Aging and stress explain the earlier start of leaf senescence in trees in warmer years: translating the latest findings on senescence regulation into the DP3 model (v1.1)

5 Michael Meier*[1,2], Christof Bigler[3], Isabelle Chuine[1]

[1] CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

[2] Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland

[3] Forest Ecology, Department of Environmental Systems Science, ETH Zurich, Zurich, Switzerland

10 *Correspondence to*: Michael Meier, (michael.meier.1@unil.ch)

Short summary. The DP3 model of leaf coloring, formulated according to the leaf development process, considerably contrasts previous models and allows to set up new hypotheses, e.g., regarding earlier onset and longer duration of senescence predicted for warmer conditions. Comparing the accuracy of the DP3 model to that of previous models and 15 the Null model (average observed date of leaf coloring) suggests that leaf coloring data are noisy, which is why observation protocols and methods should be revised.

Abstract. Leaf senescence ends the growing season of deciduous trees, affecting the amount of atmospheric CO_2 sequestered by forests. Therefore, some climate models integrate projected leaf senescence dates to simulate the carbon 20 cycle. Here, we developed a process-oriented model of leaf senescence (the 'DP3 model') by testing 34 formulations of the leaf development process based on the latest findings on the regulation of leaf aging and senescence. The period between leaf unfolding and leaf senescence was separated into the subsequent young, mature, and old leaf phases, with particular reactions to leaf aging and cold stress, photoperiod stress, and dry stress. The DP3 model simulates daily rates of aging and stress to predict dates of transition from young to mature to old leaf, senescence induction dates, and leaf 25 senescence dates. This allows new hypotheses regarding the regulation of leaf senescence to be tested. For example, the DP3 model predicted an earlier onset of senescence in warmer conditions, likely due to earlier leaf unfolding (aging) and increased cold and dry stress in spring, together with longer-lasting senescence, likely due to the later accumulation of photoperiod stress relative to leaf development and decreased cold stress in summer and fall, which can be validated through experiments and in situ observations. The DP3 model and compared previous models were equally accurate, 30 but less accurate than the Null model (average senescence date observed in the calibration sample). This lower accuracy of the DP3 and compared models is likely due to noise in the visually observed leaf senescence data, which blurs the signal of the leaf senescence process, and to incorrect model formulations. The model errors were similarly affected by climate conditions and location among compared models (including the Null model) and varied mostly due to the leaf senescence data. Noisy leaf senescence data likely force the models to resort to the mean observation, impeding infer-35 ences from accuracy-based model comparisons about the leaf senescence process. This calls for revised observation protocols and methods that measure rather than estimate different senescence stages, such as senescence induction and 50% of the leaves have changed color, e.g., based on greenness, involving digital cameras and automated image assessment.

40 **Keywords:** Leaf phenology, process-based model, senescence induction, leaf coloring, leaf fall, phenological data, observation protocols, observer bias, sample bias, bias towards the mean

1   Introduction

Leaf senescence involves several processes and regulation pathways, but the most important process is the degradation of chlorophyll and breakdown of chloroplasts to retrieve nutrients, especially nitrogen, and to mobilize them in new leaves in spring (Cooke and Weih, 2005; Keskitalo et al., 2005; Lim et al., 2007; Rogers, 2017). A side effect of this nu-

45 trient resorption is the change in leaf color from green to yellow, orange, or red (Keskitalo et al., 2005; but see Wheeler and Dietze, 2023). There have been many studies of how the timing of leaf coloring is influenced by climatic conditions (e.g., Bigler and Vitasse, 2021; Liu et al., 2018; Meier et al., 2021). As these studies usually used the term 'leaf coloring' or 'leaf senescence' to refer to a particular stage of leaf senescence, we use 'leaf senescence' as a collective term for the stages when a given relative amount of leaves have changed color or have fallen, unless stated otherwise.

50 Leaf senescence marks the end of a process that has been better understood over the last ten years, mainly thanks to studies in cell and molecular biology and in environmental sciences. These studies have shown that leaf senescence relates to leaf development state (e.g., Jan et al., 2019; Jibran et al., 2013; Lim et al., 2007). On the one hand, the development state of leaves depends on their age and thus on the time since leaf unfolding and the state of carbohydrate sinks (Jibran et al., 2013), which relates to photosynthetic activity and nutrient availability (Paul and Foyer, 2001).

55 While earlier leaf unfolding was related to earlier leaf senescence (Fu et al., 2014, 2019), an intense discussion has started about the possibility of earlier leaf senescence due to increased photosynthetic activity (Kloos et al., 2024; Lu and Keenan, 2022; Marqués et al., 2023; Norby, 2021; Zohner et al., 2023). On the other hand, the development state of leaves is influenced by hormone levels (Addicott, 1968; Jan et al., 2019; Jibran et al., 2013; Lim et al., 2007), which are, among others, stimulated by environmental stress caused by cold (Kloos et al., 2024; Wang et al., 2022; Xie et al.,

60 2015, 2018), drought (Bigler and Vitasse, 2021; Mariën et al., 2021; Tan et al., 2023; but see Kloos et al., 2024; Xie et al., 2015, 2018), heat (Bigler and Vitasse, 2021; Mariën et al., 2021; Tan et al., 2023; Xie et al., 2015, 2018), heavy rain (Kloos et al., 2024; Xie et al., 2015, 2018), short days (Addicott, 1968; Keskitalo et al., 2005; Singh et al., 2017; Tan et al., 2023; Wang et al., 2022), and lack of nutrients (Fu et al., 2019; Tan et al., 2023). In the early phase of leaf development ('young leaf'), senescence cannot be induced, whereas aging and stress induce it in later phases ('mature leaf' and

65 'old leaf') and regulate the rate of senescence (Fig. 1; Jan et al., 2019; Jibran et al., 2013; Lim et al., 2007; Paul and Foyer, 2001; Tan et al., 2023).
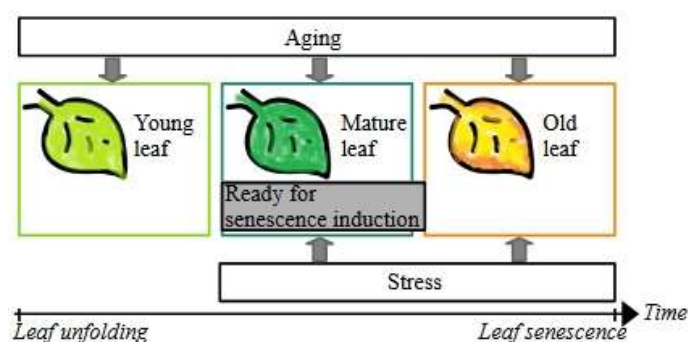
Figure 1. Leaf development. Starting with leaf unfolding, the young leaf develops first into a mature leaf and then into an old leaf. During the three phases of the young, mature, and old leaf, the leaf ages continuously. With the transition from young to mature leaf,

the leaf becomes ready for senescence induction through environmental stress (e.g., cold days). If senescence is not induced through stress in the mature leaf, it certainly is through aging with the transition from mature to old leaf. Thus, senescence cannot be induced during the phase of the young leaf and the onset of senescence (i.e., the senescence induction date) depends on the time since leaf unfolding and on the environmental conditions since the transition from young to mature leaf. Adapted from Figure 1 in Jibran et al. (2013).

As senescence induction depends on environmental conditions, leaf senescence of deciduous trees shifts as climate changes, which influences the timing and length of their growing season and thus affects the amount of CO_2 absorbed from the atmosphere (Meier et al., 2021; Menzel et al., 2020; Piao et al., 2019; but see Mariën et al., 2021). This links the feedback loop between atmospheric CO_2 concentration and climate to the feedback loop between climate and forests and more generally to terrestrial ecosystems (Luo, 2007; Richardson et al., 2013). Further, the amount of absorbed CO_2 relates to the amount of sugars available for tree growth, defense, and reproduction (Herms and Mattson, 1992; Tan et al., 2023). Therefore, accurate projections of leaf senescence dates under a changing climate are necessary for accurate forecasts of both climate change and future species composition of temperate forests.

Leaf senescence dates are often projected using process-oriented models. These models are usually based on the results of experiments testing the effect of various environmental cues, that are translated mathematically (Chuine et al., 2013; Chuine and Régnière, 2017). Various process-oriented models of leaf senescence have been proposed over the last twenty years (Liu et al., 2020; Meier and Bigler, 2023). They generally formulate leaf senescence as a one-way process that starts shortly after summer solstice by accumulating a daily rate of senescence until a threshold is reached (but see Wheeler and Dietze, 2023). The daily rate is usually dependent on temperature and day length, and the threshold is either a constant or depends on leaf unfolding dates or on environmental conditions during the growing season (e.g., Delpierre et al., 2009; Keenan and Richardson, 2015; Liu et al., 2019; Zani et al., 2020).

Previous studies have shown that these leaf senescence models are heavily biased towards the mean of the calibration sample (Meier et al., 2023) and are less efficient relatively to leaf unfolding models (e.g., Liu et al., 2020; Meier and Bigler, 2023). However, it is not yet clear whether this is due to noisy phenological data and/or an incomplete process formulation.

The phenological data used to train leaf senescence models have often been recordings of visual observations, which cover long time periods and are species-specific (e.g., ongoing since 1951 in the Swiss phenology network, 2025). However, the observations are noisy due to different observers and small sample sizes. For leaf senescence, Liu et al. (2021) showed for example that the observer bias was 15 days (d; median) and the sampling bias was 10 d (median) for 10 trees observed per population. These biases not only lead to noise between sites, but also within sites when observers and samples change. Such changes can lead to sudden changes in the mean in the time series, as was found for some Swiss sites (Auchmann et al., 2018; Swiss phenology network, 2025). Moreover, the observation protocols may differ between the meteorological institutes and citizen science based networks that are responsible for the recording in the different European countries (Menzel, 2013).

Current models formulate leaf senescence as the result of an accumulated stress caused by cold and short days after summer solstice (Delpierre et al., 2009; Dufrêne et al., 2005; Keenan and Richardson, 2015; Lang et al., 2019; Liu et al., 2019; Zani et al., 2020). Two models further consider environmental conditions before summer solstice, either through temperature and precipitation during the growing season (Liu et al., 2019) or through photosynthetic activity during the growing season (Zani et al., 2020), while one model considers age through leaf unfolding dates (Keenan and Richardson, 2015). However, in these models, environmental conditions and age affect the amount of

stress needed for leaf senescence rather than senescence induction. In other words, according to current models, senescence induction depends only on stress caused by cold and short days after summer solstice, which considerably contrasts with current knowledge (see above). None of the current models allows for senescence induction due to aging or stress before summer solstice. While two models consider stress that occurred before summer solstice, senescence is always induced after summer solstice. Finally, we are unaware if aging, stress caused by other than cold and short days, and different stress effects among the phases of leaf development have been tested, as none of the corresponding studies mentioned tested but discarded model formulations.

Here, we developed a new process-oriented model that predicts leaf senescence dates based on the latest knowledge of the physiological processes and drivers of leaf senescence. Leaf senescence was formulated through a leaf development process that starts at leaf unfolding and is driven by aging and various types of abiotic stress. We tested 34 model formulations of this process. Finally, the most accurate formulation was evaluated with a particular focus on the differences between the predicted and observed dates (i.e., 'model errors'). We addressed the following research questions:

(1) Which model formulation yields the most accurate predictions of leaf senescence dates?

(2) How accurately does this model predict leaf senescence dates compared to previous models?

(3) How do the model errors relate to the phenological data, climate, and site conditions?

2   Data and methods

2.1   Phenological data

The model was developed and evaluated with leaf phenology data of common beech (*Fagus sylvatica*, L.), which was visually observed in Austria, Germany, Great Britain, and Switzerland between 1950 and 2022 (Fig. 2, Table 1, Sect. S1.1; PEP725, 2024; Swiss phenology network, 2025; Templ et al., 2018). We used the phenological stages 50% of the leaves have unfolded as well as 50% and 100% of the leaves have changed color or have fallen (hereafter referred to as 'leaf unfolding' [LU], 'leaf senescence$_{50}$' [LS$_{50}$], and 'leaf senescence$_{100}$' [LS$_{100}$], respectively; corresponding to BBCH15, BBCH95, and BBCH97 according to Meier, 2018). The LS$_{100}$ data were recorded in Austria and Great Britain only.

We checked all site-years with regards to the order and completeness of the phenological observations. Observations of LS$_{50}$ and LS$_{100}$ that occurred between the day of year (doy) 60 and 151 were discarded, as were observations of LU that occurred after doy 180 or after LS$_{50}$ or LS$_{100}$. Thus, we considered only site-years with an observation for LU that was followed by either LS$_{50}$ or LS$_{100}$, or by both LS$_{50}$ and later LS$_{100}$, leaving 5018 sites.

From these sites, we made a pre-selection so that the phenological and geographical range of the LS$_{50}$ observations was evenly covered and all LS$_{100}$ observations were included. This involved splitting all 5018 sites into 8–10 bins with equal spans for the average and standard deviation of LS$_{50}$ as well as for latitude, longitude, and elevation, so that each bin contained at least two sites (e.g., the range between doy 232 and 328 for the average LS$_{50}$ was split into ten bins of 9.7 days). From each bin, we chose the site with the most LS$_{50}$ observations, with random choice if this applied to more than one site. These sites were completed by all sites with an LS$_{100}$ observation, resulting in a pre-selection of 7137 LS$_{50}$ and 850 LS$_{100}$ observations recorded at 244 and 106 sites, respectively.
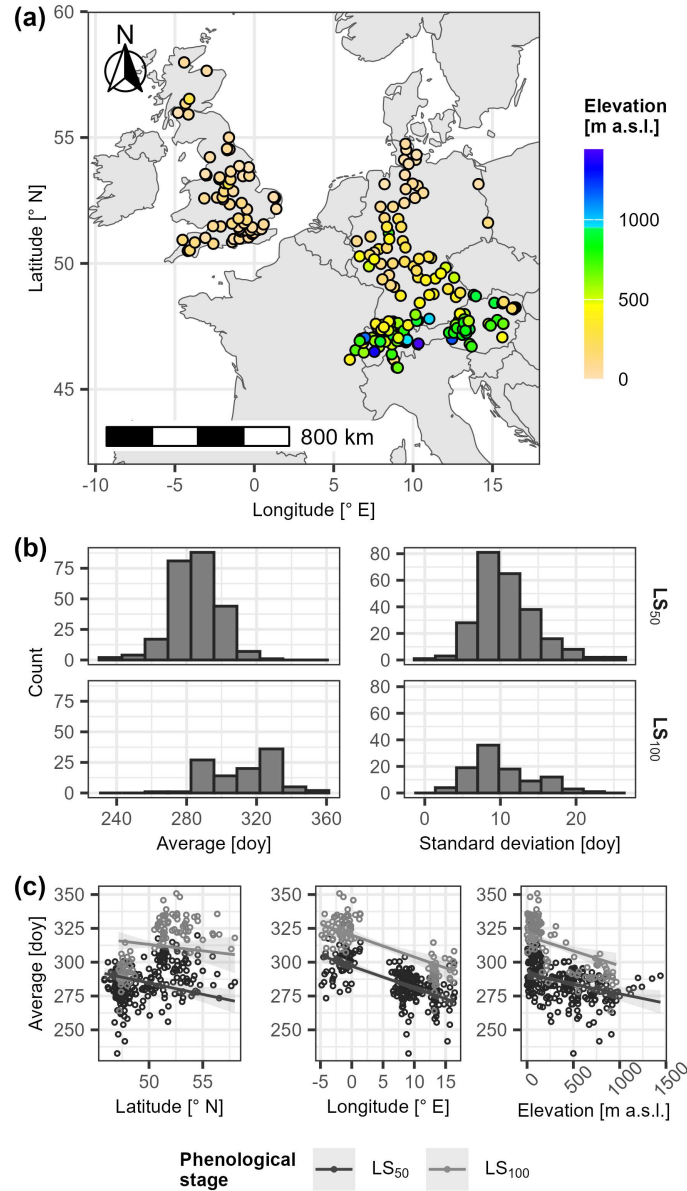
Figure 2. Selected phenological sites. Panel (a) locates the selected sites and indicates corresponding elevation [meters above sea level (m a.s.l.)]. In (b), the histograms illustrate the distributions of the site-specific average day of year (doy; left) and corresponding standard deviation (right) per leaf senescence stage [i.e., 50% and 100% of the leaves have changed color or have fallen (LS_{50} and LS_{100}, respectively); rows]. Panel (c) plots the site-specific average doy of LS_{50} and LS_{100} (grey and black circles, respectively) in relation to site latitude [°N] (left), longitude [°E] (middle), and elevation [m a.s.l.] (right), together with the linear regression lines and corresponding 99% confidence intervals. The linear regression explained site-specific average LS_{50} and average LS_{100} by latitude, longitude, and elevation (Sect. S1.1.2). Corresponding estimates were plotted against latitude for mean longitude and mean elevation (left), against longitude for mean latitude and mean elevation (middle), and against elevation for mean latitude and mean longitude.

Table 1. Observations of spring and fall leaf phenology.

Stage	Country	Sites	Total number of site-years (min.–max. per site)	Observation period	Range of observations [doy]	Source
LS_{50}	Austria	51	1011 (5–54)	1950–2015	209–321	PEP725
	Germany	68	3238 (14–65)	1951–2015	196–331	PEP725
	Great Britain	64	303 (2–6)	1999–2005	258–337	PEP725
	Switzerland	61	2585 (6–72)	1951–2022	197–344	SPN
LS_{100}	Austria	43	578 (1–34)	1950–1986	263–335	PEP725
	Great Britain	63	272 (1–6)	1999–2005	286–365	PEP725

LU	Austria	51	1020 (5–54)	1950–2015	80–166	PEP725
	Germany	68	3238 (14–65)	1951–2015	80–175	PEP725
	Great Britain	64	331 (5–6)	1999–2005	85–140	PEP725
	Switzerland	61	2585 (6–72)	1951–2022	67–161	SPN

Note: LU refers to the stage when 50% of the leaves are unfolded. LS_{50} and LS_{100} refer to the stages when 50% and 100% of the leaves, respectively, have changed color or have fallen. The timing of these stages is given by the day of year (doy). A site-year is a year for which an observation of both LU and LS_{50} or LS_{100} was recorded at a given site. Two data sources were considered: PEP725 (Templ et al., 2018) and the Swiss phenological network (SPN; Swiss phenology network, 2025).

2.2   Driver data

140   For each phenological site, weather variables, elevation, and the leaf area index (LAI) were approximated by the weighted averages from octagons with a radius of 2.5 km around the phenological sites, and combined with the atmospheric CO_2 concentration. Daily weather variables and elevation were derived for each site from the E-OBS dataset (Copernicus Climate Change Service, Climate Data Store, 2020; Cornes et al., 2018), which contains interpolated data from a 100-member ensemble driven with meteorological observations. We extracted and approximated site elevation,

145   maximum temperature, mean temperature, minimum temperature, precipitation, relative humidity, and surface short-wave down welling radiation for 1950–2022. These temperature variables were corrected through day- and site-specific lapse rates to account for elevational differences between the octagon averages and sites (i.e., the elevation according to the phenology datasets or, if missing, according to EU-DEM, 2024, with a resolution of 25 m, and the location according to the phenology datasets). These laps rates were linearly regressed from the grid cell of a particular site and the

150   eight neighboring grid cells, assuming an elevation of 0 meters above sea level (m a.s.l.) for grid cells over the sea. Occasional gaps in the regressed lapse rates were interpolated with site-specific cubic splines. LAI per site was taken from the remote sensed monthly LAI (1981–2015) in the GIMMS-LAI3g dataset (version 2; Mao and Yan, 2019). LAI is averaged among years in this dataset, and thus we also used these monthly LAI values for the years 1950–1980. Atmospheric CO_2 concentrations were taken from a reconstructed dataset for the years 1950–2013 and a remote sensed

155   dataset for the years 2002–2022 (Cheng et al., 2022; Copernicus Climate Change Service, Climate Data Store, 2018). Both datasets provide monthly data, which we distilled into annual averages. These averages were combined through weighted means over the years 2002–2013 to assure a smooth transition between the datasets. As some monthly CO_2 observations between 2002–2022 were missing, we used modeled CO_2 values derived from site-specific cubic splines based on the remote sensed data (Copernicus Climate Change Service, Climate Data Store, 2018).

160   We further calculated for each site day length, daily photosynthetic activity, and the daily Keetch and Byram drought index (KBDI). Day length was calculated following Brock (1981), using the latitude of each site (Sect. S1.2.1). Daily sink limited photosynthetic activity was calculated following Farquhar et al. (1980) and Collatz et al. (1991), using daily surface shortwave down welling radiation, day length, and mean temperature together with monthly LAI averaged among years and annual atmospheric CO_2 concentration (Sect. S1.2.2). The daily KBDI was calculated following

165   Keetch and Byram (1968), using precipitation and maximum temperature (Sect. S1.2.3).

2.3   Model conceptualization

Based on the process of leaf development according to Jibran et al. (2013), we defined our model as a one-way process that may be formulated with either two or three phases of leaf development, namely either the phases mature and old leaf or the phases young, mature, and old leaf (Figs. 1 and 3). After leaf unfolding, the young leaf is insensitive to stress

170 and ages until it becomes a mature leaf (Fu et al., 2014; Jibran et al., 2013; Keenan and Richardson, 2015). The mature leaf can be affected by stress and ages until it becomes an old leaf (Jan et al., 2019; Jibran et al., 2013; Lim et al., 2007). Senescence induction may be caused by stress in the mature leaf or by aging, in the case of which it coincides with the transition from mature to old leaf, causing the leaf to change color and to fall off (Jan et al., 2019; Jibran et al., 2013; Lim et al., 2007).

175
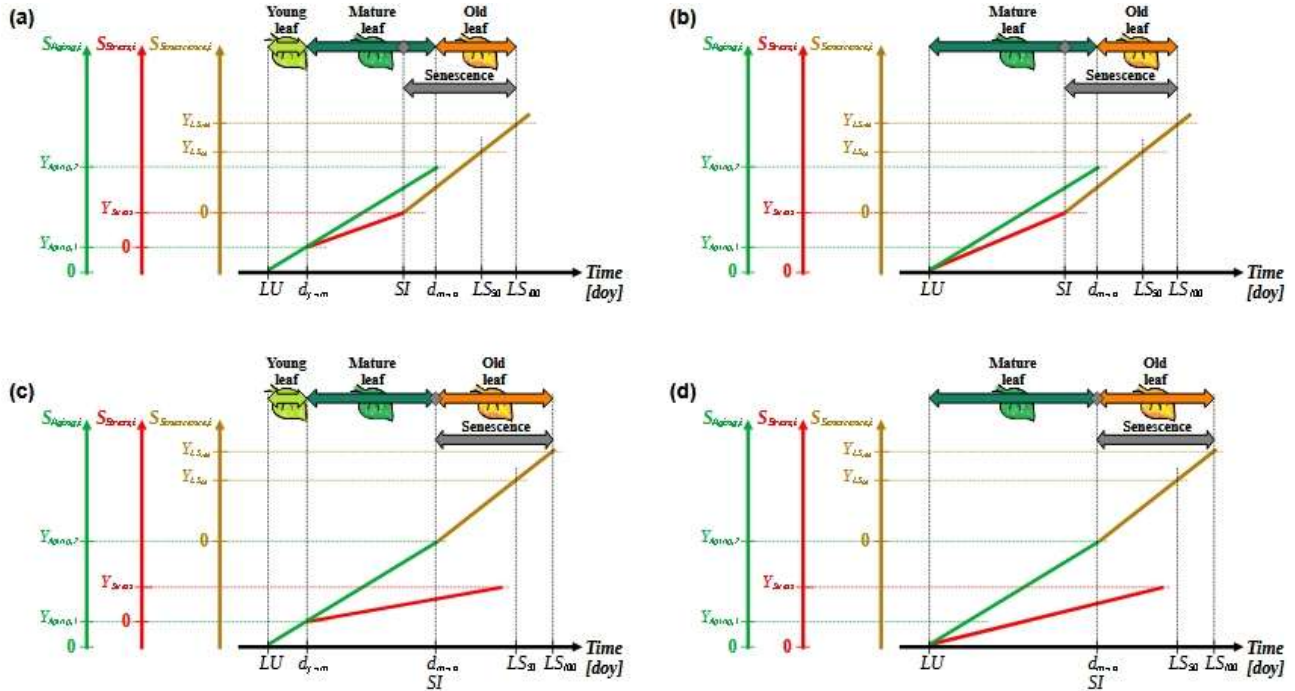
Figure 3. Conceptualization of the leaf development model. The process of leaf development is defined by three subsequent phases of leaf development, i.e., 'young leaf', 'mature leaf', and 'old leaf' [light green, dark green, and orange horizontal arrows, respectively; panel (a) and (c)]. Alternatively, the process is simplified into two subsequent development phases, i.e., 'mature leaf', and 'old leaf' (panel b and d). Senescence may be induced by stress during the phase of the mature leaf [grey rhombuses; panels (a) and (b)] or by aging on the day of transition from the mature to the old leaf [$d_{m \to o}$; panels (c) and (d), respectively]. The state of aging, stress, and senescence (y-axes; $S_{Aging,i}$, $S_{Stress,i}$, and $S_{Senescence,i}$; Eq. 1; solid green, red, and brown lines, respectively) for day i are derived from the corresponding daily rates (Eqs. 3, 4, and 8) accumulated over time [x-axis; day of year (*doy*)]. Starting from the leaf unfolding date (*LU*), these states simulate the leaf development, marked by transitions from the young to the mature leaf ($d_{y \to m}$) and $d_{m \to o}$ as well as by the dates of senescence induction (*SI*) and of the phenological stages 50% and 100% leaf coloring or fall (LS_{50} and LS_{100}, respectively). These transitions and stages occur when $S_{Aging,i}$, $S_{Stress,i}$, and $S_{Senescence,i}$ reach corresponding thresholds ($Y_{Aging,1}$, $Y_{Aging,2}$, Y_{Stress}, Y_{LS50}, and Y_{LS100}). *SI* is defined as the first day on which either Y_{Stress} or $Y_{Aging,2}$ is reached [panels (a) and (b) versus panels (c) and (d), respectively] and marks the onset of senescence (grey horizontal arrow), during which the daily senescence rate accumulates. If *SI* results from $Y_{Aging,2}$ being reached, it coincides with $d_{m \to o}$. Dotted lines are auxiliary lines.

Based on these definitions, we formulated the leaf development process under following assumptions. Aging may be simulated either by photosynthetic activity (Jibran et al., 2013; Paul and Foyer, 2001; Zohner et al., 2023) or more simply by a number of days. Stress may be simulated by a combination of sudden or gradual responses to the stressors cold, shortening day length, drought, heat, frost, heavy rain, and nutrient depletion (Bigler and Vitasse, 2021;

180 Jan et al., 2019; Jibran et al., 2013; Kloos et al., 2024; Mariën et al., 2021; Tan et al., 2023; Wang et al., 2022; Xie et al., 2015, 2018; Zohner et al., 2023). Senescence may be simulated in linear, convex, or sigmoidal dependence on combined aging and stress (Tan et al., 2023; Xie et al., 2015).

All formulations are based on daily states of aging, stress, and senescence (Eq. 1), which are compared to corresponding thresholds (Eq. 2):

$$S_{k,j} = \sum_{i=t_{0,k}}^{j} R_{k,i} \tag{1}$$

$$S_{k,j} \geq Y_k \tag{2}$$

Here, $S_{k,j}$ is the state on day j of either aging, stress, or senescence (k). This state is formulated as the sum of the corresponding rates on day i ($R_{k,i}$), which accumulated between the starting day $t_{0,k}$ and j, until the threshold Y_k is reached. In other words, the daily aging rate ($R_{Aging,i}$) accumulates from LU ($t_{0,Aging} = $ LU). The transition from young leaf to mature leaf occurs when $S_{Aging,j}$ reaches $Y_{Aging,1}$. Thus, day j becomes $t_{0,Stress}$ and the accumulation of the daily stress rate ($R_{Stress,i}$) starts, while $R_{Aging,i}$ continues to accumulate. While the transition from mature leaf to old leaf occurs when $S_{Aging,j}$ reaches $Y_{Aging,2}$, senescence is either induced with this transition or already earlier due to $S_{Stress,j}$ reaching Y_{Stress}. Upon senescence induction, day j becomes $t_{0,Senescence}$ and the daily senescence rate ($R_{Senescence,i}$) starts to accumulate. Eventually, $S_{Senescence,j}$ reaches $Y_{LS_{50}}$ and $Y_{LS_{100}}$, and respective LS_{50} and LS_{100} are marked by the corresponding days j.

$R_{Aging,i}$ was either set equal to the daily net photosynthetic activity or to one (i.e., A_{net} [mol C d^{-1}] or 1 [d d^{-1}], respectively), depending on the formulation (Eq. 3):

$$R_{Aging,i} = \begin{cases} A_{net,i} \\ 1 \end{cases} \tag{3}$$

$R_{Stress,i}$ was formulated as the sum of three to seven weighted stressors (D_{stress}; Eqs. 4–6), always considering (1) cold days (derived from minimum temperature; Tn [°C]), (2) shortening days (derived from the difference in day length; δL [h], with $\delta L_i = L_i - L_{i-1}$), and (3) dry days (approximated by KBDI; Q). In addition, some formulations of R_{Stress} also considered (4) periods of heavy rainfall (approximated by the five-days precipitation; $P5$ [mm], with $P5_i$ being the sum of P_{i-4} to P_i), (5) heat days (derived from maximum temperature; Tx [°C]), (6) nutrient depletion (approximated by the accumulated A_{net} since LU, due to the absence of soil data), and/or (7) frost days (through a response to Tn with lower thresholds than for cold days; Table S3):

$$R_{Stress,i} = \sum w_{D_{Stress}} \times f(D_{Stress,i}) \tag{4}$$

$$D_{Stress,i} \in \left\{ Tn_i, \quad \delta L_i, \quad Q_i, \quad P5_i, \quad Tx_i, \quad \sum_{l=d_{LU}}^{i} A_{net,l}, \quad Tn_i \right\} \tag{5}$$

$$f(x) = \begin{cases} g(x) \\ h(x) \end{cases} \tag{6}$$

Here, $w_{D_{Stress}}$ is the weight for the response [$f(x)$] to D_{Stress}, calculated according to $g(x)$ or $h(x)$ (Eqs. 7 and 8):

$$g(x)=\begin{cases}1 & , \ if \ x\geq a \\ 0 & , \ if \ x<a\end{cases} \tag{7}$$

220

$$h(x)=\begin{cases}1 & , \ if \ x<b_0 \\ \dfrac{b_1-x}{b_1-b_0} & , \ if \ b_0\leq x\leq b_1 \\ 0 & , \ if \ x>b_1\end{cases} \tag{8}$$

While a marks the sudden boundary between an unstressed and stressed state, b_0 and b_1 mark the lower and upper bounds, respectively, between which stress gradually increases (Fig. 4). Because stress results when $x \geq a$ and

225     $x \geq b_0$, the response to δL and Tn was formulated as $g(-\delta L)$ and $g(-Tn)$ as well as $h(-\delta L)$ and $h(-Tn)$. This translates in stress if $\delta L \leq -a$, $\delta L \leq -b_0$, $Tn \leq -a$, and $Tn \leq -b_0$. For example, if stress occurs suddenly or gradually when $\delta L \leq -0.01$ h, then $a = 0.01$ h and $b_0 = 0.01$ h, respectively. Note that these are examples, see Table 3 for the calibrated values.
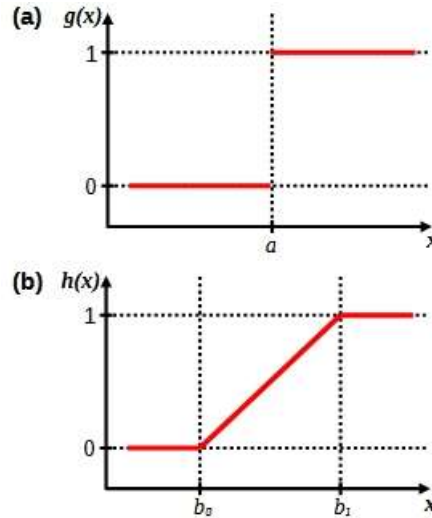
Figure 4. Response functions (solid red lines) of g(x) and h(x). In (a), a marks the boundary value of x at which $g(x)$ suddenly changes from 0 to 1 (i.e., from no effect to an effect). In (b), b_0 and b_1 mark the lower and upper bounds of x, respectively, between which $f(x)$ gradually increases from 0 to 1. Dotted lines are auxiliary lines.

$R_{Senescence,i}$ was either formulated as the sum, product, or exponential function of $R_{Aging,i}$ and $R_{Stress,i}$ or of $S_{Aging,i}$

230   and $R_{Stress,i}$, which yield linear, convex, and sigmoid curves, respectively (Eq. 9):

$$R_{Senescence,i}=\begin{cases}w_A R_{Aging,i}+w_S R_{Stress,i} \\ s_X\left(R_{Aging,i}\times R_{Stress,i}^{x_S}\right) \\ s_X\dfrac{1}{e^{c S_{Aging,i}(d-R_{Stress,i})}}\end{cases} \tag{9}$$

w_A and w_S are the weights of R_{Aging} and R_{Stress}, respectively, and s_X is a scaling factor, all of which allowed us to

235   hard code $Y_{LS_{50}} = 1$, x_S is the range bounded exponent of R_{Stress}, while c and d are the parameters of the sigmoid curve that relates R_{Stress} and S_{Aging} (Lang et al., 2019).

2.4   Model calibration and validation

We selected the observations for the calibration and validation samples with different procedures. To have a low risk of overfitting (i.e., the bias–variance trade-off; Sect. 2.2.2 in James et al., 2017), each calibration sample contained at least ten observations per calibrated parameter (Meier and Bigler, 2023). We defined two calibration datasets: one to calibrate a model that predicts both LS_{50} and LS_{100} simultaneously, and one to calibrate a model that predicts LS_{50} only. For the two datasets, we identified the site-years with the most extreme conditions during the growing season, i.e., the hottest, coldest, driest ten day periods observed between LU and LS_{50} as well as the shortest and longest growing season observed in the pre-selected data (Sect. 2.1). For the first dataset, hereafter called 'LS_{50}-LS_{100} sample', we randomly selected 250 of these site-years containing an observation for both LS_{50} and LS_{100}. For the second dataset, hereafter referred to as 'LS_{50} sample', we randomly selected 250 of these site-years containing observations for LS_{50}. These calibration samples were paired with validation samples that contained all remaining LS_{50} and LS_{100} observations or all remaining LS_{50} observations, respectively. We drew twice both the LS_{50} and LS_{50}-LS_{100} samples. While model development was based on the LS_{50}-LS_{100} samples, model evaluation was based on the LS_{50} sample to allow for a comparison with previously published models. All models were calibrated five times per drawn sample (i.e., ten 'calibration runs' per model and LS_{50} sample or LS_{50}-LS_{100} sample) by minimizing the root mean squared error (RMSE; Eq. S44) with generalized simulated annealing and optimal, model-specific controls (see Sect. S2.2; Xiang et al., 1997, 2017).

2.5   Model development

We based our model on the most accurate formulation of the leaf development process after testing different formulations in several iterations (Fig. 5; see Table S3 for parameter ranges). First, we defined the process structure based on sudden responses [$g(x)$; Eq. 7] to the stressors for cold days, shortening days, and dry days. In iteration 1, we tested the definition of the aging rate (R_{Aging}) and of the senescence rate ($R_{Senescence}$) based on the two phases of leaf development 'mature leaf' and 'old leaf' (Figs. 3b and 3c). R_{Aging} was formulated as a function of either the net photosynthetic activity (A_{net}) or of the number of days (Eq. 3). $R_{Senescence}$ was formulated in linear, convex, or sigmoidal dependence on combined aging and stress [through R_{Aging} or the state of aging (S_{Aging}) and through the stress rate (R_{Stress})] in either a sum, product, or exponential function (Eq. 9). In iteration 2, we tested the number of phases of leaf development and added the phase 'young leaf' (Figs. 3a and 3d). Thus, we formulated R_{Stress} with with gradual responses [$h(x)$; Eq. 8] to the initial stressors and a forward selection of additional stressors. In iteration 3, we considered each stressor for cold days, shortening days, and dry days through $h(x)$ rather than $g(x)$. In iteration 4, we considered one additional stressor, i.e., heavy rain periods, heat days, nutrient depletion, or frost days through $g(x)$. In iteration 5, we considered the additional stressor through $h(x)$ rather than $g(x)$. In iteration 6, the procedure of iterations 4 and 5 was repeated as long as they resulted in a formulation that was selected for further development.

         The formulations to be further developed were selected according to the accuracy of the corresponding model in predicting LS_{50} and LS_{100}, i.e., through calibration with the LS_{50}-LS_{100} sample. This accuracy was assessed with the Akaike information criterion corrected for small samples (AICc; Eq. S41; Akaike, 1974; Burnham and Anderson, 2004), which accounts for both the goodness-of-fit between the predicted and observed leaf senescence dates and the number of free parameters. The AICc was calculated for each calibration run (see Sect. 2.4) and the run with the highest AICc per model was excluded. We further developed the formulations of the two models with the lowest median AICc across the given and all previous iterations. Finally, the model with the lowest median AICc was selected and further evaluated.
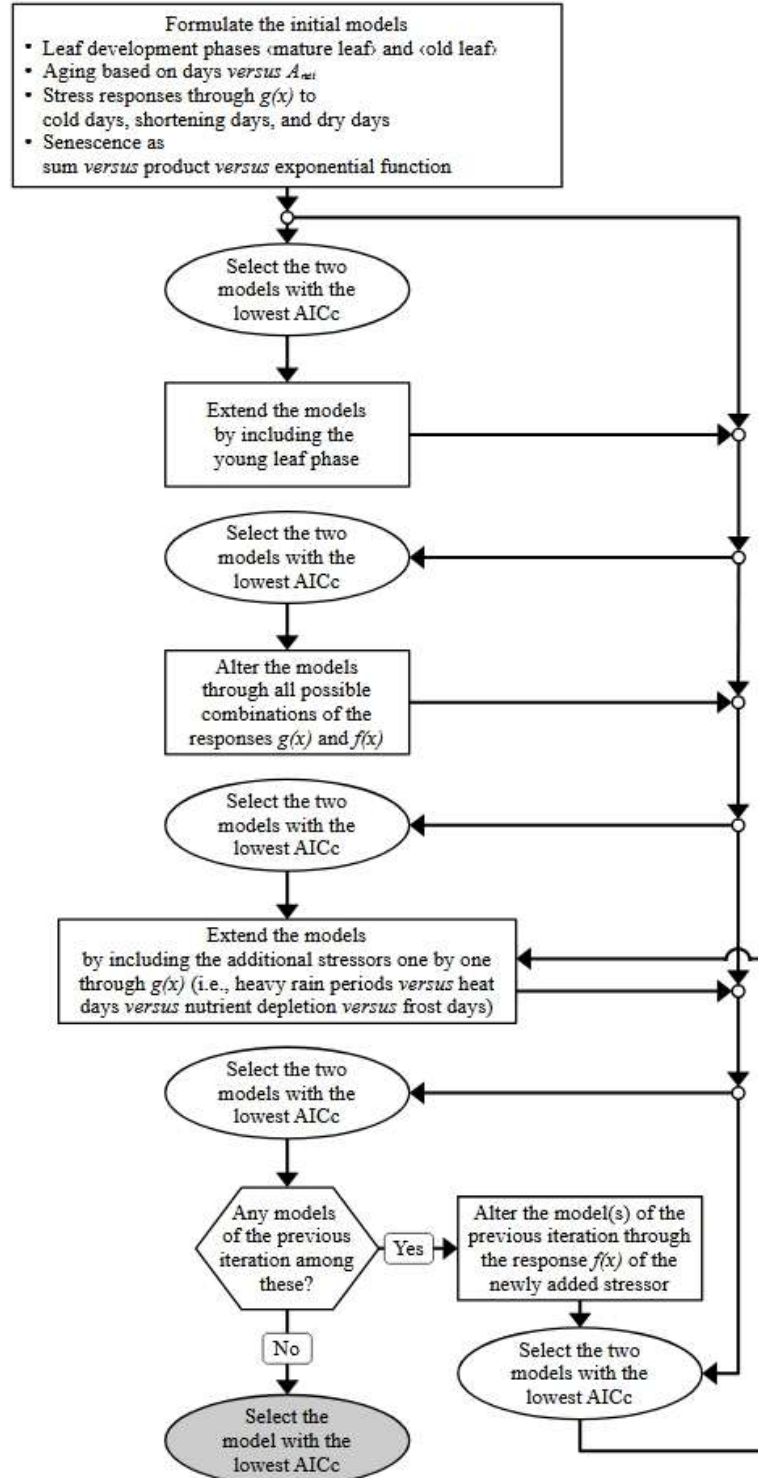
Figure 5. Model development. The iterations of model development are symbolized with rectangles. Selection of the best formulated models (ellipses) was based on the Akaike information criterion corrected for small samples (AICc; Eq. S41; Akaike, 1974; Burnham and Anderson, 2004) and the final selection is marked grey. For the response functions $g(x)$ and $h(x)$, see Eqs. 7 and 8.

2.6 Model evaluation

First, we evaluated the functionality of the selected model. We were particularly interested in the causes of senescence induction that could be due to aging or stress (Fig. 3). We counted how often aging versus stress induced senescence, and we quantified the relative amount of accumulated stress caused by each stressor at the time of senescence induction. We compared both aging- and stress-induced senescence as well as the relative amounts of stress across mean annual temperature (MAT; °C), mean annual KBDI (MAQ), latitude (LAT; °N), and elevation (ELV; m a.s.l.) for the given year and site. While MAT and MAQ were assumed to directly affect cold and dry stress, LAT relates to day length through the inclination angle of the Earth (Brock, 1981), and ELV relates to dry stress through decreasing nutrients with elevation (Huber et al., 2007; Loomis et al., 2006). The evaluation was based on the calibration runs that resulted in the highest modified Kling-Gupta efficiency (KGE'; Eq. S45; Gupta et al., 2009; Kling et al., 2012), which combines bias, variability, and correlation of the predicted and observed leaf senescence dates.

Second, we compared the accuracy between the selected model and three previously published models, namely the CDD, DM2, and PIA model. Because these models simulate only one stage of leaf senescence, which usually is LS_{50}, we based our comparison on this stage (Delpierre et al., 2009; Dufrêne et al., 2005; Zani et al., 2020). The CDD model determines LS_{50} by the time the cold degree-days reaches a particular threshold (Dufrêne et al., 2005). The DM2 model accumulates the product of temperature differences and day length ratios to corresponding thresholds until the threshold that determines LS_{50} is reached (Delpierre et al., 2009). The PIA model accumulates temperatures and day lengths that are combined in an exponential function, and derives the threshold to determine LS_{50} from the photosynthetic activity during the growing season (Zani et al., 2020). All these models were compared based on the calibration run that resulted in the highest KGE'. Further, we compared the RMSE and AICc as well as the Pearson correlation (ρ) across the entire validation sample ($\rho_{Overall}$), across space ($\rho_{Spatial}$), and across time ($\rho_{Temporal}$). $\rho_{Spatial}$ was calculated across sites based on their mean predicted and observed LS_{50}. $\rho_{Temporal}$ was calculated for each site, based on the yearly predicted and observed LS_{50}.

Third, we estimated the extent to which the model error (i.e., predicted minus observed LS_{50}) was affected by data structure as well as by climatic and spatial deviations from the LS_{50} calibration sample, using a linear mixed-effects model (LMM; Pinheiro and Bates, 2000) and an analysis of variance (ANOVA; Sect. S2.4; Fox, 2016). In the LMM, the response variable 'model error' was explained by the factor variable 'country' as well as the interaction of the factor variable 'model' with each of the differences between a site-year and the average of the calibration sample in MAT (δMAT), MAQ (δMAQ), the accumulated A_{net} between LU and summer solstice (δA_{net}), latitude (δLAT), and elevation (δELV). The random intercept was grouped by 'site'. The LMM was fitted with fast restricted maximum likelihood (Wood, 2011), and served as basis for the ANOVA. This type-III ANOVA (Yates, 1934) quantified the impact of the explanatory variables on the variance of the model error that was explained by the LMM. The impact attributable to data structure was caused by the fixed effects of 'country' and the standard deviation in the random intercepts grouped by 'site', while the impacts attributable to climatic versus spatial deviations from the calibration sample was caused by the effects of δMAT, δMAQ, and δA_{net} versus the effects of δLAT and δELV, respectively.

2.7 Statistical software and reporting of results

We used the programming language R, together with the R package data.table for data processing (Barrett et al., 2024). In R, data from xslx files were extracted with the R package readr (Wickham et al., 2024), and data from netCDF files were extracted and averaged with the R packages ncdf4 (Pierce, 2023), raster (Hijmans, 2023), sf (Pebesma, 2018;

Pebesma and Bivand, 2023), and sp (Bivand et al., 2013; Pebesma and Bivand, 2005). Leap years were identified with the function leap_year in the R package lubridate (Grolemund and Wickham, 2011). Gaps in the regressed lapse rates were filled with the function na.spline in zoo (Zeileis and Grothendieck, 2005). Seasonal splines of atmospheric CO_2 concentrations were calculated with the function sm in npreg (Helwig, 2024). The leaf senescence models were cali-
320   brated with the R package GenSA (Xiang et al., 2013), while the LMM was fitted with the R package mgcv (Wood, 2017) and the ANOVA was calculated with the R package stats (R Core Team, 2025). LMM estimates and 99% confidence intervals (i.e., significance level $α = 0.01$) for combined coefficients, e.g., the effect of δMAT for a given model, were calculated with the Delta method (Chpt. 5.1.4 in Fox and Weisberg, 2019; Chpt. 9.9 in Wasserman, 2004) through the function deltaMethod in the R package car (Fox and Weisberg, 2019). For each LMM coefficient and ANOVA im-
325   pact, we expressed the most optimistic change of odds between the null hypothesis (being zero; H_0) and alternative hypothesis (being different from zero or greater than zero, respectively; H_1) with the minimum Bayes factor ($\underline{BF}_{01}$), labeling H_0:H_1 ratios of 1/1000 and 1/100 as 'decisive' and 'very strong', respectively (Held and Ott, 2018; Johnson, 2005). $\underline{BF}_{01}$ was calculated from the p-values and number of data with the function tCalibrate in the R package pCalibrate (Held and Ott, 2018). For the visualizations, we used the R packages ggplot and ggpubr (Kassambara, 2020; Wickham,
330   2016), as well as the R packages ggspatial and rnaturalearth for the maps (Dunnington, 2023; Massicotte and South, 2023).

3   Results

3.1   Model formulation – the DP3 model

We tested 34 formulations of the leaf development process through 1428 calibration runs, and found that three subsequent leaf development phases resulted in the most accurate model (according to the AICc; Figs. 3a, 3c, 6 and S1–S2).
335   In this model, the phase 'young leaf' starts with leaf unfolding. As a daily aging rate R_{Aging} accumulates (Eq. 10), the simulated state of aging increases by one day per day. When this state reaches the threshold $Y_{Aging,1}$ (Eqs. 1 and 2), the phase 'mature leaf' begins. During this phase, the leaf continues to age and is also sensitive to stress caused by cold days, shortening days, and dry days, to which we hereafter refer to as 'cold stress', 'photoperiod stress', and 'dry stress', respectively. This stress is summarized in a daily stress rate (R_{Stress}; Eq. 11) and thus accumulated to determine the state
340   of stress. The first day that either the state of stress or the state of aging reaches the respective thresholds Y_{Stress} or $Y_{Aging,2}$ (Eqs. 1 and 2), senescence is induced, while the phase 'old leaf' starts only when the state of aging reaches $Y_{Aging,2}$. Once senescence is induced, a daily senescence rate ($R_{Senescence}$) in convex dependence on stress accumulates (Eq. 12) and determines the state of senescence. The days this state reaches the thresholds $Y_{LS_{50}}$ and $Y_{LS_{100}}$ (Eqs. 1 and 2) correspond to the predicted dates of LS_{50} and LS_{100}, respectively. Hereafter, we refer to this model as 'DP3' model (Tables 2 and 3;
345   Meier, 2025b, coded in R).

$$R_{Aging,i} = 1 \tag{10}$$

$$R_{Stress,i} = w_C\,g(-Tn_i) + w_P\,g(-\delta L_i) + w_D\,g(Q_i) \tag{11}$$

350

$$R_{Senescence,i} = s_X\,R_{Stress,i}^{x_S} \tag{12}$$

Here, w_C, w_P, and w_D are the weights for the response functions $g(x)$ (Eq. 7) to the minimum temperature (Tn),
difference in day length (δL), and KBDI (Q) on day i, respectively [e.g., $w_P\ g(-\delta L_i)$ results in the photoperiod stress on
355   day i]. s_X is the scaling factor for R_{Stress}, which is shaped by x_S.
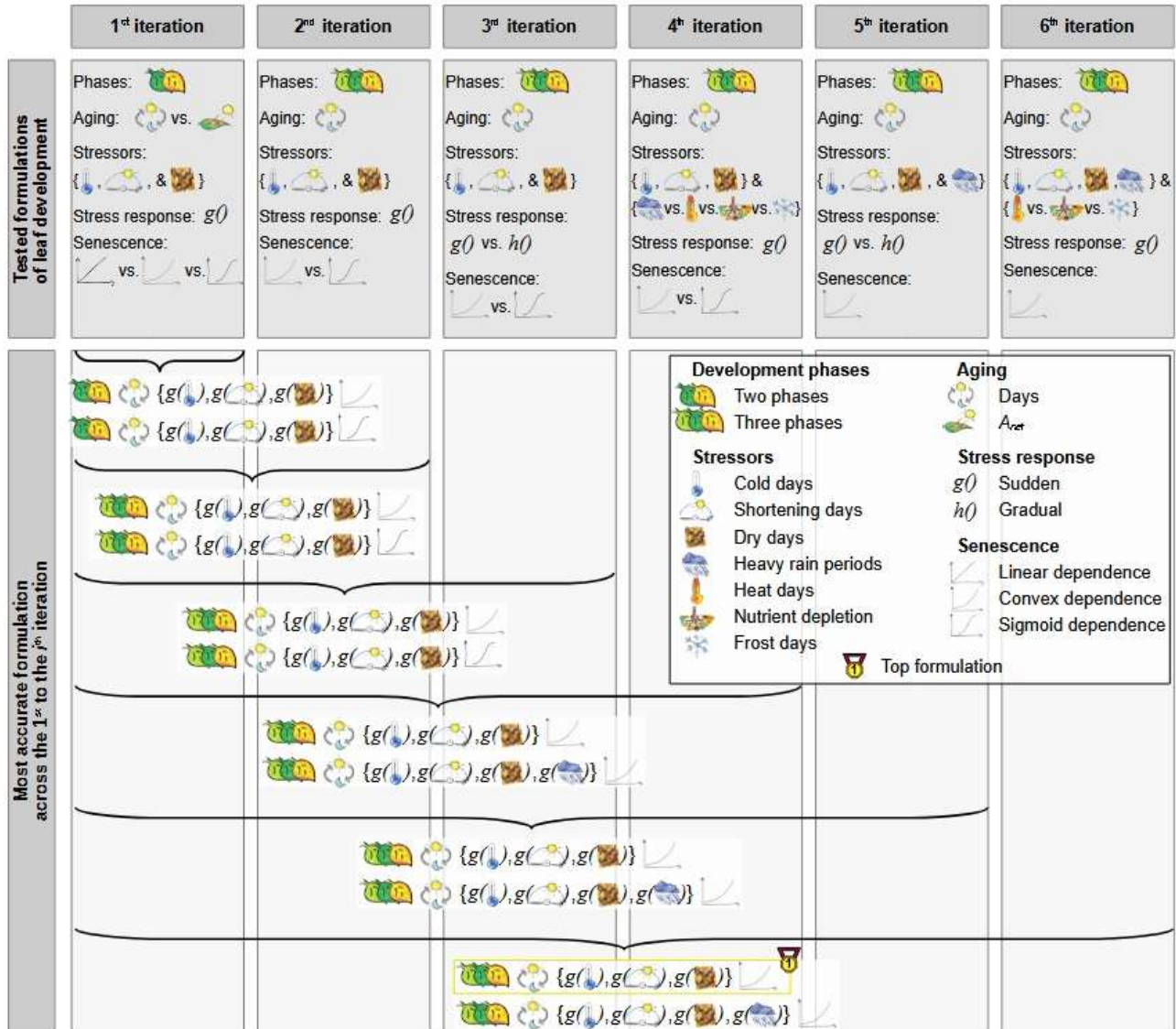
Figure 6. Tested model formulations. The tested formulations differed in their number of leaf development phases (i.e., two or three phases), in their driver of the aging rate [i.e., days or photosynthetic activity (A_{net})], their stress rate in response [i.e., $g(x)$ or $h(x)$] to the stressors cold, shortening, dry, heat, and frost days, heavy rain periods, and nutrient depletion, and their dependence of the senescence rate on aging and stress (i.e., linear, convex, or sigmoid dependence as a the result of a sum, product, or exponential function, respectively). After each iteration, we identified the two most accurate formulations across the given and all previous iterations (Fig. 5, Sect. 2.5). These formulation were further developed through the next iteration. As soon as an iteration did not produce any new model formulations, we selected the more accurately formulated model ('top formulation'; i.e., the 'DP3' model). All formulations were tested for beech based on the LS_{50}-LS_{100} sample (Sect. 2.4).

Table 2. Input and output variables of the DP3 model

	Collective list	Name	Definition	Unit	Format
Input	-	*par*	Model parameters (see Table 3)	-	Vector
	data	*LU*	Observed leaf unfolding dates	doy	Vector
		id	Unique identifier of each *LU* (character)	-	Vector

		D_i	Daily number of days (i.e., 1 per day)	-	Matrix
		Tn_i	Daily minimum temperature	°C	Matrix
		δL_i	Daily difference in day length to previous day	h	Matrix
		Q_i	Daily Keetch and Byram drought index	-	Matrix
	-	*stages*	Leaf senescence stages to be predicted (character, defaults to $LS_{Default}$)	-	Vector
Output	-	LS	Predicted leaf senescence dates, including senescence induction (SI)	doy	Matrix
	transitions	$d_{y \to m}$	Predicted timing of transition from young to mature leaf	doy	Vector
		$d_{m \to o}$	Predicted timing of transition from mature to old leaf	doy	Vector
	aging	$R_{Aging,i}$	Daily rate of aging	-	Matrix
		$S_{Aging,i}$	State of aging (i.e., accumulated $R_{Aging,i}$ since LU)	-	Matrix
	stress	$X_{Cold,i}$	Daily cold stress [i.e., $w_C\, g(-Tn_i)$]	-	Matrix
		$X_{Photoperiod,i}$	Daily photoperiod stress [i.e., $w_P\, g(-\delta L_i)$]	-	Matrix
		$X_{Dry,i}$	Daily dry stress [i.e., $w_D\, g(-Q_i)$]	-	Matrix
		$R_{Stress,i}$	Daily rate of stress	-	Matrix
		$S_{Stress,i}$	State of stress (i.e., accumulated $R_{Stress,i}$ since $d_{y \to m}$)	-	Matrix
	senescence	$R_{Senescence,i}$	Daily rate of senescence	-	Matrix
		$S_{Senescence,i}$	State of senescence (i.e., accumulated $R_{Senescence,i}$ since SI)	-	Matrix

Note: Daily variables refer to day i, and accumulated variables refer to the period until day i. The vector *par* contains the model parameters listed in Table 3. In the collective lists data, aging, stress, and senescence, the rows of the matrices refer to the days of the year, while the columns refer to site-years and are ordered identically between all matrices. This order matches the order of the vectors in the collective lists data and transition. For the LS matrix, the rows refer to the site-years and the columns refer to the senescence induction date and the dates of the leaf senescence stages indicated by the vector *stages*. LU, LS, $d_{y \to m}$, and $d_{m \to o}$ are given in day of year (doy; Meier, 2025b).

Table 3. Fitted parameters of the DP3 model

		Fitted value	
Parameter	Meaning	LS_{50}-LS_{100}	LS_{50}
$-a_C$	Boundary below which cold stress is 1 versus 0 (referring to Tn_i)	2.55 °C	0.06 °C
$-a_P$	Boundary below which photoperiod stress is 1 versus 0 (referring to δL_i)	−0.0587 h	−0.0016 h
a_D	Boundary above which dry stress is 1 versus 0 (referring to Q_i)	176.94	183.82
w_C	Weight of cold stress	0.14	0.29
w_P	Weight of photoperiod stress	0.02	0.52
w_D	Weight of dry stress	0.22	0.05
s_X	Scaling factor of the senescence rate	0.59	0.35
x_S	Shape parameter of the stress rate	0.21	5.67
$Y_{Aging,1}$	Age threshold for the transition from young to mature leaf	41.59 d	1.57 d
$Y_{Aging,2-Aging,1}$	The threshold of aging during the mature leaf phase	137.31 d	71.58 d
$Y_{Aging,2}$	Theoretical age threshold for the transition from mature to old leaf	178.90 d	73.14 d
$Y_{LS_{100}}$	Senescence threshold for LS_{100} (all leaves have changed color or have fallen)	5.95	-

Note: The parameters refer to the equations 7 and 9–11 and were fitted for beech with the LS_{50} and LS_{50}-LS_{100} samples (Sect. 2.4). All parameters were calibrated within the initial ranges (Table S3) to their fitted value. To avoid fitted values of $Y_{Aging,1} > Y_{Aging,2}$, we used and calibrated $Y_{Aging,2-Aging,1}$ instead of $Y_{Aging,2}$. The theoretical threshold $Y_{Aging,2}$ was not calibrated but calculated from $Y_{Aging,1} + Y_{Aging,2-Aging,1}$ and displayed for easier interpretation. The thresholds for stress (Y_{Stress}) and LS_{50} (Y_{LS50}; i.e., the time when 50% of the leaves have changed color or have fallen) were hard coded with $Y_{Stress} = 1$ and $Y_{LS50} = 1$. The shortening of day length of 0.0016 h (a_P; corresponding to 0.1 minutes) based on the LS_{50} calibration is reached on doy 175, 174, and 174 (i.e., June 24, 23, and 23) at the exemplary minimum, median, and maximum latitudes of our samples (i.e., 45.9° N, 47.8° N, and 58.0° N), respectively. Alternatively, the shortening of 0.0587 h (3.5 minutes) based on the LS_{50}-LS_{100} calibration is reached on doy 252 and 202 (i.e., September 9 and July 21) at the median and maximum latitudes of our samples, respectively, whereas it is never reached at the minimum latitude.

According to the DP3 model, leaf senescence was generally induced earlier during warmer years and at lower elevations (Fig. 7; Tables S5–S8). In average, senescence was induced a month earlier when mean annual temperatures were 13–15 °C than when they were 4–6 °C (i.e., May 29 versus June 22 and April 20 versus May 12 when the DP3

360   model was calibrated with the LS_{50}-LS_{100} and LS_{50} samples, respectively; hereafter referred to as 'DP3$_{LS_{50}-LS_{100}}$ model' and 'DP3$_{LS_{50}}$ model'; Sect. 2.4). Accordingly, senescence induction was 20 days earlier below 288 m a.s.l. than above 1150 m a.s.l. (i.e., June 5 versus June 25 and April 26 versus May 16 based on the DP3$_{LS_{50}-LS_{100}}$ and DP3$_{LS_{50}}$ model, respectively). Both the DP3$_{LS_{50}-LS_{100}}$ model and DP3$_{LS_{50}}$ model predicted generally longer-lasting senescence (i.e., the duration between senescence induction and LS_{50} or LS_{100}) during years of higher mean annual temperatures (Fig. S3; Tables

365   S5–S8).

Figure 7. Senescence induction. Panel (a) and (b) are based on simulations by the DP3 model calibrated with the LS_{50}-LS_{100} versus LS_{50} samples, respectively (Sect. 2.4). The top row of each panel shows the number of site-years in the bins, which were equally distributed among mean annual temperature (MAT; °C), mean annual Keetch and Byram drought index (MAQ), latitude (LAT; °N), and elevation (ELV; m a.s.l.). The second row of each panel visualizes the senescence induction (SI) dates in day of year (doy). While the mean and median dates are marked with black dots and grey lines, respectively, the most extreme values are indicated with dots if outside ±1.5 times the inner quartile range from the 1st and 3rd quartile, and with whiskers otherwise. The third row of each panel illustrates the relative number of site-years during which senescence was induced by stress versus aging or by both stress and aging (i.e., both the accumulated stress and aging rates reached their thresholds for SI on the same date). The bottom row of each panel shows the relative amounts of cold stress, photoperiod stress (Photop.), and dry stress that accumulated at stress-caused SI.

Stress induced senescence two times and 40 times more often than aging according to the $DP3_{LS_{50}-LS_{100}}$ and $DP3_{LS_{50}}$ model, respectively (Fig. 7, Tables S5–S8). Thus, while aging was of negligible importance to senescence induction according to the $DP3_{LS_{50}}$ model, it mattered according to the $DP3_{LS_{50}-LS_{100}}$ model, particularly during years of medium mean annual temperature (6–13° C) as well as at sites of medium latitude (48.3–55.6 °N) and of low elevation (below 576 m a.s.l.). At the time of senescence induction due to stress, the amounts of accumulated photoperiod stress and cold stress relative to total stress were 56% versus 44% ($DP3_{LS_{50}-LS_{100}}$ model) and 77% versus 23% ($DP3_{LS_{50}}$ model), respectively, while the corresponding amounts of dry stress were 0.5% and 0.0%. Photoperiod stress dominated mostly in warm years and medium-elevation sites according to the $DP3_{LS_{50}-LS_{100}}$ model, whereas it did so in cool years and high-elevation sites according to the $DP3_{LS_{50}}$ model. In summary, photoperiod stress rather than cold and dry stress induced leaf senescence, but the importance of these stressors and their dependency on climatic conditions and location differed between the $DP3_{LS_{50}-LS_{100}}$ and $DP3_{LS_{50}}$ model.

Accordingly, the relative importance of these stressors for the duration of senescence differed between the $DP3_{LS_{50}-LS_{100}}$ and $DP3_{LS_{50}}$ model (Fig. S3; Tables S5–S8). Photoperiod stress clearly dominated the progress from senescence induction to LS_{50} according to the $DP3_{LS_{50}}$ model. However, according to the $DP3_{LS_{50}-LS_{100}}$ model and especially during cool years, cold stress was most important between senescence induction and LS_{50}, whereas photoperiod stress was most important between senescence induction and LS_{100}.

3.2 Model accuracy

Leaf senescence dates were predicted with similar accuracy by the DP3 model as by previous models (Fig. 8; Table 4). All models calibrated with the LS_{50} sample resulted in an RMSE of ~15 d, with the lowest RMSE for the Null model (i.e., constant prediction of each the average LS_{50} and LS_{100} observations in the calibration sample). The $LS_{50}-LS_{100}$ sample yielded considerable higher RMSE for both the $DP3_{LS_{50}-LS_{100}}$ and Null model, namely 23–25 d and 18–21 d, respectively. Nevertheless, the $DP3_{LS_{50}-LS_{100}}$ model resulted in the highest overall correlation ($\rho_{Overall}$ of 0.2 for LS_{100}). The highest correlation across space was obtained with the PIA model ($\rho_{Spatial}$ of 0.4), while the DP3 model resulted in the highest correlation across time (average $\rho_{Temproal}$ of 0.05 according to $DP3_{LS_{50}}$ and according to $DP3_{LS_{50}-LS_{100}}$ for LS_{100}).

Table 4. Model accuracy

Model	Sample	Stage	KGE'	RMSE	AICc	$\rho_{Overall}$	$\rho_{Spatial}$	$\overline{\rho}_{Temporal}$	n
CDD	LS_{50}	LS_{50}	-0.13	16.1	57797	0.01	-0.09	0.04	6887
DM2	LS_{50}	LS_{50}	-0.26	15.0	56862	0.02	-0.12	0.00	6887
PIA	LS_{50}	LS_{50}	-0.19	14.8	56701	0.10	0.44	-0.04	6887
$DP3_{LS_{50}}$	LS_{50}	LS_{50}	-0.23	15.2	57083	0.02	-0.02	0.05	6887
Null	LS_{50}	LS_{50}	NA	14.8	NA	NA	NA	NA	6887
$DP3_{LS_{50}-LS_{100}}$	$LS_{50}-LS_{100}$	LS_{50}	-0.01	25.0	63911	0.04	-0.06	0.03	6887
$DP3_{LS_{50}-LS_{100}}$	$LS_{50}-LS_{100}$	LS_{100}	0.14	23.2	NA	0.22	0.17	0.05	600
Null	$LS_{50}-LS_{100}$	LS_{50}	NA	18.1	NA	NA	NA	NA	6887
Null	$LS_{50}-LS_{100}$	LS_{100}	NA	21.7	NA	NA	NA	NA	600

Note: The Null model constantly predicts the average observation in the calibration sample (i.e., either the stage when 50% or 100% of the leaves have changed color or have fallen; LS_{50} or LS_{100}, respectively). The modified Kling-Gupta efficiency (KGE'), root mean squared error (RMSE), Akaike information criterion for small samples (AICc), and Pearson correlation overall, across space, and across time [$\rho_{Overall}$, $\rho_{Spatial}$, and average $\rho_{Temporal}$ ($\overline{\rho}_{Temporal}$), respectively] are explained in Sect. 2.6, S2.1, and S2.2. All these metrics were calculated for the predicted and observed dates in the validation samples LS_{50} and $LS_{50}-LS_{100}$ (Sect. 2.4). Except the RMSE, they result in NA if the variance of the predicted dates is zero, which is the case for the Null model. In addition, the AICc for the stage LS_{100} according to the model $DP3_{LS_{50}-LS_{100}}$ was omitted because n, i.e., the number of observations in the validation sample, differed between LS_{100} and LS_{50}.
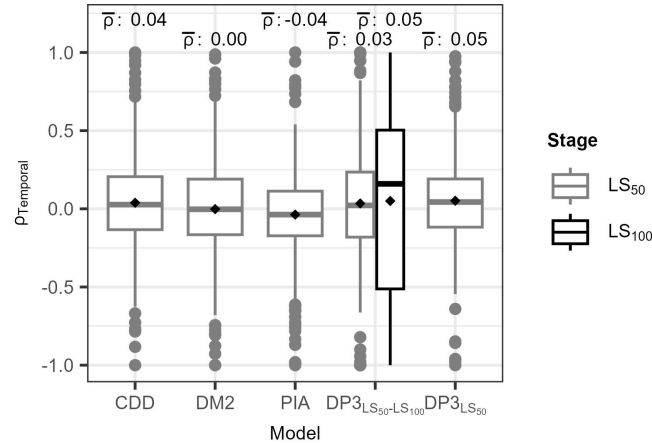
Figure 8. Temporal Pearson correlation ($\rho_{Temporal}$). The distribution of $\rho_{Temporal}$ per site between predicted and observed leaf senescence [the dates when 50% and 100% of the leaves have changed color or have fallen (LS$_{50}$ and LS$_{100}$, respectively)] is displayed for each model. The DP3 model was calibrated twice, namely with the LS$_{50}$-LS$_{100}$ sample (DP3$_{LS_{50}-LS_{100}}$) and with the LS$_{50}$ sample (DP3$_{LS_{50}}$; Sect. 2.4), the latter of which was also used to calibrate the CDD, DM2, and PIA model. The mean $\rho_{Temporal}$ (black rhombuses) is indicated above each box ($\bar{\rho}$). The boxes indicate the inner quartile range and the median (middle line). The most extreme values are indicated with dots if outside ±1.5 times the inner quartile range from the 1st and 3rd quartile, and with whiskers otherwise.

3.3 Model error

The model errors according to the DP3 model and previous models were similarly affected by data structure and climatic and spatial deviations from the calibration sample as the Null model (Fig. 9a). The data structure was described by the fixed effects of countries and the random intercepts grouped by sites. The countries altered the model error by –18 to +8 d, depending on the model (Tables S9–S10). The standard deviation in the model error due to the random intercepts was 9 d. Depending on the model, the fixed effects of the climatic deviations ranged from –22 to –19 d 10°C^{-1} (δMAT), from +3.6 to +9.0 d 100^{-1} (δMAQ), and from +4.1 to +4.6 d 10mol C^{-1} (δA_{net}), respectively. The model-specific effects of the spatial deviations δLAT and δELV ranged from +2.0 to +2.1 d °N^{-1} and from +1.0 to +1.1 d 100m^{-1}, respectively. While the evidence in the data was decisive ($\underline{BF}_{01}$ < 1/1000; Sect. 2.7) and significant ($p < 0.005$) for an effect of the CDD model on the model error different from zero as well as for the individual climatic deviations and δLAT, it was significant for corresponding effects of the DM2 model as well as for δELV and all individual countries. The evidence was neither decisive nor significant for any effect different from zero of the interaction terms between the models and the climatic or spatial deviations. The LMM explained the model error with an adjusted R^2 of 0.44. Differences among sites attributed for 92% of the variance in the model error explained by the LMM, followed by the effects of δA_{net} and δMAT (6% and 2%, respectively), whereas the effects of the models accounted for 0.3% (Fig. 9b; Table S11). In general, the model errors according to the DP3 model and previous models behaved as those of the Null model and mainly varied due to data structure.

Earlier onset of leaf senescence in warmer years
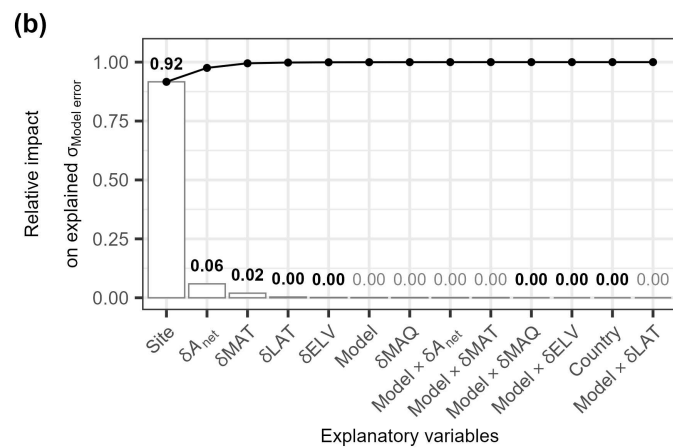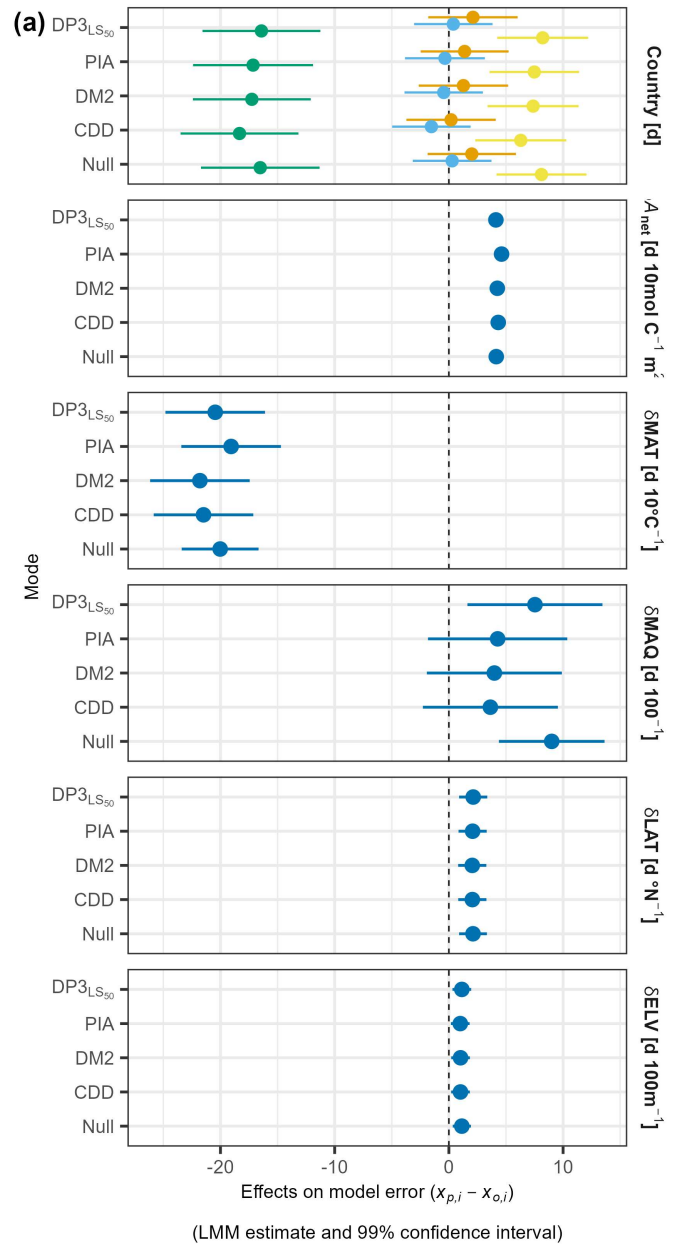
Figure 9. Model error versus data structure and climatic and spatial deviations. Panel (a) visualizes the LMM-based, model-specific estimated fixed effects (dots) and 99% confidence intervals (bars) of data structure described by 'country', climatic deviations described by mean annual temperature (δMAT; d $10°C^{-1}$), mean annual Keetch and Byram drought index (δMAQ; d 100^{-1}), and accumulated net photosynthetic activity between leaf unfolding and summer solstice (δA_{net}; d $10mol\ C^{-1}$), and spatial deviations described by latitude (δLAT; d $°N^{-1}$) and elevation (δELV; d $100m^{-1}$). These deviations were calculated as the difference between a given site-year and the average in the calibration sample. The colors indicate the countries Austria (AUT), Great Britain (GBR), Germany (GER), and Switzerland (SUI) as well as estimates across countries (AC). The model error was calculated as the predicted minus the observed timing ($x_{p,i} - x_{o,i}$). Panel (b) shows, based on the ANOVA, the relative impact of the explanatory variables on the variance in the model error as explained by the LMM. The random intercepts in the LMM were grouped by 'site', also describing data structure. The bars indicate the impact of individual variables, while the connected dots show the accumulated impact. The numbers above each bar state the impact, being bold in case of combined significance and decisiveness (i.e., $p \leq 0.01$ and $\underline{BF}_{01} \leq 1/1000$).

4   Discussion

4.1   Model formulation

410   The DP3 model predicts leaf senescence dates through a novel formulation that differs considerably from the formulation of current models. This novel formulation may change the way we see leaf senescence, namely as a consequence of leaf development that relates to both aging and stress. Current models start their simulation with the onset of senescence, i.e., the senescence induction date, which they determine from day length and temperature (e.g., Delpierre et al., 2009; Dufrêne et al., 2005; Keenan and Richardson, 2015; Lang et al., 2019; Liu et al., 2019; Zani et al., 2020). This

415   date is calibrated such that leaf senescence dates are predicted most accurately. In the DP3 model and in addition to this prerequisite, accumulated aging or accumulated stress since leaf unfolding must have reached a given threshold (i.e., $Y_{Aging,2}$ and Y_{Stress} as well as Figs. 3c and 3a, respectively; Table 3). In other words, while current models define the senescence induction date backward, the DP3 model defines it both backward and forward, arguably resulting in a more robust definition. Moreover, as current models generally ignore aging (but see the model by Keenan and Richardson,

420   2015, which considers the leaf unfolding date in the stress threshold for leaf senescence), their formulation partially ignores current knowledge (e.g., Field and Mooney, 1983; Guo et al., 2021; Jibran et al., 2013; Lim et al., 2007). In addition, the models by Liu et al. (2019) and Zani et al. (2020) postulate an effect of the conditions before senescence induction on senescence duration, which remains speculative. However, these conditions likely affect senescence induction dates, possibly through photosynthetic activity (Zohner et al., 2023) or through aging and stress (DP3 model).

425   The novel formulation of the DP3 model supports the advancement of leaf senescence research by postulating new hypotheses. To our knowledge, it is the first process-based leaf senescence model that (a) predicts leaf senescence dates through daily leaf development status, (b) starts the simulation with leaf unfolding, (c) differentiates between daily aging and stress rates, and (d) predicts the dates of transition between the leaf developmental phases young, mature, and old leaf as well as the senescence induction date. This allows the development of several new hypotheses

430   (Carley, 1999; Hauke et al., 2020), which may relate to the currently disputed effect of climate change on productivity (Lu and Keenan, 2022; Norby, 2021; Zani et al., 2020; Zohner et al., 2023) and can be tested by controlled experiments. In particular, these hypotheses may concern (1) the duration of the young leaf phase during which stress cannot induce senescence, (2) the timing and cause (i.e., aging versus stress) of senescence induction, and (3) the relative importance of the stressors in relationship to climate and location, all of which are further elaborated here below.

435   The duration of the young leaf phase differed considerably between the $DP3_{LS_{50}-LS_{100}}$ and $DP3_{LS_{50}}$ model (i.e., the DP3 models calibrated with the LS_{50}-LS_{100} versus LS_{50} samples; Sect. 2.4), namely 41 d versus 1 d, respectively. Be-

cause the DP3 assumes that stress during this phase is irrelevant for senescence induction, the duration of this phase affects the induction and end of senescence (see below). Moreover, corresponding projections under future climate scenarios are also likely affected, as the probability of late spring frost events likely will change under climate warming
440 (Bigler and Bugmann, 2018; Meier et al., 2018; Sangüesa-Barreda et al., 2021). Therefore, duration and characteristic of this young leaf phase should be examined further, e.g., with controlled experiments that apply continuous stress right after leaf unfolding to determine until when stress is either completely irrelevant for senescence induction or accumulates without inducing senescence.

Senescence was induced in late spring/early summer and more often by stress than by aging, but the induc-
445 tion dates and the stress:aging ratios differed notably between the $DP3_{LS_{50}-LS_{100}}$ and $DP3_{LS_{50}}$ model. Senescence induction dates predicted by the $DP3_{LS_{50}-LS_{100}}$ and $DP3_{LS_{50}}$ models differed by 40 d, which matches the difference in the predicted duration of the young leaf phase (see above). As stress during the young leaf phase does not affect the predicted leaf senescence dates by definition (Figs. 1, 3a, and 3c), this result illustrates the importance of studying the effects of stress after leaf unfolding (see above). It also shows that different combinations of calibrated model parameters eventually yield
450 similar predictions. Such compensating effects between different model parameters have also been reported in previous studies (Chuine and Régnière, 2017; Van der Meersch and Chuine, 2025), and explain the different stress:aging ratios as well as the earlier senescence induction during warmer years and at lower elevations. On the one hand, the ratio of stress to aging induced senescence shifts in favor of stress with shorter young leaf phases, because more cold stress can accumulate in spring. On the other hand, senescence must be induced earlier when senescence lengthens to predict leaf
455 senescence dates that are close to the average observation in the calibration sample, as suggested by model accuracy and model error (see below). Longer senescence, in turn, results from arguably reduced cold stress in fall in warmer years and at lower sites. Nevertheless, earlier induction and longer duration of senescence in warmer years may also be a valid description of reality (Yu et al., n.d.; Zohner et al., 2023). However, Zohner et al. (2023) argued that senescence induction dates relate negatively to pre-solstice productivity (see also Zani et al., 2020), whereas we showed that these
460 dates relate to particular interactions between aging and stress rather than to productivity (see below; Eqs. 3 & 10; Lu and Keenan, 2022; Marqués et al., 2023; Norby, 2021). Because such different mechanisms very likely affect leaf senescence projections under climate warming, they certainly need further investigations.

How do aging and stress interact to predict earlier induction and longer duration of senescence in warmer years and at lower sites? The aging requirement for the transition from mature to old leaf (i.e., $Y_{Aging,2}$; Table 3) repre-
465 sents the longest possible duration from leaf unfolding to senescence induction. Earlier senescence induction is only possible through sufficient stress between the transition from young to mature leaf and senescence induction (i.e., Y_{Stress}; Table 3), whereas stress after senescence induction relates negatively to the duration of senescence. Thus, according to the DP3 model, both earlier leaf unfolding and increased stress in spring advance senescence induction, while reduced stress in summer and fall lengthens senescence, which corresponds to observed patterns. Leaves unfold earlier at lower
470 sites in general (Vitasse et al., 2009, 2013) and in warm springs in particular (given that the buds have been sufficiently chilled; Asse et al., 2018; Meier et al., 2021; Menzel et al., 2020). Warmer years have been shown to increase cold stress in spring (i.e., through leaves unfolding overly early in comparison to late frost; Asse et al., 2018; Meier et al., 2018; Sangüesa-Barreda et al., 2021) and relate positively to dry stress (i.e., through evapotranspiration; Allen et al., 1994; Berdanier and Clark, 2018; Wu et al., 2022), but leave photoperiod stress unaffected (Brock, 1981). Thus, earlier senes-
475 cence induction is likely due to earlier leaf unfolding, and thus aging, and increased cold and dry stress in spring, while

the longer senescence duration is likely due to the later accumulation of photoperiod stress relative to leaf development and reduced cold stress in summer and fall.

480 Surprisingly at first, the $DP3_{LS_{50}-LS_{100}}$ model postulated photoperiod rather than cold and dry stress of being the most important stressor for senescence induction during warmer years, whereas the $DP3_{LS_{50}}$ model saw photoperiod stress of being most important during cooler years. By definition, stress only matters when it occurs after the transition from young to mature leaf (Figs. 1, 3a, and 3c). Photoperiod stress occurs, if at all, after this transition (i.e., after July 21 to September 9 in the $DP3_{LS_{50}-LS_{100}}$ model and after June 23–24 in the $DP3_{LS_{50}}$ model, depending on latitude; Table 3), and gains in importance quickly after the first occurrence unless senescence is induced soon by either cold or dry stress. Cold stress likely occurs in spring and fall, both before and after the transition from young to mature leaf. Thus, cold

485 days in spring are less likely to affect senescence induction the later the young leaf phase ends, and vice versa. In addition, more cold days in fall can be accumulated the later the young leaf phase ends and the later photoperiod stress starts to accumulate, and vice versa. Therefore, on the one hand, the long young leaf phase and late accumulation of photoperiod stress in the $DP3_{LS_{50}-LS_{100}}$ model favor the accumulation of cold stress in fall, which arguably decreases in warmer years, making photoperiod stress relatively more important. On the other hand, the short young leaf phase in the $DP3_{LS_{50}}$

490 model favors the accumulation of cold stress in spring, which likely decreases in cooler years through leaves unfolding overly late in comparison to late frost (Asse et al., 2018; Meier et al., 2018; Sangüesa-Barreda et al., 2021), making photoperiod stress relatively more important.

4.2 Model accuracy

We compared the DP3 model to three previous models of leaf senescence (i.e., the models CDD, DM2, and PIA;

495 Delpierre et al., 2009; Dufrêne et al., 2005; Zani et al., 2020) based on the LS_{50} calibration sample and found the RMSE of all compared models to be above the RMSE for the Null model. This may be explained by unrealistic model formulations, poor model calibrations, and noisy data to drive and calibrate the models, all of which we discuss here below.

While the formulations of the compared models differ, they all build on the results of previous studies. For example, according to all compared models, the leaf senescence date advances due to cold temperatures, which was also

500 observed by Kloos et al. (2024), Wang et al. (2022), Wang and Liu (2023), and Xie et al. (2015, 2018). Moreover, in all but one model, shorter days cause earlier leaf senescence, which is is in agreement with Addicott (1968), Keskitalo et al. (2005), Singh et al. (2017), Tan et al. (2023), and Wang et al. (2022). Therefore, while the Null model predicted the leaf senescence dates more accurately according to the RMSE, it is unlikely that it is more realistically formulated than the compared models. The currently most realistic model is arguably the DP3 model (Jan et al., 2019; Jibran et al., 2013;

505 Lim et al., 2007), which makes it the first choice to study the leaf senescence process (see above). Moreover, while the Null model could be a good choice for predictions of leaf senescence dates (i.e., accuracy), the most suited models for predictions of leaf senescence trends (i.e., precision) may have to be identified yet.

We calibrated the compared models with the generalized simulated annealing algorithm and with model-specific controls (Sect. 2.4 and S2.1; Xiang et al., 1997, 2017). Algorithm and controls affect the accuracy of the calibrated

510 models (Meier and Bigler, 2023). Therefore, we used generalized simulated annealing, which was shown to yield accurate models of leaf phenology (Chuine et al., 1998; Meier and Bigler, 2023) and has been used by many studies to calibrate such models (e.g., Basler, 2016; Liu et al., 2019; Meier et al., 2018; Zani et al., 2020). In addition, we used model-specific controls selected to most accurately predict leaf senescence dates for same the validation samples (Sect. S2.2) as the comparison to the Null model was based on. Possible overfitting (James et al., 2017) through this procedure

515   would have benefited the compared models and is unlikely, as the number of observations in the calibration samples was large enough (Sect. 2.4; Jenkins and Quintana-Ascencio, 2020; Meier and Bigler, 2023). Therefore, it is highly improbable that this procedure caused the models to be calibrated so poorly that they are outperformed by the Null model.

All compared models were driven with daily weather data from the E-OBS dataset (Cornes et al., 2018) and calibrated and validated with leaf senescence data from the datasets of Meteo Swiss and PEP725 (Swiss phenology net-

520   work, 2025; Templ et al., 2018). The E-OBS dataset has been used by many studies (e.g., Bowling et al., 2024; Meng et al., 2021; Schwaab et al., 2021; Zeng and Wolkovich, 2024), and we are unaware of any difficulties concerning the daily weather data used here. The Meteo Swiss and PEP725 datasets, however, compile visually observed leaf senescence data, and such data is noisy due to different observers and small sample sizes (Liu et al., 2021): estimates of the leaf senescence dates for individual trees varied by 15 d (median, spreading from 2–53 d) between observers, and in-

525   creased to 28 d (median) for different samples of ten trees. The data become even noisier if the observers follow different protocols from various institutions and countries (Menzel, 2013), eventually blurring the signal of the leaf senescence process. Arguably the more this signal is blurred, the closer the simulations will follow the mean observation in the data. Here, we used leaf senescence data from 244 sites (i.e., probably at least 244 observers) and four countries (Sect. 2.1), which implies considerable noise and thus a blurred signal of the leaf development process. This data very

530   likely forced the compared models to predict leaf senescence dates close to the mean observation, impairing their accuracy.

4.3   Model error

While climatic and spatial deviations from the calibration sample generally affected the model error, their model-specific effects only differed insignificantly from the Null model. In other words, the model error in the compared models

535   reacted similarly to climatic and spatial deviations as the model error of the Null model. This implies that the compared models predicted leaf senescence dates closely to the mean observation of the calibration sample and thus were heavily biased to the mean (i.e., as the Null model). Possible explanations for this are unrealistic model formulations, poor model calibrations, and noisy data. Interestingly, Meier et al. (2023), who reported a heavy bias towards the mean for 21 process-oriented models of leaf senescence, based their study on leaf senescence data from 500 sites (i.e., probably at

540   least 500 observers) and at least three countries from the PEP725 dataset (Templ et al., 2018). This supports our inference that the compared models resorted to the mean observation due to the used leaf senescence data rather than to model formulations and model calibrations.

Moreover, leaf senescence data was most relevant for the model error in the compared models, which was illustrated by the fixed effects of countries and the variation caused by the random intercepts grouped by sites. These ef-

545   fects of countries differed considerably between countries, demonstrating how different observation protocols (see above; Menzel, 2013) add noise to leaf senescence data, which to our knowledge has not yet been investigated. The random intercepts grouped by sites varied considerably, and corresponding differences among sites were attributed to a substantial amount of the explained variance in the model error (Chpt. 23.3.2 in Fox, 2016). Meier et al. (2023) also noted a large amount of the explained variance in the RMSE being attributed to differences among sites. They reasoned

550   that this was caused by, among others, noisy leaf senescence data (see above) and different inter-annual variability of observations between the sites (Cole and Sheldon, 2017; Čufar et al., 2015; Li et al., 2022; Liu et al., 2020). It remains to be seen if such site-specific inter-annual variability as well as inter-site variability in leaf senescence dates would be predicted correctly by models calibrated with noise-free data.

4.4 Ways forward

555 While the DP3 model is likely the currently most realistic process-oriented model of leaf senescence, it may be developed further by (1) testing other drought indices, (2) considering nutrient depletion in combination with drought, and (3) ameliorating the formulation of the senescence rate. First, while various indices summarize drought differently (Speich, 2019; Zargar et al., 2011), the KBDI used here can be calculated from few data, being based on precipitation and temperature. It should be tested, however, if other indices, such as the standardized precipitation evapotranspiration index

560 (based on precipitation and temperature; Vicente-Serrano et al., 2010) or the ratio of actual to potential evapotranspiration (based on precipitation, temperature, and soil moisture; Bugmann and Cramer, 1998), may approximate the effects of dry stress on leaf senescence more accurately. Second, despite more accurate predictions of LS_{50} and LS_{100} when nutrient depletion was disregarded (Figs. 6 and S1), model errors indicated earlier observed than predicted LS_{50} and LS_{100} dates due to nutrient depletion as approximated by elevation (Fig. 9; Tables S9–S10). This can be explained by higher

565 elevation relating to increased nutrient depletion, which in turn fuels dry stress (Fu et al., 2014; Huber et al., 2007; Loomis et al., 2006; Tan et al., 2023). Consequently, drought indices that consider nutrient depletion should be tested. Third, the $DP3_{LS_{50}-LS_{100}}$ model was considerably less accurate than the $DP3_{LS_{50}}$ model, implying difficulties in the accurate, simultaneous prediction of LS_{50} and LS_{100}. This points to an incorrectly formulated dependency of the senescence rate on aging and stress (Eqs. 1 and 12), and corresponding new formulations should be evaluated.

570 In addition, because noisy data blur the signal of the leaf development process, (1) alternative data may be used, (2) observation protocols may be revised, and (3) visually observed data may be carefully selected. First, alternative data to calibrate and validate models of leaf senescence include data recorded with phenocams and remote sensed data in which leaf senescence dates are identified through the measured greenness, machine learning algorithms, and vegetation indices (Donnelly et al., 2022; Dronova and Taddeo, 2022; Gong et al., 2024; Richardson, 2023; Zeng et al.,

575 2020). While these data are species-specific if recorded with phenocams, this may not be the case for remote sensed data (Joiner et al., 2016; Tang et al., 2016). Second, revised observation protocols should describe how to determine dates of leaf senescence stages (i.e., senescence induction, LS_{50}, and LS_{100} at least) based on the measured, rather than estimated, state of leaf senescence. Such a measurement could be based on the greenness derived from images taken with consumer-grade digital cameras (Ide and Oguma, 2013; Richardson et al., 2018; Toomey et al., 2015; Zimmerman

580 and Richardson, 2024). Moreover, a given observational time series should be based on at least 25 trees which are measured every other week (Liu et al., 2021; Morellato et al., 2010). Third, visually observed leaf senescence data should be selected primarily from the point of view of precision, for example by ensuring identical observation protocols and by sampling from cleaned data with a minimum of breakpoints (i.e., sudden changes in the mean). For this, the time series may be cleaned from outliers (Schaber et al., 2010) and separated through a breakpoint analysis (Auchmann et al.,

585 2018), before being sampled, preferably through spatially and climatologically stratified sampling, according to the research focus (e.g., gaining insight in the underlying processes or producing most accurate or most precise predictions Meier and Bigler, 2023).

5 Conclusion

The DP3 model builds on three subsequent phases of leaf development: the young, mature, and old leaf phase. The young leaf is insensitive to stress and transfers into a mature leaf solely due to aging. The mature leaf answers to aging

590 and stress, both of which may induce senescence. While aging induces senescence with the transition from mature to old leaf, stress may already does so during the mature leaf phase through combining cold stress, photoperiod stress, and

595 dry stress. The output of the DP3 model includes daily rates of aging rates as well as of cold, photoperiod, and dry stress along with the dates of transition from young to mature to old leaf, senescence induction dates, and the leaf senescence dates. Thus, the DP3 model allows to develop testable hypotheses about the leaf senescence process, for example regarding the effect of site conditions on the induction and duration of senescence: the DP3 model predicted earlier onset of senescence (i.e., senescence induction) under warmer conditions, likely due to earlier leaf unfolding and thus aging, and increased cold and drought stress in spring, as well as longer-lasting senescence, likely due to the later accumulation of photoperiod stress relative to leaf development and reduced cold stress in summer and fall. Both these predictions and their implied relationships with aging and stress can be tested through experiments and in situ observations.

600 This makes the DP3 model an important tool in the research of leaf senescence.

The accuracy of the DP3 model and of previous models of leaf senescence was lower than the accuracy of the Null model (i.e. the constant prediction of the average observation in the calibration sample). This was probably due to model formulations that do not fully reflect the leaf senescence process and, more importantly, to the leaf senescence data used for calibration and validation. Visually observed leaf senescence data are susceptible to observer bias and

605 based on observation protocols that are partly inconsistent between countries. Such noisy data blur the signal of the leaf senescence process, thereby probably forcing the models to resort to the average observation. This leads to low accuracy, regardless from the model formulation, which hinders the necessary further development of process-oriented models of leaf senescence.

The model error of the compared models was similarly affected by climatic and spatial deviations from the

610 calibration sample across models, and varied mainly due to the leaf senescence data. The similar effect of climatic and spatial deviations on the model error across models (including the Null model) illustrates that these models were heavily biased towards the mean. Moreover, the degree of noise in the used leaf senescence data is exemplified by these data accounting for over 90% of the explained variance in the model error. Therefore, these data should be selected with particular attention to precision, e.g., by using time series without sudden changes in the mean. Moreover, revised observation

615 protocols should include senescence induction dates and rely on measurements rather than visual estimates. Such measurements may be based on the greenness of leaves to identify the degree of color change, involving digital cameras and automated image assessment.

Code and data availability

The R code for the DP3 model is openly available on Zenodo (Meier, 2025b, https://doi.org/10.5281/zenodo.14749339), together with the R code for the 2-phased version of the DP3 model ('DP2 model'), i.e., the DP3 model without young

620 leaf phase. While all raw data used are publicly available and referenced in section 2, the predicted leaf senescence dates analyzed and compared are openly accessible under https://doi.org/10.5061/dryad.tht76hf97 (Meier, 2025a).

Author contributions

MM, IC, and CB initialized the study and the model development, while the final study and model development were conceptualized MM and IC. MM designed the methodology and created the models, analyzed the models with input from IC and CB, visualized the results with input from IC, and wrote the draft with contributions from IC and CB. All

625 authors approved the final manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

We thank the two anonymous reviewers and Shilong Ren from the Egusphere community for the critical reading of our manuscript and the helpful comments. Further acknowledgment is given to the E-OBS dataset and the data providers in the ECA&D project (https://www.ecad.eu). All leaf senescence models during model development and for model evaluation were calibrated and validated on the high performance computing cluster at the University of Lausanne. This study was funded by the Swiss National Science Foundation SNSF (project 210744).

630

References

Addicott, F. T.: Environmental factors in the physiology of abscission, Plant Physiol., 43, 1471–9, https://doi.org/PMID: 16657013, PMCID: PMC1087141., 1968.

Akaike, H.: A new look at the statistical model identification, IEEE Trans. Autom. Control, 19, 716–723, https://doi.org/10.1109/TAC.1974.1100705, 1974.

Allen, R. G., Smith, M., Perrier, A., and Pereira, L. S.: An update for the definition of reference evapotranspiration, ICID Bull., 43, 1–34, 1994.

Asse, D., Chuine, I., Vitasse, Y., Yoccoz, N. G., Delpierre, N., Badeau, V., Delestrade, A., and Randin, C. F.: Warmer winters reduce the advance of tree spring phenology induced by warmer springs in the Alps, Agric. For. Meteorol., 252, 220–230, https://doi.org/10.1016/j.agrformet.2018.01.030, 2018.

Auchmann, R., Brugnara, Y., Brönnimann, S., Gehrig, R., Pietragalla, B., Begert, M., Sigg, C., Knechtl, V., Calpini, B., and Konzelmann, T.: Quality Analysis and Classification of Data Series from the Swiss Phenology Network, MeteoSwiss, Zürich-Flughafen, 2018.

Barrett, T., Dowle, M., Srinivasan, A., Gorecki, J., Chirico, M., and Hocking, T.: data.table: Extension of `data.frame`, 2024.

Basler, D.: Evaluating phenological models for the prediction of leaf-out dates in six temperate tree species across central Europe, Agric. For. Meteorol., 217, 10–21, https://doi.org/10.1016/j.agrformet.2015.11.007, 2016.

Berdanier, A. B. and Clark, J. S.: Tree water balance drives temperate forest responses to drought, Ecology, 99, 2506–2514, https://doi.org/10.1002/ecy.2499, 2018.

Bigler, C. and Bugmann, H.: Climate-induced shifts in leaf unfolding and frost risk of European trees and shrubs, Sci Rep, 8, 9865, https://doi.org/10.1038/s41598-018-27893-1, 2018.

Bigler, C. and Vitasse, Y.: Premature leaf discoloration of European deciduous trees is caused by drought and heat in late spring and cold spells in early fall, Agric. For. Meteorol., 307, 108492, https://doi.org/10.1016/j.agrformet.2021.108492, 2021.

Bivand, R. S., Pebesma, E., and Gomez-Rubio, V.: Applied spatial data analysis with R, Second edition., Springer, New York, USA, 2013.

Bowling, D. R., Schädel, C., Smith, K. R., Richardson, A. D., Bahn, M., Arain, M. A., Varlagin, A., Ouimette, A. P., Frank, J. M., Barr, A. G., Mammarella, I., Šigut, L., Foord, V., Burns, S. P., Montagnani, L., Litvak, M. E., Munger, J. W., Ikawa, H., Hollinger, D. Y., Blanken, P. D., Ueyama, M., Matteucci, G., Bernhofer, C., Bohrer, G., Iwata, H., Ibrom, A., Pilegaard, K., Spittlehouse, D. L., Kobayashi, H., Desai, A. R., Staebler, R. M., and Black, T. A.: Phenology of Photosynthesis in Winter-Dormant Temperate and Boreal Forests: Long-Term Observations From Flux Towers and Quantitative Evaluation of Phenology Models, J. Geophys. Res. Biogeosciences, 129, e2023JG007839, https://doi.org/10.1029/2023JG007839, 2024.

Brock, T. D.: Calculating solar radiation for ecological studies, Ecol. Model., 14, 1–19, https://doi.org/10.1016/0304-3800(81)90011-9, 1981.

Bugmann, H. and Cramer, W.: Improving the behaviour of forest gap models along drought gradients, For. Ecol. Manag., 103, 247–263, https://doi.org/10.1016/s0378-1127(97)00217-x, 1998.

Burnham, K. P. and Anderson, D. R.: Multimodel Inference: Understanding AIC and BIC in Model Selection, Sociol. Methods Res., 33, 261–304, https://doi.org/10.1177/0049124104268644, 2004.

Carley, K. M.: On generating hypotheses using computer simulations, Syst. Eng., 2, 69–77, https://doi.org/10.1002/(SICI)1520-6858(1999)2:2<69::AID-SYS3>3.0.CO;2-0, 1999.

Cheng, W., Dan, L., Deng, X., Feng, J., Wang, Y., Peng, J., Tian, J., Qi, W., Liu, Z., Zheng, X., Zhou, D., Jiang, S., Zhao, H., and Wang, X.: Global monthly gridded atmospheric carbon dioxide concentrations under the historical and future scenarios, Sci. Data, 9, 83, https://doi.org/10.1038/s41597-022-01196-7, 2022.

Chuine, I. and Régnière, J.: Process-based models of phenology for plants and animals, Annu. Rev. Ecol. Evol. Syst., 48, 159–182, https://doi.org/10.1146/annurev-ecolsys-110316-022706, 2017.

Chuine, I., Cour, P., and Rousseau, D. D.: Fitting models predicting dates of flowering of temperate-zone trees using simulated annealing, Plant Cell Environ., 21, 455–466, https://doi.org/10.1046/j.1365-3040.1998.00299.x, 1998.

Chuine, I., de Cortazar-Atauri, I. G., Kramer, K., and Hänninen, H.: Plant development models, in: Phenology: An integrative environmental science, edited by: Schwartz, M. D., Springer Netherlands, Dordrecht, 275–293, 2013.

Cole, E. F. and Sheldon, B. C.: The shifting phenological landscape: Within- and between-species variation in leaf emergence in a mixed-deciduous woodland, Ecol. Evol., 7, 1135–1147, https://doi.org/10.1002/ece3.2718, 2017.

Collatz, G. J., Ball, J. T., Grivet, C., and Berry, J. A.: Physiological and environmental-regulation of stomatal conductance, photosynthesis and transpiration - A model that includes laminar boundary-layer, Agric. For. Meteorol., 54, 107–136, https://doi.org/10.1016/0168-1923(91)90002-8, 1991.

Cooke, J. E. K. and Weih, M.: Nitrogen storage and seasonal nitrogen cycling in Populus: bridging molecular physiology and ecophysiology, New Phytol., 167, 19–30, https://doi.org/10.1111/j.1469-8137.2005.01451.x, 2005.

Copernicus Climate Change Service, Climate Data Store: Carbon dioxide data from 2002 to present derived from satellite observations. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), last access: 30 April 2024, https://doi.org/10.24381/cds.f74805c8, 2018.

Copernicus Climate Change Service, Climate Data Store: E-OBS daily gridded meteorological data for Europe from 1950 to present derived from in-situ observations. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), last access: 10 June 2023, https://doi.org/10.24381/cds.151d3ec6, 2020.

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M., and Jones, P. D.: An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets, J. Geophys. Res. Atmospheres, 123, 9391–9409, https://doi.org/10.1029/2017JD028200, 2018.

Čufar, K., De Luis, M., Prislan, P., Gricar, J., Crepinsek, Z., Merela, M., and Kajfez-Bogataj, L.: Do variations in leaf phenology affect radial growth variations in Fagus sylvatica?, Int J Biometeorol, 59, 1127–1132, https://doi.org/10.1007/s00484-014-0896-3, 2015.

Delpierre, N., Dufrene, E., Soudani, K., Ulrich, E., Cecchini, S., Boe, J., and Francois, C.: Modelling interannual and spatial variability of leaf senescence for three deciduous tree species in France, Agric. For. Meteorol., 149, 938–948, https://doi.org/10.1016/j.agrformet.2008.11.014, 2009.

Donnelly, A., Yu, R., Jones, K., Belitz, M., Li, B., Duffy, K., Zhang, X., Wang, J., Seyednasrollah, B., Gerst, K. L., Li, D., Kaddoura, Y., Zhu, K., Morisette, J., Ramey, C., and Smith, K.: Exploring discrepancies between in situ phenology and remotely derived phenometrics at NEON sites, Ecosphere, 13, e3912, https://doi.org/10.1002/ecs2.3912, 2022.

Dronova, I. and Taddeo, S.: Remote sensing of phenology: Towards the comprehensive indicators of plant community dynamics from species to regional scales, J. Ecol., 110, 1460–1484, https://doi.org/10.1111/1365-2745.13897, 2022.

Dufrêne, E., Davi, H., Francois, C., le Maire, G., Le Dantec, V., and Granier, A.: Modelling carbon and water cycles in a beech forest: Part I: Model description and uncertainty analysis on modelled NEE, Ecol. Model., 185, 407–436, https://doi.org/10.1016/j.ecolmodel.2005.01.004, 2005.

Dunnington, D.: ggspatial: Spatial data framework for ggplot2, 2023.

PEP725: http://www.pep725.eu/, last access: 25 June 2024.

EU-DEM: https://ec.europa.eu/eurostat/web/gisco/geodata/digital-elevation-model/eu-dem, last access: 25 July 2024.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species, Planta, 149, 78–90, https://doi.org/10.1007/bf00386231, 1980.

Field, C. and Mooney, H. A.: Leaf age and seasonal effects on light, water, and nitrogen use efficiency in a California shrub, Oecologia, 56, 348–355, https://doi.org/10.1007/BF00379711, 1983.

Fox, J.: Applied regression analysis and generalized linear models, 3rd ed., Sage Publications, Thousand Oaks, California, USA, 2016.

Fox, J. and Weisberg, S.: An R Companion to Applied Regression, Third., SAGE, Los Angeles, 2019.

Fu, Y. H., Campioli, M., Vitasse, Y., De Boeck, H. J., Van den Berge, J., AbdElgawad, H., Asard, H., Piao, S. L., Deckmyn, G., and Janssens, I. A.: Variation in leaf flushing date influences autumnal senescence and next year's flushing date in two temperate tree species, Proc. Natl. Acad. Sci. U. S. A., 111, 7355–7360, https://doi.org/10.1073/pnas.1321727111, 2014.

Fu, Y. H., Piao, S. L., Delpierre, N., Hao, F. H., Hanninen, H., Geng, X. J., Penuelas, J., Zhang, X., Janssens, I. A., and Campioli, M.: Nutrient availability alters the correlation between spring leaf-out and autumn leaf senescence dates, Tree Physiol., 39, 1277–1284, https://doi.org/10.1093/treephys/tpz041, 2019.

Gong, Z., Ge, W., Guo, J., and Liu, J.: Satellite remote sensing of vegetation phenology: Progress, challenges, and opportunities, ISPRS J. Photogramm. Remote Sens., 217, 149–164, https://doi.org/10.1016/j.isprsjprs.2024.08.011, 2024.

Grolemund, G. and Wickham, H.: Dates and Times Made Easy with lubridate, J. Stat. Softw., 40, 1–25, https://doi.org/10.18637/jss.v040.i03, 2011.

Guo, Y., Ren, G., Zhang, K., Li, Z., Miao, Y., and Guo, H.: Leaf senescence: progression, regulation, and application, Mol. Hortic., 1, 5, https://doi.org/10.1186/s43897-021-00006-9, 2021.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hauke, J., Achter, S., and Meyer, M.: Theory Development Via Replicated Simulations and the Added Value of Standards, J. Artif. Soc. Soc. Simul., 23, 12, https://doi.org/10.18564/jasss.4219, 2020.

Held, L. and Ott, M.: On p-Values and Bayes Factors, Annu. Rev. Stat. Its Appl., 5, 393–419, https://doi.org/10.1146/annurev-statistics-031017-100307, 2018.

Helwig, N. E.: npreg: Nonparametric Regression via Smoothing Splines, 2024.

Herms, D. A. and Mattson, W. J.: The dilemma of plants - To grow or defend, Q. Rev. Biol., 67, 283–335, https://doi.org/10.1086/417659, 1992.

Hijmans, R. J.: raster: Geographic Data Analysis and Modeling, 2023.

Huber, E., Wanek, W., Gottfried, M., Pauli, H., Schweiger, P., Arndt, S. K., Reiter, K., and Richter, A.: Shift in soil–plant nitrogen dynamics of an alpine–nival ecotone, Plant Soil, 301, 65–76, https://doi.org/10.1007/s11104-007-9422-2, 2007.

Ide, R. and Oguma, H.: A cost-effective monitoring method using digital time-lapse cameras for detecting temporal and spatial variations of snowmelt and vegetation phenology in alpine ecosystems, Ecol. Inform., 16, 25–34, https://doi.org/10.1016/j.ecoinf.2013.04.003, 2013.

James, G., Witten, D., Hastie, T., and Tibishirani, R.: An introduction to statistical learning: with applications in R, Springer, New York a.o., 2017.

Jan, S., Abbas, N., Ashraf, M., and Ahmad, P.: Roles of potential plant hormones and transcription factors in controlling leaf senescence and drought tolerance, Protoplasma, 256, 313–329, https://doi.org/10.1007/s00709-018-1310-5, 2019.

Jenkins, D. G. and Quintana-Ascencio, P. F.: A solution to minimum sample size for regressions, PLoS One, 15, e0229345, https://doi.org/10.1371/journal.pone.0229345, 2020.

Jibran, R., Hunter, D. A., and Dijkwel, P. P.: Hormonal regulation of leaf senescence through integration of developmental and stress signals, Plant Mol. Biol., 82, 547–561, https://doi.org/10.1007/s11103-013-0043-2, 2013.

Johnson, V. E.: Bayes Factors based on test statistics, J. R. Stat. Soc. Ser. B Stat. Methodol., 67, 689–701, 2005.

Joiner, J., Yoshida, Y., Guanter, L., and Middleton, E. M.: New methods for the retrieval of chlorophyll red fluorescence from hyperspectral satellite instruments: simulations and application to GOME-2 and SCIAMACHY, Atmospheric Meas. Tech., 9, 3939–3967, https://doi.org/10.5194/amt-9-3939-2016, 2016.

Kassambara, A.: ggpubr: "ggplot2" Based Publication Ready Plots, 2020.

Keenan, T. F. and Richardson, A. D.: The timing of autumn senescence is affected by the timing of spring phenology: implications for predictive models, Glob Chang Biol, 21, 2634–2641, https://doi.org/10.1111/gcb.12890, 2015.

Keetch, J. J. and Byram, G. M.: A Drought Index for Forest Fire Control, Department of Agriculture, Forest Service, Southeastern Forest Experiment Station, Asheville, NC: U.S., 1968.

Keskitalo, J., Bergquist, G., Gardestrom, P., and Jansson, S.: A cellular timetable of autumn senescence, Plant Physiol., 139, 1635–1648, https://doi.org/10.1104/pp.105.066845, 2005.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, J. Hydrol., 424–425, 264–277, https://doi.org/10.1016/j.jhydrol.2012.01.011, 2012.

Kloos, S., Klosterhalfen, A., Knohl, A., and Menzel, A.: Decoding autumn phenology: Unraveling the link between observation methods and detected environmental cues, Glob. Change Biol., 30, e17231, https://doi.org/10.1111/gcb.17231, 2024.

Lang, W., Chen, X., Qian, S., Liu, G., and Piao, S.: A new process-based model for predicting autumn phenology: How is leaf senescence controlled by photoperiod and temperature coupling?, Agric. For. Meteorol., 268, 124–135, https://doi.org/10.1016/j.agrformet.2019.01.006, 2019.

Li, S., Wang, Y., Ciais, P., Sitch, S., Sato, H., Shen, M., Chen, X., Ito, A., Wu, C., Kucharik, C. J., and Yuan, W.: Deficiencies of Phenology Models in Simulating Spatial and Temporal Variations in Temperate Spring Leaf Phenology, J. Geophys. Res. Biogeosciences, 127, e2021JG006421, https://doi.org/10.1029/2021JG006421, 2022.

Lim, P. O., Kim, H. J., and Gil Nam, H.: Leaf senescence, Annu. Rev. Plant Biol., 58, 115–136, https://doi.org/10.1146/annurev.arplant.57.032905.105316, 2007.

Liu, G., Chen, X. Q., Zhang, Q. H., Lang, W. G., and Delpierre, N.: Antagonistic effects of growing season and autumn temperatures on the timing of leaf coloration in winter deciduous trees, Glob. Change Biol., 24, 3537–3545, https://doi.org/10.1111/gcb.14095, 2018.

Liu, G., Chen, X. Q., Fu, Y. S., and Delpierre, N.: Modelling leaf coloration dates over temperate China by considering effects of leafy season climate, Ecol. Model., 394, 34–43, https://doi.org/10.1016/j.ecolmodel.2018.12.020, 2019.

Liu, G., Chuine, I., Denechere, R., Jean, F., Dufrene, E., Vincent, G., Berveiller, D., and Delpierre, N.: Higher sample sizes and observer inter-calibration are needed for reliable scoring of leaf phenology in trees, J. Ecol., 109, 2461–2474, https://doi.org/10.1111/1365-2745.13656, 2021.

Liu, Q., Piao, S. L., Campioli, M., Gao, M. D., Fu, Y. S. H., Wang, K., He, Y., Li, X. Y., and Janssens, I. A.: Modeling leaf senescence of deciduous tree species in Europe, Glob. Change Biol., 26, 4104–4118, https://doi.org/10.1111/gcb.15132, 2020.

Loomis, P. F., Ruess, R. W., Sveinbjörnsson, B., and Kielland, K.: Nitrogen cycling at treeline: Latitudinal and elevational patterns across a boreal landscape, Ecoscience, 13, 544–556, https://doi.org/10.2980/1195-6860(2006)13[544:NCATLA]2.0.CO;2, 2006.

Lu, X. and Keenan, T. F.: No evidence for a negative effect of growing season photosynthesis on leaf senescence timing, Glob. Change Biol., 28, 3083–3093, https://doi.org/10.1111/gcb.16104, 2022.

Luo, Y.: Terrestrial carbon-cycle feedback to climate warming, Annu. Rev. Ecol. Evol. Syst., 38, 683–712, https://doi.org/10.1146/annurev.ecolsys.38.091206.095808, 2007.

Mao, J. and Yan, B.: Global monthly mean leaf area index climatology, 1981-2015 (1), https://doi.org/10.3334/ORNLDAAC/1653, 2019.

Mariën, B., Dox, I., De Boeck, H. J., Willems, P., Leys, S., Papadimitriou, D., and Campioli, M.: Does drought advance the onset of autumn leaf senescence in temperate deciduous forest trees?, Biogeosciences, 18, 3309–3330, https://doi.org/10.5194/bg-18-3309-2021, 2021.

Marqués, L., Hufkens, K., Bigler, C., Crowther, T. W., Zohner, C. M., and Stocker, B. D.: Acclimation of phenology relieves leaf longevity constraints in deciduous forests, Nat. Ecol. Evol., 7, 198–204, https://doi.org/10.1038/s41559-022-01946-1, 2023.

Massicotte, P. and South, A.: rnaturalearth: World Map Data from Natural Earth, 2023.

Meier, M.: Simulated and observed timings of leaf senescence for Fagus sylvatica (L.) at selected European sites, https://doi.org/10.5061/dryad.tht76hf97, 2025a.

Meier, M.: The DP3 model (v1.1) - Simulating the timing of leaf senescence, , Zenodo, https://doi.org/10.5281/zenodo.14749339, 2025b.

Meier, M. and Bigler, C.: Process-oriented models of autumn leaf phenology: ways to sound calibration and implications of uncertain projections, Geosci. Model Dev., 16, 7171–7201, https://doi.org/10.5194/gmd-16-7171-2023, 2023.

Meier, M., Fuhrer, J., and Holzkaemper, A.: Changing risk of spring frost damage in grapevines due to climate change? A case study in the Swiss Rhone Valley, Int. J. Biometeorol., 62, 991–1002, https://doi.org/10.1007/s00484-018-1501-y, 2018.

Meier, M., Vitasse, Y., Bugmann, H., and Bigler, C.: Phenological shifts induced by climate change amplify drought for broad-leaved trees at low elevations in Switzerland, Agric. For. Meteorol., 307, 108485, https://doi.org/10.1016/j.agrformet.2021.108485, 2021.

Meier, M., Bugmann, H., and Bigler, C.: Process-oriented models of leaf senescence are biased towards the mean: Impacts on model performance and future projections, Glob. Change Biol., 30, https://doi.org/10.1111/gcb.17099, 2023.

Meier, U.: Growth stages of mono- and dicotyledonous plants: BBCH Monograph, Open Agrar Repositorium, Quedlinburg, 2018.

Meng, L., Zhou, Y. Y., Gu, L. H., Richardson, A. D., Penuelas, J., Fu, Y. S., Wang, Y. Q., Asrar, G. R., De Boeck, H. J., Mao, J. F., Zhang, Y. G., and Wang, Z. S.: Photoperiod decelerates the advance of spring phenology of six deciduous tree species under climate warming, Glob. Change Biol., 27, 2914–2927, https://doi.org/10.1111/gcb.15575, 2021.

Menzel, A.: Europe, in: Phenology: An Integrative Environmental Science, edited by: Schwartz, M. D., Springer Netherlands, Dordrecht, 53–65, https://doi.org/10.1007/978-94-007-6925-0_4, 2013.

Menzel, A., Yuan, Y., Matiu, M., Sparks, T., Scheifinger, H., Gehrig, R., and Estrella, N.: Climate change fingerprints in recent European plant phenology, Glob. Change Biol., 26, 14, https://doi.org/10.1111/gcb.15000, 2020.

Swiss phenology network: https://www.meteoswiss.admin.ch/weather/measurement-systems/land-based-stations/swiss-phenology-network.html, last access: 29 January 2025.

Morellato, L. P. C., Camargo, M. G. G., D'Eça Neves, F. F., Luize, B. G., Mantovani, A., and Hudson, I. L.: The Influence of Sampling Method, Sample Size, and Frequency of Observations on Plant Phenological Patterns and Interpretation in Tropical Forest Trees, in: Phenological Research: Methods for Environmental and Climate Change Analysis, edited by: Hudson, I. L. and Keatley, M. R., Springer Netherlands, Dordrecht, 99–121, 2010.

Norby, R. J.: Comment on "Increased growing-season productivity drives earlier autumn leaf senescence in temperate trees," Science, 371, eabg1438, https://doi.org/10.1126/science.abg1438, 2021.

Paul, M. J. and Foyer, C. H.: Sink regulation of photosynthesis, J. Exp. Bot., 52, 1383–1400, https://doi.org/10.1093/jexbot/52.360.1383, 2001.

Pebesma, E.: Simple features for R: Standardized support for spatial vector data, R J., 10, 439–446, https://doi.org/10.32614/RJ-2018-009, 2018.

Pebesma, E. and Bivand, R. S.: Classes and methods for spatial data in R, R News, 5, 9–13, 2005.

Pebesma, E. and Bivand, R. S.: Spatial Data Science: With Applications in R, Chapman HallCRC, https://doi.org/10.1201/9780429459016, 2023.

Piao, S. L., Liu, Q., Chen, A. P., Janssens, I. A., Fu, Y. S., Dai, J. H., Liu, L. L., Lian, X., Shen, M. G., and Zhu, X. L.: Plant phenology and global climate change: Current progresses and challenges, Glob. Change Biol., 25, 1922–1940, https://doi.org/10.1111/gcb.14619, 2019.

Pierce, D.: ncdf4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files, 2023.

Pinheiro, J. C. and Bates, D. M.: Mixed-effects models in S and S-PLUS, Springer, New York, 528 S. pp., 2000.

R Core Team: R: A language and environment for statistical computing, 2025.

Richardson, A. D.: PhenoCam: An evolving, open-source tool to study the temporal and spatial variability of ecosystem-scale phenology, Agric. For. Meteorol., 342, 109751, https://doi.org/10.1016/j.agrformet.2023.109751, 2023.

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., and Toomey, M.: Climate change, phenology, and phenological control of vegetation feedbacks to the climate system, Agric. For. Meteorol., 169, 156–173, https://doi.org/10.1016/j.agrformet.2012.09.012, 2013.

Richardson, A. D., Hufkens, K., Milliman, T., Aubrecht, D. M., Chen, M., Gray, J. M., Johnston, M. R., Keenan, T. F., Klosterman, S. T., Kosmala, M., Melaas, E. K., Friedl, M. A., and Frolking, S.: Tracking vegetation phenology across diverse North American biomes using PhenoCam imagery, Sci. Data, 5, 24, https://doi.org/10.1038/sdata.2018.28, 2018.

Rogers, H. J.: Leaf senescence, in: Encyclopedia of Applied Plant Sciences (Second Edition), vol. 1, edited by: Thomas, B., Murray, B. G., and Murphy, D. J., Academic Press, Amsterdam, 308–314, 2017.

Sangüesa-Barreda, G., Di Filippo, A., Piovesan, G., Rozas, V., Di Fiore, L., García-Hidalgo, M., García-Cervigón, A. I., Muñoz-Garachana, D., Baliva, M., and Olano, J. M.: Warmer springs have increased the frequency and extension of late-frost defoliations in southern European beech forests, Sci. Total Environ., 775, 145860, https://doi.org/10.1016/j.scitotenv.2021.145860, 2021.

Schaber, J., Badeck, F., Doktor, D., and von Bloh, W.: Combining messy phenological time series, in: Phenological research: methods for environmental and climate change analysis, edited by: Hudson, I. L. and Keatley, M. R., Springer, Dordrecht Heidelberg London New York, 147–158, 2010.

Schwaab, J., Meier, R., Mussetti, G., Seneviratne, S., Bürgi, C., and Davin, E. L.: The role of urban trees in reducing land surface temperatures in European cities, Nat. Commun., 12, 6763, https://doi.org/10.1038/s41467-021-26768-w, 2021.

Singh, R. K., Svystun, T., AlDahmash, B., Jönsson, A. M., and Bhalerao, R. P.: Photoperiod- and temperature-mediated control of phenology in trees – a molecular perspective, New Phytol., 213, 511–524, https://doi.org/10.1111/nph.14346, 2017.

Speich, M. J. R.: Quantifying and modeling water availability in temperate forests: a review of drought and aridity indices, Iforest-Biogeosciences For., 12, 1–16, https://doi.org/10.3832/ifor2934-011, 2019.

Tan, S., Sha, Y., Sun, L., and Li, Z.: Abiotic Stress-Induced Leaf Senescence: Regulatory Mechanisms and Application, Int. J. Mol. Sci., 24, https://doi.org/10.3390/ijms241511996, 2023.

Tang, J. W., Körner, C., Muraoka, H., Piao, S. L., Shen, M. G., Thackeray, S. J., and Yang, X.: Emerging opportunities and challenges in phenology: a review, Ecosphere, 7, 17, https://doi.org/10.1002/ecs2.1436, 2016.

Templ, B., Koch, E., Bolmgren, K., Ungersbock, M., Paul, A., Scheifinger, H., Rutishauser, T., Busto, M., Chmielewski, F. M., Hajkova, L., Hodzic, S., Kaspar, F., Pietragalla, B., Romero-Fresneda, R., Tolvanen, A., Vucetic, V., Zimmermann, K., and Zust, A.: Pan European Phenological database (PEP725): A single point of access for European data, Int J Biometeorol, 62, 1109–1113, https://doi.org/10.1007/s00484-018-1512-8, 2018.

Toomey, M., Friedl, M. A., Frolking, S., Hufkens, K., Klosterman, S., Sonnentag, O., Baldocchi, D. D., Bernacchi, C. J., Biraud, S. C., Bohrer, G., Brzostek, E., Burns, S. P., Coursolle, C., Hollinger, D. Y., Margolis, H. A., McCaughey, H., Monson, R. K., Munger, J. W., Pallardy, S., Phillips, R. P., Torn, M. S., Wharton, S., Zeri, M., and Richardson, A. D.: Greenness indices from digital cameras predict the timing and seasonal dynamics of canopy-scale photosynthesis, Ecol. Appl., 25, 99–115, https://doi.org/10.1890/14-0005.1, 2015.

Van der Meersch, V. and Chuine, I.: Can inverse calibration help improving process-explicit species distribution models?, Ecol. Model., 506, 111132, https://doi.org/10.1016/j.ecolmodel.2025.111132, 2025.

Vicente-Serrano, S. M., Begueria, S., and Lopez-Moreno, J. I.: A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index, J. Clim., 23, 1696–1718, https://doi.org/10.1175/2009jcli2909.1, 2010.

Vitasse, Y., Delzon, S., Dufrene, E., Pontailler, J. Y., Louvet, J. M., Kremer, A., and Michalet, R.: Leaf phenology sensitivity to temperature in European trees: Do within-species populations exhibit similar responses?, Agric. For. Meteorol., 149, 735–744, https://doi.org/10.1016/j.agrformet.2008.10.019, 2009.

Vitasse, Y., Hoch, G., Randin, C. F., Lenz, A., Kollas, C., Scheepens, J. F., and Korner, C.: Elevational adaptation and plasticity in seedling phenology of temperate deciduous tree species, Oecologia, 171, 663–678, https://doi.org/10.1007/s00442-012-2580-9, 2013.

Wang, H., Gao, C., and Ge, Q.: Low temperature and short daylength interact to affect the leaf senescence of two temperate tree species, Tree Physiol., 42, 2252–2265, https://doi.org/10.1093/treephys/tpac068, 2022.

Wang, J. and Liu, D.: Larger diurnal temperature range undermined later autumn leaf senescence with warming in Europe, Glob. Ecol. Biogeogr., 32, 734–746, https://doi.org/10.1111/geb.13674, 2023.

Wasserman, L.: All of Statistics. A Concise Course in Statistical Inference, 1st ed., Springer New York, NY, 2004.

Wheeler, K. I. and Dietze, M. C.: A trigger may not be necessary to cause senescence in deciduous broadleaf forests. (Preprint), bioRxiv, 2023.06.07.544057, https://doi.org/10.1101/2023.06.07.544057, 2023.

Wickham, H.: ggplot2: Elegant graphics for data analysis, Springer-Verlag, New York, 2016.

Wickham, H., Hester, J., and Bryan, J.: readr: Read Rectangular Text Data, 2024.

Wood, S. N.: Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models, J. R. Stat. Soc. Ser. B Stat. Methodol., 73, 3–36, https://doi.org/10.1111/j.1467-9868.2010.00749.x, 2011.

Wood, S. N.: Generalized additive models: An introduction with R, 2nd edition., Chapman and Hall/CRC, New York, 2017.

Wu, C., Peng, J., Ciais, P., Peñuelas, J., Wang, H., Beguería, S., Andrew Black, T., Jassal, R. S., Zhang, X., Yuan, W., Liang, E., Wang, X., Hua, H., Liu, R., Ju, W., Fu, Y. H., and Ge, Q.: Increased drought effects on the phenology of autumn leaf senescence, Nat. Clim. Change, 12, 943–949, https://doi.org/10.1038/s41558-022-01464-9, 2022.

Xiang, Y., Sun, D. Y., Fan, W., and Gong, X. G.: Generalized Simulated Annealing algorithm and its application to the Thomson model, Phys. Lett. A, 233, 216–220, https://doi.org/10.1016/s0375-9601(97)00474-x, 1997.

Xiang, Y., Gubian, S., Martin, F., Suomela, B., and Hoeng, J.: Generalized Simulated Annealing for Global Optimization: The GenSA Package, R J., 5, 13–28, https://doi.org/10.32614/RJ-2013-002, 2013.

Xiang, Y., Gubian, S., and Martin, F.: Generalized Simulated Annealing, in: Computational Optimization in Engineering - Paradigms and Applications, 25–46, 2017.

Xie, Y., Wang, X. J., and Silander, J. A.: Deciduous forest responses to temperature, precipitation, and drought imply complex climate change impacts, Proc. Natl. Acad. Sci. U. S. A., 112, 13585–13590, https://doi.org/10.1073/pnas.1509991112, 2015.

Xie, Y., Wang, X. J., Wilson, A. M., and Silander, J. A.: Predicting autumn phenology: How deciduous tree species respond to weather stressors, Agric. For. Meteorol., 250, 127–137, https://doi.org/10.1016/j.agrformet.2017.12.259, 2018.

Yates, F.: The analysis of multiple classifications with unequal numbers in the different classes, J. Am. Stat. Assoc., 29, 51–66, https://doi.org/10.2307/2278459, 1934.

Yu, H., Mo, Z., Tan, T., Qiu, L., and Guo, L.: Increased discrepancy between autumn leaf senescence and the end of photosynthesis influenced by climate and spring phenology.

Zani, D., Crowther, T. W., Mo, L., Renner, S. S., and Zohner, C. M.: Increased growing-season productivity drives earlier autumn leaf senescence in temperate trees, Science, 370, 1066–1071, https://doi.org/10.1126/science.abd8911, 2020.

Zargar, A., Sadiq, R., Naser, B., and Khan, F. I.: A review of drought indices, Environ. Rev., 19, 333–349, https://doi.org/10.1139/a11-013, 2011.

Zeileis, A. and Grothendieck, G.: zoo: S3 Infrastructure for Regular and Irregular Time Series, J. Stat. Softw., 14, 1–27, https://doi.org/10.18637/jss.v014.i06, 2005.

Zeng, L., Wardlow, B. D., Xiang, D., Hu, S., and Li, D.: A review of vegetation phenological metrics extraction using time-series, multispectral satellite data, Remote Sens. Environ., 237, 111511, https://doi.org/10.1016/j.rse.2019.111511, 2020.

Zeng, Z. A. and Wolkovich, E. M.: Weak evidence of provenance effects in spring phenology across Europe and North America, New Phytol., 242, 1957–1964, https://doi.org/10.1111/nph.19674, 2024.

Zimmerman, O. R. and Richardson, A. D.: Near-Surface Sensor-Derived Phenology, in: Phenology: An Integrative Environmental Science, edited by: Schwartz, M. D., Springer Nature Switzerland, Cham, 461–478, https://doi.org/10.1007/978-3-031-75027-4_20, 2024.

Zohner, C. M., Mirzagholi, L., Renner, S. S., Mo, L., Rebindaine, D., Bucher, R., Palouš, D., Vitasse, Y., Fu, Y. H., Stocker, B. D., and Crowther, T. W.: Effect of climate warming on the timing of autumn leaf senescence reverses after the summer solstice, Science, 381, eadf5098, https://doi.org/10.1126/science.adf5098, 2023.