Unusually warm winter seasons may compromise the performance of current phenology models – Predicting bloom dates in young apple trees with PhenoFlex

Eduardo Fernandez [a,b,*], Katja Schiffers [a], Carsten Urbach [c], Eike Luedeling [a]

[a] *Department of Horticultural Science, Institute of Crop Science and Resource Conservation (INRES), University of Bonn, 53121 Bonn, Germany*
[b] *Escuela de Agronomía, Pontificia Universidad Católica de Valparaíso, Casilla 4-D, Quillota, Chile*
[c] *Helmholtz Institut für Strahlen- und Kernphysik (HISKP), University of Bonn, Germany*

A B S T R A C T

Phenology models are crucial tools for assessing climate change impacts in forestry, ecology and agriculture. Such models are typically calibrated with observational or experimental data and validated with a set of independent observations. While there have been extensive discussions about validation approaches, systematic studies assessing the effects of the calibration data on the predictive performance of the fitted model are scarce. We evaluated the impact of marginal seasons in the calibration data set on the predictive power of an integrated modeling framework (PhenoFlex) that was recently proposed to predict spring phenology in temperate trees. We calibrated PhenoFlex with phenology records of apple trees from a multi-season experiment (59 experimental seasons) that included five unusually warm winter seasons. For comparison, we excluded these marginal seasons in a second version of the analysis. We fitted the 12 model parameters to data, assessed model performance using a common validation data set and evaluated the chill and heat responses during dormancy for both versions. Despite high overall accuracy, our results indicated a better model performance (Root Mean Square Errors of 2.3 versus 5.5 days) when excluding the marginal seasons. We observed a similar shape for the chill response curve across versions but a greater chill effectiveness when including the marginal seasons. Fitted parameters suggest a hard drop in heat efficiency beyond the optimum temperature when including the marginal seasons, probably highlighting the need for more moderate conditions during model calibration. Our results demonstrate a good performance of PhenoFlex when calibration and validation data were comparable, but they also indicate risks involved in using the framework to project phenology under conditions that differ strongly from those used for calibration. Further evaluation and validation under experimentally or naturally occurring warm conditions may improve our understanding of the response of temperate trees to mild winter conditions.

1. Introduction

Temperature is among the strongest environmental cues modulating development stages in many temperate plant species, including deciduous fruit and nut trees (Chuine and Regniere, 2017). The continuous increase in temperature due to global climatic change has already modified the phenology (i.e. the occurrence of development stages) of many temperate tree species in various regions (Menzel et al., 2006; Walther et al., 2002). Analyzing historic records of 542 plant species collected in 21 European countries (observations of about 10+ years between 1951-2000 depending on the species), Menzel et al. (2006)

reported an advance of spring and summer phases by up to 4.6 days per °C. Some evidence suggests, however, that for some phenology events such as the spring leaf unfolding of European tree species, the rate of advance in response to warmer conditions has decreased (Fu et al., 2015). In line with previous studies (Cook et al., 2012; Luedeling et al., 2013), an analysis conducted in China by Guo et al. (2015) reported that apricot trees cultivated in warm climates are likely to suffer a delay in the timing of spring events due to global warming. Many authors have suggested that spring phenology in temperate species results from the combined effects of both autumn/winter cold and spring heat (Campoy et al., 2011; Cook et al., 2012; Guo et al., 2015; Luedeling, 2012). For

---

future scenarios, the potentially dual response of trees to temperature variation may become a challenge for scientists studying temperate trees as well as for practitioners growing such species. To accurately forecast the impacts of climate change on tree phenology, modelling approaches must consider the combined effects of environmental conditions observed during autumn/winter and spring.

Temperate tree species experience a dormancy phase during autumn/winter, which allows these trees to endure the cold conditions usually observed in their habitat of origin (Fadón et al., 2020; Vegis, 1964). According to Lang et al. (1987), tree dormancy can be classified into para-, endo- and eco-dormancy, with the last two frequently associated with winter dormancy and implying different sub-processes at tree and bud level (see Fadón et al. (2020) for a review). Whereas the need for cold temperatures has been proposed as characteristic of the endo-dormancy phase (represented as Chill Requirement – CR; Luedeling (2012)), warm conditions have been suggested as the major driver of bud development during the eco-dormancy period (described as Heat Requirement – HR; Luedeling (2012)). The need for cold and subsequent warm conditions interact to modulate the breaking of dormancy as well as bud burst and bloom dates in temperate trees.

Phenology models usually attempt to represent the processes of chill and heat accumulation as well as their interaction during dormancy. The extent and nature of this interaction, however, remain mostly unclear. Three options representing the chilling/forcing relationship have been proposed in the past: "sequential" as suggested by Ashcroft et al. (1977), "parallel" as suggested by Landsberg (1974) and "overlapping" as suggested by Cannell and Smith (1983). Among these options, the "overlapping" structure has gained relevance due to its more biologically credible nature (Harrington et al., 2010; Pope et al., 2014). An "overlapping" between the chilling and forcing phases of dormancy may indicate that chill and heat can compensate for each other under particular circumstances, resulting in similar bloom outcomes for different combinations of chill and heat (Pope et al., 2014). Whereas phenology models considering a "sequential" or "parallel" relationship inherit an evident fixed structure (e.g. heat only accumulates after CR is reached in the case of the "sequential" model), the "overlapping" approach offers an opportunity to make the link between both phases flexible. While several studies have not found the chill overlap approach to be an improvement (e.g. Darbyshire et al. (2020); Prats-Llinàs et al. (2019)), Darbyshire et al. (2016) reported that a model with 75% overlap outperformed a "sequential" approach in explaining apple phenology records from different regions of Australia. This and other former "overlapping" approaches attempting to model bloom dates have, however, failed to add a truly flexible link between phases, instead assuming fixed overlapping phases between the chilling and forcing periods (Darbyshire et al., 2016; Pope et al., 2014). Using a model structure that determines the nature of the interaction as well as the shape of the transition from endo- to eco-dormancy based on observations can be expected to improve the accuracy of predicting spring phenology events in deciduous fruit trees. The applicability of such a model structure, however, would be restricted to the species and conditions used in model development, potentially hampering the transferability of modelling outcomes. A trade-off between expected accuracy and transferability of results needs to be considered when developing phenology models.

In a recent study, Luedeling et al. (2021) proposed a new dormancy modelling framework – the PhenoFlex framework – to forecast bud burst and bloom dates in temperate fruit trees. Compared to similar approaches reported in the past, PhenoFlex is considerably more flexible in its assumptions regarding the dormancy breaking process. PhenoFlex uses the Dynamic model (Erez et al., 1990; Fishman et al., 1987a; Fishman et al., 1987b) as sub-model to account for the process of chill accumulation during the endo-dormancy phase and the Growing Degree Hours model (Anderson et al., 1986) as sub-model to represent the process of heat accumulation during the eco-dormancy phase. To describe the interaction between the chilling and forcing phases, both of

which contribute to the dormancy breaking process, the authors allowed the framework to determine the extent of overlap between phases based on temperature and phenology observations. This structure allows the model the possibility to represent any of the three formerly reported interactions between the chilling and forcing phases ("sequential", "parallel" and "overlapping"). Using 60 years of phenology records for apple and pear trees cultivated in a temperate climate in Germany, Luedeling et al. (2021) demonstrated that PhenoFlex outperformed several other prediction models, including the StepChill model derived from the Unified model (Chuine, 2000) and a naïve machine learning algorithm based on a Gaussian process with a linear kernel. Despite the promising results obtained when predicting bloom dates, the authors reported and warned about some implausible fitted parameters for apple (e.g. chill response for temperatures up to 30°C), probably resulting from a lack of relevant weather situations (such as greater temperature variation during the chilling phase) in the data set used for model calibration (Luedeling et al., 2021). Implementing the PhenoFlex modelling framework with phenology records obtained from warmer environments (from mild-winter locations or greenhouse experiments) is therefore likely to help understand the origin of implausible temperature response curves as well as identify reliable model parameters.

The main aim we set for this work was to evaluate the performance of the PhenoFlex modelling framework (Luedeling et al., 2021) in response to the temperature range covered by the calibration data. We used phenology data from 59 experimental seasons (or trials) with varying temperature profiles spanning a wide range of environmental settings to assess the performance of PhenoFlex when used to forecast bloom dates of young apple trees. We evaluated the impact of marginal winter conditions in the calibration data set on model performance by using two versions of the analysis. Whereas in version 1 (hereafter PhenoFlex$_{all}$) we considered all experimental seasons, in version 2 (hereafter PhenoFlex$_{excluded}$) we removed from the calibration data five marginal experimental seasons that still generated bloom records but appeared unlikely to be observed in most places where apple trees are cultivated. We compared the model performance for both versions of the analysis and estimated the response curves for chill and heat accumulation using the fitted parameters.

2. Materials and methods

2.1. Weather and phenology records

We analyzed a subset of phenology and weather records reported on in an earlier study by Fernandez et al. (2021b). In brief, we collected data for 59 experimental seasons (Fig. 1; see supplementary materials for additional description) by frequently transferring young potted trees across various environments over two consecutive winters (2018/2019 and 2019/2020). These environments were a heated greenhouse (set to maintain a temperature between 5 and 25°C), an unheated greenhouse, three chambers covered with different materials (ethylene-tetrafluoroethylene-copolymer, float glass and frosted glass) and field conditions at Campus Klein-Altendorf and Endenich of the University of Bonn, Germany. We used 177 three-year-old potted trees of apple (*Malus domestica* Borkh.) cultivar "Elstar" grafted onto "M9" rootstock. We collected air temperature records (on an hourly basis) for each experimental season using portable data loggers (Tinytag TGP-4500, Tinytag TGU-4500 and EasyLog USB 31), as well as fixed devices in the case of the heated greenhouse (RAM 224.401) and field conditions at Campus Klein-Altendorf (Wilmers NDL485). To obtain phenology records, we classified tree phenology twice a week according to the BBCH scale (Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie) for pome fruit (Meier, 2001). We recorded full bloom as the moment we observed at least 50% of flowers open with first petals falling (Meier, 2001). We obtained the final bloom dates by computing the median across three replicates per experimental season.
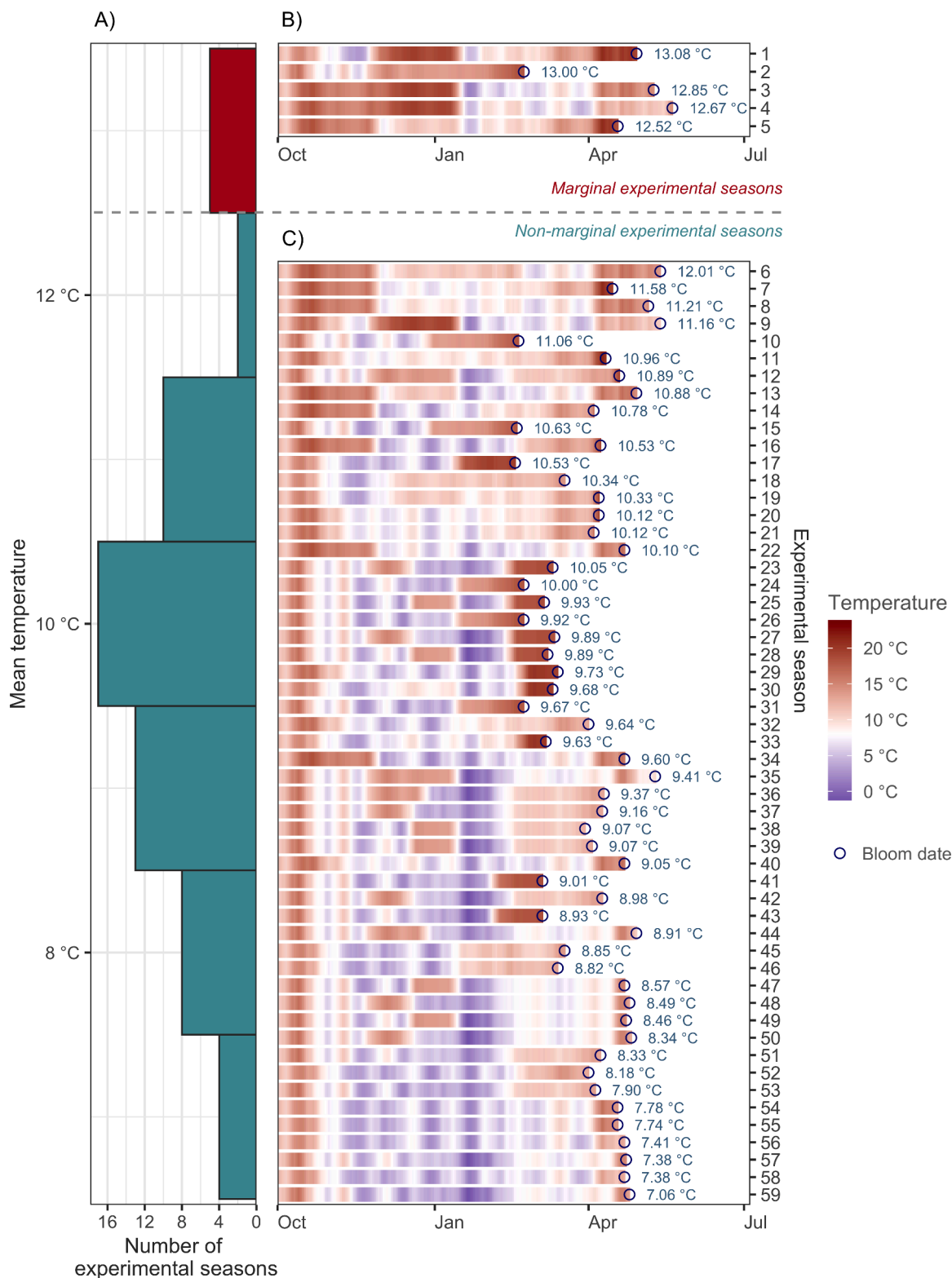
Fig. 1. Schematic illustration of the experimental seasons used for assessing the performance of the PhenoFlex modelling framework. In panel "A", we show the histogram of mean temperature (computed between October 1 and the date of bloom) across 59 experimental seasons. We show the number of experimental seasons in each bin (width of 1°C) on the inverted x-axis (increasing from right to left) and the mean temperature on the y-axis. The bin color represents the experimental seasons used in different versions of the analysis, with red showing the experimental seasons removed from the calibration in PhenoFlex$_{excluded}$. In "B" (marginal experimental seasons) and "C" (non-marginal experimental seasons), we show the specific experimental seasons (y-axis on the right) and the temperature experienced by the trees during the experiment (bar color). The date of bloom is represented by blue open circles and the mean temperature for each experimental season is shown in blue after the circle.

2.2. Evaluation of the impact of marginal conditions on model performance

To assess the impact of marginal conditions during winter on the performance of the PhenoFlex modelling framework, we defined two versions of the analysis. Whereas in PhenoFlex_all we included all available experimental seasons, in PhenoFlex_excluded we removed five experimental seasons that may have featured thermal conditions that were marginal for overcoming the dormancy of apple trees. We identified these seasons (Fig. 1B) according to mean temperature by selecting a small cluster of the 5 warmest seasons for exclusion (Fig. 1A). It should be noted, however, that some trees among these marginal experimental seasons (all obtained during the second year) showed few flower buds and slightly irregular bloom. We were unable to evaluate whether this behavior resulted specifically from restrictive thermal conditions or from any additional factor (e.g. bud initiation and differentiation in summer), since trees from other non-marginal experimental seasons showed a similar pattern during the second year of the experiment.

We started the analysis by defining a common set of seasons for PhenoFlex_all and PhenoFlex_excluded for model calibration purposes. To this end, we randomly selected 40 experimental seasons from the set of seasons numbered 6 to 59 in Fig. 1C. We used these 40 seasons to calibrate the PhenoFlex_excluded version of the framework. For PhenoFlex_all, we calibrated the model using a random subset of 35 experimental seasons drawn from the seasons selected in the first step, plus the five marginal seasons (Fig. 1B). We therefore obtained 40 seasons for calibration under both versions of our analysis. We used the remaining 14 experimental seasons to validate the phenology modelling approach under both versions of the analysis. In this validation, we estimated bloom dates using the set of parameters obtained by calibrating the model with the different versions of the analysis.

To provide insights on the importance of using the modelling framework with comparable data for calibration and validation, we conducted an additional validation of the PhenoFlex_excluded version using the five marginal seasons. With this validation, we aimed to assess the performance of the framework when used to extrapolate to conditions that differed strongly from those used for calibration. Since we were unable to systematically compare the results of this analysis with a framework calibrated and validated with marginal seasons (due to data limitations), we present and discuss major results of this analysis in the main manuscript and provide a detailed report on these results in the supplementary materials accompanying this manuscript.

2.3. Calibration of the PhenoFlex modelling framework

A set of twelve parameters needs to be fitted to data to calibrate the PhenoFlex model (Fig. 2; Table 1; see PhenoFlex vignette in Urbach et al. (2021) for full description). These parameters include six values for the Dynamic model (chill sub-model), three values for the Growing Degree Hours model (heat sub-model) and three parameters to link both sub-models. We followed the procedure described by Luedeling et al. (2021) to implement the calibration of the PhenoFlex modelling framework. Apart from phenology and weather records as inputs, the fitting function requires initial estimates for all parameters as well as lower and upper bounds for these parameters. We initialized the fitting procedure by using the original set of parameters reported by the authors of both the chill and heat sub-models (Anderson et al., 1986; Erez et al., 1990; Fishman et al., 1987a; Fishman et al., 1987b). For the three remaining parameters y_c, z_c and s_1 (with y_c being the chill requirement, z_c being the heat requirement and s_1 being the slope for the transition between the chilling and forcing phases; Table 1), we selected the values proposed by Urbach et al. (2021) in the PhenoFlex vignette. Similarly, we selected the values proposed in the vignette to define the initial values for the lower and upper bounds for all parameters except for y_c and z_c, for which we used considerably wider ranges (20 to 80 for y_c and 100 to 500 for z_c). This allowed the model the possibility to find the best requirements of chill and heat based on the data used for calibration, with rather small restrictions (compared to using narrower ranges).

In previous similar studies (Chuine et al., 2016; Egea et al., 2021; Luedeling et al., 2021), the identification of the best fitted parameters was achieved after between 10 and 30 iterations of the optimization procedure. In our case, we iterated the fitting procedure ten times for each version of the analysis. We defined an iteration as successful when the Root Mean Square Error (RMSE) of the new fitted model was smaller than the RMSE value from the previous version. As input for each new iteration, we used the set of parameters fitted in the previous run as well as the lower and upper bounds after adjusting the values, when the estimated parameter was close to the boundaries (see supplementary materials for the specific sets of parameters used in each iteration). As final model parameters, we selected the set of values obtained after the
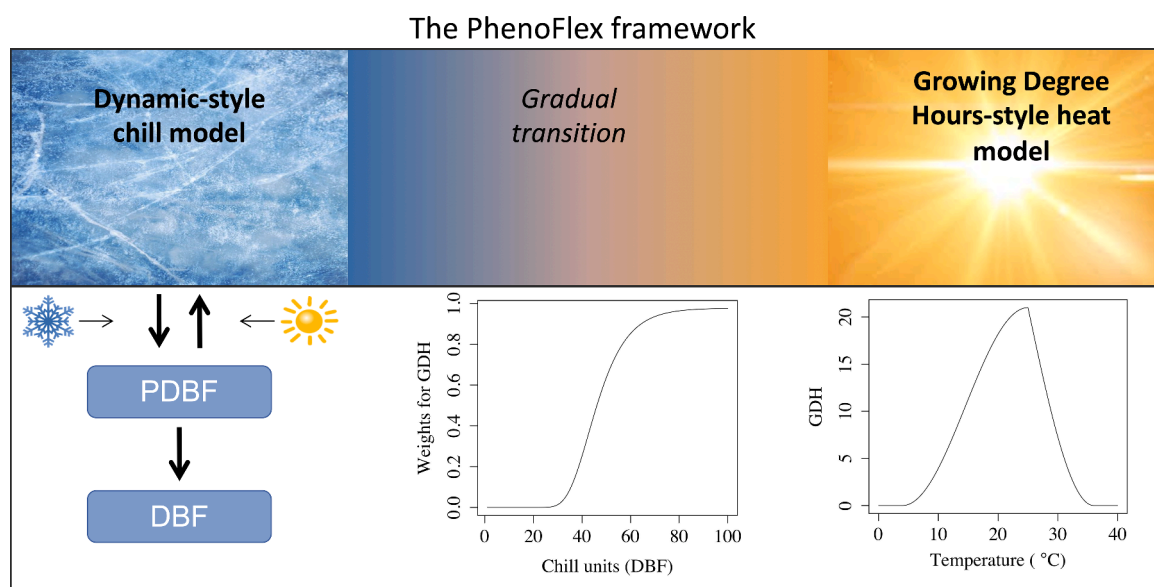
The PhenoFlex framework

Fig. 2. Schematic illustration of the PhenoFlex modelling framework. PhenoFlex consists of two sub models ("Dynamic-style" chill model and "Growing Degree Hours-style" heat model), which are linked in a gradual transition to represent the dormancy phase of temperate trees. In the figure, PDBF, DBF, and GDH stand for "Precursor of Dormancy Breaking Factor", "Dormancy Breaking Factor" and "Growing Degree Hour", respectively.

Table 1

Description, default values and best fit parameters for the two versions of the PhenoFlex framework that we implemented. The values after the $\pm$ signs represent the standard errors estimated by bootstrapping. PhenoFlex$_{all}$ included 5 marginal experimental seasons in addition to 35 experimental seasons randomly selected at the beginning of the analysis. For detailed description of the parameters, we refer to Luedeling et al. (2021).

Model parameter	Description	Default value	PhenoFlex$_{all}$	PhenoFlex$_{excluded}$
y_c	Chill requirement (value defining the end of chill accumulation)	40	68.28 ± 0.82	33.70 ± 1.67
z_c	Heat requirement (value defining the end of heat accumulation)	190	319.10 ± 9.52	370.97 ± 23.00
s_1	Slope for the transition between chilling and forcing phases	0.5	0.85 ± 0.35	0.13 ± 0.02
T_u (°C)	Optimal temperature for the heat model	25	27.89 ± 0.30	24.95 ± 0.34
E_0 (K)	Activation energy for intermediate compound formation (chill model)	3,372.8	$3,310.35 \pm 0.03$	$3,371.00 \pm 1.22$
E_1 (K)	Activation energy for intermediate compound destruction (chill model)	9,900.3	$9,901.64 \pm 0.20$	$9,901.25 \pm 2.86$
A_0 (h^{-1})	Amplitude for intermediate compound formation (chill model)	6,319.5	$6,396.17 \pm 25.03$	$6,214.57 \pm 55.81$
A_1 (h^{-1})	Amplitude for intermediate compound destruction (chill model)	$5.94 \bullet 10^{13}$	$5.94 \bullet 10^{13} \pm 1.83 \bullet 10^8$	$5.94 \bullet 10^{13} \pm 2.08 \bullet 10^8$
T_f (°C)	Temperature for the transition of the sigmoidal function (chill model)	4	6.48 ± 0.94	1.74 ± 0.44
T_c (°C)	Upper threshold temperature for the heat model	36	27.91 ± 1.34	53.34 ± 0.35
T_b (°C)	Base temperature for the heat model	4	5.59 ± 0.42	4.01 ± 0.69
s (K^{-1})	Sigmoidal function slope determining the production of Chill Portions	1.6	1.39 ± 12.15	3.17 ± 0.20

tenth model run, which resulted in the lowest RMSE of our optimization procedure.

2.4. Parameter uncertainty estimation and model validation

Following the procedure proposed by Luedeling et al. (2021), we used bootstrapping to assess the uncertainty of our parameter estimations. In brief, we randomly sampled (with replacement) the residuals for bloom dates estimated during the calibration of the PhenoFlex model and added the sampled residuals to the observed bloom dates of our calibration data set. This step allowed us to obtain an additional data set with 40 new bloom dates. In a second step, we reimplemented the fitting procedure with the new bloom dates (observed records plus the sampled residuals) to estimate a new set of parameters. We iterated this procedure ten times and saved the fitted parameters after each run. To express the uncertainty in our parameter estimation, we computed the standard deviation across the ten bootstrapping replications as well as the 16th and 84th percentiles, which can be used to characterize the standard error in non-normally distributed data (i.e. equivalent to $\mu \pm \sigma$ in a normal distribution). We then used the different sets of parameters to estimate bloom dates (10 times per experimental season according to the bootstrapping replicates) using the validation data set. We estimated the uncertainty around the estimated bloom dates by computing the standard deviation across the ten replications.

We assessed the performance of the PhenoFlex framework for the two versions of the analysis by computing the Root Mean Square Error (RMSE) as well as the Ratio of Performance to Interquartile range (RPIQ). Whereas RMSE can be considered a good indicator for comparing model performance (with lower values indicating better performance), this metric may be greatly influenced by the distribution of the samples in validation data sets (Bellon-Maurel et al., 2010). RPIQ on the other hand (with greater values indicating better performance), offers a standardization of standard errors by considering the distribution of the population in validation data sets (Bellon-Maurel et al., 2010). For calculation of both metrics, we used the respective functions contained in the chillR package (version 0.72.4; Luedeling (2021)). In addition to RMSE and RPIQ, we computed the Akaike Information Criterion corrected for small sample size (AICc; Burnham and Anderson (2003)) to compare the results of the fitting procedure (only for model calibration) in PhenoFlex$_{all}$ and PhenoFlex$_{excluded}$. For the specific equation implemented to compute AICc, we refer to the original PhenoFlex manuscript (Luedeling et al., 2021). In addition, we compared the performance of both versions of PhenoFlex with a null model that predicts bloom to always fall on the mean bloom date of the calibration dataset. To this end, we estimated the RMSE between the average bloom date in the calibration data sets (i.e. across 40 experimental seasons) and the bloom date in the validation data sets used in PhenoFlex$_{all}$ and PhenoFlex$_{excluded}$.

2.5. Fitted chill and heat response curves

Since the procedure implemented to fit PhenoFlex to data generates parameters for the chill and heat sub models that differ from the values reported in their original versions, we characterized the idealized response curve of these sub models to variation in temperature. To this end, we computed the response of the Dynamic model, using the parameters obtained in the calibration procedure, to 1,200 h at temperatures between -5°C and 25°C using a 0.1°C interval. This approach allowed us to estimate the accumulation of chill (in arbitrary units) to all temperatures within the specified range as well as to characterize chill effectiveness using the newly fitted parameters for the chill sub model. In the case of the Growing Degree Hours model, we estimated the heat accumulation efficiency (between 0 and 1) following the same approach but using temperatures between -5°C and 60°C with a 0.1°C interval. We implemented the same characterization for each version of the analysis and compared the response curves graphically.

2.6. Reproducibility, tools for data preparation, model implementation and figure generation

Data curation and preparation, model implementation, model performance analysis and figure generation were done in the R programming environment (R Core Team, 2021). For agro-climatic analyses and implementation of PhenoFlex we used the chillR package (version 0.72.4; Luedeling (2021)). For data preparation and figure generation we mainly used libraries within the tidyverse framework (Wickham et al., 2019). For reproducibility, we deposited all data used in this study as well as procedures to implement the analysis in a public GitHub repository (https://github.com/EduardoFernandezC/phenoflex _exp_data).

3. Results

3.1. Model parameters

We observed considerable differences for some of the 12 fitted parameters between the two versions of the PhenoFlex framework (Table 1). Whereas in PhenoFlex$_{all}$ we estimated a chill requirement of 63.28 ± 0.82 units, in PhenoFlex$_{excluded}$ this parameter reached a much smaller value of 33.70 ± 1.67 units of chill. Regarding the heat requirement, we observed greater values in PhenoFlex$_{excluded}$ (370.97 ± 23.00 heat units) compared to PhenoFlex$_{all}$ (319.10 ± 9.52 heat units). Note, however, that these values cannot be interpreted in absolute terms, but have to be considered relative to the efficiency of the chill and heat response (shown in section "3.2. Fitted chill and heat response curves"). When excluding the marginal experimental seasons from the calibration procedure (PhenoFlex$_{excluded}$), our analysis suggested a

reduction in the slope for the transition between the chilling and forcing periods (s_1 parameter) from 0.85 ± 0.35 to 0.13 ± 0.02. We observed comparable results for most parameters associated with the Dynamic model except for the slope (parameter s) of the sigmoidal function, which determines what fraction of the intermediate compound is converted into a Chill Portion, and the temperature modulating this transition (parameter T_f). The T_f parameter reached a value about 4 times smaller in PhenoFlex$_{excluded}$ compared to PhenoFlex$_{all}$ and the s parameter reached a value about 2.5 times greater in the second version of the analysis (PhenoFlex$_{excluded}$; Table 1). While we observed comparable results among versions for the parameters T_u and T_b (optimum and base temperature for the Growing Degree Hours model, respectively), the parameter for the upper threshold in the GDH model (parameter T_c) differed greatly between PhenoFlex$_{all}$ (27.91 ± 1.34°C) and PhenoFlex$_{excluded}$ (53.34 ± 0.18°C).

3.2. Fitted chill and heat response curves

Despite some differences in the specific values for optimal temperature between PhenoFlex$_{all}$ and PhenoFlex$_{excluded}$, our analysis produced comparable response curves for chill and heat effectiveness (Fig. 3). In PhenoFlex$_{all}$, we found noticeable chill effectiveness (>1 unit of chill) for temperatures above 0.8°C, although some chill response (<0.01 units of chill) was observed at temperatures below 0°C. In PhenoFlex$_{excluded}$, noticeable chill effectiveness of 2.12 units was found to

occur at temperatures of -0.5°C, increasing rapidly for temperatures above 0°C (8.14 units of chill). Maximum chill efficiency of 36.6 units of chill occurred in PhenoFlex$_{all}$ at 8.8°C, whereas PhenoFlex$_{excluded}$ showed a lower maximum chill efficiency of 26.3 units at 4.0°C. Fitted parameters indicated differing ranges of effective temperatures for chill accumulation among versions of the analysis. In PhenoFlex$_{all}$, chill accumulation was registered at temperatures up to 14.1°C, whereas in PhenoFlex$_{excluded}$ only temperatures up to 11.1°C were found effective.

Regarding heat effectiveness for temperatures between 0°C and 20°C, our results suggest a similar response among versions of the analysis (despite PhenoFlex$_{excluded}$ showing a slightly lower base temperature for development). We observed major differences for temperatures >20°C, with an observed maximum heat efficiency at 27.8°C in PhenoFlex$_{all}$ and 24.9°C in PhenoFlex$_{excluded}$. While we observed a rapid drop to 0 heat units after reaching the maximum efficiency in PhenoFlex$_{all}$, the heat efficiency in PhenoFlex$_{excluded}$ showed a more gradual decline towards 0 heat units for temperatures above 53.3°C (Fig. 3).

3.3. Model performance

Including the five marginal experimental seasons (PhenoFlex$_{all}$) in the calibration data set appeared to reduce the performance of the PhenoFlex framework (Fig. 4). Compared to PhenoFlex$_{excluded}$, the results of PhenoFlex$_{all}$ showed a greater RMSE and AICc, and a smaller RPIQ for the calibration data set (Fig. 4).
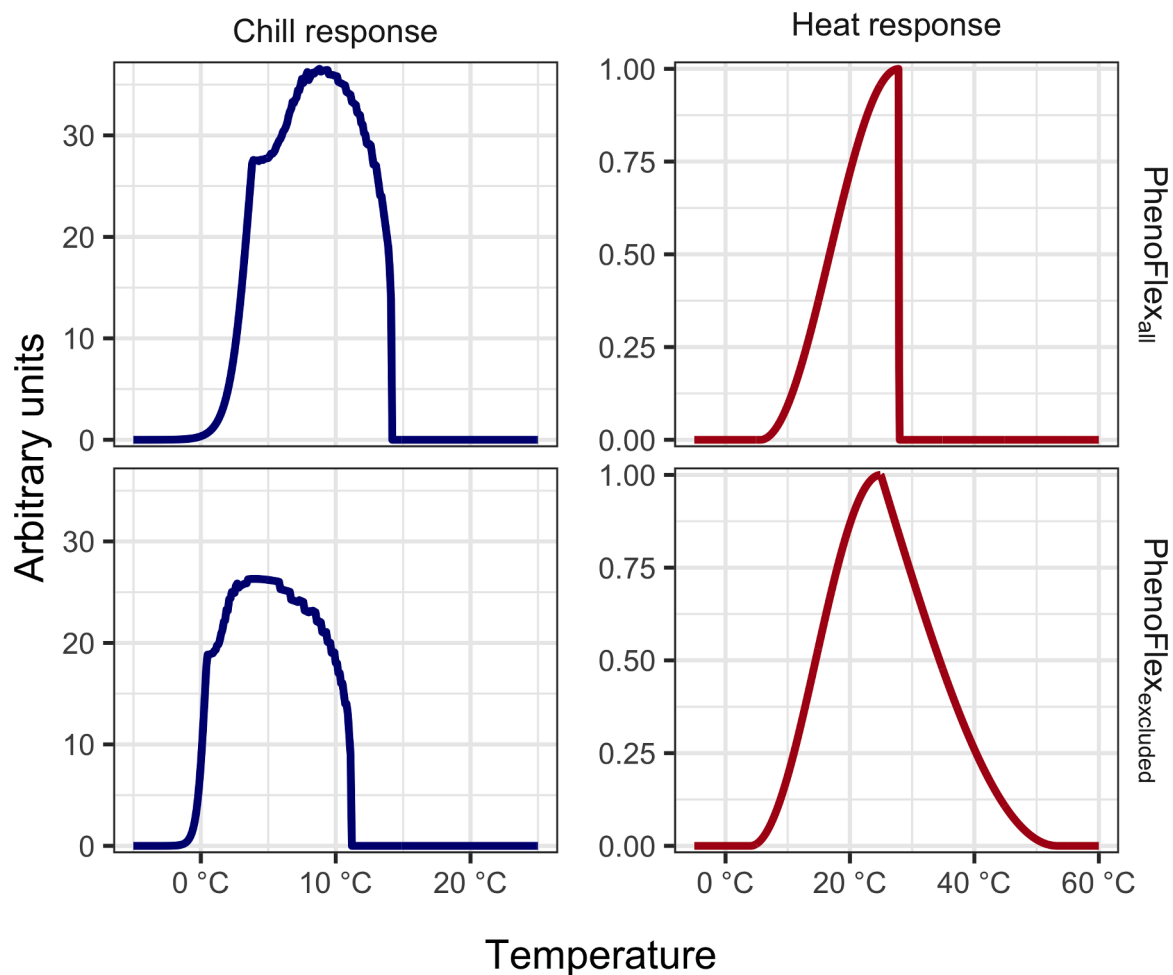

Fig. 3. Chill and heat response curves for the two versions of the analysis. On the y-axis of the left panels, we show absolute arbitrary units for chill effectiveness after 1,200 hours at a constant temperature between -5°C and 25°C. On the y-axis of the right panels, we show heat efficiency (between 0 and 1) for constant temperatures between -5°C and 60°C. The two versions of the analysis differ in the array of experimental seasons used for calibrating the PhenoFlex framework, with PhenoFlex$_{all}$ considering 5 seasons with marginal temperature for overcoming dormancy.

Regarding the validation of the PhenoFlex framework, our results generated an RMSE value of 5.5 days in PhenoFlex$_{all}$ versus an RMSE of 2.3 days in PhenoFlex$_{excluded}$. In contrast to the results of PhenoFlex$_{excluded}$, our RPIQ estimate in PhenoFlex$_{all}$ was lower compared to the value estimated for calibration. Our RPIQ results for validation were 5.7 in PhenoFlex$_{all}$ and 13.8 in PhenoFlex$_{excluded}$. Overall, the bootstrapping procedure suggested small standard errors in the estimation of bloom dates (Fig. 4). Validating the performance of the null model (predicting that bloom always falls on the mean bloom date of the calibration set) in predicting bloom dates of the validation set resulted in RMSE values of 18.47 days for the seasons used in PhenoFlex$_{all}$ and 18.94 days for the seasons used in PhenoFlex$_{excluded}$.

When using the PhenoFlex$_{excluded}$ version to predict bloom in the five marginal seasons, our results suggest an overall poor performance of the framework (Fig. S4). In this validation, RMSE and RPIQ reached values of 61.2 days and 0.3, respectively. It should be noted, however, that these results are greatly influenced by one outlier showing 127 days of error (observed versus predicted). When removing the outlier from the marginal seasons, the prediction error ranged between 9.3 and 44.4 days (Fig. S4).

The residuals obtained when validating the PhenoFlex framework depended on the version of the analysis (Fig. 5). In PhenoFlex$_{all}$, we observed a tendency towards underestimation of bloom dates (a bias towards early dates), whereas in PhenoFlex$_{excluded}$ the model was slightly biased towards late bloom. The median for the residuals was

-4.06 days in PhenoFlex$_{all}$ and 0.71 days in PhenoFlex$_{excluded}$. We observed a greater absolute error in PhenoFlex$_{all}$ (with a mean of 4.84 days) compared to the second version of the analysis (mean of 1.93 days). Finally, we observed a relatively wide distribution of residuals in PhenoFlex$_{all}$ compared to a narrow distribution in PhenoFlex$_{excluded}$ (Fig. 5). The interquartile ranges for the residuals in PhenoFlex$_{all}$ and PhenoFlex$_{excluded}$ were 6.17 and 1.92 days, respectively.

4. Discussion

4.1. Performance of PhenoFlex in comparison with previous approaches

Although comparisons with earlier phenology model assessments are greatly affected by factors such as the methods used for data collection, calibration and validation and possibly the species, some prominent former studies may serve as benchmarks to evaluate our results. Prediction errors for both the PhenoFlex$_{excluded}$ (RMSE of 2.3 days) and PhenoFlex$_{all}$ (RMSE of 5.5 days) models were much lower than those of the naïve null model based on average bloom dates (RMSE of 18-19 days), indicating their usefulness for forecasting bloom dates in young apple trees. Our results mirror the values obtained by Luedeling et al. (2021), who validated the same modelling framework with historic data from mature apple trees cv. "Boskoop" in Germany (RMSE of 3.82 days) but are considerably lower than the values reported by Hoffmann and Rath (2013), who developed 6 phenology models to predict bloom and
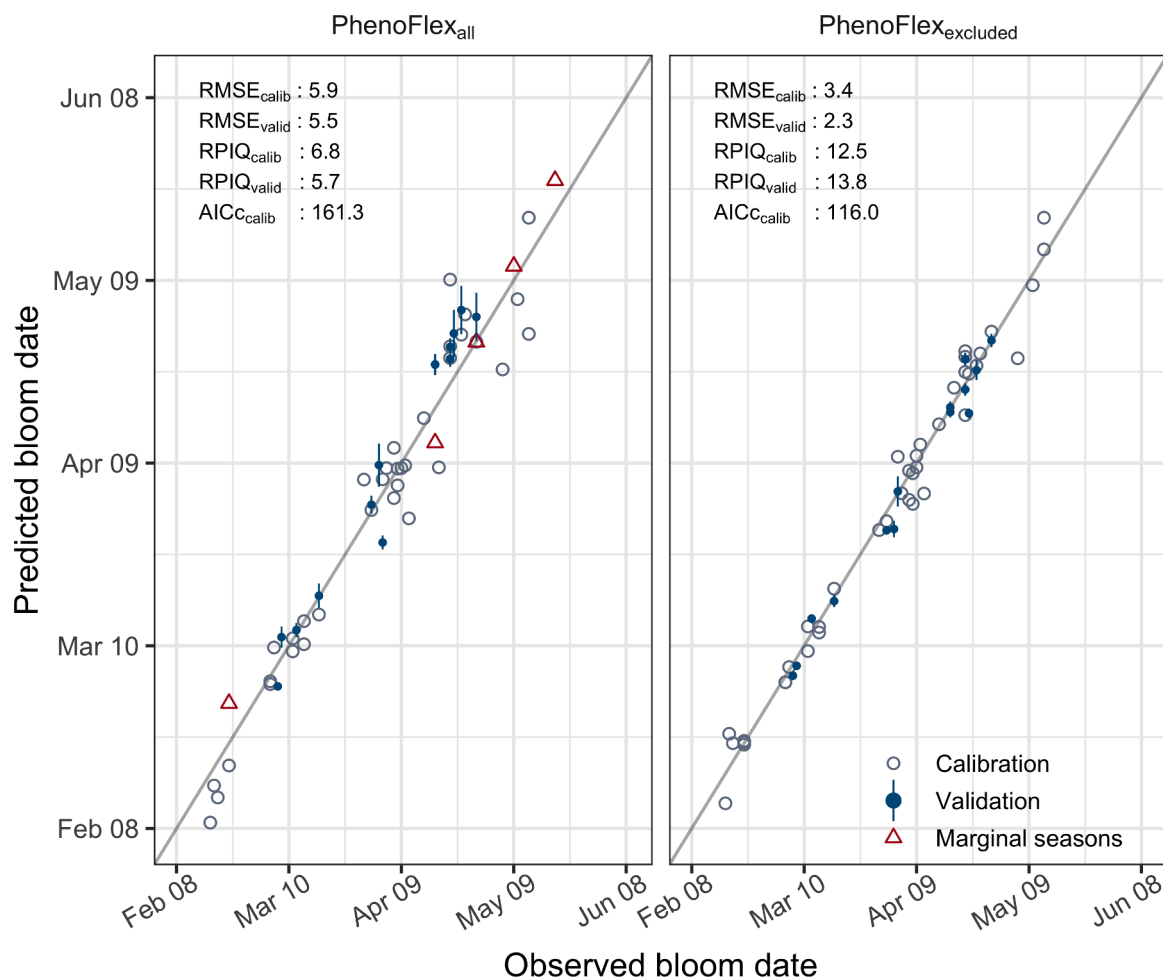

Fig. 4. Observed versus predicted bloom dates for calibration (open circles and triangles) and validation (solid blue circles) data sets and model performance metrics for the two versions of the analysis. We show the five marginal experimental seasons included in the analysis in PhenoFlex$_{all}$ using red open triangles. The remaining experimental seasons used for calibration in both versions are represented by the blue open circles. The solid grey line represents a perfect match between observed and predicted bloom dates. The vertical lines in the validation circles represent the error estimated by bootstrapping.
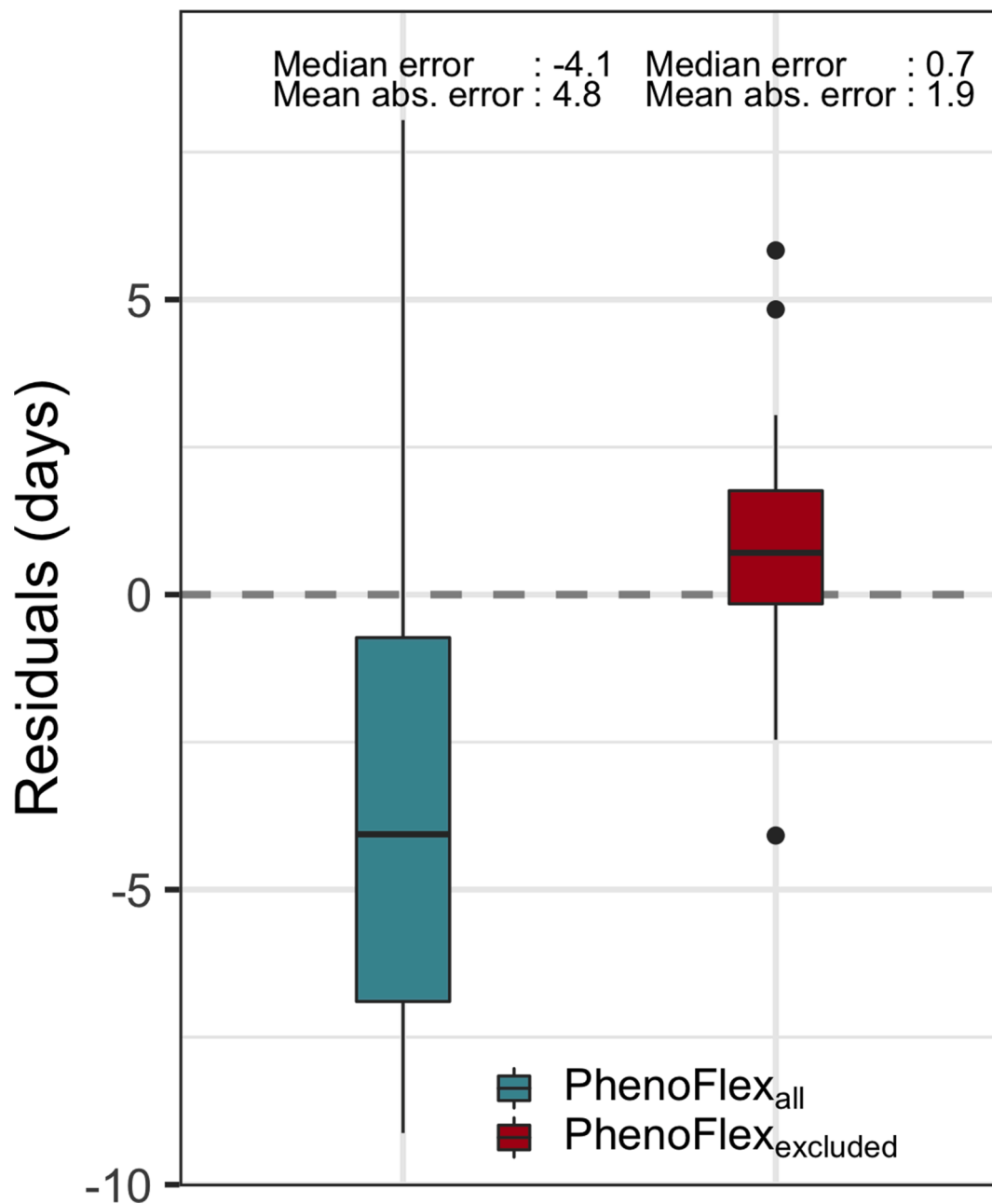
Fig. 5. Distribution of residuals for bloom date projections (14 experimental seasons) obtained with the validation data set for the two versions of our PhenoFlex modelling framework. The versions represent different approaches for calibrating the framework (with PhenoFlex$_{excluded}$ excluding five marginal seasons from the calibration data set).

frost risk for apple trees in Germany (RMSE of 5.9 days on average). Along the same lines, our RMSE results are lower compared to the results reported by Darbyshire et al. (2017), who evaluated the phenology of apple cv. "Golden Delicious" in 14 locations across the globe and reported an RMSE between 6.7 and 17.1 days for a sequential approach and between 5.5 and 8.8 days for a chill overlap model. Our RMSE results are also lower than the values reported by Darbyshire et al. (2016), who analyzed phenology data of apple cv. "Cripps Pink" in Australia using a fixed-overlap structure (RMSE of 5.8 days). Darbyshire et al. (2016), hypothesized that chill accumulated after fulfilling the CR can modify the HR for a certain share of the forcing phase, suggesting an overlap of 75% as best suited for explaining the observations. Compared to using a hard threshold for the overlap, the more flexible structure used in PhenoFlex to determine the shape of the transition between the chilling and forcing phases may explain its better performance relative

to the approach taken by Darbyshire et al. (2016). With adequate data for calibration and after assessing the framework with data from different species and environments, the flexible structure included in PhenoFlex may offer researchers a good option for improving the accuracy of current and future phenology assessments.

4.2. Impacts of marginal seasons and considerations for future model implementation

We observed a considerable improvement in the performance of PhenoFlex when excluding the five marginal experimental seasons from the calibration data set. Only to provide a context in terms of growing conditions, we estimated a median chill accumulation of 24 Chill Portions (CP) for these marginal seasons (using as reference the chilling phase statistically delineated by Fernandez et al. (2021b) for apple trees

using the same experimental design), which is substantially lower compared to reported chill requirements of about 50 CP for this cultivar (Fernandez et al., 2021b; Fernandez et al., 2020a). We hypothesize that the difference in model performance resulted from including somewhat extreme values in the calibration procedure of PhenoFlex$_{all}$, a situation that potentially hampered the identification of adequate model parameters. Accordingly, we obtained greater error estimates for the version of the analysis that included the marginal seasons for the parameter s_1 (0.35 versus 0.02), which defines the slope for the chilling/forcing transition. To some extent, the good performance of the framework we observed when excluding the marginal seasons may challenge the hypothesis by Luedeling et al. (2021), who suggested the need for including relevant environmental conditions in the analysis to reach adequate model fits. We highlight, however, that the five marginal seasons considered in PhenoFlex$_{all}$ may have been too warm for apple trees to overcome dormancy through its regular dormancy release mechanisms. We believe the remaining seasons featured an adequate inter-seasonal variation that helped the fitting procedure identify an appropriate set of parameters and response curves as suggested by Luedeling et al. (2021), ensuring a good performance of the framework under the kind of conditions that are usually faced by apple trees during the dormancy season.

Even though we were unable to validate both versions of the framework using the marginal seasons, the results of our supplementary analysis (i.e. validation of PhenoFlex$_{excluded}$ with the five marginal seasons; supplementary materials) highlight the importance of explicitly testing the predictive accuracy of the model under environmental conditions outside the range of the calibration data. When comparable temperature conditions are used for calibration and validation, the model is – despite potentially incorrect assumptions – sufficiently flexible to be well fitted to data and shows good projection quality. However, when validation data contain conditions outside the calibration range, the conceivably inaccurate model structure cannot be compensated by parameter fitting and finds expression in high residual errors. The lower performance of PhenoFlex$_{excluded}$ when used to predict bloom under marginal conditions may also suggest a limitation of our current understanding of the dormancy release process. A related plausible explanation is that under extreme conditions, some alternative mechanisms to the regular dormancy release processes, which are potentially ignored in the current version of the PhenoFlex framework, may be involved in breaking dormancy in temperate trees. Further research is required to systematically assess the extent of this apparent limitation of the framework and improve the applicability of PhenoFlex.

4.3. Effectiveness of chill and heat models with fitted parameters

Despite slight differences between PhenoFlex$_{all}$ and PhenoFlex$_{excluded}$, our analysis outlined generally plausible response curves for chill and heat accumulation. We observed the start of chill accumulation at temperatures near 0°C in both versions of the analysis, with an upper limit around 15°C in PhenoFlex$_{all}$ and 11°C in PhenoFlex$_{excluded}$. Whereas the lower threshold we observed for chill accumulation is comparable to the 2°C reported by Luedeling et al. (2021), the upper bounds are clearly different from the >30°C obtained in the former study. The inclusion of relevant environmental conditions (e.g. more inter-seasonal variation) in our data set may have helped the model identify a more reasonable response to chilling conditions (Luedeling et al., 2021). The results for chill accumulation we observed for PhenoFlex$_{excluded}$ are comparable to the outputs reported by Egea et al. (2021), who recently explored a re-parameterization of the Dynamic model using experimental data for apricot in Spain and who reported chill accumulation for temperatures between 0°C and 12.1°C. For PhenoFlex$_{all}$, the overall shape of the chill response curve resembles the pattern reported in the original Dynamic model study (Fishman et al., 1987a).

Regarding the heat response, our results indicate comparable base

(5.6°C in PhenoFlex$_{all}$ and 4.0°C in PhenoFlex$_{excluded}$) and optimum (27.9°C in PhenoFlex$_{all}$ and 25.0°C in PhenoFlex$_{excluded}$) temperature thresholds for both versions of the analysis. For the upper bound of the heat response, however, the results differed widely across the versions. In PhenoFlex$_{excluded}$, we observed a smooth decline for above-optimal temperatures towards null efficiency for temperatures above 53.3°C, mirroring the shape of the response reported by Luedeling et al. (2021) for apple cv. "Boskoop". While the overall shape of the heat response curve observed in PhenoFlex$_{excluded}$ seems plausible, the apparently extreme upper limit (i.e. 53.3°C) may indicate a need for further research on heat efficiency during eco-dormancy. In PhenoFlex$_{all}$, on the other hand, we observed a sharp drop beyond the maximum heat efficiency, which mirrors the shape of the response reported in the original PhenoFlex study for pear cv. "Alexander Lucas" (Luedeling et al., 2021). The sharp drop observed in PhenoFlex$_{all}$ conflicts with our understanding of biological processes, which rarely feature such abrupt responses to temperature variation (Hatfield et al., 2011). In this regard, the five additional seasons included in PhenoFlex$_{all}$ may have been unusually marginal, suggesting a need for more intermediate environmental conditions to decipher whether the hard drop is a model artifact or a genuine plant response during eco-dormancy.

4.4. Model parameters fitted under marginal and non-marginal conditions

We can provide little comparison regarding the specific parameters fitted by the two versions of the analysis. On the one hand, PhenoFlex as well as other re-parameterization studies involving the Dynamic and/or GDH models are rather recent (Egea et al., 2021; Luedeling et al., 2021), and on the other hand, model parameterization is highly sensitive to data collection and calibration procedures. The comparisons we conduct only aim to contextualize our results. The chill requirement in PhenoFlex$_{all}$ was about twice as great as in PhenoFlex$_{excluded}$, probably because of a lower chill efficiency at similar temperature for the version excluding the marginal seasons. It should be noted, however, that the chill units in PhenoFlex are arbitrary and not necessarily comparable across analyses. A similar pattern was reported by Luedeling et al. (2021) when comparing the PhenoFlex$_{fitted}$ (all parameters fitted to data) versus PhenoFlex$_{fixed}$ (chill sub-model parameters from the original Dynamic model) approaches. For the remaining parameters E_0, E_1 and A_0, which define the Dynamic model, our fitting procedure settled at slightly lower values than those reported in the original PhenoFlex article (Luedeling et al., 2021). For A_1 (the amplitude of the intermediate compound destruction), the results among studies are perfectly aligned. We observed strong differences when comparing our values for the same parameters to the values reported by Egea et al. (2021), who re-parametrized the Dynamic model with experimental data for apricot in Spain. The similarities between our results and the results by Luedeling et al. (2021), who also analyzed apple, and the differences between our results and the results by Egea et al. (2021), highlight the importance of considering the species and even cultivar when characterizing the chill accumulation process in temperate trees.

4.5. PhenoFlex outlook and challenges for phenology modelling approaches

Many previous studies have relied on the original sets of parameters reported for the Dynamic and GDH models to assess climate change impacts on temperate fruit production for particular locations (Buerkert et al., 2020; Darbyshire et al., 2013; del Barrio et al., 2021; Delgado et al., 2021; Fernandez et al., 2020b) or regions (Benmoussa et al., 2020; Fernandez et al., 2021a; Rodríguez et al., 2019). Such general assessments can now be fine-tuned with species and cultivar-sensitive phenology modelling frameworks (e.g. PhenoFlex). However, implementing a framework such as PhenoFlex requires a considerable quantity of phenology data, which are scarce in many regions and for many cultivars. In line with a previous study by Fernandez et al. (2021b), we

demonstrated that multi-environment experiments can be a useful strategy to overcome data limitations. Such analyses, however, are based on the use of artificial conditions (e.g. heated greenhouses) and young potted trees. While artificial environments may offer an opportunity to assess the phenological response of trees under warmer scenarios, there is little certainty on the probability of observing the same settings under field conditions in the future. Along the same lines, we acknowledge that young potted trees may show a slightly different response to temperature compared to mature trees in the orchard. By using this approach, on the other hand, we were able to generate phenology records that correspond to 59 seasons under normal conditions, a situation that may represent a promising alternative for further studies in the context of climate change.

Apart from being quite data-hungry, PhenoFlex (as well as other phenology modeling approaches) strongly relies on sub-models that hardly represent the state-of-the-art on processes associated with chill and heat accumulation. As suggested by Chuine et al. (2016), including the transition point between endo- and eco-dormancy, which is modulated by numerous physiological processes (see Fadón et al. (2020) for a review), may greatly improve the robustness of PhenoFlex or any alternative framework. Accounting for the endo-dormancy release date during the calibration of phenology models will enable researchers to generate accurate projections for future scenarios, particularly in warm-winter regions (Chuine et al., 2016). Such accurate projections are expected to support farmers in making informed decisions in adapting their orchards to meet future challenges.

5. Conclusions

Overall, our experimental seasons helped improve the performance of PhenoFlex when compared to the original study. The improved performance probably resulted from including relevant environmental conditions (e.g. high inter-seasonal variation) in the calibration data set. On the other hand, including five marginal seasons (characterized by warm conditions during winter) in the calibration data set reduced the performance of the model in forecasting bloom dates for the validation data set. We hypothesize that these five seasons prevented the model from defining a reasonable set of parameters by featuring extreme values for calibration. This clearly suggests that researchers and practitioners using PhenoFlex may need to include seasons with comparable thermal conditions for calibration and validation to maximize the performance of the modeling framework.

In general, we conclude that PhenoFlex outlined reasonable temperature responses for both chill and heat accumulation, despite a sharp drop in the heat response when the five marginal seasons were included and an apparently extreme upper limit in the heat curve when removing those seasons. These results indicate that dormancy researchers may focus on elucidating whether the sharp drop in PhenoFlex$_{all}$ and the apparently extreme upper threshold in PhenoFlex$_{excluded}$, both associated with the heat response, are model artifacts or a genuine plant response during eco-dormancy.

Although the approach may be greatly improved by developing enhanced sub-models accounting for chill and heat accumulation, PhenoFlex offers an opportunity to assemble our current knowledge regarding dormancy modelling and accurately predict bloom dates in temperate trees. To provide useful insights that allow additional improvements of PhenoFlex or alternative phenology modelling approaches, researchers and practitioners should validate the framework in different environments and for different species.

Author contributions

EF conceptualized the research idea with EL supervising and guiding EF. EF implemented the experiment, collected the data and conducted the analyses. EL, KS and CU contributed to analyzing and discussing the results. EF wrote the first draft of the manuscript with all authors commenting on and editing the document.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2022.109020.

References

Anderson, J.L., Richardson, E.A., Kesner, C.D., 1986. Validation of chill unit and flower bud phenology models for 'montmorency' sour cherry. Acta Hort 184, 71–78.

Ashcroft, G.L., Richardson, E.A., Seeley, S.D., 1977. A statistical method of determining chill unit and growing degree hour requirements for deciduous fruit trees. HortScience 12 (4), 347. -248.

Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.M., McBratney, A., 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. TrAC - Trends Anal. Chem. 29 (9), 1073–1081.

Benmoussa, H., Luedeling, E., Ghrab, M., Ben Mimoun, M., 2020. Severe winter chill decline impacts Tunisian fruit and nut orchards. Clim. Change 162 (3), 1249–1267.

Buerkert, A., Fernandez, E., Tietjen, B., Luedeling, E., 2020. Revisiting climate change effects on winter chill in mountain oases of northern Oman. Clim. Change 162, 1399–1417.

Burnham, K.P., Anderson, D.R., 2003. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer Science & Business Media, New York, USA.

Campoy, J.A., Ruiz, D., Egea, J., 2011. Dormancy in temperate fruit trees in a global warming context: a review. Sci. Hortic. 130 (2), 357–372.

Cannell, M.G.R., Smith, R.I., 1983. Thermal Time, Chill Days and prediction of budburst in Picea sitchensis. J. Appl. Ecol. 20 (3), 951–963.

Chuine, I., Regniere, J., 2017. Process-Based models of phenology for plants and animals. Annu. Rev. Ecol., Evol. Syst. 48, 159–182.

Chuine, I., et al., 2016. Can phenological models predict tree phenology accurately in the future? The unrevealed hurdle of endodormancy break. Global Change Biol 22 (10), 3444–3460.

Chuine, I., 2000. A unified model for budburst of trees. J. Theor. Biol. 207 (3), 337–347.

Cook, B.I., Wolkovich, E.M., Parmesan, C., 2012. Divergent responses to spring and winter warming drive community level flowering trends. Proc. Natl. Acad. Sci. 109 (23), 9000–9005.

Core Team, R, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Darbyshire, R., Webb, L., Goodwin, I., Barlow, E.W.R., 2013. Impact of future warming on winter chilling in Australia. Int. J. Biometeorol. 57 (3), 355–366.

Darbyshire, R., Pope, K., Goodwin, I., 2016. An evaluation of the chill overlap model to predict flowering time in apple tree. Sci. Hortic. 198, 142–149.

Darbyshire, R., et al., 2017. A global evaluation of apple flowering phenology models for climate adaptation. Agric. For. Meteorol. 240, 67–77.

Darbyshire, R., Lopez, J.N., Song, X.X., Wenden, B., Close, D., 2020. Modelling cherry full bloom using 'space-for-time' across climatically diverse growing environments. Agric. For. Meteorol. 284, 9.

del Barrio, R., et al., 2021. Climate change impacts on agriculture's southern frontier - Perspectives for farming in North Patagonia. Int. J. Climatol. 41 (1), 726–742.

Delgado, A., Dapena, E., Fernandez, E., Luedeling, E., 2021. Climatic requirements during dormancy in apple trees from northwestern Spain - global warming may threaten the cultivation of high-chill cultivars. Eur. J. Agron. 130.

Egea, J.A., Egea, J., Ruiz, D., 2021. Reducing the uncertainty on chilling requirements for endodormancy breaking of temperate fruits by data-based parameter estimation of the dynamic model: a test case in apricot. Tree Physiol 41, 644–656.

Erez, A., Fishman, S., Linsley-Noakes, G.C., Allan, P., 1990. The dynamic model for rest completion in peach buds. Acta Hort 276, 165–174.

Fadón, E., Fernandez, E., Behn, H., Luedeling, E., 2020. A conceptual framework for winter dormancy in deciduous trees. Agronomy 10 (2), 241.

Fernandez, E. et al., 2020a. Mild water stress makes apple buds more likely to flower and more responsive to artificial forcing — Impacts of an unusually warm and dry summer in Germany Agronomy, 10(2): 274.

Fernandez, E., Whitney, C., Cuneo, I.F., Luedeling, E., 2020b. Prospects of decreasing winter chill for deciduous fruit production in Chile throughout the 21st century. Clim. Change 159 (3), 423–439.

Fernandez, E., Caspersen, L., Illert, I., Luedeling, E., 2021a. Warm winters challenge the cultivation of temperate species in South America - a spatial analysis of chill accumulation. Clim. Change 169, 28.

Fernandez, E., et al., 2021b. Boosting statistical delineation of chill and heat periods in temperate fruit trees through multi-environment observations. Agric. For. Meteorol. 310, 108652.

Fishman, S., Erez, A., Couvillon, G.A., 1987a. The temperature-dependence of dormancy breaking in plants - computer-simulation of processes studied under controlled temperatures. J. Theor. Biol. 126 (3), 309–321.

Fishman, S., Erez, A., Couvillon, G.A., 1987b. The temperature-dependence of dormancy breaking in plants - mathematical-analysis of a 2-step model involving a cooperative transition. J. Theor. Biol. 124 (4), 473–483.

Fu, Y.S.H., et al., 2015. Declining global warming effects on the phenology of spring leaf unfolding. Nature 526 (7571), 104–107.

Guo, L., Dai, J.H., Wang, M.C., Xu, J.C., Luedeling, E., 2015. Responses of spring phenology in temperate zone trees to climate warming: a case study of apricot flowering in China. Agric. For. Meteorol. 201, 1–7.

Harrington, C.A., Gould, P.J., St Clair, J.B., 2010. Modeling the effects of winter environment on dormancy release of Douglas-fir. For. Ecol. Manage. 259 (4), 798–808.

Hatfield, J.L., et al., 2011. Climate impacts on agriculture: Implications for crop production. Agron. J. 103 (2), 351–370.

Hoffmann, H., Rath, T., 2013. Future bloom and blossom frost risk for *Malus domestica* considering climate model and impact model uncertainties. PLoS One 8 (10).

Landsberg, J.J., 1974. Apple fruit bud development and growth - analysis and an empirical model. Ann. Bot. 38 (158), 1013–1023.

Lang, G.A., Early, J.D., Martin, G.C., Darnell, R.L., 1987. Endo-, para-, and ecodormancy: physiological terminology and classification for dormancy research. HortScience 22, 371–377.

Luedeling, E., Guo, L., Dai, J.H., Leslie, C., Blanke, M.M., 2013. Differential responses of trees to temperature variation during the chilling and forcing phases. Agric. For. Meteorol. 181, 33–42.

Luedeling, E., Schiffers, K., Fohrmann, T., Urbach, C., 2021. PhenoFlex - an integrated model to predict spring phenology in temperate fruit trees. Agric. For. Meteorol. 307, 108491.

Luedeling, E., 2021. chillR: Statistical methods for phenology analysis in temperate fruit trees.

Luedeling, E., 2012. Climate change impacts on winter chill for temperate fruit and nut production: A review. Sci. Hortic. 144, 218–229.

Meier, U., 2001. Growth stages of mono-and dicotyledonous plants. BBCH Monograph. Federal Biological Research Centre for Agriculture and Forestry, Oxford.

Menzel, A., et al., 2006. European phenological response to climate change matches the warming pattern. Global Change Biol 12 (10), 1969–1976.

Pope, K.S., Da Silva, D., Brown, P.H., DeJong, T.M., 2014. A biologically based approach to modeling spring phenology in temperate deciduous trees. Agric. For. Meteorol. 198, 15–23.

Prats-Llinàs, M.T., DeJong, T.M., Jarvis-Shean, K.S., Girona, J., Marsal, J., 2019. Performance of a chill overlap model for predicting budbreak in Chardonnay grapevines over a broad range of growing conditions. Am. J. Enol. Vitic. 70 (1), 50–59.

Rodríguez, A., et al., 2019. Chilling accumulation in fruit trees in Spain under climate change. Nat. Hazards Earth Syst. Sci. 19 (5), 1087–1103.

Urbach, C., Luedeling, E. and Schiffers, K., 2021. PhenoFlex. https://cran.r-project. org/web/packages/chillR/vignettes/PhenoFlex.html. Last accessed on September 20, 2021.

Vegis, A., 1964. Dormancy in higher plants. Annu. Rev. Plant Physiol. 15, 185–224.

Walther, G.R., et al., 2002. Ecological responses to recent climate change. Nature 416 (6879), 389–395.

Wickham, H., et al., 2019. Welcome to the Tidyverse. J. Open Source Softw. 4 (43), 1686.