https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

An example of how data quality hinders progress: translating the latest findings on the regulation of leaf senescence timing in trees into the DP3 model (v1.0)

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

[1] Centre d'Ecologie Foncitonelle et Evolutive, CNRS, Montpellier, France

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

Correspondence to: Michael Meier, (michael.meier@cefe.cnrs.fr)

Short summary. Formulated according to the leaf development process, the DP3 model of leaf coloring considerably contrasts previous models and allows to set up new hypotheses, e.g. on aging versus stress caused color changes. The DP3 model was as accurate as previous models and a comparison to the constant simulation of the mean date of leaf coloring indicated that noisy leaf coloring data forced the models to resort to this mean, which hinders model evaluation.

Abstract. The timing of leaf senescence ends the growing season of deciduous trees, affecting the amount of atmospheric CO_2 sequestered by forests. Some climate models integrate the timing of leaf senescence, which can be simulated with process-oriented models. Here, we developed a process-oriented model of leaf senescence (the 'DP3 model') by testing 34 formulations of the leaf development process. The period between leaf unfolding and leaf senescence was separated into three subsequent phases with particular reactions to aging and stress, (sum of cold, photoperiod, and dry stress). The DP3 model and the compared previous models were equally accurate, but less accurate than the Null model (i.e., constant simulation of the mean observation of the calibration sample). This lower accuracy was very likely due to noise in the visually observed leaf senescence data, which blurred the signal of the process of leaf senescence, and incorrect model formulations. The leaf senescence data were attributed to most of the variation in the model error of the models compared, which was similarly affected by climatic and spatial deviations from the calibration sample across models. The DP3 model considerably contrasts previous models, allowing the development of new hypotheses, e.g. on the cause of senescence induction. Independently from model formulation, noisy leaf senescence data likely force the models to resort to the mean observation, impeding inferences from accuracy-based model comparisons about the process of leaf senescence. This implies the usage of data from as few sites as possible to minimize the noise due to different observers and small sample sizes when evaluating and further developing models of leaf senescence. Moreover, revised observation protocols should explain how to measure rather than to estimate the timing leaf senescence, e.g., based on greenness, involving digital cameras and automated image assessment.

Keywords: DP3 model, process-based, leaf senescence, observation protocols, observer bias, sample bias, bias towards the mean

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

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 nutrient retraction 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 the term 'leaf senescence' to refer to the stage when a given relative amount of leaves have changed color or have fallen, unless stated otherwise.

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 the timing of leaf senescence under a changing climate are necessary for accurate forecasts of both climate change and future species composition of temperate forests.

The timing of leaf senescence is 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 the timing of leaf unfolding, or on temperature, precipitation, and photosynthetic activity 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 individuals 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 breaks 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).

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

The processes relevant to leaf senescence have 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). 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., 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, senescence cannot be induced, whereas aging and stress induce it in later phases and regulate the rate of senescence (Jan et al., 2019; Jibran et al., 2013; Lim et al., 2007; Paul and Foyer, 2001; Tan et al., 2023).

Here, we developed a new process-oriented model that simulates the timing of leaf senescence based on the latest knowledge of the physiological processes and drivers of leaf senescence. The timing of 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 simulated and observed values (i.e., 'model errors'). We addressed the following research questions:

(1) Which model formulation most accurately simulates the relationship between leaf development and the timing of leaf senescence?
(2) How accurately does this model simulate leaf senescence 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, Switzerland, and the United Kingdom between 1950 and 2022 (Fig. 1, Table 1; PEP725, 2024; Swiss phenology network, 2025; Templ et al., 2018). We used the phenological stages 50% of the leaves are 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). The LS_{100} data were recorded in Austria and the United Kingdom 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.

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

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 232 and 328 days 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.

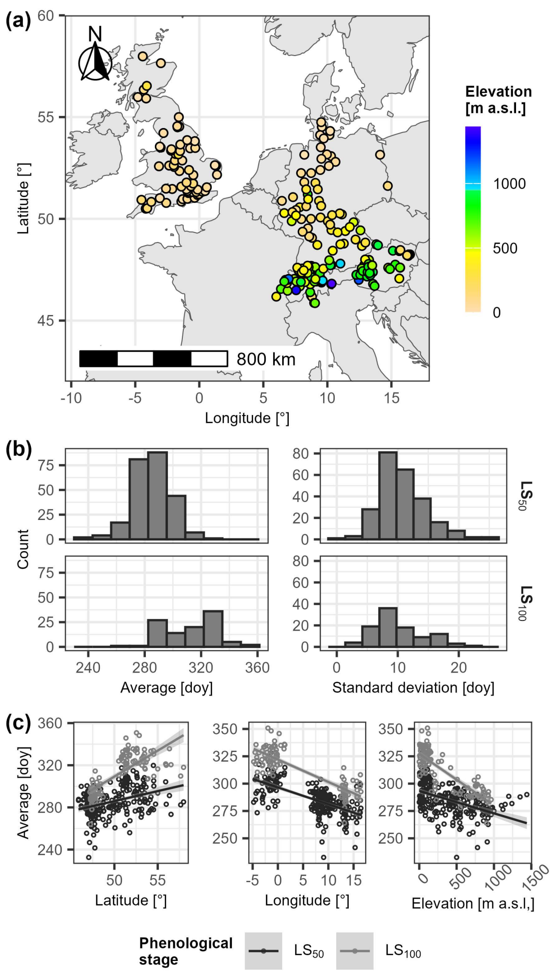


Figure 1. Selected phenological sites. Panel (a) locates the selected sites and indicates corresponding elevation [meters above seal level (m a.s.l.)]. In (b), the histograms illustrate the distributions of the site-specific average day of year (left) and corresponding standard deviation (right) per phenological 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 day of year of LS_{50} and LS_{100} (grey and black circles, respectively) in relation to site latitude [°] (left), longitude [°] (middle), and elevation [m a.s.l.] (right), together with the linear regression line and corresponding 95% confidence interval.

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

Table 1. Observations of spring and autumn 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
	Switzerland	61	2585 (6–72)	1951–2022	197–344	SPN
	United Kingdom	64	303 (2–6)	1999–2005	258–337	PEP725
LS_{100}	Austria	43	578 (1–34)	1950–1986	263–335	PEP725
	United Kingdom	63	272 (1–6)	1999–2005	286–365	PEP725

Note: 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. LS_{50} and LS_{100} refer to the stages when 50% and 100% of the leaves, respectively, have changed color or fallen. The timing of these stages is given by the day of year (doy). Two data sources were considered: PEP725 (Templ et al., 2018) and the Swiss phenological network (SPN; Swiss phenology network, 2025).

2.2 Driver data

We derived daily weather variables and the elevation for each site from the E-OBS dataset (Cornes et al., 2018), approximating site elevation, maximum temperature, mean temperature, minimum temperature, precipitation, relative humidity, and surface shortwave down welling radiation for 1950–2022 (Copernicus Climate Change Service, Climate Data Store, 2020) by the weighted averages from octagons with a radius of 2.5 km around the sites. The temperature variables were further corrected for the 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 corresponding weighted average elevation (i.e., according to the E-OBS dataset) with day- and site-specific lapse rates. We linearly regressed these lapse rates from the grid cell of a particular site and the eight neighboring grid cells, assuming an elevation of 0 m a.s.l. for grid cells over the sea. Occasional gaps in the regressed lapse rates were interpolated with site-specific cubic splines. Atmospheric CO_2 concentrations were taken from a reconstructed and a remote sensed dataset for the years 1950–2013 and 2002–2022, respectively (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. Due to missing observations for 2002–2022, we used modeled data derived from site-specific cubic splines based on the remote sensed data. The weighted average leaf area index (LAI) per site was taken from the remote sensed monthly mean LAI (1981–2015) in the GIMMS-LAI3g dataset (version 2; Mao and Yan, 2019).

We further calculated for each site day length, daily photosynthetic activity, and the daily Keetch and Byram drought index. 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 surface shortwave down welling radiation, day length, mean temperature, atmospheric CO_2 concentration, and leaf area index (Sect. S1.2.2). The daily Keetch and Byram drought index (KBDI) was calculated following 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 ‘young leaf’, ‘mature leaf’, and ‘old leaf’ (Fig. 2). After leaf unfolding, the young leaf is insensitive to stress and reaches the phase of mature leaf

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

through aging only (Fu et al., 2014; Jibran et al., 2013; Keenan and Richardson, 2015). The mature leaf continues to age and can be affected by stress until it reaches the phase of an old leaf (Jan et al., 2019; Jibran et al., 2013; Lim et al., 2007). The old leaf finally senescences until it changes color and eventually falls off (Jan et al., 2019; Jibran et al., 2013; Lim et al., 2007).

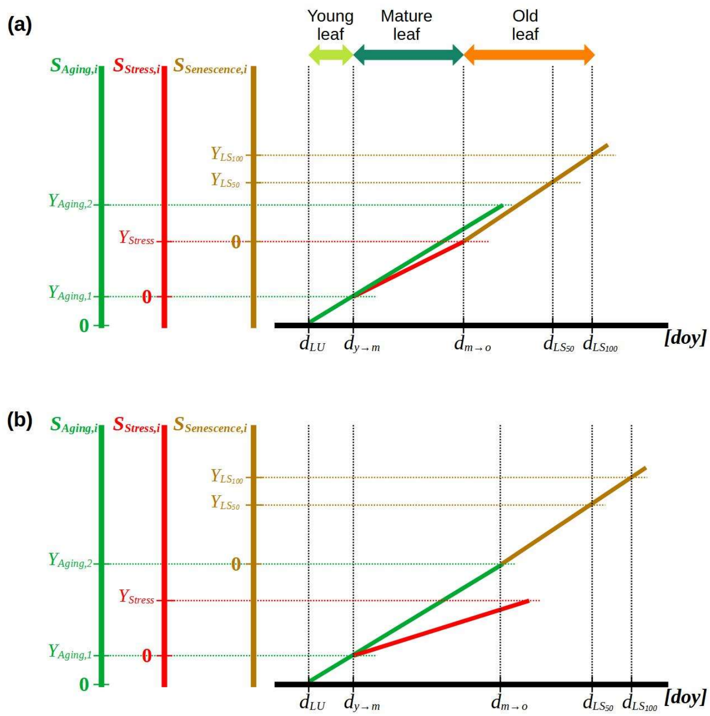


Figure 2. The leaf development process. This process is defined by three subsequent phases of leaf development, i.e., ‘young leaf’, ‘mature leaf’, and ‘old leaf’, with particular responses to aging and stress, both of which may induce leaf senescence [panel (a) illustrates induction due to stress and panel (b) due to aging]. 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 day of leaf unfolding (d_{LU}), these states simulate the leaf development, marked by transitions from the young to the mature leaf ($d_{y \rightarrow m}$) and from the mature to the old leaf ($d_{m \rightarrow o}$) as well as by the phenological stages 50% leaf coloring ($d_{LS_{50}}$) and 100% leaf coloring or fall ($d_{LS_{100}}$). These transitions and stages occur when $S_{Aging,i}$, $S_{Stress,i}$, and $S_{senescence,i}$ breach corresponding thresholds ($Y_{Aging,1}$, $Y_{Aging,2}$, Y_{Stress}, $Y_{LS_{50}}$, and $Y_{LS_{100}}$). $d_{m \rightarrow o}$ is defined as the first day on which either Y_{Stress} or $Y_{aging,2}$ is breached [panel (a) and (b), respectively] and marks the beginning of senescence during which the daily senescence rate is accumulated. Dotted lines are auxiliary lines.

We constructed and tested (see Sect. 2.2.3) several formulations of leaf development by combining the following assumptions. We considered that aging could be modeled 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 modeled by a combination of the stressors cold, shortening day length, drought, heat, frost, heavy rain, and nutrient depletion (Bigler and Vitasse, 2021; Jan et al., 2019; Jibran et al., 2013; Kloos et al., 2024; Mariën et al., 2021; Tan et al., 2023; Wang et al.,

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

2022; Xie et al., 2015, 2018; Zohner et al., 2023). Finally, we considered that leaf senescence could result as a combination of 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):

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

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

Here, $S_{k,j}$ is the state at day j of either aging, stress, or senescence (k) that are formulated as the sum of the corresponding rates at day i ($R_{k,i}$), which accumulated between the starting day s_k and j, until the threshold Y_k is reached. In other words, the daily aging rate ($R_{Aging,i}$) accumulates from the day of LU (s_{Aging}). The transition from young leaf to mature leaf occurs when $S_{Aging,j}$ reaches $Y_{Aging,1}$. Thus, day j becomes s_{Stress} and the accumulation of the daily stress rate ($R_{Stress,i}$) starts, while $R_{Aging,i}$ continues to accumulate. The transition from mature leaf to old leaf occurs when either $S_{Aging,j}$ reaches $Y_{Aging,2}$ or $S_{Stress,j}$ reaches Y_{Stress}. Now, day j becomes $s_{Senescence}$ 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 the Keetch and Byram drought index [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 to P_{i-4}), (5) heat days (derived from maximum temperature; Tx [°C]), (6) nutrient depletion (approximated by the accumulated A_{net} since d_{LU}, due to the absence of soil data), and/or (7) frost days (derived from minimum temperature; Tn [°C]):

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

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

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

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

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):

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

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

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. 3). Because $x \geq a$ and $x \geq b_0$ result in stress, the response to δL and Tn was formulated as $g(-\delta L)$ and $g(-Tn)$ as well as $h(-\delta L)$ and $h(-Tn)$, which translates in stress if $\delta L \leq -a \vee -b_0$ and $Tn \leq -a \vee -b_0$ (e.g., if stress occurs suddenly or gradually when $\delta L \leq -0.01$ h, then $a = 0.01$ h or $b_0 = 0.01$ h, respectively).

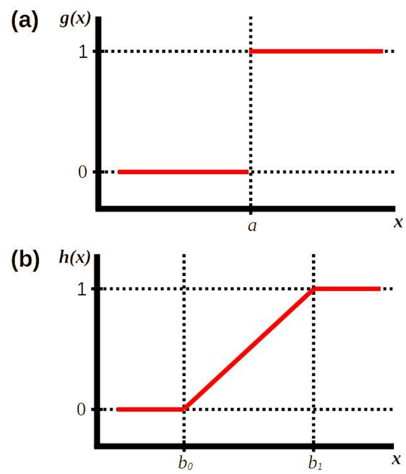


Figure 3. Response functions (solid red lines) of g(x) and h(x). In (a), a marks the boundary value of x at which $g(x)$ suddenly changes from 0 to 1 (i.e., from no effect to an effect). In (b), b_0 and b_1 mark the lower and upper bounds of x, respectively, between which $f(x)$ gradually increases from 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 $S_{Aging,i}$ and $R_{Stress,i}$ (Eq. 9):

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

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

w_A and w_S are the weights of R_{Aging} and R_{Stress}, respectively, s_X is a scaling factor, all of which allowed us to 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 simulates both LS_{50} and LS_{100} simultaneously, and one to calibrate a model that simulates LS_{50} only. For the two datasets, we selected site-years from those with the most extreme conditions during the growing season, i.e., the hottest, coldest, driest ten day periods between LU and LS_{50} as well as the shortest and longest growing season. For the first dataset, hereafter called 'LS_{50}-LS_{100} sample', we 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 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. S43) with generalized simulated annealing and optimal, model-specific controls (see Sect. S2.1; 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. First, we defined the process structure. In iteration 1, we tested the definition of the aging rate (R_{Aging}) as either a function of net photosynthetic activity (A_{net}) or the number of days (Eq. 3), and the definition of the senescence rate ($R_{Senescence}$) as a combination of the stress rate (R_{stress}) and either R_{Aging} or the state of aging (S_{Aging}) in either a sum, product, or exponential function (Eq. 9). In iteration 2, we tested the number of phases of leaf development, i.e., either two phases 'mature leaf' and 'old leaf', or three phases 'young leaf', 'mature leaf', and 'old leaf'. Thus, we formulated R_{stress}. In iteration 3, we tested the effect of cold, shortening, and dry days as either a threshold response function [*g(x)*] or a gradual response function [*h(x)*; Eq. 6]. In iteration 4, we tested additional drivers, i.e. heavy rainfall, heat days, nutrient depletion, and frost days (Eq. 5). In iteration 5, we tested the effect of the additional driver as either *g(x)* or *h(x)* (Eq. 6). In the subsequent iterations, 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 simulating LS_{50} and LS_{100}, i.e., through calibration with the LS_{50}-LS_{100} sample. This accuracy was assessed with the

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

Akaike information criterion corrected for small samples (AICc; Eq. S40; Akaike, 1974; Burnham and Anderson, 2004), which accounts for both the goodness-of-fit between the simulated and observed values and the number of free parameters. We calculated the AICc for each calibration run (see Sect. 2.4) and excluded the run with the highest AICc per model, before identifying the two models with the lowest median AICc across the given and all previous iterations. Finally, we selected the model based on the formulation with the lowest median AICc, which was further evaluated.

2.6 Model evaluation

First, we evaluated the functionality of the selected model if calibrated for LS_{50} only (as done by previous studies, e.g., Delpierre et al., 2009; Dufrêne et al., 2005; Zani et al., 2020). We were particularly interested in the causes of senescence induction, i.e., the transition from mature leaf to old leaf that could be due to the reaching of either $Y_{Aging,2}$ or Y_{Stress}. 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; °), and elevation (ELV; m a.s.l.) for the given year and site. While MAT and MAQ are assumed to directly affect cold and dry stress, LAT relates to day length through the inclination angle of the Earth, 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. S44; Gupta et al., 2009; Kling et al., 2012), which combines bias, variability, and correlation of the simulated 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. The CDD model determines LS_{50} by the time the cold degree-days day length 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 across the entire validation sample (ρ_{Overall}), across space (ρ_{Spatial}), and across time (ρ_{Temporal}). ρ_{Spatial} was based on the site-specific mean observed and simulated LS_{50} across sites. ρ_{Temporal} was calculated for each site based on the yearly observed and simulated LS_{50}.

Third, we estimated the extent to which the model error was affected by data structure as well as by climatic and spatial deviations from the calibration sample, using a linear mixed-effects model (LMM; Pinheiro and Bates, 2000) and an analysis of variance (ANOVA; Sect. S2.3; 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 an ANOVA. This type-III ANOVA (Yates, 1934) quantified the impact of the explanatory variables in 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

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

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 calibrated 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, 2022). LMM estimates and 99% confidence intervals (i.e., significance level $a = 0.01$) for combined coefficients, e.g., the effect of δMAT for a given model, were calculated with the Delta method (Fox and Weisberg, 2019, Chpt. 5.1.4; Wasserman, 2004, Chpt. 9.9) through the function deltaMethod in the R package car (Fox and Weisberg, 2019). For each LMM coefficient and ANOVA impact, 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, 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. 4 and S1). 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, shortening, 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 of stress. The first day that either the state of aging or the state of stress reaches the respective thresholds $Y_{aging,2}$ or Y_{Stress} (Eqs. 1 and 2), the phase 'old leaf' starts. During this third phase, a daily senescence rate ($R_{Senescence}$) 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 simulated dates of LS_{50} and LS_{100}, respectively. Hereafter, we refer to this model as 'DP3' model (Tables 2 and 3; Meier, 2025b, coded in R).

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

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

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

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

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 the Keetch and Byram drought index (*Q*) at day *i*, respectively [e.g., w_P g($-\delta L_i$) results in the photoperiod stress at day *i*]. s_X is the scaling factor for R_{Stress}, which is 'shaped' by x_S.

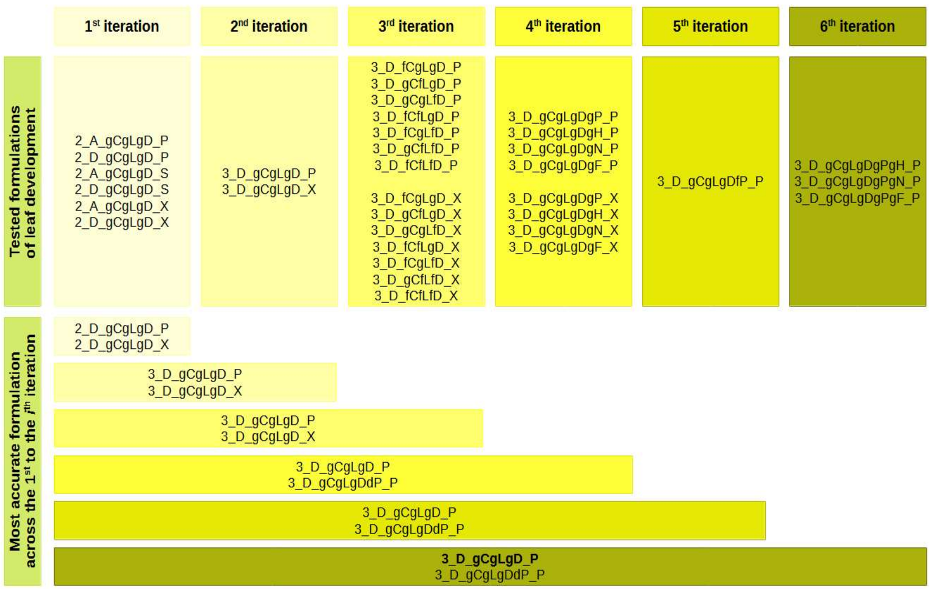


Figure 4. Formulations of leaf development tested through iterations of model development. The tested formulations were labeled as x_P_x_A_x_S_x_X, with x_P being the number of leaf development phases (i.e., 2 or 3), x_A being the driver of the aging rate (i.e., A or D for photosynthesis or days, respectively), x_S being the stress rate in response (i.e., g or h for *g(x)* or *h(x)*) to the stressors cold (C), shortening (L), dry (D), heat (H), and frost (F) days, heavy rainfall (P), and nutrient depletion (N), and x_X indicating the formulation of the senescence rate (i.e., S, P, or X when formulated as a sum, product, or exponential function of aging and stress, respectively). After each iteration, we identified the two most accurate formulations across the given and all previous iterations. These formulation were further developed through the next iteration. As soon as two subsequent iterations resulted in the same two most accurate formulations, we selected the more accurately formulated model (bold; i.e., the 'DP3' model). All formulations were tested for beech based on the LS_{50}-LS_{100} sample (Sect. 2.4).

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

Table 2. Input and output variables of the DP3 model

	Name	Definition	Unit	Format
Input	LU	Observed timing of leaf unfolding	doy	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
Output	LS	Simulated timing of leaf senescence	doy	Vector
	$d_{y\to m}$	Simulated timing of transition from young to old leaf	doy	Vector
	$d_{m\to o}$	Simulated timing of transition from old to mature leaf	doy	Vector
	$R_{Aging,i}$	Daily rate of aging	-	Matrix
	$S_{Aging,i}$	Accumulated rate of aging since LU	-	Matrix
	$X_{Cold,i}$	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
	$S_{Stress,i}$	Accumulated rate of stress since $d_{m\to o}$	-	Matrix
	$R_{Senescence,i}$	Daily rate of senescence	-	Matrix
	$S_{Senescence,i}$	Accumulated rate of senescence since $d_{m\to o}$	-	Matrix

Note: Daily variables refer to day i, and accumulated variables refer to the period until day i. 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. The order of the vectors matches these order of the matrix columns.

Table 3. Fitted parameters of the DP3 model

Parameter	Meaning	Initial boundaries	Fitted value
$-a_C$	Boundary below which cold stress is 1 versus 0 (referring to Tn_i)	0–30 °C	0.06 °C
$-a_P$	Boundary below which photoperiod stress is 1 versus 0 (referring to δL_i)	–0.25–0.25 h	–0.0016 h
a_D	Boundary above which dry stress is 1 versus 0 (referring to Q_i)	0–800	183.82
w_C	Weight of cold stress	0–1	0.29
w_P	Weight of photoperiod stress	0–1	0.52
w_D	Weight of dry stress	0–1	0.05
s_X	Scaling factor of the senescence rate	0–1	0.35
x_S	Shape parameter of the stress rate	0–10	5.67
$Y_{Aging,1}$	Age threshold for the transition from young to mature leaf	0–50 d	1.57 d
$Y_{Aging,2-Aging,1}$	The threshold of aging during the mature leaf phase	0–250 d	71.57 d
$Y_{Aging,2}$	Theoretical age threshold for the transition from mature to old leaf	-	73.14 d

Note: The parameters refer to the equations 7 and 9–11 and were fitted for beech with the LS_{50} sample (Sect. 2.4). All parameters were calibrated within the initial boundaries 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} ($Y_{LS_{50}}$; i.e., the time when 50% of the leaves have changed color or fallen) were hard coded with $Y_{Stress} = 1$ and $Y_{LS_{50}} = 1$.

In the DP3 model, leaf senescence starts with the transition from 'mature leaf' to 'old leaf', being induced by either the state of aging or the state of stress. Stress induced senescence 40 times more often than aging (Fig. 5a, Table S3). On average, stress rather than aging induced leaf senescence at cooler and dryer sites as well as at higher latitudes and higher elevations. At the time of senescence induction due to stress, the amount of accumulated photoperiod and cold stress was 77% and 23%, respectively, while dry stress was 0% (Fig. 5b, Table S4). Photoperiod stress was generally more important than cold stress, especially in cool, high-elevation sites. Note that dry stress had some importance in inducing senescence according to the LS_{50}-LS_{100} sample, which was used for model development and explains the consideration of dry stress in the model (results not shown). In summary, while senescence was mainly induced by

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

stress rather than aging and mainly by photoperiod stress rather than cold and dry stress, the importance of these stressors depended on the climatic conditions and location.

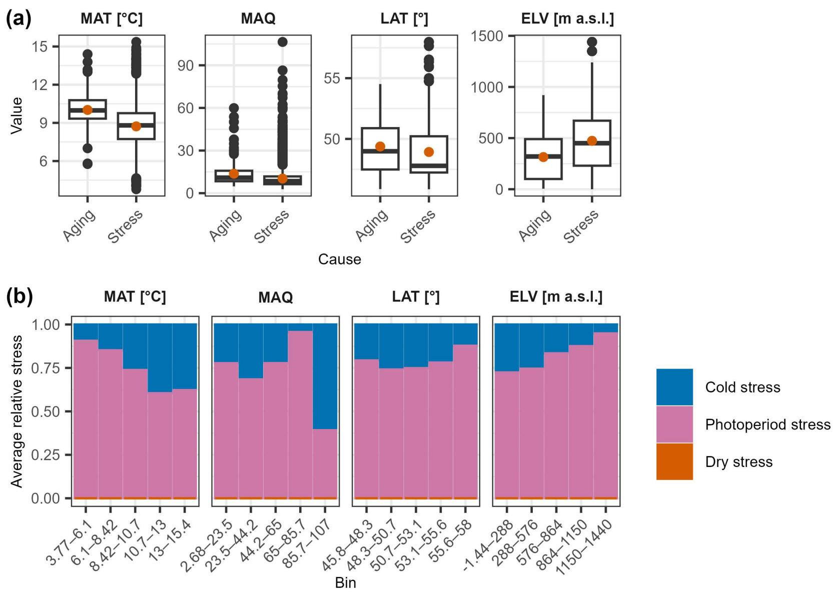


Figure 5. Senescence induction by aging versus stress. Panel (a) shows, how the site years are distributed across mean annual temperature (MAT), mean annual KBDI (MAQ), latitude (LAT), and elevation (ELV). The boxes indicate the inner quartile range and the median (middle line). The most extreme values are indicated with dots if outside ±1.5 times the inner quartile range from the 1st and 3rd quartile, and with whiskers otherwise. Orange dots show the mean. Panel (b) shows the relative amount of cold, photoperiod, and dry stress that accumulated at the time of senescence induction by stress, averaged per bin of equal width across the dimensions MAT, MAQ, LAT, and ELV.

3.2 Model accuracy

The DP3 model simulates leaf senescence with similar accuracy as previous models (Fig. 6; Table 4). All models resulted in an RMSE of ~15 d, with the lowest RMSE for the Null model (i.e., constant simulation of the average observation in the calibration sample). The CDD model resulted in the highest KGE'. While the simulations correlated best with the observations when based on the PIA model, especially across space ($\rho_{Spatial}$ of 0.4), they correlated best across time when based on the DP3 model (average $\rho_{Temproal}$ of 0.05).

Table 4. Model accuracy

Model	KGE'	RMSE	AIC	AICc	$\rho_{Overall}$	$\rho_{spatial}$	$\overline{\rho}_{Temporal}$	n
CDD	-0.13	16.07	57797	57797	0.01	-0.09	0.04	6887
DM2	-0.26	15.01	56862	56862	0.02	-0.12	-0.00	6887
PIA	-0.19	14.83	56701	56701	0.10	0.44	-0.04	6887
DP3	-0.23	15.24	57083	57083	0.02	-0.02	0.05	6887
Null	NA	14.81	NA	NA	NA	NA	NA	6887

Note: The Null model constantly simulates the average observation in the calibration sample. The modified Kling-Gupta efficiency (KGE'), root mean squared error (RMSE), Akaike information criterion (AIC), AIC for small samples (AICc), and Pearson correlation overall, across space, and across time ($\rho_{Overall}$, $\rho_{Spatial}$, and average $\rho_{Temporal}$, i.e., $\overline{\rho}_{Temporal}$, respectively) are explained in Sect. 2.6,

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

S2.1, and S2.2. All these metrics were calculated for the simulations and observations of the validation sample. Except the RMSE, they result in NA if the variance of the simulated values is zero, which is the case for the Null model. n indicates the number of observations in the validation sample.

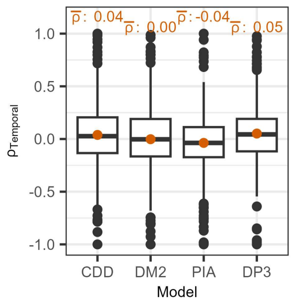
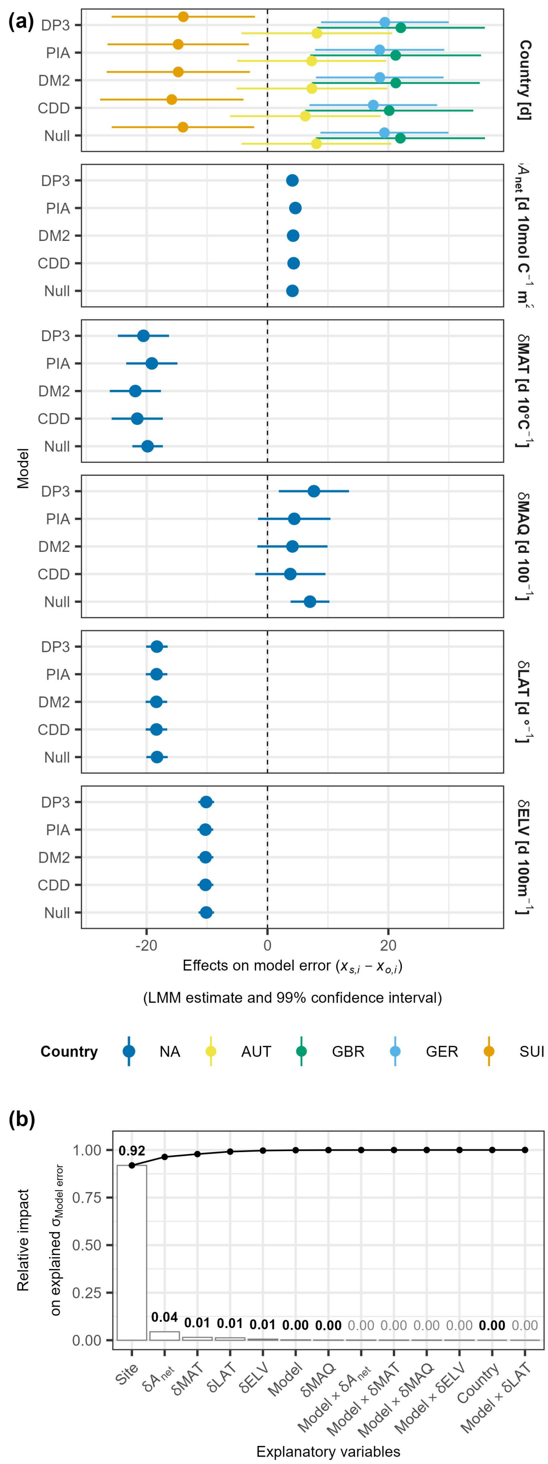


Figure 6. Temporal Pearson correlation ($\rho_{Temporal}$). The distribution ot the Pearson correlation between simulated and observed leaf senescence within site ($\rho_{Temporal}$) is displayed for each model. The mean is indicated in orange above each box ($\bar{\rho}$). See Fig. 5 for the interpretation of the boxes, middle lines, whiskers, and dots.

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. 7). Data structure was described by the fixed effects of countries and the random intercepts grouped by sites. The countries altered the model error by –16 to +19 d (Tables S5 and S6). The standard deviation in the model error due to the random intercepts was 31 d. Depending on the model, the fixed effects of the climatic deviations, ranged from –19 to –16 d 10°C^{-1} (δMAT), from 1.3 to 5.2 d 100^{-1} (δMAQ), and from 3.9 to 4.4 d 10mol C^{-1} (δA_{net}), respectively. The model-specific effects of the spatial deviations δLAT and δELV ranged from ~–18 d °$^{-1}$ and ~–10 d 100m^{-1}, respectively. While the evidence in the data was decisive (BF_{01} < 1/1000) for the effect of the CDD model, the evidence was significant ($p < 0.005$) for the effects of all previous models. The evidence was neither decisive nor significant for any effect 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 0.43. 90% of the variance within this error was attributable to the differences among sites, followed by the effects of δA_{net} and δMAT (4% and 2%, respectively), whereas the effects of the models accounted for less than 0.5% (Table S7). 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.

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025



https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

Figure 7. 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 °$^{-1}$) 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), United Kingdom (GBR), Germany (GER), and Switzerland (SUI). The model error was calculated as the simulated minus the observed timing ($x_{s,i} - x_{o,i}$). Panel (b) shows 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 minimum Bayes factor $\leq$ 1/1000).

4 Discussion

4.1 Model formulation

The DP3 model simulates the timing of leaf senescence through a novel formulation, supporting the advancement of leaf senescence research. To our knowledge, it is the first process-based leaf senescence model that (1) simulates the timing of leaf senescence through daily leaf development status, (2) starts the simulation with leaf unfolding, (3) differentiates between daily aging and stress rates, and (4) predicts the dates of transition between the leaf developmental phases young leaf, mature leaf, and old leaf. This allows the development of several new hypotheses, e.g. on the timing and cause (i.e. aging versus stress) of senescence induction marked by the transition to the old leaf (Carley, 1999; Hauke et al., 2020). As these hypotheses can be tested by controlled experiments, the DP3 model is likely to become an important source for expanding knowledge about leaf senescence.

For example, the DP3 model proposed that stress rather than aging induced leaf senescence at cooler and higher sites (Fig. 5), which seems reasonable. In comparison to warmer and lower sites, leaves probably unfolded later in the course of the year at cooler and higher sites. Consequently the accumulation of both aging and stress started later, too. Thus, the leaves had aged less, when the threshold for photoperiod stress was breached for the first time, which occurred on the first day that was ~5 seconds shorter than the previous day (i.e., just after summer solstice; Table 3) and started the daily accumulation of photoperiod stress. As two days of photoperiod stress induced senescence by themselves, leaves started to senescence no later than on the second day after summer solstice, while they potentially started earlier, particularly at warmer, lower sites. This could be tested in both in situ and controlled experiments.

Moreover, the DP3 model postulated a greater importance of photoperiod stress than cold stress in inducing senescence at cooler and higher sites, which may be puzzling at first. However, due to senescence induction no later than on the second day after summer solstice (see above), cold stress relevant for senescence induction was accumulated in spring rather than in midsummer and corresponded to late frost events [i.e., frost events that occurred on the second day after leaf unfolding or later (Sect. 3.1; Table 3)]. This implies that such frost events were more frequent at warmer and lower sites than at cooler and higher sites, which agrees with previous studies (Asse et al., 2018; Meier et al., 2018; Sangüesa-Barreda et al., 2021), but contrasts other studies (Bigler and Bugmann, 2018; Zohner et al., 2020), and should be studied further.

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

4.2 Model accuracy

We compared the DP3 model to three previous models of leaf senescence (Delpierre et al., 2009; Dufrêne et al., 2005; Zani et al., 2020) and found the RMSE of all compared models to be above the RMSE for the Null model (i.e. the constant simulation of the average observation in the calibration sample). This my be explained by unrealistic model formulations, poor model calibrations, and noisy data to drive and calibrate the models.

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 timing of leaf senescence advances due to cold temperatures. This positive correlation between the timing of leaf senescence and temperature was 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, the 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, it is unlikely that the Null model is more realistically formulated than the compared models developed in four separate studies, but these models almost certainly do not fully reflect the process of leaf senescence either, so they should be tested in different ways and further developed.

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). The used algorithm and controls affect the accuracy of the calibrated models (Meier and Bigler, 2023). Therefore, we used a well established optimization algorithm. (Generalized) simulated annealing 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 simulate leaf senescence dates for the validation samples (Sect. S2.1). Possible overfitting (James et al., 2017) through this procedure 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). Moreover, the compared models would have benefited from overfitting, as the comparison to the Null model was based on the same validation samples as the selection of the controls. Therefore, it is very 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 network, 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 timing of leaf senescence for individual trees varied by 15 d (median, spreading from 2–53 d) between observers, and increased 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 process of leaf senescence. 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., at least 244 observers) and four countries (Sect. 2.1), which implies considerable noise. Such noisy leaf senescence data very likely forced the compared models to simulate the timing of leaf senescence close to the mean observation, impairing their accuracy.

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

4.3 Model error

While the model error was generally affected by climatic and spatial deviations from the calibration sample, their model-specific effects only differed insignificantly from the Null model. In other words, the model error in the compared models reacted similarly to climatic and spatial deviations as the model error of the Null model. This implies that the compared models simulated the timing of leaf senescence closely to the mean observation of the calibration sample. Possible explanations for such a bias are likely the same as listed above, i.e., 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., at 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.

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 effects of countries have, to our knowledge, not been studied yet, and differed considerably between countries, which demonstrates the noise added to leaf senescence data by different observation protocols (see above; Menzel, 2013). 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 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 the timing of leaf senescence would be simulated correctly by models calibrated with noise-free data.

4.4 Ways forward

We have discussed, how noisy leaf senescence data force models to simulate the mean observation, resulting in low accuracy regardless of model formulation. On the one hand, this calls into question the current practice of comparing the accuracy of models and then drawing inferences about the process of leaf senescence from the formulation of the most accurate model. On the other hand, this leads to the question of how to proceed.

To proceed, alternative data may be used, observation protocols may be revised, and visually observed data may be carefully selected. Alternative data to calibrate and validate models of leaf senescence include data recorded with phenocams and remote sensed data in which the timing of leaf senescence is 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., 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). Revised observation protocols should describe the measurement of the timing of leaf senescence, e.g. through the greenness derived from images taken with consumer-grade digital cameras (Ide and Oguma, 2013; Richardson et al., 2018; Toomey et al., 2015; Zimmerman and Richardson, 2024). 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 using as few sites as possible. However, the use of such data could lead to a model that does not represent the entire population. To mitigate this risk, the few selected sites should be far apart both spatially and climatologically.

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

5 Conclusion

The DP3 model builds on three subsequent phases of leaf development. During these phases, leaf development answers to aging, aging and stress, and senescence as a result of stress, respectively, with stress combining cold, photoperiod, and dry stress. The output variables of the DP3 model include the corresponding daily rates of aging, of combined stress, and of cold, photoperiod, and dry stress, together with the timing of the transitions between the phases of leaf development and the timing of leaf senescence. As this allows to develop testable hypotheses about leaf senescence, for example regarding site conditions and the relative importance of cold, photoperiod, and dry stress for senescence induction, the DP3 model likely becomes 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 simulation of the average observation in the calibration sample). This was probably due to model formulations that do not fully reflect the process of leaf senescence and, more importantly, to the leaf senescence data used for calibration and validation. Visually observed leaf senescence data consists of individual estimates of the timing of leaf senescence, which are based on observation protocols that are partly inconsistent between countries. Such noisy data blurs the signal of the process of leaf senescence, thereby probably forcing the models to resort to the mean observation and causing low accuracy, independently from model formulation. Therefore, noisy leaf senescence data likely impede inferences from accuracy-based comparisons of model formulations about the process of leaf senescence, 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 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 90% of the explained variance in the model error. Therefore, these data should be selected with particular attention to precision, e.g. by using as few sites with identical observation protocols as possible. Moreover, revised observation protocols should explain how to measure rather than estimate the timing leaf senescence. 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.14749340). While all raw data used are publicly available and referenced in section 2, the simulated leaf senescence data analyzed is 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. MM and IC conceptualized the final study and model development. 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 authors approved the final manuscript.

Competing interests

The authors declare that they have no conflict of interest.

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

Acknowledgments

We acknowledge 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).

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.

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.

Baayen, R. H., Davidson, D. J., and Bates, D. M.: Mixed-effects modeling with crossed random effects for subjects and items, J. Mem. Lang., 59, 390–412, https://doi.org/10.1016/j.jml.2007.12.005, 2008.

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.

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.

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.

Candelieri, A.: A gentle introduction to Bayesian Optimization, 2021 Winter Simulation Conference (WSC), Phoenix, AZ, and virtual, 1–16, https://doi.org/10.1109/WSC52266.2021.9715413, 2021.

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.

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

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.

Fisher, R. A. and Russell, E. J.: On the mathematical foundations of theoretical statistics, Philos. Trans. R. Soc. Lond. Ser. Contain. Pap. Math. Phys. Character, 222, 309–368, https://doi.org/10.1098/rsta.1922.0009, 1997.

Foster, G. R., McCool, D. K., Renard, K. G., and Moldenhauer, W. C.: Conversion of the universal soil loss equation to SI metric units, J. Soil Water Conserv., 36, 355–359, 1981.

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.

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

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.

Geodesy: Approximate formulas for the transformation between Swiss projection coordinates and WGS84, 2016.

Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., and Sitch, S.: Terrestrial vegetation and water balance - hydrological evaluation of a dynamic global vegetation model, J. Hydrol., 286, 249–270, https://doi.org/10.1016/j.jhydrol.2003.09.029, 2004.

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.

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.

Haxeltine, A. and Prentice, I. C.: BIOME3: An equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability, and competition among plant functional types, Glob. Biogeochem. Cycles, 10, 693–709, https://doi.org/10.1029/96gb02344, 1996.

Haxeltine, A., Prentice, I. C., and Creswell, D. I.: A coupled carbon and water flux model to predict vegetation structure, J. Veg. Sci., 7, 651–666, https://doi.org/10.2307/3236377, 1996.

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.

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

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.

Maes, F., Wehenkel, L., and Ernst, D.: Meta-learning of exploration/exploitation strategies: The Multi-armed bandit case, Agents and Artificial Intelligence, 100–115, 2013.

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, Dryad, https://doi.org/10.5061/dryad.tht76hf97, 2025a.

Meier, M.: The DP3 model (v1.0) - Simulating the timing of leaf senescence, Zenodo, https://doi.org/10.5281/zenodo.14749340, 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.

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.

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. R Foundation for Statistical Computing, 2022.

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.

https://doi.org/10.5194/egusphere-2025-460
Preprint. Discussion started: 14 April 2025

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.

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.

Shaw, H. G.: Centigrade-Fahrenheit temperature conversion, J. Chem. Educ., 8, 727, 1931.

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.

Sitch, S., Prentice, I. C., Smith, B., Cramer, W., Kaplan, J. O., Lucht, W., Sykes, M. T., Thonicke, K., and Venevsky, S.: LPJ - A Coupled Model Of Vegetation Dynamics And The Terrestrial Carbon Cycle. https://www.researchgate.net/publication/37456884, 2000.

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.

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.

Wohlfahrt, G. and Gu, L.: The many meanings of gross photosynthesis and their implication for photosynthesis research from leaf to globe, Plant Cell Environ., 38, 2500–2507, https://doi.org/10.1111/pce.12569, 2015.

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.

Woods, H. W.: Centigrade-fahrenheit temperature conversion, J. Chem. Educ., 8, 370, 1931.

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.

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.

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., Mo, L., Sebald, V., and Renner, S. S.: Leaf-out in northern ecotypes of wide-ranging trees requires less spring warming, enhancing the risk of spring frost damage at cold range limits, Glob. Ecol. Biogeogr., 29, 1065–1072, https://doi.org/10.1111/geb.13088, 2020.

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.