Phenological Model Intercomparison for Estimating Grapevine Budbreak Date \((Vitis\ vinifera\ L.)\) in Europe

Luisa Leolini\textsuperscript{1,}*\textsuperscript{1}, Sergi Costafreda-Aumedes\textsuperscript{1}\textsuperscript{1}, João A. Santos\textsuperscript{2}\textsuperscript{1}, Christoph Menz\textsuperscript{3}, Helder Fraga\textsuperscript{2}\textsuperscript{2}, Daniel Molitor\textsuperscript{4}\textsuperscript{1}, Paolo Merante\textsuperscript{5}, Jürgen Junk\textsuperscript{4}\textsuperscript{4}, Thomas Kartschall\textsuperscript{3}, Agnès Destrac-Irvine\textsuperscript{6}, Cornelis van Leeuwen\textsuperscript{6}, Aureliano C. Malheiro\textsuperscript{2}\textsuperscript{1}, José Eiras-Dias\textsuperscript{7}, José Silvestre\textsuperscript{7}, Camilla Dibari\textsuperscript{1}, Marco Bindi\textsuperscript{1}\textsuperscript{1} and Marco Moriondo\textsuperscript{8}\textsuperscript{1}

\textsuperscript{1} Department of Agriculture, Food, Environment and Forestry, DAGRI, University of Florence, UNIFI, 50144 Florence, Italy; sergi.costafreda@unifi.it (S.C.-A.); camilla.dibari@unifi.it (C.D.); marco.bindi@unifi.it (M.B.)
\textsuperscript{2} Centre for the Research and Technology of Agro-Environmental Sciences, CITAB, Universidade de Trás-os-Montes e Alto Douro, UTAD, 5000-801 Vila Real, Portugal; jsantos@utad.pt (J.A.S.); hfraga@utad.pt (H.F.); amalheir@utad.pt (A.C.M.)
\textsuperscript{3} Potsdam Institute for Climate Impact Research, PIK, 14473 Potsdam, Germany; menz@pik-potsdam.de (C.M.); thomas@pik-potsdam.de (T.K.)
\textsuperscript{4} Luxembourg Institute of Science and Technology, LIST, 4422 Luxembourg, Luxembourg; daniel.molitor@list.lu (D.M.); juergen.junk@list.lu (J.J.)
\textsuperscript{5} Tropical Plant Production and Agricultural Systems Modelling—TROPAGS, University of Göttingen, 37077 Göttingen, Germany; paolo.merante@uni-goettingen.de
\textsuperscript{6} EGFV, Bordeaux Sciences Agro, INRAE, University of Bordeaux, ISVV, F-33882 Villenave-d’Ornon, France; agnes.destrac-irvine@inra.fr (A.D.-I.); vanleeuwen@agro-bordeaux.fr (C.v.L.)
\textsuperscript{7} Instituto Nacional de Investigação Agrária e Veterinária, INIAV, Estação Vitivinícola Nacional, Quinta da Almoinhã, 2565-191 Dois Portos, Portugal; eiras.dias@iniav.pt (J.E.-D.); jose.silvestre@iniav.pt (J.S.)
\textsuperscript{8} Institute of Bioeconomy, National Research Council, CNR-IBE, 50019 Florence, Italy; marco.moriondo@cnr.it

* Correspondence: luisa.leolini@unifi.it; Tel.: +39-055-2755742

Received: 16 April 2020; Accepted: 22 May 2020; Published: 29 May 2020

Abstract: Budbreak date in grapevine is strictly dependent on temperature, and the correct simulation of its occurrence is of great interest since it may have major consequences on the final yield and quality. In this study, we evaluated the reliability for budbreak simulation of two modeling approaches, the chilling-forcing (CF), which describes the entire dormancy period (endo- and eco-dormancy) and the forcing approach (F), which only describes the eco-dormancy. For this, we selected six phenological models that apply CF and F in different ways, which were tested on budbreak simulation of eight grapevine varieties cultivated at different latitudes in Europe. Although none of the compared models showed a clear supremacy over the others, models based on CF showed a generally higher estimation accuracy than F where fixed starting dates were adopted. In the latter models, the accurate simulation of budbreak was dependent on the selection of the starting date for forcing accumulation that changes according to the latitude, whereas CF models were independent. Indeed, distinct thermal requirements were found for the grapevine varieties cultivated in Northern and Southern Europe. This implies the need to improve modeling of the dormancy period to avoid under- or over-estimations of budbreak date under different environmental conditions.

Keywords: chilling-forcing models; dormancy; forcing models; phenology
1. Introduction

Grapevine phenological phases have a profound influence on the final grape and wine quality [1–3]. In particular, changes in the current climate conditions are strongly affecting the occurrence of the developmental stages and length of the grapevine growing cycle in the most renowned wine areas [4,5].

Several authors have reported the increasing temperature trend across the most traditional wine areas of the Mediterranean basin during the last decades. For example, Tomasi et al. [6] showed a 2.3 °C-temperature increase over the grape growing season in Veneto region (Italy) during the period 1964-2009, while van Leeuwen et al. [7] evidenced a 2.0 °C increase from 1951 to 2016 in the Bordeaux wine production area. Moreover, Ramos et al. [8] observed an overall warming from 1.0 to 2.2 °C during the grape growing season in north-eastern Spain. These changes are expected to greatly influence grapevine varietal suitability across the European region [9–12].

Since grapevine phenology is mainly temperature-driven [4], the current warming is affecting the phenological stages and shifting the annual growing cycle [13–16]. Among the most important developmental phases, budbreak represents the starting point of vegetative growth. A temperature increase can have a dual effect on this phase (anticipating or delaying its occurrence), with strong consequences on the following reproductive stages. Indeed, a shift in budbreak may determine poor shoot growth together with fewer clusters and an inhomogeneous grape development [17,18].

Under increasing temperature, a delay in budbreak phase can be determined by insufficient chilling during the dormancy period [19]. Indeed, the bud vegetative growth is limited by a first sub-period induced by physiological conditions (endo-dormancy), and then by a second sub-period in which bud opening is affected by the environmental conditions (eco-dormancy [17,20]). In this context, some studies showed that dormancy release is regulated by bud exposure period to temperatures between 0 to 10 °C [19,21], until fulfilment of the chilling requirement, and a subsequent period of temperatures (from 0 to 25 °C), until the reaching of the forcing requirement [22].

Budbreak occurrence being strongly influenced by the dormancy period, several studies attempted the implementation of phenological models for improving the prediction of this developmental stage [23–28]. Phenological models can be classified in two main categories: forcing and chilling-forcing models. Forcing models are simply based on the forcing unit accumulation, starting, in general, from a fixed date in the year. These models only focus on the description of the eco-dormancy period assuming that the chilling unit accumulation requirement was previously fulfilled and the endo-dormancy period had concluded. Instead, chilling-forcing models can explain both the endo- and eco-dormancy periods by taking into consideration the chilling unit and the forcing accumulation with respect to specific temperature thresholds [24,29,30].

Hence, the selection of one of these two categories for an accurate budbreak simulation often gave rise to a wide debate [29,30]. Chilling-forcing models can take into account the influences of temperature on the entire dormancy period, evidencing not only the shortening effect of warmer temperatures during eco-dormancy but also a possible delay of budbreak caused by a failed chilling requirement [13,31]. However, the limited improvement in the performances of these models and the easier implementation provided by forcing models, focus on the issue about the most suitable modeling approach for budbreak estimation. In modeling studies, the choice of phenological model plays an important role in the use of crop simulation models and in determining potential risk of late spring frost events. In crop simulation models, phenology defines the plant growing cycle with important consequences on future biomass partitioning [32]. In this context, an accurate estimation of budbreak stage may avoid randomly selecting a starting point for vegetative biomass accumulation and over- or under-detection of late spring frost events under different climate scenarios [31,33].

Thus, the phenological intercomparison may be useful to highlight the performances of different models in simulating budbreak [29,30]. In order to take into account the influences of the grapevine precocity cycle and the effects of environmental conditions on budbreak estimation, model intercomparison experiments should consider multiple grapevine varieties and study sites. Indeed, grapevine varieties are characterized by various thermal requirements and focusing on a specific variety...
does not allow the range of different physiological responses derived from the impact of temperature on dormancy period to be highlighted [30]. In addition, the use of a single date as a starting point for the chilling and forcing accumulation [29] does not allow the influence of temperature on the beginning of endo- or eco-dormancy periods to be appreciated, especially when the model intercomparison is performed with phenological data of grapevine varieties collected in different environments.

The objective of this study is to evaluate the performances of six phenological models (BRIN tested using hourly and daily temperature) at reproducing the budbreak date of eight grapevine varieties collected at different latitudes in Europe. In order to consider the influences of temperature during the dormancy period on model representations, two starting dates were set during chilling-forcing and forcing models’ simulation, respectively. The analysis of budbreak date estimation is thus addressed to obtain responses of grapevine varieties’ behavior under different environments.

2. Materials and Methods

2.1. Observations

The phenological data were collected from 15 experimental vineyards in three European countries (France, Luxembourg and Portugal, see Figure 1 and Table 1). The weather data (i.e., daily maximum and minimum air temperature, °C) were freely available from weather services. More specifically, for France, the land-based daily weather data were collected by the National Centers for Environmental Information and then archived at NOAA (https://www.ncdc.noaa.gov/ [34]) and from a weather station near the VitAdapt project [35]. Hourly temperatures, rainfall, wind speed and reference evapotranspiration were recorded and stored in Climatik, a database managed by INRA, US 1116 Agroclim research station, 84914 Avignon, France (https://intranet.inra.fr/climatik_v2; limited access). For Luxembourg, the weather data were collected by the Administration des Services Techniques de l’Agriculture at the Institut Viti-Vinicole (https://agriculture.public.lu/) while the weather data for Portugal derived from the Instituto Nacional de Investigação Agrária e Veterinária (INIAV http://www.iniav.pt/) and the Associação para o Desenvolvimento da Viticultura Duriense (ADVID https://www.advid.pt/). The weather data were obtained from the closest and more representative stations for the vineyard weather conditions (Figure 1 and Table 1).

![Figure 1. Map of the vineyards and weather stations situated in three European countries (Luxembourg, France and Portugal). The vineyards are displayed with triangles and points in relation to the use of phenological data for model fitting or application. The arrows indicate that both vineyards and weather stations are located in the same site. 10 weather stations and 10 vineyards are situated in Douro valley (Portugal).](image-url)
Table 1. Observed weather (daily maximum and minimum air temperature, °C) and budbreak data of eight grapevine varieties used for model fitting, validation and application in different locations across Europe. The geographical positions of all weather stations and vineyards are identified with the latitude (Lat) and longitude (Long) coordinates (degrees).

<table>
<thead>
<tr>
<th>Variety × Site</th>
<th>Weather Data</th>
<th>Budbreak Data</th>
<th>Procedure</th>
<th>Variety</th>
<th>Country</th>
<th>Weather Station</th>
<th>Period</th>
<th>Vineyard</th>
<th>Period</th>
<th>Cases</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Validation</td>
<td>Chardonnay</td>
<td>Luxembourg</td>
<td>49.50</td>
<td>6.35</td>
<td>2016–2018</td>
<td>49.50</td>
<td>6.35</td>
<td>2017–2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gewürztraminer</td>
<td>Luxembourg</td>
<td>49.50</td>
<td>6.35</td>
<td>2010–2018</td>
<td>49.50</td>
<td>6.35</td>
<td>2011–2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Grenache</td>
<td>France</td>
<td>43.58</td>
<td>3.96</td>
<td>1950–2017</td>
<td>43.33</td>
<td>3.56</td>
<td>1951–2012</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>49.50</td>
<td>6.35</td>
<td>2010–2018</td>
<td>49.50</td>
<td>6.35</td>
<td>2011–2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>49.50</td>
<td>6.35</td>
<td>2017–2018</td>
<td>49.50</td>
<td>6.35</td>
<td>2018–2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>49.50</td>
<td>6.35</td>
<td>2011–2018</td>
<td>49.50</td>
<td>6.35</td>
<td>2012–2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>49.50</td>
<td>6.35</td>
<td>1972–2018</td>
<td>49.50</td>
<td>6.35</td>
<td>1973–2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>49.50</td>
<td>6.35</td>
<td>1971–2018</td>
<td>49.50</td>
<td>6.35</td>
<td>1972–2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>49.50</td>
<td>6.35</td>
<td>1972–2018</td>
<td>49.50</td>
<td>6.35</td>
<td>1973–2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>49.50</td>
<td>6.35</td>
<td>1974–2018</td>
<td>49.50</td>
<td>6.35</td>
<td>1975–2018</td>
</tr>
</tbody>
</table>

The phenological data were collected for eight grapevine varieties (Cabernet Sauvignon, Chardonnay, Gewürztraminer, Grenache, Pinot Gris, Riesling, Touriga Franca and Touriga Nacional). For France, these data were retrieved from PERPHECLIM platform (https://www6.inrae.fr/projet-accaf-perpheclim_eng/The-PERPHECLIM-project [36]) and VitAdapt project [35]. For Luxembourg, the data were obtained from the Administration des Services Techniques de l’Agriculture at the Institut Viti-Vinicole. Finally, the data for the Lisbon wine region (Portugal) were obtained from the national ampelographic collection at the INIAV, located in Dois Portos. The data for the Douro wine region were collected by the associates from ADVID. Most of these grapevine varieties are for high-quality wine production derived from the world’s benchmark regions and can be grouped on the basis of average growing season temperatures and ripening potential [37]. The phenological observations were collected by different research institutions and farmers at budbreak date which typically corresponded to BBCH 09 [38], except for Cabernet Sauvignon, Touriga Franca and Touriga Nacional collected in France (Bordeaux site; BBCH 07). We assumed that the distance between BBCH 07 and 09 is covered in 2 days as reported in [27].

2.2. Phenological Models

In total, we tested 6 models for estimating the budbreak date, divided into forcing (F) and chilling-forcing (CF) models. We used Growing Degree Days (GDD; [39]), WANG [40] and UNIFORC [24] for F models and BRIN [29] (Hourly and Daily approach), UNICHILL and UNIFIED [24] for CF models.
2.2.1. GDD

The GDD model [39] is an F model that describes the eco-dormancy period based on accumulation of the daily average temperature \(T_{\text{avg}}\) (°C) above a specific threshold \(T_b\) (°C). The GDD accumulation \(\text{Force}_{\text{GDD}}\) starts at a fixed date of the year \(t_0\) and is accumulated until the GDD requirement is satisfied at budbreak date \(\text{Bud}\); Equation (1)).

\[
\text{Force}_{\text{GDD}} = \sum_{t_0}^{\text{Bud}} (T_{\text{avg}} - T_b),
\]

2.2.2. WANG

The WANG model [40] is a temperature response function used, in this study, to describe the eco-dormancy period from a fixed date of the year \(t_0\) and not to calculate the chilling unit accumulation as reported for other similar functions [30]. The temperature function \(\text{Force}_{\text{WANG}}\) allows the phenological development of grapevine to be evaluated using a non-linear approach based on three cardinal temperatures: \(T_{\text{opt}}\) (optimum temperature, °C), \(T_{\text{max}}\) (maximum temperature, °C) and \(T_{\text{min}}\) (minimum temperature, °C; Equations (2a) and (2b)).

\[
\text{Force}_{\text{WANG}} = \sum_{t_0}^{\text{Bud}} \max\left\{ \frac{2 \cdot (T_{\text{avg}} - T_{\text{min}})^{\alpha} \cdot (T_{\text{opt}} - T_{\text{min}})^{\alpha} - (T_{\text{avg}} - T_{\text{min}})^{2\alpha}}{(T_{\text{opt}} - T_{\text{min}})^{2\alpha}}, 0 \right\} \hspace{1cm} (2a)
\]

\[
\alpha = \frac{\ln(2)}{\ln\left(\frac{T_{\text{max}} - T_{\text{min}}}{T_{\text{opt}} - T_{\text{min}}}\right)} \hspace{1cm} (2b)
\]

where \(T_{\text{avg}}\) is the daily average temperature (°C) and \(\text{Bud}\) is the budbreak date.

2.2.3. UNIFORC

The UNIFORC model [24] is an F model based on a sigmoid curve for calculating the forcing unit accumulation \(\text{Force}_{\text{UNIFORC}}\) from a fixed day of the year \(t_0\) until budbreak occurrence \(\text{Bud}\); Equation (3)):

\[
\text{Force}_{\text{UNIFORC}} = \sum_{t_0}^{\text{Bud}} \frac{1}{1 + e^{d(T_{\text{avg}} - e)}},
\]

where \(T_{\text{avg}}\) is daily average temperature (°C), while \(d\) and \(e\) \((d < 0\) and \(e > 0\)) are empirical parameters of the function.

2.2.4. BRIN

The BRIN model [29] is a CF model based on the Bidabe’s Cold Action Model [41,42] in which the endo-dormancy period is evaluated through the accumulation of chilling units \(\text{Chill}_{\text{BRIN}}\); Equation (4)). The eco-dormancy period is based on forcing unit accumulation \(\text{Force}_{\text{BRIN}}\) calculated with the Richardson equation [23]. The \(\text{Force}_{\text{BRIN}}\) is computed using both a daily (Equation (5)) and hourly (Equations (6a) and (6b)) temperature accumulation approach.

\[
\text{Chill}_{\text{BRIN}} = \sum_{t_0}^{\text{Cendoo}} \left( Q_{10c} \frac{T_o(n)}{10} + Q_{10c} \frac{T_o(n)}{10} \right),
\]

\[
\text{Force}_{\text{BRIN}} \hspace{1cm} (\text{Daily}) = \sum_{\text{Cendoo}} \text{Richardson}_{\text{Daily}}(n),
\]

\[
\text{Force}_{\text{BRIN}} \hspace{1cm} (\text{Daily}) = \sum_{\text{Cendoo}} \text{Richardson}_{\text{Daily}}(n),
\]
Forc\textsubscript{BRINHourly} = \sum_{C_{\text{endo}}}^{Bud} Richardson\textsubscript{Hourly}(n) \hspace{1cm} (6a)

Richardson = \text{Max}(\text{Min}(T_{\text{avg}} - T_t, T_h - T_t), 0), \hspace{1cm} (6b)

where \(t_0\) is the simulation starting date, \(T_{\text{avg}}\) is daily average temperature (°C), \(C_{\text{endo}}\) is the date at which the chilling requirement is satisfied, \(Bud\) is budbreak date, \(Q_{10c}\) is the geometric progression rate of the thermal dormancy response, \(T_m\) and \(T_{\text{t}}\) are daily minimum and maximum temperatures (°C), respectively. \(Richardson\textsubscript{Daily}\) and \(Richardson\textsubscript{Hourly}\) are daily and hourly Richardson equations used for representing the eco-dormancy period [29]. The Richardson equations are limited by two cardinal temperatures: a lower \((T_l, °C)\) and an upper limit temperature \((T_h, °C)\).

2.2.5. UNICHILL

The UNICHILL model [24] is a CF two-phase sequential model constituted by two functions describing chilling accumulation \((\text{Chill}_{\text{UNICHILL}})\) during endo-dormancy (Equation (7)) and forcing unit accumulation \((\text{Forc}_{\text{UNICHILL}})\) during eco-dormancy period (Equation (8)). After chilling requirement is satisfied at a specific date \((C_{\text{endo}})\), the forcing units start accumulating until the forcing requirement is reached, leading to the occurrence of the budbreak phase \((Bud)\). The two functions are defined as followed:

\[
\text{Chill}_{\text{UNICHILL}} = \sum_{t_0}^{C_{\text{endo}}} \frac{1}{1 + e^{(T_{\text{avg}} - c)^2 + b(T_{\text{avg}} - c)}}, \hspace{1cm} (7)
\]

\[
\text{Forc}_{\text{UNICHILL}} = \sum_{C_{\text{endo}}}^{Bud} \frac{1}{1 + e^{(T_{\text{avg}} - c)}}, \hspace{1cm} (8)
\]

where \(t_0\) is the simulation starting date, \(T_{\text{avg}}\) is daily average temperature (°C) and \(a, b, c, d,\) and \(e\) are empirical parameters of the function.

2.2.6. UNIFIED

The UNIFIED model [24] is a CF overlapping model based on the same equations for chilling and forcing unit accumulation used in the UNICHILL model. However, UNIFIED model includes the effect of chilling units on forcing accumulation through implementation of the exponential decrease curve (Equation (9)) in which \(\text{ForcReq}_{\text{UNIFIED}}\) (\(F_{\text{crit}}\)) is the critical amount of forcing and \(\text{Chill}_{\text{UNIFIED}}\) is the total amount of chilling at day \(t_d\). The budbreak date is estimated when both chilling and forcing requirements are satisfied:

\[
\text{ForcReq}_{\text{UNIFIED}} = w e^{-z\text{Chill}_{\text{UNIFIED}}(t_d)}, \hspace{1cm} (9)
\]

where \(w\) and \(z\) are empirical parameters of the function.

2.3. Model Fitting and Validation

2.3.1. Outliers Removal

The model fitting, validation and application were performed after removing outliers from the observed dataset. This procedure consists in the identification of possible anomalies by comparing the observed and simulated results obtained by a preliminary models’ fitting against the original datasets for each variety. The outliers were identified by using the Tukey method [43], based on the non-parametric measure of the statistical dispersion of a given distribution quantified using the interquartile range (IQR). For each model and starting date, the outliers represent the values in which the difference between the observed and simulated budbreak Days of Year (DOYs) are outside the whiskers of the distribution (below \(Q1 - 1.5 \text{ IQR}\) or above \(Q3 + 1.5 \text{ IQR}\) values). An outlier of a specific variety is removed from the dataset when it is observed in at least 4 combinations (models x starting
The percentage of outliers removed from the fitting, validation and application datasets is: 4.17% (Cabernet Sauvignon), 5.36% (Chardonnay), 1.85% (Gewurztraminer and Riesling), 3.70% (Grenache), 1.52% (Pinot Gris), 10% (Touriga Franca and Nacional). These values were generally observed after extreme weather conditions (e.g., frosts of 1956, 1963 in France) or due to other influences (e.g., agronomic practices or errors in data collection) which cannot be reproduced during model simulation.

2.3.2. Parameters’ Range and Starting Date

The parameters’ range of variation in the 6 phenological models was defined according to the literature [29,30,44–46]. For the GDD model, the fitted parameter was $T_b$ which ranges from 0 to 10 °C. In this study, 0 °C was considered the limit temperature for bud growth while 10 °C was the maximum base temperature adopted for describing the eco-dormancy period as found in other studies [19,22,29,44]. For the WANG model, this function being generally used for wheat simulation [47], the parameter range of the cardinal temperatures was established considering the values obtained by other similar functions (Fbeta) implemented for simulating budbreak date in grapevine [30]. Thus, $T_{\min}$ and $T_{\max}$ were set at 0 and 45 °C respectively, while $T_{opt}$ ranged in a ±30% of the value reported in Fila et al. [30]. For UNIFORC model, the $d$ and $e$ parameter values varied in an ±30% interval with respect to the average value reported for different grapevine varieties (Glera, Cabernet Sauvignon, Chardonnay and Merlot [46]). For both daily and hourly approaches of BRIN model, the $Q_{10}$ parameter range was defined in a ±30% interval of the value estimated for several grapevine varieties [29]. Moreover, the lower limit temperature ($T_l$) was fitted in the range 0–10 °C, as reported for GDD model while the upper limit temperature ($T_h$) for post dormancy was fixed at 25 °C as indicated in Garcia de Cortazar-Atauri et al. [29]. For the UNICHILL model, the parameters $a$, $b$, $c$, $d$ and $e$ were in a range of ±30% of the overall minimum and maximum value reported by Fila [46] for Glera, Cabernet Sauvignon, Chardonnay and Merlot varieties (field calibration). However, in the specific case of Cabernet Sauvignon and Chardonnay, the parameter range of $a$, $b$, $c$, $d$ and $e$ varied in ±30% interval with respect to the parameter values found for these varieties. Similarly, the range adopted for $a$, $b$, $c$, $d$ and $e$ parameters in UNIFIED model was identified by considering ±30% of the minimum and maximum values reported by Fila et al. [46] (field calibration). Based on the same study, the parameters $w$ and $z$ were fixed by considering the average values reported for field calibration of Sangiovese and Montepulciano.

Concerning the use of fixed starting dates $t_0$ during model simulation, we set the 1st of January [30] and 1st of March [27] of the current year for the F models and the 1st September [24] and 1st of August [29] of the previous year for the CF models.

2.3.3. Model Fitting, Validation and Application

The methodology used for model fitting, validation and application is displayed in Figure 2. After outliers removal (Section 2.3.1), the first step was to randomly split the observed data (budbreak DOY per year) into 10 random subsamples (RS) in which 60% of the observed data was used for model fitting whereas the remaining 40% was retained for validation. Once the 10 random subsamples were prepared, the model fitting was performed for any combination of varieties, models and starting dates using the Phenological Modeling Platform (PMP) software (version 5.5) [48] and setting the parameter ranges as described in Section 2.3.2. The best fitting for any combination was found by exploring the range of each parameter through the Metropolis-Hastings algorithm [49–51] of PMP. Each fitted model was validated on the remaining 40% of the data and then applied on the validation datasets of the other subsamples by using the R software environment (version 3.5.0) [52]. Thus, the parameter set of each model combination (model x variety x starting date), which provided the highest-performance (HP) during fitting and validation procedures, was selected. Moreover, the average parameter sets (AV) derived from all fitted subsamples was also tested on the entire distribution of validation datasets. Finally, the HP and AV parameter sets were applied to the independent datasets of other study sites to evaluate the behavior of different parameterizations under different environmental conditions.
2.4. Statistical Analysis

The statistical indices used for evaluating the phenological model performances during fitting, validation and application were: R-squared ($R^2$, Equation (10), [53]); Akaike information criterion (AIC, Equation 11, [54]); Root Mean Squared Error (RMSE, Equation (12), [55]) and modeling efficiency coefficient (EF, Equation (13), [56]), each defined as follows:

$$R^2 = 1 - \frac{\sum_i (O_i - P_i)^2}{\sum_i (O_i - \bar{O})^2},$$  \hspace{1cm} (10)

$$\text{AIC} = 2k - 2 \ln(L),$$  \hspace{1cm} (11)

$$\text{RMSE} = \left[ \frac{\sum_i (P_i - O_i)^2}{n} \right]^{0.5},$$  \hspace{1cm} (12)

$$\text{EF} = 1 - \frac{\sum_i (P_i - O_i)^2}{\sum_i (O_i - \bar{O})^2},$$  \hspace{1cm} (13)

where $O_i$ is the observed value, $\bar{O}$ is the average of the observed values, $P_i$ is the predicted value, $n$ is the number of observations, $k$ is the number of parameters in the model, and $L$ is the maximized value of the likelihood function of the predicted model.

Finally, the minimum, 10% quantile, median, mean, 90% quantile and maximum values of the statistics were extracted for describing the simulation results on all validation datasets.

3. Results

3.1. Overall Model Behavior in Fitting and Validation

The results of the phenological model intercomparison were described comparing the performances of the different modeling approaches (e.g., F vs. CF) and starting dates (1st January/March vs. 1st September/August) during model fitting and validation. Initially, the model fitting and validation results evidenced higher average performances for CF ($R^2 = 0.74$; AIC = 150.42; RMSE = 4.22, Val, ...
$R^2 = 0.61; \text{AIC } = 111.64; \text{RMSE } = 5.79$) compared to F models ($Fit, R^2 = 0.61; \text{AIC } = 160.62; \text{RMSE } = 5.51; \text{Val, } R^2 = 0.59; \text{AIC } = 112.01; \text{RMSE } = 6.10$). Among all models (Figure 3 and Table S1), the highest performances were achieved by UNIFIED model ($Fit, R^2 = 0.77; \text{AIC } = 147.60; \text{RMSE } = 3.96; \text{Val, } R^2 = 0.64; \text{AIC } = 110.74; \text{RMSE } = 5.42$), while the lowest were obtained by the WANG model ($Fit, R^2 = 0.61; \text{AIC } = 161.15; \text{RMSE } = 5.60$) and GDD model ($\text{Val, } R^2 = 0.58; \text{AIC } = 112.59; \text{RMSE } = 6.25$). Moreover, higher performances were found for those models in which the starting date was fixed at 1st September in CF models ($Fit, R^2 = 0.74; \text{AIC } = 150.24; \text{RMSE } = 4.21; \text{Val, } R^2 = 0.61; \text{AIC } = 111.59; \text{RMSE } = 5.75$) and at 1st January in F models ($Fit, R^2 = 0.66; \text{AIC } = 159.20; \text{RMSE } = 5.53; \text{Val, } R^2 = 0.63; \text{AIC } = 111.10; \text{RMSE } = 6.15$). Finally, concerning grapevine varieties, Chardonnay was best fitted by the models ($Fit, R^2 = 0.90; \text{AIC } = 138.66; \text{RMSE } = 4.90; \text{Val, } R^2 = 0.89; \text{AIC } = 101.54; \text{RMSE } = 5.97$), while the poorest results were obtained for Touriga Nacional ($Fit, R^2 = 0.42; \text{AIC } = 147.09; \text{RMSE } = 6.23; \text{Val, } R^2 = 0.31; \text{AIC } = 105.40; \text{RMSE } = 7.74$) and Touriga Franca in terms of AIC values ($Fit = 180.19; \text{Val } = 127.35$).

Figure 3. Statistical performances ($R^2, \text{AIC, RMSE, EF}$) of all combinations of the six phenological models, eight grapevine varieties and different starting dates during fitting and validation. The results of forcing (F) and chilling-forcing (CF) models (different bar colours) are displayed at the top and bottom of the figure, respectively. The two bars identify the model fitting (no bar texture) and validation (line bar texture). CS = Cabernet Sauvignon, CH = Chardonnay, GE = Gewürztraminer, GR = Grenache, PG = Pinot Gris, RS = Riesling, TF = Touriga Franca, TN = Touriga Nacional.
3.2. HP versus AV

In order to compare the results of different models’ parameterizations, we selected the highest performance parameter set (HP, Table 2) from the 10 fitting and validation subsamples for each combination of model, starting date and variety. We then selected the average parameter set (AV, Table 3) from the 10 fitting subsamples for each combination model, starting date and variety. On these bases, Figures S1–S7 show the performances ($R^2$ and AIC) of the model fitting and validation (symbols) and the applications of the different parameterizations on all validation subsamples (boxplot) obtained by considering all grapevine varieties and starting dates. In detail, the selected parameter sets with highest performances (HP; filled symbols with different shapes according to the starting dates) were the result of the highest rank value of $R^2$ and AIC and the minimum discrepancy between fitting (red) and validation (blue) models. By contrast, the AV parameter sets (boxplots in shadow area) were the application of the average parameterization obtained for all subsamples. Despite that similar performances for HP and AV parameter sets were evidenced in Figures S1–S7, higher discrepancy was found in the behavior of CF and F models. In Figures S1–S4 (CF models), no differences were observed when different starting dates were set during model fitting and application. On the other hand, high discrepancy was found in model performances when different starting dates were adopted in Figures S5–S7 (F models). This trend is also evidenced by the correlations between observed and simulated DOYs in Figure 4, Figures S8 and S9. The results of model fitting showed accurate performances for all selected HP, with more discrepancy between starting dates of F models (yellow and green correlations) compared to those used in CF models (blue and red correlations). The values of $R^2$ also highlighted the higher performances obtained with some varieties compared to others (Chardonnay vs. Touriga Nacional).

Figure 4. Correlations between observed and simulated budbreak Days of Year (DOY) of the GDD and UNIFIED models. The results were obtained for all grapevine varieties and starting dates using the HP parameter set.
Table 2. HP (highest-performance) parameter sets for each combination of models, grapevine varieties and starting dates. \( C_{\text{crit}} \) = critical amount of chilling and \( F_{\text{crit}} \) = critical amount of forcing.

<table>
<thead>
<tr>
<th>Variety</th>
<th>GDD</th>
<th>WANG</th>
<th>UNIFORC/UNICHILL/UNIFIED</th>
<th>BRIN h/BRIN d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( t_0 )</td>
<td>( F_{\text{crit}} )</td>
<td>( t_h )</td>
<td>( F_{\text{crit}} )</td>
</tr>
<tr>
<td>Cabernet Sauvignon</td>
<td>1</td>
<td>763.47</td>
<td>0</td>
<td>67.32</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>206.89</td>
<td>1.18</td>
<td>19.49</td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>433.23</td>
<td>2.34</td>
<td>14.63</td>
</tr>
<tr>
<td>Chardonnay</td>
<td>60</td>
<td>49.18</td>
<td>7.21</td>
<td>4.00</td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>645.68</td>
<td>0</td>
<td>47.27</td>
</tr>
<tr>
<td>Grenache</td>
<td>60</td>
<td>85.11</td>
<td>5.99</td>
<td>5.17</td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>137.17</td>
<td>7.10</td>
<td>12.11</td>
</tr>
<tr>
<td>Gewürztraminer</td>
<td>60</td>
<td>255.89</td>
<td>3.61</td>
<td>9.17</td>
</tr>
<tr>
<td>Pinot Gris</td>
<td>60</td>
<td>122.16</td>
<td>7.39</td>
<td>12.13</td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>250.48</td>
<td>3.91</td>
<td>9.36</td>
</tr>
<tr>
<td>Riesling</td>
<td>60</td>
<td>156.14</td>
<td>6.74</td>
<td>12.60</td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>279.34</td>
<td>3.46</td>
<td>11.80</td>
</tr>
<tr>
<td>Touriga Franca</td>
<td>60</td>
<td>743.93</td>
<td>0</td>
<td>36.53</td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>149.08</td>
<td>2.86</td>
<td>7.61</td>
</tr>
<tr>
<td>Touriga Nacional</td>
<td>60</td>
<td>688.00</td>
<td>1.06</td>
<td>35.57</td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>199.44</td>
<td>0.80</td>
<td>8.19</td>
</tr>
</tbody>
</table>
**Table 3.** AV (average) parameter sets for each combination of models, grapevine varieties and starting dates. $C_{crit}$ = critical amount of chilling and $F_{crit}$ = critical amount of forcing.

<table>
<thead>
<tr>
<th>Variety</th>
<th>GDD</th>
<th>WANG</th>
<th>UNIFOR/UNCHILL/UNIFIED</th>
<th>BRIN h/BRIN d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t_0$</td>
<td>$F_{crit}$</td>
<td>$T_0$</td>
<td>$C_{crit}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$F_{crit}$</td>
<td>$T_{opt}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$T_{min}$</td>
<td>$T_{max}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$a$</td>
<td>$b$</td>
<td>$c$</td>
<td>$d$</td>
</tr>
<tr>
<td>Cabernet Sauvignon</td>
<td>1</td>
<td>765.16</td>
<td>0</td>
<td>68.54</td>
</tr>
<tr>
<td></td>
<td>-121</td>
<td></td>
<td></td>
<td>152.62</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>225.23</td>
<td>3.16</td>
<td>11.83</td>
</tr>
<tr>
<td></td>
<td>-152</td>
<td></td>
<td></td>
<td>180.78</td>
</tr>
<tr>
<td>Chardonnay</td>
<td>1</td>
<td>527.48</td>
<td>0.91</td>
<td>22.21</td>
</tr>
<tr>
<td></td>
<td>-121</td>
<td></td>
<td></td>
<td>130.52</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>47.34</td>
<td>7.24</td>
<td>4.10</td>
</tr>
<tr>
<td></td>
<td>-152</td>
<td></td>
<td></td>
<td>163.81</td>
</tr>
<tr>
<td>Grenache</td>
<td>1</td>
<td>653.93</td>
<td>0</td>
<td>54.66</td>
</tr>
<tr>
<td></td>
<td>-121</td>
<td></td>
<td></td>
<td>151.30</td>
</tr>
<tr>
<td>Gewürztraminer</td>
<td>1</td>
<td>314.55</td>
<td>2.55</td>
<td>12.43</td>
</tr>
<tr>
<td></td>
<td>-121</td>
<td></td>
<td></td>
<td>176.80</td>
</tr>
<tr>
<td>Pinot Gris</td>
<td>1</td>
<td>126.62</td>
<td>7.39</td>
<td>12.26</td>
</tr>
<tr>
<td></td>
<td>-121</td>
<td></td>
<td></td>
<td>173.20</td>
</tr>
<tr>
<td>Riesling</td>
<td>1</td>
<td>156.13</td>
<td>6.78</td>
<td>12.43</td>
</tr>
<tr>
<td></td>
<td>-121</td>
<td></td>
<td></td>
<td>166.51</td>
</tr>
<tr>
<td>Touriga Franca</td>
<td>1</td>
<td>719.41</td>
<td>0.07</td>
<td>36.93</td>
</tr>
<tr>
<td></td>
<td>-121</td>
<td></td>
<td></td>
<td>137.41</td>
</tr>
<tr>
<td>Touriga Nacional</td>
<td>1</td>
<td>710.02</td>
<td>0.77</td>
<td>37.39</td>
</tr>
<tr>
<td></td>
<td>-121</td>
<td></td>
<td></td>
<td>108.85</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>176.18</td>
<td>1.89</td>
<td>9.42</td>
</tr>
<tr>
<td></td>
<td>-152</td>
<td></td>
<td></td>
<td>139.52</td>
</tr>
</tbody>
</table>
A deeper analysis of the model validation showed that the normalized difference between HP and AV, \((\text{HP-AV}/\text{HP}) \times 100\), was higher in CF \((R^2 = 11.99\%, AIC = 3.36\%, \text{RMSE} = 22.20\%)\) compared to F models \((R^2 = 5.73\%, AIC = 1.51\%, \text{RMSE} = 8.16\%)\). Among CF models, UNICHILL model showed the lowest difference between HP and AV in terms of \(R^2\) and AIC \((R^2 = 6.20\%, AIC = 2.70\%)\), while it evidenced the highest difference in terms of RMSE \((40.02\%)\). Considering F models, WANG showed the highest difference \((R^2 = 7.07\%, AIC = 2.25\%, \text{RMSE} = 9.38\%)\) and UNIFORC, the lowest \((R^2 = 4.21\%, AIC = 0.84\%, \text{RMSE} = 7.13\%)\). Following the same trend, the starting dates 1st September and 1st August \((R^2 = 3.89\%, AIC = 0.86\%, \text{RMSE} = 15.62\%)\) showed bigger differences compared to 1st January and 1st March \((R^2 = 2.96\%, AIC = 0.72\%, \text{RMSE} = 7.62\%)\). In particular, 1st September showed a lower difference in terms of AIC and RMSE \((AIC = 0.86\%, \text{RMSE} = 13.53\%)\) compared to 1st August in CF models while 1st January showed the highest difference in F models \((R^2 = 2.88\%, AIC = 0.65\%, \text{RMSE} = 7.23\%)\). Finally, regarding grapevine varieties, the highest difference was found for Touriga Nacional \((R^2 = 23.96\%, \text{RMSE} = 24.20\%)\) while the lowest was shown by Chardonnay in terms of \(R^2\) \((1.89\%)\) and by Touriga Franca in terms of AIC and RMSE \((1.10\% \text{ and } 9.55\%, \text{respectively})\).

### 3.3. Model Applications on Other Independent Datasets

The statistical performances of the phenological models, applied on the respective independent datasets, are reported in Figure 5 and Figures S10–S16 while the comparison between CF and F models during fitting and applications is shown in Figure 6 and Figure S17. In general, higher performances were found for CF compared to F models during fitting and applications (Figure 5), except for those varieties characterized by few observed data in the model application (Gewürztraminer, Riesling; Figure 6 and Figure S17).

![Figure 5. Highest-performance (HP) and average (AV) parameters’ set applied on the independent datasets by considering all phenological models and starting dates. The example of Chardonnay variety.](image-url)
In particular, the model application showed similar performances for both HP (\(R^2 = 0.55\), AIC = 57.97, RMSE = 9.41) and AV (\(R^2 = 0.54\), AIC = 58.37, RMSE = 9.26) parameter sets. However, this difference slightly increased when CF (\(R^2 = 0.56\), AIC = 59.05, RMSE = 8.75) and F models (\(R^2 = 0.52\), AIC = 57.01, RMSE = 10.01) were compared without distinguishing between HP and AV parameterizations. Therefore, CF models showed better results in terms of \(R^2\) (+8%) and RMSE (−15%) compared to F models while lower results of AIC value were obtained for CF compared to F models (+3.44%). A detailed analysis revealed that the CF models with highest performances (\(R^2\) and AIC) were UNICHILL (\(R^2 = 0.64\), AIC = 56.74) and UNIFIED (\(R^2 = 0.58\), AIC = 56.61) models, while, in F models, the better results were provided by UNIFORC (\(R^2 = 0.54\), AIC = 55.98). However, in terms of RMSE, highest performances were found for BRIN hourly (8.37, CF) and GDD (9.67, F). Concerning the starting dates, similar results were found when the 1st September (\(R^2 = 0.58\), AIC = 58.92, RMSE = 8.55) and 1st of March (\(R^2 = 0.57\), AIC = 59.17, RMSE = 8.95) were analyzed, with the 1st September showing better performance (\(R^2 = +1.83\), AIC = -0.44% and RMSE = -4.68%). Instead, higher differences were found when the 1st January (\(R^2 = 0.55\), AIC = 60.47, RMSE = 7.74) and 1st of March (\(R^2 = 0.49\), AIC = 52.57, RMSE = 12.47) were compared, evidencing the most relevant performances for 1st January (\(R^2 = +10.83\) and RMSE = −61.23%, with the exception of AIC = +14.15%).

Moreover, the behavior of grapevine varieties in response to specific climate conditions was noticed during model application. The higher performances of CF compared to F, as well the relevant
differences between starting dates when F models were used, may be explained by the influence of temperatures on the dormancy period. The end of the first phase of dormancy (endo-dormancy) in CF models was generally simulated between the end of December (e.g., Touriga Nacional: DOY 356 for $t_0 = -121$ and $-152$, respectively) and end of February (e.g., Riesling: DOY 56 and 57 for $t_0 = -121$ and $-152$, respectively). More specifically, the period of endo-dormancy fulfilment was similarly reached regardless of starting date (1st September/August), allowing the next forcing unit accumulation during the eco-dormancy period to be correctly satisfied (differences in days from 0 to 3 days). The analysis further highlighted that the endo-dormancy release occurs later at northern than at southern latitudes and this is consistent with results obtained using F approach. These demonstrate that a late starting date (1st of March) for forcing unit accumulation is more suitable for budbreak simulation of varieties collected in Northern Europe (e.g., Gewürztraminer, Riesling and Pinot Gris), whose endo-dormancy release is simulated by CF late in the season (54 DOY on average). Conversely, an early starting date (1 January) provided the best results in simulating budbreak for grapevine varieties collected in Southern Europe (e.g., Touriga Franca and Nacional, Cabernet Sauvignon, Grenache and Chardonnay), which are characterized by an early end of endo-dormancy as simulated by CF models (13 DOY on average).

4. Discussion

This study compares the performance of six phenological models in simulating the budbreak date of eight grapevine varieties across different European sites. We found that the model performances characterized by a Chilling-Forcing approach were generally higher compared to those obtained for Forcing approach when different starting dates were set (Figure 6 and Figure S17). Although model intercomparison studies on the estimation of budbreak date in grapevine have already been conducted by other authors [15,29,30], this study focused on evaluation of the models’ reliability to reproduce budbreak in different environmental conditions, by considering the impact of temperature during the chilling and forcing accumulation period. In contrast to other works [27,30,57,58], this model intercomparison considers the use of a long-term series of observed phenology, which include grapevine varieties with different precocity levels collected at various latitudes in three European countries (France, Luxembourg and Portugal, Figure 1 and Table 1). The extensive observed dataset coupled to the methodology applied in this study allowed to evaluate the influence of study site-temperature variability on the grapevine development. Moreover, the cross-validation scheme adopted in this model intercomparison experiment provided an exhaustive analysis of the bias and variance of the CF and F models while the comparison between two different types of parameterization (HP or AV) pointed out the issue to select a specific parameter set for model applications (Results section, Figure 5). This in-depth investigation of model fitting, validation and application was particularly useful for evaluating the influence of temperature during the dormancy period. In this context, the different model responses during budbreak predictions are related to the physiological requirements needed for dormancy fulfilment of the grapevine varieties. Differently from the CF models, which describe the entire dormancy period (endo- and eco-dormancy), the F models only describe the eco-dormancy period through forcing units accumulation, under the assumption that the chilling requirement was already satisfied [24]. Although the debate on which model type provides the highest reliability is still open [29,30], our analysis shows that CF models, considering the influences of winter and spring temperatures on the accumulation of chilling and forcing units, may avoid the under- or over-estimation at the beginning of eco-dormancy, which occur when F models are set on fixed dates. In this study, we estimated a longer endo-dormancy for the vines from the Northern European sites (e.g., Gewürztraminer), which occurs later than the endo-dormancy of vines from the Southern European sites (e.g., Touriga Nacional), irrespective of the precocity level of the different varieties. In other terms, grapevine varieties such as Gewürztraminer (early budbreak) and Riesling (late budbreak) collected in Northern Europe showed a late endo-dormancy release (higher chilling requirement) with respect to Chardonnay (early budbreak) and Cabernet Sauvignon (late budbreak)
mainly collected in Southern Europe (lower chilling requirement). Although this trend was examined in this study only from a modeling point of view without any physiological evidence on the dormancy period, our results suggested that plants cultivated at different latitudes might show distinct dormancy period lengths. The higher accuracy of CF compared to F models is likely linked to the ability of these models to estimate the endo-dormancy release and to better account for the effect of temperature in this phase, avoiding the use of a fixed a priori value for the beginning of the eco-dormancy period as in F models [27,29]. Indeed, the goodness of F models is shown to significantly increase when the starting date is optimized against observed experimental data [30,45], demonstrating that the beginning of eco-dormancy period is dependent to the origin of the observed dataset and, thus, on the fulfilment of specific physiological requirements that change in relation to the climate conditions of the different environments. In fact, although no relevant differences were found for budbreak date estimation of grapevine varieties using CF models set on 1st of September/August starting dates, the discrepancy of the F model simulations was evident when 1st of January/March dates were used. Higher accuracy of F models to predict the budbreak date of grapevine varieties collected in Northern European countries was observed using the 1st of March starting date as also confirmed by [27], while the 1st January seems to be a more suitable starting date for representing budbreak date of the grapevine varieties from Southern Europe (Figures S5–S7). In this case, a compromise between both starting dates may be represented by the beginning of February [59]. However, the adoption of different starting dates in F models highlights the role of temperature on bud growth and the need to satisfy specific dormancy requirements for varieties cultivated at different latitudes before budbreak occurrence. These results were supported by the work of Reis-Pereira et al. [60] on the grapevine flowering stage. The authors pointed out that the selection of a priori value for the starting of thermal unit accumulation in F models may lead to an under- or over-estimation of the flowering phase, demonstrating the importance of a correct evaluation of the previous phenological stage in respect to the subsequent one. These findings were also observed in relation to the flowering stage of other species such as olive tree e.g., [61]). In this context, we infer that the dormancy period of grapevine may play a relevant role for predicting budbreak date especially under future scenarios. Indeed, more accurate estimations of budbreak date using CF models are provided when the endo-dormancy release is calibrated against observed data of perennial crops. This behavior is particularly evident under warmer temperatures (after 2050) where CF without endo-dormancy calibration and F models showed poor performances of phenology prediction [62]. The accurate identification of the endo-dormancy release avoids selecting a fixed a priori date, especially when this may have strong consequences on the estimation of budbreak date (as in F models).

The impact of warmer temperatures after the endo-dormancy release showed higher effect on the budbreak phase compared to the influence of temperatures during chilling accumulation, as reported by Martinez-Lüscher et al. [63] in a two-year ad hoc experiment with different grapevine varieties and temperature regimes. However, the starting date for thermal unit accumulation depends on the physiological features of the grapevine variety and their interaction with the environment. Indeed, we found that the effect of temperature during the last part of winter and beginning of spring was also evident when models were tested on external sites (independent datasets, Figure 5). Although most of the models showed satisfactory results during application (R^2 and AIC), the average error (RMSE) between observed and simulated data was relatively high (from 4 to 19 days). This situation was exacerbated in F models where the reliability in simulating budbreak is dependent to the use of a specific starting date (1st of January/1st of March) and thus the original observed dataset. The discrepancy between the two starting dates raises the question of which initial date to choose for forcing unit accumulation when simulations are conducted on a European scale [31,64].
The implications of this study are particularly relevant for growers especially when considering the impact of climate change. Our study, in line with a large part of literature, confirmed that chill units accumulations, the amount of which varies according to variety and location, is required for exiting from endo-dormancy phase. In a warmer climate, as expected in the future, this amount may not fully satisfy, leading to developmental issues during the growing season as outlined in the introduction of this paper. Some authors already highlighted that, on a global scale, wine areas may shift to match their specific climate niches in case of no adaptation [11,65], while new varietal selections and breeding programmes are needed for improving the adaptability of the grapevine varieties and maintaining the high quality production [12,59,66]. In this context, our parameterizations may be used as metric, in combination with climate simulations under future scenarios, to highlight which regions and varieties are expected to be most exposed to the risk of budbreak failure. Accordingly, the use of alternative varieties, better adapted to the expected future climate, may be proposed or the direction of specific breeding programmes can be indicated to develop varieties with a lower chill units requirement as requested in a warmer climate. Moreover, we should also take into account that, besides thermal requirements, dormancy phase may be influenced by the photoperiod [67]. Since each grapevine variety requires critical thresholds of day length for the onset of dormancy [68,69], this regulating factor should be considered when CF models are adopted for simulations at different latitudes. As such, comprehensive datasets of phenological observations and ad hoc experiments for evaluating the influence of more factors on dormancy [30,46,69] may be useful to take into account for improving model reliability under future climate scenarios and different environments [31,64,70].

5. Conclusions

This model intercomparison evidenced the different performance of the phenological models in reproducing the budbreak date of eight grapevine varieties collected in different European regions. The reliability of two main phenological modeling approaches (CF and F) was evaluated considering different starting dates (1st January/March or 1st September/August) and the selection methodologies of the parameter set (HP and AV) for model application. The results showed, in general, higher performances of CF compared to F models where fixed starting dates were adopted, even if neither modeling approach highlighted a clear supremacy over the other. For this reason, CF and F models' behavior should be further investigated through ad hoc experiments for evaluating the role of dormancy on budbreak occurrence. Indeed, the findings of our model simulations evidenced distinct thermal requirements for grapevine varieties cultivated at different latitudes (i.e., Northern Europe vs. Southern Europe). The use of different starting dates highlighted that, unlike CF models, the accurate estimation of budbreak in F models is affected by the selection of a starting date for beginning of the eco-dormancy period. This implies that, if this phase is not accurately defined, F models may lead to over- or under-estimations of the budbreak date under different environmental conditions. In this context, comprehensive phenological datasets (e.g., grapevine phenological data collected at different latitudes) are needed to thoroughly investigate the effect of temperature variability on dormancy period and budbreak estimation. Finally, it is proposed the development and implementation of improved phenological models, which consider the effect of other climate (e.g., radiation) and soil (e.g., soil temperature) variables as well as some regulating factors of the endo-dormancy phase (e.g., day length).
Supplementary Materials: The following are available online at http://www.mdpi.com/2076-3417/10/11/3800/s1.

Figure S1: $R^2$ and AIC values of the BRIN Daily model obtained for all fitted and average models, grapevine varieties and starting dates. The symbols (points and triangles) show the model results at two starting dates (in this example $t_0 = -121$ and $-152$ DOY), while the two colours represent fitting (red) and validation (blue) procedure, respectively. The filled symbols constitute the $R^2$ and AIC values of the models obtained by the selected HP parameter sets, while the dashed (fitting) and dotted (validation) lines depict the average $R^2$ and AIC values for both starting dates (in this example orange line = $-121$ DOY and green line = $-152$ DOY). The boxplots represent the $R^2$ and AIC values distribution for HP (orange and light blue: HP model application on all validation subsamples for $t_0 = -121$ and $t_0 = -152$ DOY, respectively) and AV (boxplot in the shaded area, orange and light blue: AV application on all validation subsamples for $t_0 = -121$ and $t_0 = -152$ DOY, respectively). The outliers represent all values out of the 1.5*InterQuartile Range (IQR) of the boxplot. Figure S2: $R^2$ and AIC values of the BRIN Hourly model obtained for all fitted and average models, grapevine varieties and starting dates. The symbols (points and triangles) show the model results at two starting dates (in this example $t_0 = -121$ and $-152$ DOY), while the two colours represent fitting (red) and validation (blue) procedure, respectively. The filled symbols constitute the $R^2$ and AIC values of the models obtained by the selected HP parameter sets, while the dashed (fitting) and dotted (validation) lines depict the average $R^2$ and AIC values for both starting dates (in this example orange line = $-121$ DOY and green line = $-152$ DOY). The boxplots represent the $R^2$ and AIC values distribution for HP (orange and light blue: HP model application on all validation subsamples for $t_0 = -121$ and $t_0 = -152$ DOY, respectively) and AV (boxplot in the shaded area, orange and light blue: AV application on all validation subsamples for $t_0 = -121$ and $t_0 = -152$ DOY, respectively). The outliers represent all values out of the 1.5*InterQuartile Range (IQR) of the boxplot. Figure S3: $R^2$ and AIC values of the UNICHIll model obtained for all fitted and average models, grapevine varieties and starting dates. The symbols (points and triangles) show the model results at two starting dates (in this example $t_0 = -121$ and $-152$ DOY), while the two colours represent fitting (red) and validation (blue) procedure, respectively. The filled symbols constitute the $R^2$ and AIC values of the models obtained by the selected HP parameter sets, while the dashed (fitting) and dotted (validation) lines depict the average $R^2$ and AIC values for both starting dates (in this example orange line = $-121$ DOY and green line = $-152$ DOY). The boxplots represent the $R^2$ and AIC values distribution for HP (orange and light blue: HP model application on all validation subsamples for $t_0 = -121$ and $t_0 = -152$ DOY, respectively) and AV (boxplot in the shaded area, orange and light blue: AV application on all validation subsamples for $t_0 = -121$ and $t_0 = -152$ DOY, respectively). The outliers represent all values out of the 1.5*InterQuartile Range (IQR) of the boxplot. Figure S4: $R^2$ and AIC values of the UNIFIED model obtained for all fitted and average models, grapevine varieties and starting dates. The symbols (points and triangles) show the model results at two starting dates (in this example $t_0 = -121$ and $-152$ DOY), while the two colours represent fitting (red) and validation (blue) procedure, respectively. The filled symbols constitute the $R^2$ and AIC values of the models obtained by the selected HP parameter sets, while the dashed (fitting) and dotted (validation) lines depict the average $R^2$ and AIC values for both starting dates (in this example orange line = $-121$ DOY and green line = $-152$ DOY). The boxplots represent the $R^2$ and AIC values distribution for HP (orange and light blue: HP model application on all validation subsamples for $t_0 = -121$ and $t_0 = -152$ DOY, respectively) and AV (boxplot in the shaded area, orange and light blue: AV application on all validation subsamples for $t_0 = -121$ and $t_0 = -152$ DOY, respectively). The outliers represent all values out of the 1.5*InterQuartile Range (IQR) of the boxplot. Figure S5: $R^2$ and AIC values of the UNIFORCl model obtained for all fitted and average models, grapevine varieties and starting dates. The symbols (points and triangles) show the model results at two starting dates (in this example $t_0 = 1$ and $60$ DOY), while the two colours represent fitting (red) and validation (blue) procedure, respectively. The filled symbols constitute the $R^2$ and AIC values of the models obtained by the selected HP parameter sets, while the dashed (fitting) and dotted (validation) lines depict the average $R^2$ and AIC values for both starting dates (in this example orange line = $1$ DOY and green line = $60$ DOY). The boxplots represent the $R^2$ and AIC values distribution for HP (orange and light blue: HP model application on all validation subsamples for $t_0 = 1$ and $t_0 = 60$ DOY, respectively) and AV (boxplot in the shaded area, orange and light blue: AV application on all validation subsamples for $t_0 = 1$ and $t_0 = 60$ DOY, respectively). The outliers represent all values out of the 1.5*InterQuartile Range (IQR) of the boxplot. Figure S6: $R^2$ and AIC values of the GDD model obtained for all fitted and average models, grapevine varieties and starting dates. The symbols (points and triangles) show the model results at two starting dates (in this example $t_0 = 1$ and $60$ DOY), while the two colours represent fitting (red) and validation (blue) procedure, respectively. The filled symbols constitute the $R^2$ and AIC values of the models obtained by the selected HP parameter sets, while the dashed (fitting) and dotted (validation) lines depict the average $R^2$ and AIC values for both starting dates (in this example orange line = $1$ DOY and green line = $60$ DOY). The boxplots represent the $R^2$ and AIC values distribution for HP (orange and light blue: HP model application on all validation subsamples for $t_0 = 1$ and $t_0 = 60$ DOY, respectively) and AV (boxplot in the shaded area, orange and light blue: AV application on all validation subsamples for $t_0 = 1$ and $t_0 = 60$ DOY, respectively). The outliers represent all values out of the 1.5*InterQuartile Range (IQR) of the boxplot. Figure S7: $R^2$ and AIC values of the WANG model obtained for all fitted and average models, grapevine varieties and starting dates. The symbols (points and triangles) show the model results at two starting dates (in this example $t_0 = 1$ and $60$ DOY), while the two colours represent fitting (red) and validation (blue) procedure, respectively. The filled symbols constitute the $R^2$ and AIC values of the models obtained by the selected HP parameter sets, while the dashed (fitting) and dotted (validation) lines depict the average $R^2$ and
AIC values for both starting dates (in this example orange line = 1 DOY and green line = 60 DOY). The boxplots represent the R² and AIC values distribution for HP (orange and light blue: HP model application on all validation subsamples for t₀ = 1 and t₀ = 60 DOY, respectively) and AV (boxplot in the shaded area, orange and light blue: AV application on all validation subsamples for t₀ = 1 and t₀ = 60 DOY, respectively). The outliers represent all values out of the 1.5 InterQuartile Range (IQR) of the boxplot. Figure S8: Correlations between observed and simulated budbreak Days of Year (DOY) of the UNIFORC and UNICHILL models. The results were obtained for all grapevine varieties and starting dates using the HP parameter set. Figure S10: Highest-Performance (HP) and average (AV) parameters’ set applied on the independent datasets by considering all phenological models and starting dates. The example of Cabernet Sauvignon variety. Figure S11: Highest-Performance (HP) and average (AV) parameters’ set applied on the independent datasets by considering all phenological models and starting dates. The example of Gewürztraminer variety. Figure S12: Highest-Performance (HP) and average (AV) parameters’ set applied on the independent datasets by considering all phenological models and starting dates. The example of Grenache variety. Figure S13: Highest-Performance (HP) and average (AV) parameters’ set applied on the independent datasets by considering all phenological models and starting dates. The example of Pinot Gris variety. Figure S14: Highest-Performance (HP) and average (AV) parameters’ set applied on the independent datasets by considering all phenological models and starting dates. The example of Riesling variety. Figure S15: Highest-Performance (HP) and average (AV) parameters’ set applied on the independent datasets by considering all phenological models and starting dates. The example of Touriga Franca variety. Figure S16: Highest-Performance (HP) and average (AV) parameters’ set applied on the independent datasets by considering all phenological models and starting dates. The example of Touriga Nacional variety. Figure S17: Statistical results (R², AIC and RMSE) of the comparison between CF (black bars) and F (red bars) models in model fitting and application for all grapevine varieties using AV parameters’ set. The differences between CF and F models in model fitting and applications were displayed using green and yellow bars, respectively. CS = Cabernet Sauvignon, CH = Chardonnay, GE = Gewürztraminer, GR = Grenache, PG = Pinot Gris, RS = Riesling, TF = Touriga Franca, TN = Touriga Nacional. Table S1: Statistical indices (RMSE, EF, R², AIC) obtained for all models, varieties and starting dates after model fitting and validation on all subsamples.


Funding: This research was funded by the European Union’s Horizon 2020 Research and Innovation Programme, under the Clim4Vitis project “Climate change impact mitigation for European viticulture: knowledge transfer for an integrated approach”, grant agreement no. 810176. It was also supported by FCT-Portuguese Foundation for Science and Technology, under the project UIDB/04033/2020 and the French National Research Agency (ANR) in the frame of the Investments for the Future Program, within the cluster of excellence COTE (ANR-10-LABX-45).

Acknowledgments: We wish to acknowledge the Clim4Vitis project “Climate change impact mitigation for European viticulture: knowledge transfer for an integrated approach”, funded by European Union’s Horizon 2020 Research and Innovation Programme, under grant agreement no. 810176 that supported this study. H.F. thanks the FCT for contract CEECIND/00447/2017 and project COA/CAC/0050/2019.

Conflicts of Interest: The authors declare no conflict of interest.

References
2. Van Leeuwen, C.; Seguin, G. The concept of terroir in viticulture. J. Wine Res. 2006, 17, 1–10. [CrossRef]
16. Molitor, D.; Junk, J. Climate change is implicating a two-fold impact on air temperature increase in the ripening period under the conditions of the Luxembourgish grapegrowing region. OENO One 2019, 53, 409–422. [CrossRef]


60. Reis Pereira, M.; Ribeiro, H.; Abreu, I.; Eiras-Dias, J.; Mota, T.; Cunha, M. Predicting the flowering date of Portuguese grapevine varieties using temperature-based phenological models: A multi-site approach. J. Agric. Sci. 2018, 156, 865–876. [CrossRef]