 3; yellow), *Betula pendula* (n_{KS} = 4, n_{SJA} = 3 and n_{NO} = 3; purple) and *Populus tremula* (n_{BB} = 4 and n_{EP} = 2; light green) trees at the sampling sites of the Klein Schietveld (KS), Park of Brasschaat (PB) and Fortress of Borsbeek (BB) in Belgium, El Puig (EP), Sant Joan de les Abadesses (SJA) and Fogars de Montclús (FM) in Spain, and the farm of Hoxmark (NO) in Norway from 2017 to 2020. The second derivative's transition date represent here (I) the moment when the acceleration in the decline of the chlorophyll content (CCI) graphs was the highest or (II) a local minima in the second derivative of any of the CCI graphs. As the curvature (κ) of the CCI graphs was approximated here by the second derivative, we can also interprete the second derviative's transition date as (III) the moment when the change in curvature of any of the CCI graphs was the highest or (IV) the moment when the declination of the CCI graphs concavity was highest. The density plots therefore indicate here the range of dates that constitute the variation surrounding the actual timing on the onset of senescence (defined as the moment when the acceleration in the CCI's decline occurs most rapidly, or the curvature of the CCI's graph is highest). A priori, these plots best characterize the process behind the onset of senescence.

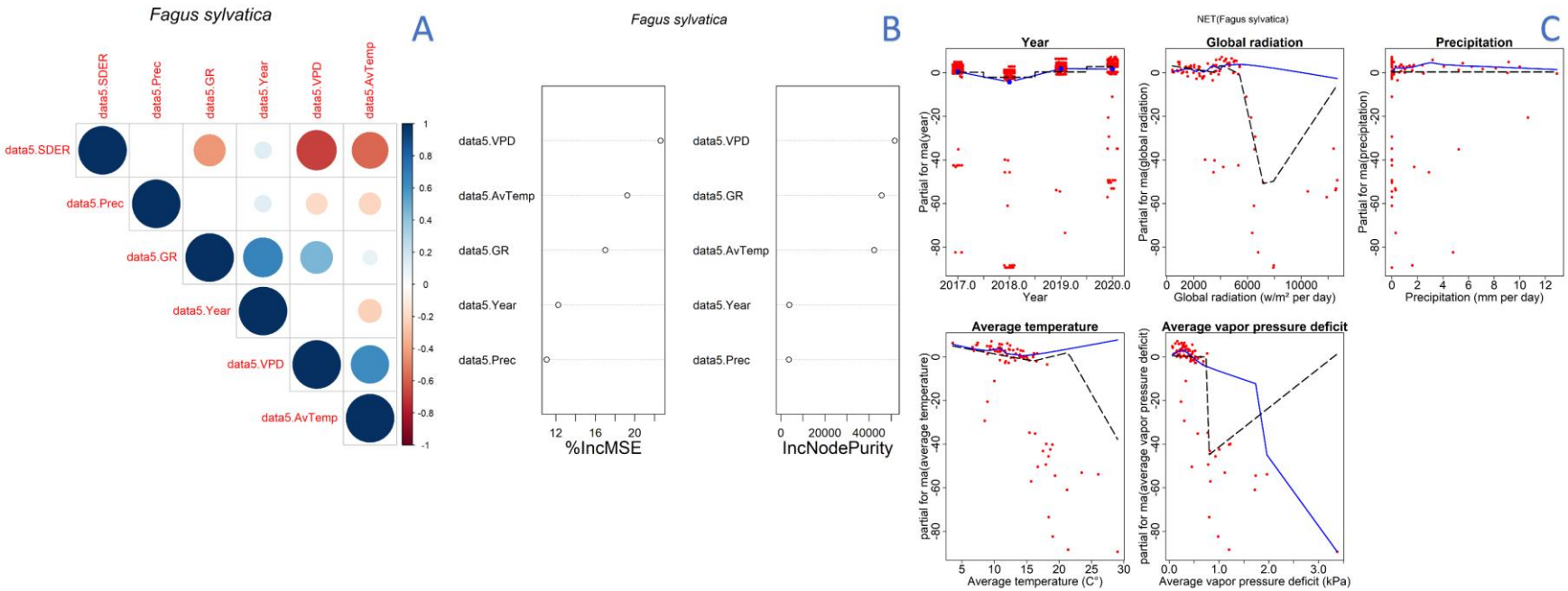


Fig. 4: Assessment of the effects that five environmental variables can have on the second derivative's transition date (SDER) in mature *Fagus sylvatica* (n_{KS} = 8) trees at the sampling sites of the Klein Schietveld and Park of Brasschaat from 2017 to 2019. The five potential environmental drivers are the daily average temperature (C°; AvTemp), the daily sum of precipitation (mm; Prec), the daily sum of global radiation (w/m²; GR) between 7 a.m. to 7 p.m., the daily average vapor pressure deficit (kPa; VPD) between 7 a.m. to 7 p.m. and the year (Year). The year represents here both time and any other potential time-dependent factor (e.g. herbivory effects). Panel A shows correlation matrices indicating Pearson's correlation coefficient. Panel B shows the result of random forest models indicating the percentage increase in the mean square error rate (%IncMSE) and the increase in node purity (IncNodePurity). Panel C shows the response of GAMLSS models with multivariate adaptive regression splines (MARS) that performed enhanced adaptive regressions through hinges (EARTH). Alternatively, panel C shows poor man's partial dependence plots that indicate the model's response when one predictor variable is varying while the other predictor variables medians are kept constant. Red dots indicate the response points, while the black and blue dashed lines indicate nonlinear functions through hinges and smooth lines, respectively. The effect of the year and precipitation is not considered significant for *Fagus sylvatica*.

Tables

Table 1: Overview of the mean, median, mode, standard deviation and standard errors (SE) of the second derivative's transition dates' distributions for the mature *Fagus sylvatica* (n_{KS} = 4, n_{PB} = 4, n_{FM} = 4 and n_{NO} = 3), *Quercus robur* (n_{PB} = 4 and n_{NO} = 3), *Betula pendula* (n_{KS} = 4, n_{SJA} = 3 and n_{NO} = 3) and *Populus tremula* (n_{BB} = 4 and n_{EP} = 2) trees at the sampling sites of the Klein Schietveld (KS), Park of Brasschaat (PB) and Fortress of Borsbeek (BB) in Belgium, El Puig (EP), Sant Joan de les Abadesses (SJA) and Fogars de Montclús (FM) in Spain, and the farm of Hoxmark (NO) in Norway from 2017 to 2020.

Species	Site	Year	Mean (Doy)	Mean (Date)	SE Mean (Days)	Median (Doy)	Median (Date)	SE Median (Days)	Standard deviation (Days)	Mode (Doy)	Mode (Date)
Betula pendula	KS	2017	281	8 October 2017	2	283	10 October 2017	2	20	283	10 October 2017
Betula pendula	KS	2018	300	27 October 2018	1	301	28 October 2018	1	11	301	28 October 2018
Betula pendula	KS	2019	288	15 October 2019	1	279	6 October 2019	1	15	279	6 October 2019
Betula pendula	KS	2020	267	23 September 2020	2	287	13 October 2020	3	32	233	20 August 2020
Betula pendula	NO	2017	248	5 September 2017	1	248	5 September 2017	1	13	231	19 August 2017
Betula pendula	NO	2018	262	19 September 2018	0	262	19 September 2018	0	1	262	19 September 2018
Betula pendula	NO	2019	240	28 August 2019	1	240	28 August 2019	1	9	240	28 August 2019
Betula pendula	SJA	2018	258	15 September 2018	0	258	15 September 2018	0	1	258	15 September 2018
Betula pendula	SJA	2019	286	13 October 2019	1	287	14 October 2019	1	11	287	14 October 2019
Betula pendula	SJA	2020	272	28 September 2020	1	281	7 October 2020	2	19	240	27 August 2020
Fagus sylvatica	FM	2018	286	13 October 2018	1	291	18 October 2018	2	20	291	18 October 2018
Fagus sylvatica	FM	2019	287	14 October 2019	1	294	21 October 2019	2	19	293	20 October 2019
Fagus sylvatica	FM	2020	288	14 October 2020	0	290	16 October 2020	1	7	291	17 October 2020
Fagus sylvatica	KS	2017	297	24 October 2017	1	299	26 October 2017	1	15	299	26 October 2017
Fagus sylvatica	KS	2018	288	15 October 2018	2	295	22 October 2018	2	25	295	22 October 2018
Fagus sylvatica	KS	2019	299	26 October 2019	0	299	26 October 2019	0	4	299	26 October 2019
Fagus sylvatica	KS	2020	300	26 October 2020	1	301	27 October 2020	1	11	301	27 October 2020
Fagus sylvatica	NO	2017	268	25 September 2017	1	270	27 September 2017	1	7	269	26 September 2017
Fagus sylvatica	NO	2018	278	5 October 2018	1	280	7 October 2018	1	9	280	7 October 2018
Fagus sylvatica	NO	2019	279	6 October 2019	1	280	7 October 2019	1	10	281	8 October 2019
Fagus sylvatica	PB	2017	296	23 October 2017	1	299	26 October 2017	1	11	299	26 October 2017
Fagus sylvatica	PB	2018	298	25 October 2018	0	298	25 October 2018	0	1	298	25 October 2018
Fagus sylvatica	PB	2019	299	26 October 2019	1	300	27 October 2019	1	9	300	27 October 2019
Fagus sylvatica	PB	2020	299	25 October 2020	1	301	27 October 2020	1	12	301	27 October 2020
Populus tremula	BB	2018	273	30 September 2018	3	294	21 October 2018	3	33	296	23 October 2018
Populus tremula	BB	2019	286	13 October 2019	2	296	23 October 2019	3	23	296	23 October 2019
Populus tremula	BB	2020	279	5 October 2020	1	281	7 October 2020	2	18	270	26 September 2020
Populus tremula	EP	2018	297	24 October 2018	0	297	24 October 2018	0	4	297	24 October 2018
Populus tremula	EP	2019	301	28 October 2019	2	309	5 November 2019	2	21	309	5 November 2019
Populus tremula	EP	2020	271	27 September 2020	1	273	29 September 2020	1	9	278	4 October 2020
Quercus robur	NO	2017	237	25 August 2017	1	234	22 August 2017	1	10	231	19 August 2017
Quercus robur	NO	2018	242	30 August 2018	2	223	11 August 2018	2	22	221	9 August 2018
Quercus robur	NO	2019	254	11 September 2019	0	252	9 September 2019	1	7	248	5 September 2019
Quercus robur	PB	2017	291	18 October 2017	2	300	27 October 2017	3	26	301	28 October 2017
Quercus robur	PB	2018	293	20 October 2018	2	298	25 October 2018	2	19	298	25 October 2018
Quercus robur	PB	2019	257	14 September 2019	4	240	28 August 2019	5	49	208	27 July 2019
Quercus robur	PB	2020	299	25 October 2020	1	303	29 October 2020	1	12	303	29 October 2020

Table 2: Overview of the variation (i.e. R^2) in the transition dates explained by each variable in the Random Forest models for *Fagus sylvatica*, *Quercus robur*, *Betula pendula and Populus tremula*. %IncMSE and IncNodePurity stand for the percentage increase in the mean square error and the node purity of the predictions, respectively.

	Fagus sylvatica		***Quercus robur***		***Betula pendula***		***Populus tremula***	
	%IncMSE	**IncNodePurity**	**%IncMSE**	**IncNodePurity**	**%IncMSE**	**IncNodePurity**	**%IncMSE**	**IncNodePurity**
Year	12	3816	17	22215	22	18372	23	7588
Average temperature	19	42270	27	145023	21	57090	25	47115
Global radiation	17	45654	21	85850	21	46866	25	35847
Vapor pressure deficit	23	51638	19	75337	21	55008	24	43001
Precipitation	11	3632	18	38492	18	23974	19	8148
% Total variance explained	92	NA	98	NA	97	NA	89	NA

Table 3: Overview of the GAMLSS model response indicators for the fitted GAMLSS models with multivariate adaptive regression splines for *Fagus sylvatica*, *Quercus robur*, *Betula pendula* and *Populus tremula* (Models S3 – S5). Family represent the normal-exponential-student-*t* (NET), generalized gamma (GG) and zero-inflated Box-Cox Cole and Green (BCCGo) distribution families that are used to model the response variable. The distribution's intercept coefficients for the mean (μ), variation (σ), skewness (ν) and kurtosis (τ) distribution parameters are also given. Furthermore we represent the generalized R^2 of Nagelkerke, a summary of the quantile residual values, the degrees of freedom (Df) of the model fit, the degrees of freedom of the residuals, the global deviance, Akaike Information Criterion (AIC) and normalized root mean square error (RMSE). Indicative for a 'good' model fit is a mean, variance, skewness, kurtosis and Filliben correlation coefficient of 0,1,0,3 and 1, respectively (Hohberg et al., 2020)

Species	Family	Algorithm	R^2 - Cox Snell	R^2 - Cragg Uhler	Summary of the Quantile Residuals					Distribution's intercept coef.				Df fit	Df residuals	Global Deviance	AIC	Normalized RMSE
					Mean	Variance	Coef. of skewness	Coef. of kurtosis	Filliben correlation coef.	μ	σ	ν	τ					
Fagus sylvatica	NET	RS	0.83	0.83	-0.0879	1.4681	-0.2429	3.8768	0.9835	297.4	-0.291	NA	NA	17	1559	5809	5843	0.01
Quercus robur	GG	RS	0.93	0.93	-0.0070	1.0060	0.1860	13.5411	0.9086	5.639	-4.057	-31.81	NA	18	592	3729	3765	0.009
Betula pendula	GG	RS	0.93	0.93	-0.0001	1.0013	-0.0192	23.7295	0.8250	5.64	-4.03	-8.296	NA	18	658	4101	4137	0.009
Populus tremula	BCCGo	RS	0.92	0.92	-0.0034	1.0260	-0.0193	5.2149	0.9756	5.628	-3.977	20.24	NA	19	442	2902	2940	0.01