Calibration and validation of grapevine budburst models using growth-room experiments as data source

Gianni Fila\textsuperscript{a,*,b}, Bruno Di Lena\textsuperscript{b,d}, Massimo Gardiman\textsuperscript{a}, Paolo Storchi\textsuperscript{c}, Diego Tomasi\textsuperscript{a}, Oriana Silvestroni\textsuperscript{d}, Andrea Pitacco\textsuperscript{e}

\textsuperscript{a} Centro di Ricerca per la Viticoltura (CRA-VIT), I-31015, Conegliano (TV), Italy
\textsuperscript{b} Regione Abruzzo, ARSA, Centro Agrometeorologico Regionale, I-66020, Scerni (CH), Italy
\textsuperscript{c} Unità di Ricerca per la Viticoltura (CRA-VIC), I-52100, Avezzano, Italy
\textsuperscript{d} Università Politecnica delle Marche, Dipartimento di Scienze Ambientali e Produzioni Vegetali, I-60131, Ancona, Italy
\textsuperscript{e} Università di Padova, Dipartimento di Agronomia Animali Alimenti Risorse Naturali e Ambiente - DAFNE, Agripolis, I-35020, Legnaro (PD), Italy

**A R T I C L E   I N F O**

Article history:
Received 8 September 2011
Received in revised form 13 February 2012
Accepted 6 March 2012

Keywords:
Grapevine
Budburst
Phenology
Modelling
Dormancy
Climate change

**A B S T R A C T**

Robust calibration of phenological models requires long term field observations, which are not always available or sufficiently widespread. This has motivated the evaluation of short-term experiments using cuttings under semi-controlled conditions as an alternative data source. Single-node cuttings from two grapevine cultivars were exposed to variable chilling durations and allowed to sprout in a growth room. The observed budburst dates and temperature series were used to calibrate two budburst models, which were validated against a 39-year field observation dataset by means of a fuzzy-logic based integrated index (FI). Satisfying validation scores were obtained, ranging from 0.262 to 0.411 on a 0–1 scale (best-worst response).

The experiment was then inverted, using field data for calibration and cuttings for validation, and FI scores ranging between 0.352 and 0.495 were obtained. On this occasion however, the models were not able to estimate budburst occurring after short chilling exposures, where they returned either high over-estimations or failed completely. This was due to the narrow winter length variability in the field dataset, which made the optimization algorithm converge towards unrealistically high chilling requirements and artificial durations of the temperature effects on dormancy. Cutting-based calibration on the other hand produced parameterizations that were more consistent with available experimental knowledge.

Despite this difference between the two approaches proved to be equivalent under the climatic conditions present, but not when tested on projected scenarios of climate change over the period 1990–2090, where cutting-calibrated models, which are more sensitive to decreasing winter length, predicted higher variations of the budburst dates.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Models of grapevine phenology are becoming popular tools for assessing the impact of climate change on viticulture (Webb et al., 2007; Kwon et al., 2008; Caffarra and Eccel, 2011; Duchène et al., 2010) and for supporting GIS-assisted zoning studies to identify the most suitable areas for specific cultivars (Bois et al., 2008; Scaglione et al., 2008). Given the growing interest in these kinds of applications, which are characterized by high degrees of extrapolation across many environmental conditions on both spatial and time scales, model robustness represents a crucial issue (Caffarra and Eccel, 2010).

The more up-to-date phenology models of perennial plants combine the description of dormancy dynamics in overwintering buds with a heat-sensitive development phase, often termed “forcing phase”, leading to budburst (e.g. Richardson et al., 1974; Cannell and Smith, 1983; Cesaraccio et al., 2004; Chuiue, 2000; de Cortázar-Atauri et al., 2009). They are considered an evolution from the more traditional ones, based only on the accumulation of heat units from a fixed date, which are collectively referred to as “Thermal Time models” (e.g. Cannell and Smith, 1983; Linkosalo et al., 2008) or “Spring Warming models” (e.g. Hunter and Lechowicz, 1992; Parker et al., 2011). These latter models differ in how heat units are calculated: Growing Degree Day (GDD, Winkler et al., 1974; Bonhomme, 2000), or variously defined functions of daily temperature (e.g. Chuiue et al., 1999b; Parker et al., 2011).

Compared to Thermal Time models, those accounting both for chilling and forcing adapt more easily to environment and genotype variability, but due to increased complexity they require...
considerable amounts of experimental data for parameterization and validation to reach adequate robustness and to avoid overfitting.

It is precisely this that represents one of the biggest limitations in phenology research. Since development stages can be observed in the field once a year only, it takes decades to assemble a sufficiently large dataset. But even when long historical data series are available, they may be biased by heterogeneities in plant material, or in the criteria for identifying the phenological stage, not to mention eventual gaps in meteorological data.

In this work we have evaluated the possibility of using repeated growth room experiments with grapevine cuttings to build a dataset for the calibration and validation of budburst models. Evaluated against field observations this approach presents a number of advantages. First of all, vine cuttings require small areas and can be induced to budburst in only a few weeks, therefore many experiments can be concentrated over a relatively short period of time whilst ensuring a high variability of conditions. Furthermore, budburst dates and temperatures can be recorded with high homogeneity and accuracy. In a short time it should therefore be possible to build a phenological dataset sufficiently large to alleviate the lack of field observations or to integrate them if they are scarce. Another advantage is the possibility of observing plant behaviour under a wider range of climatic conditions, which can be varied at will. This allows for minimization of local specificity which translates into a higher level of generalization and greater robustness, both desirable features for applications involving spatial estimates and/or projections into the future.

Many studies in the past have already used cuttings under controlled conditions for investigating the effect of temperature on grapevine budburst (Pouget, 1967; Weaver et al., 1975; Calò et al., 1976; Dokoozlian, 1999), but none of the ensuing information has been incorporated into a mathematical model to gain a comprehensive understanding of experimental results and to facilitate their practical exploitation.

Our objectives were (i) to assess whether budburst data derived from growth room experiments can effectively replace long term field datasets used for model calibration, (ii) to analyse the consistency of the obtained parameterization, and (iii) to evaluate its application to long-term predictions related to climate change.

## 2. Materials and methods

### 2.1. Cutting sampling and handling

Dormant one-bud cuttings from the cultivars ‘Montepulciano’ and ‘Sangiovese’ were repeatedly excised from a vineyard located at Scerni, in the Abruzzo region (Central Italy, 42.10N, 14.57E) during three consecutive Autumn/Winter seasons between 2007 and 2010. Vines were grafted on Koher 5BB rootstock and trained to a spur-pruned cordon system with 3 m × 1.5 m plant spacing. Samplings started when at least seven days with daily mean temperatures below 10 °C were recorded. This condition occurred in mid October, 2007 and 2009, and at the end of November in 2008. In the first two seasons ten and nine samples were taken, respectively, up until the month of March, whilst only two samplings in January and February were executed in the third season (Table 1).

Each sample (n = 90) was put into a growth room under semi-controlled conditions, where temperature maintained a diurnal variation of 8.0 ± 2.5 °C above that outside, so that during the experimental campaign maximum and minimum temperature increased from 10 to 25 °C and from 6 to 18 °C, respectively. In order to increase chilling duration variability, in some of the treatments part of the samples were kept in a separate room at 2–3 °C for 10–35 days prior to placing them under growth conditions. Budburst was recorded when 50% of the buds reached stage 09 of the BBCH scale for grape (Lorenz et al., 1994).

### 2.2. Budburst field data

Historical budburst records were provided by the ampelographicic collections from the Research Centre for Viticulture located in Susegana (North-Eastern Italy, 45.85N, 12.26E) and from the Research Unit for Viticulture located in Arezzo (Central Italy, 43.29N, 11.90E), both being structures of the Agriculture Research Council (CRA). The collections have the same planting scheme with a Sylovo training system and a 3 m × 1.5 m plant spacing. Both cultivars were grafted on SO4 rootstock.

Data from Susegana were collected from 1985 to 2010, with two missing years (2003 and 2007), whereas data from Arezzo were collected from 1996 to 2010. A total of 39 annual records were therefore available for each cultivar.

Both vineyards were equipped with on-site automatic meteorological stations, which recorded daily maximum and minimum temperatures.

### 2.3. Models

The ‘Unified’ and ‘Unichill’ models (Table 2) developed by Chuine (2000) were calibrated with the growth room results and validated with the field data. Both models assume that dormancy is a two-stage process, starting with a “rest” period which ends when daily accumulation of chilling units starting from September 1st reaches a critical sum (Ccri). At this point the second stage begins called “quiescence”, in which heat units (or ‘forcing units’) are accumulated until budburst occurs as soon as another critical sum (Fcri) is reached. Daily values of chilling and forcing units are calculated using temperature dependent functions. In the Unified model Fcri decreases during a sensitive period (Tc) according to an exponential decrease function of the total accumulated chilling, whilst in Unichill it is a fixed empirically determined parameter.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22 November (0)</td>
<td>11 November (13)</td>
<td>5 January (0)</td>
</tr>
<tr>
<td>2</td>
<td>3 December (0)</td>
<td>22 December (0)</td>
<td>5 January (10)</td>
</tr>
<tr>
<td>3</td>
<td>18 December (0)</td>
<td>2 February (0)</td>
<td>2 February (0)</td>
</tr>
<tr>
<td>4</td>
<td>3 January (0)</td>
<td>2 February (16)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>21 January (0)</td>
<td>2 February (31)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5 February (0)</td>
<td>24 February (0)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5 February (14)</td>
<td>24 February (35)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5 February (28)</td>
<td>5 March (0)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>19 February (0)</td>
<td>5 March (26)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>4 March (0)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
-budburst dates in brackets
-The duration of the cold treatment at 2–3 °C prior to the forcing treatment in growth room.

### Table 2

**Description of the ‘Unified’ and ‘Unichill’ models (further details in the text).**

<table>
<thead>
<tr>
<th>Eq. (1)</th>
<th>(Daily chilling unit (c.u.))</th>
<th>Ccri = ( \frac{1}{a_1} + \frac{b_1}{c_1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eq. (2)</td>
<td>(Daily forcing unit (f.u.))</td>
<td>Fcri = ( \frac{1}{k} \sum_{i=1}^{n} (T_{\text{min}} - T_i) )</td>
</tr>
<tr>
<td>Eq. (3)</td>
<td>(Critical f.u. sum)</td>
<td>Fcri = W ( e^{-(k/Tc)} )</td>
</tr>
</tbody>
</table>

- **a1, b1, c1** - empirical parameters of the c.u. vs. temperature function
- **b2, c2** - empirical parameters of the f.u. vs. temperature function
- Dormancy breaks when c.u. summation reaches a critical value (Ccri)
- Budburst occurs when f.u. accumulation reaches a second critical value (Fcri)
- **W, k** - empirical parameters of the Fcri adjusting function in the Unified model
- **Tc** - length of the period (days) where Eq. (3) applies
- **Tmin** - daily mean temperature
- **Sc** - totally accumulated c.u. within the Tc period
The Unified and Unichill models have 9 and 7 parameters, respectively.

2.4. Model fitting

Models were calibrated by the Simulated Annealing method (Kirkpatrick et al., 1983; Cerny, 1985), which has been already used for fitting phenology models (Chuine et al., 1998; Schaber and Badeck, 2003). Among the many existing versions of this algorithm, in this study we used that of Corana et al. (1987), also described in a modern pseudo code by Goffe et al. (1994). In our implementation, the algorithm was driven alternatively by three pseudo-random number generators: Marsaglia (Marsaglia et al., 1990), Knuth (1981) and Mersenne-Twister (Matsumoto and Nishimura, 1998).

The algorithm was programmed to minimize the root mean squared error (RMSE) between estimated and observed budburst dates:

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (E_i - M_i)^2}{n}}
\]

where \(E_i\) is the ith estimated budburst date, \(M_i\) is the ith observed budburst date and \(n\) is the number of pairs. Optimizations were performed by repeatedly running the algorithm, each time changing the initial set of parameter values and the random generator until a stable minimum RMSE was obtained.

2.5. Model performance evaluation

The external validity of a model is commonly assessed using various statistics (Martiroma and Belloccchi, 1999), such as the coefficient of determination \(R^2\) or the mean square error (RMSE).

Belloccchi et al. (2002) pointed out that analysing one or more statistics, separately and/or in an organized manner, may be inadequate. This is because each statistic evaluates only one particular aspect of the model. A model may perform well with respect to one statistic, but it may be deficient with respect to others, which are equally as important to the researcher.

To prevent such problems, Belloccchi et al. (2002) developed a fuzzy-logic based system to allow simultaneous evaluation of several statistical indices. The system consists in the aggregation of several statistics into a single index by assigning an expert weight in proportion to the relative importance assigned to each input statistic.

Only a brief and intuitive description of the method will be given here, as all theory and methodology is detailed in the seminal work, and applications of the system can be found in Rivington et al. (2005), Donatelli et al. (2005), Diodato and Bellocchi (2007), Abbara and Savage (2008) and Confalonieri et al. (2010).

We defined an indicator, termed “FI” (fuzzy indicator), whose value depended on those simultaneously taken by the following statistics: (i) root mean squared error (RMSE), (ii) the coefficient of determination \(R^2\), (iii) the absolute difference between 1 and the slope of the regression line, (iv) the absolute value of the intercept of the regression line, and the (v) modelling efficiency (EF) (Loague and Green, 1991).

The choice of statistics defines an expert strategy to rank model performance. The basic idea is that a “good” model is one having the least average difference between estimates and measurements and where the linear regression of estimates vs. measurements is as close as possible to the bisector line \(y = x\). The index EF adds further stringency and useful information because it allows the immediate identification of inefficient models. It is upper-bound by 1 and can assume negative values, which indicate that the average value of all measures is a better predictor than the model used.

<table>
<thead>
<tr>
<th>Index</th>
<th>Range</th>
<th>Best value</th>
<th>Fuzzy classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0−∞</td>
<td>0</td>
<td>F&lt;1; U&gt;10</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0−1</td>
<td>1</td>
<td>F&gt;0.90; U&lt;0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3

Statistical indices aggregated into the fuzzy index (FI), and definition of the fuzzy classes. For each statistic, FI is the interval where the model response is considered optimal, while U is where the model is considered to have failed.

\(a\) Slope of the estimates vs. measures linear regression.

\(b\) Intercept of the estimates vs. measures linear regression.

FI is a dimensionless value between 0 (best model response) and 1 (worst model response). The calculating procedure, based on the multi-valued fuzzy set theory introduced by Zadeh (1965), follows the so-called Sugeno or Takagi–Sugeno–Kang method of fuzzy inference (Sugeno, 1985).

The variation interval of each statistic was divided into two membership classes according to an expert judgement, namely favourable (F) and unfavourable (U) membership. F is the interval where the statistic values are considered optimal, while if a model falls in the U interval, it is considered unsuccessful. A membership value equal to 0 or 1 is assigned, respectively, in the two aforementioned cases. In the fuzzy range, where statistics are simultaneously F and U, the membership degree takes intermediate values determined by two complementary S-shaped quadratic functions (Liao, 2002).

Once membership values have been assigned, FI is calculated. Briefly, the reasoning runs as follows: if all input variables are in the F subset, then FI is 0, while if they are all U then FI is 1. For all other cases, FI takes intermediate values calculated upon a set of decision rules based on relative weights assigned to each combination of F/U classes, depending on the importance assigned to each statistic. In this work we have assigned equal importance to all statistics. Table 3 displays the limits assigned to each statistic that makes up the indicator.

2.6. Analysis outline

In the first part of the work, we assessed whether model calibration using cutting-based experiments was effective in providing good budburst estimates, by validation on independent field data. Model calibration was then repeated against field observations, and validation was performed on the cutting data. In this way we evaluated how differently calibrated models vary in estimation accuracy and in the consistency of the quantitative description of the dormancy process.

Differences in the latter aspect are expected to affect model sensitivity to interannual variability, which is relevant in climate change impact analysis. This effect was evaluated by running models calibrated on both datasets on projected climate series and analysing how budburst date changes in respect to present climate conditions.

3. Results

3.1. Dataset variability

Data collected from growth room experiments showed a higher variability in exposure to cold temperatures. In Table 4 this duration was expressed by the number of days between September 1st and budburst with mean temperature below 10°C. This has long been considered the upper limit to chilling-effective temperatures (Pouget, 1972; Bernstein, 1984). The variation range was from 15 to 113 days in the cutting dataset and from 92 to 154 days in the field.
Table 4
 Variability of budburst date and chilling duration in the two datasets used for calibration and validation of the phenology models.

<table>
<thead>
<tr>
<th>Data source</th>
<th>Budburst date (day of the year)</th>
<th>No. of days with $T_{\text{mean}} &lt; 10^\circ$C (from 1st September to budburst)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Range of variation</td>
<td>Std. deviation</td>
</tr>
<tr>
<td><strong>Cuttings</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field</td>
<td>–11 to 106</td>
<td>29.9</td>
</tr>
<tr>
<td></td>
<td>86–120</td>
<td>8.0</td>
</tr>
</tbody>
</table>

Table 5
Fitting and validation results of the ‘Unified’ and ‘Unichill’ model on Montepulciano and Sangiovese grapevine cultivars. In Part (A) (Cutting calibration), models were fitted on the laboratory data and validated on a 39-year field data pooled from two sites. In Part (B) (Field calibration) models were fitted on field data and validated on the cutting dataset.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>‘Unified’ model</th>
<th>‘Unichill’ model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Montepulciano</td>
<td>Sangiovese</td>
</tr>
<tr>
<td></td>
<td>Fitting</td>
<td>Validation</td>
</tr>
<tr>
<td><strong>RMSE</strong></td>
<td>2.866</td>
<td>5.697</td>
</tr>
<tr>
<td><strong>$R^2$</strong></td>
<td>0.989</td>
<td>0.559</td>
</tr>
<tr>
<td>$</td>
<td>1 - \text{slope}</td>
<td>$</td>
</tr>
<tr>
<td>$</td>
<td>\text{intercept}</td>
<td>$</td>
</tr>
<tr>
<td><strong>EF</strong></td>
<td>0.989</td>
<td>0.121</td>
</tr>
<tr>
<td><strong>FI</strong></td>
<td>0.023</td>
<td>0.348</td>
</tr>
</tbody>
</table>

For field calibration, validation was only tested on a subset of data, which only included budburst events observed after the 60th day of the year (see text).

The higher chilling duration variability in cutting data was associated with a longer time span between the first and last budburst events, which was 117 days vs. 34 days in the field dataset.

3.2. Model fitting and validation

Both models showed good fit when they were calibrated on cutting data for both cultivars, with FI scores ranging from 0.019 to 0.035. At validation the best performance was obtained by Unichill on Sangiovese (FI=0.262) and the worst by the same model on Montepulciano (FI=0.411), whereas the Unified model yielded comparable scores (0.348 and 0.304) on the two cultivars (Table 5A).

Good fit was also obtained when the models were calibrated on field data. The Unified model scored FI values of 0.089 and 0.071 for Montepulciano and Sangiovese respectively, while according to Unichill the corresponding values were 0.252 and 0.121 (Table 5B). Validation statistics were only calculated for budburst recorded after about mid-February, as models failed or returned strong over-estimations before this time (Fig. 1). The thus calculated FI values were also satisfying, but higher than in cutting-based models, between 0.352 and 0.495.

3.3. Chilling and forcing requirements

In the models adopted in this study, chilling and forcing requirements are quantified by the $C_{\text{crit}}$ and $F_{\text{crit}}$ parameters, which represent the number of days the cultivar has to spend at the specific optimal temperatures to release the dormancy rest and quiescence phases.

The values taken by model parameters, including $C_{\text{crit}}$ and $F_{\text{crit}}$, are displayed in Table 6. For comparison purposes the table also shows $F_{\text{crit}}$ for the Unified model, although for this model it is not actually a parameter, but a variable depending on the total amount of chilling received. The values reported in brackets are the average values ± the standard deviation calculated on the cutting dataset.

Table 6
Parameters of Unified and Unichill models for the cultivars ‘Montepulciano’ (M) and ‘Sangiovese’ (S) as resulting from the numerical optimization against budburst data derived from growth room experiments or field data. For comparison purposes the table displays $F_{\text{crit}}$ (average ± stand. dev.) also for the Unified model, calculated on the cutting dataset. These latter values are reported in brackets because in the Unified model $F_{\text{crit}}$ is a variable and not a parameter. Cutting- and field- parameterization show relevant differences in $C_{\text{crit}}$ and $F_{\text{crit}}$ values (see text for details and discussion).

<table>
<thead>
<tr>
<th>Model/cv</th>
<th>$a_e$</th>
<th>$b_e$</th>
<th>$c_e$</th>
<th>$b_f$</th>
<th>$c_f$</th>
<th>$C_{\text{crit}}$</th>
<th>$W$</th>
<th>$K$</th>
<th>$T_c$</th>
<th>$F_{\text{crit}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cutting calibration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unified/M</td>
<td>1.154</td>
<td>–10.964</td>
<td>0.537</td>
<td>–0.264</td>
<td>12.069</td>
<td>8.5</td>
<td>45.2</td>
<td>–0.0058</td>
<td>100</td>
<td>(37.9 ± 1.44)</td>
</tr>
<tr>
<td>Unified/S</td>
<td>3.299</td>
<td>–23.197</td>
<td>0.886</td>
<td>–0.292</td>
<td>14.203</td>
<td>12.5</td>
<td>20.7</td>
<td>–0.0004</td>
<td>166</td>
<td>(20.1 ± 0.08)</td>
</tr>
<tr>
<td>Unichill/M</td>
<td>2.560</td>
<td>23.668</td>
<td>9.838</td>
<td>–0.232</td>
<td>13.135</td>
<td>9.1</td>
<td>23.2</td>
<td>11.7</td>
<td>23.6</td>
<td></td>
</tr>
<tr>
<td>Unichill/S</td>
<td>3.441</td>
<td>23.839</td>
<td>8.182</td>
<td>–0.267</td>
<td>13.642</td>
<td>11.7</td>
<td>23.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field calibration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unified/M</td>
<td>1.180</td>
<td>–28.201</td>
<td>–8.832</td>
<td>–0.201</td>
<td>15.472</td>
<td>112.7</td>
<td>17.6</td>
<td>–9.3 × 10^-5</td>
<td>193</td>
<td>(17.4 ± 0.01)</td>
</tr>
<tr>
<td>Unified/S</td>
<td>0.282</td>
<td>–10.708</td>
<td>0.214</td>
<td>–0.217</td>
<td>13.187</td>
<td>132.5</td>
<td>24.5</td>
<td>–0.0004</td>
<td>167</td>
<td>(22.9 ± 0.06)</td>
</tr>
<tr>
<td>Unichill/M</td>
<td>1.190</td>
<td>–20.322</td>
<td>–1.897</td>
<td>–0.194</td>
<td>15.041</td>
<td>114.9</td>
<td>19.0</td>
<td>22.3</td>
<td>22.3</td>
<td></td>
</tr>
</tbody>
</table>
As far as chilling requirement is concerned, in models calibrated on cutting data the optimization converged towards low values of $C_{\text{crit}}$, around 10 days on average. This differs sharply from what was obtained from field-calibrated models, where $C_{\text{crit}}$ took much higher values, all above 110 days for all cultivars and models. In both calibrations $C_{\text{crit}}$ was higher in Sangiovese, while no relevant differences were found between models.

The low $C_{\text{crit}}$ in cutting calibrated models allowed early completion of the dormancy rest phase, which ended between November 17th and 30th when calculated on the field dataset. The more than ten-fold higher chilling requirements obtained after field calibration caused the models to delay rest break until January 8th to February 9th (Table 7).

The required forcing period in cutting-based models for Montepulciano was 37.9 or 35.5 days, and 20.1 or 23.6 days for Sangiovese according to the Unified and Unichill models, respectively. Shorter durations as well as smaller differences between cultivars were obtained in field-calibrated models, where $F_{\text{crit}}$ values were about 20 days on average (Table 6).

Forcing requirements were expected to be complementary to those of chilling: for a given budburst date, the longer the chilling duration, the shorter the forcing phase must be and vice versa. This was what was effectively observed in Montepulciano, where the field-derived $F_{\text{crit}}$ were half the values of those obtained from cutting calibration, but not in Sangiovese, where $F_{\text{crit}}$ remained unaltered, with an average value of 22.2, despite a ten-fold increase in $C_{\text{crit}}$ when passing from cuttings to field calibration.

These contrasting behaviours between cutting- and field-based models were analysed by taking into account the chilling temperature intervals, which resulted very different in the two calibration types. The interval width of optimal chilling temperatures is strictly related to chilling requirement, since the former determines how fast the latter is fulfilled.

### 3.4. Chilling temperatures

In models calibrated on cuttings the temperature-dependent function of chilling units was characterized by a rectangular shape, where the range of optimal temperatures was 0–10.5 °C in Montepulciano and 0.7–8.3 °C in Sangiovese, with very little variation between models (Fig. 2, upper graphs). Field-calibrated models showed the same function shape, but with a broader optimal temperature interval, and a wider difference between cultivars. In Montepulciano the optimal ranges were −9.1 to 15.3 °C (Unified model) and −2.3 to 15.6 °C (Unichill model), while in Sangiovese they were −0.6 to 39.5 °C (Unified) and −2.0 to 39.8 °C (Unichill) (Fig. 2, bottom graphs).

### 3.5. Forcing temperatures

The optimized functions relating forcing units to temperature were plotted in Fig. 3, where they are compared to some data from Pouget (1967) and Moncur et al. (1989), the only studies, to our knowledge, that have attempted to assess a quantitative relationship between temperature and budburst rate.

![Graph showing validation of field-calibrated models on budburst data obtained with growth-room experiments.](image-url)

**Fig. 1.** Validation of field-calibrated models on budburst data obtained with growth-room experiments. Downward triangles stand for failed estimation, which were given an arbitrary out-of-range value. Double arrows indicate the variation range of the calibration dataset.

### Table 7

<table>
<thead>
<tr>
<th>Model</th>
<th>Cultivar</th>
<th>Cutting calibration</th>
<th>Field calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unified</td>
<td>Montepulciano</td>
<td>17 November</td>
<td>09 February</td>
</tr>
<tr>
<td></td>
<td>Sangiovese</td>
<td>30 November</td>
<td>15 January</td>
</tr>
<tr>
<td>Unichill</td>
<td>Montepulciano</td>
<td>20 November</td>
<td>07 February</td>
</tr>
<tr>
<td></td>
<td>Sangiovese</td>
<td>25 November</td>
<td>08 January</td>
</tr>
</tbody>
</table>
Four cultivars were examined in Pouget’s study. We examined what he found to be the earliest (‘Perle de Csaba’) and the latest budbursters (‘Ugni blanc’). In both papers the original data were expressed as budburst speed (in days⁻¹ dimensions), so they were normalised to maximum speed temperature (25°C in Pouget, and 20°C in Moncur et al.) to convert them into daily relative rate units, thus making them comparable to our data.

Since the asymptotic nature of the curves did not present distinct values which could be taken as reference limits to the interval of forcing-effective temperatures, we defined as “minimum active temperature” the one at which the forcing unit is equal to 0.05, and as “minimum optimal temperature” the one above which the forcing unit is ≥0.95.

According to this criterion, small differences were obtained within curves yielded by cutting-calibration, where the minimum
active temperature varied from 0.5 to 4.2 °C, and the minimum optimal temperature fell to between 23.3 and 25.9 °C.

Even in the field-based models the forcing functions appeared to be similar among models and cultivars, except for the Unified model on Sangiovese. With respect to those derived from cutting calibration these were broader by an average of 1.2 °C and less steep. The minimum active temperature varied from −2.7 to 0.9 °C, while the optimal values of the forcing unit started at between 26.8 and 33.0 °C.

All the curves showed good correspondence with literature point data. Cutting-derived curves were, however, closer to Moncur’s data, on the cool edge of the variation interval, whereas the field-based ones were more fitting to Pouget’s data.

3.6. Budburst predictions under climate change scenarios

3.6.1. Weather data generation

Synthetic daily time series of minimum and maximum temperature were built upon the predictions of a general circulation model, to test the behaviour of the models under changing climate conditions. The task was achieved by means of the LARS-WG5 software (Racsko et al., 1991; Semenov and Strattonvitch, 2010), a stochastic weather generator which, in its latest version incorporates predictions from 15 climate models used in the fourth IPCC Assessment Report (IPCC, 2007). Using local weather records for the period 1960–2010 to calibrate the system, predictions from the HadCM3 model were downscaled for the Susegana site, to generate daily time series for three time periods in the future: 2011–2033, 2046–2065 and 2080–2099, which will be hereafter referred to by their intermediate years, i.e. 2020, 2055 and 2090. One set of projections was generated for each of the three emission scenarios, ‘A1B’, ‘A2’ and ‘B1’.

The A1B scenario predicts rapid economic and population growth up to 2050, and a gradual decrease thereafter. The various energy sources are equilibrated and new and more efficient technologies are developed. In the A2 scenario, population increases continually, technology and per capita income changes slowly and in a fragmented way. B1 is the most ecologically-friendly scenario considered. Population and economy grows and then declines as in A1, but with rapid changes towards a service- and information-centred economy leading to reductions in material intensity and the introduction of clean and resource efficient technologies (IPCC, 2007).

The downscaled predictions from HadCM3 for the Susegana site are shown in Table 8, with respect to the average temperatures of Autumn, Winter and Spring, which are the seasons pertinent to the adopted models. Comparison to present time is made with reference to the baseline period 1980–1999, referred to with the year 1990, which is the same used in the IPCC Fourth Assessment Report. As the table shows, the emission scenario A2 is the most pessimistic, showing the highest increments, which are not, however, very different from A1B. The highest increment, predicted by 2090 is +4.4 °C for the average Spring temperature. The B1 scenario was the most optimistic one, with a maximum increment of 2.8 °C for Spring in 2090. In general, temperature increments in Autumn and Winter are lower than those predicted for Spring.

3.6.2. Model simulations

The Unified and Unichill models were run on the generated temperature series, and the dates of budburst and dormancy rest break were estimated. As expected, model application to future emission scenarios predicted forward shifts of budburst timing, which increased as projections move further into the future. Overall, A2 was the scenario with the maximum estimated changes, though they were not very far from those observed on A1B, which differed from the former by no more than 1.8 days, while that with the lowest variations was B1 (Fig. 4).

In Montepulciano the estimates derived from cutting- and field-calibrated models were coincident or with a maximum difference of 2.7 days until 2020, then they diverged, and at the 2090 projection they differed by between 5 (Unified model on B1 scenario) and 12 days (Unichill on A1B scenario). In this cultivar, cutting-calibrated models therefore predicted higher shifts of the budburst date, with a maximum of 16.5 days (Unified model on A2) vs. 12.3 (Unified on A2) predicted by their counterpart models.

In Sangiovese, the two calibration types showed a parallel variation across the projected time periods, with a maximum difference of three days in the Unified model, while with Unichill they were almost coincident (Fig. 4). The overall predicted change of the budburst date was 21.5 days, averaging over models and cultivars in the A2 and A1B scenarios, while in B1 it was 13 days.

The difference in cultivar response is more evident when looking at the predicted dates of rest release (Fig. 5). Whilst climate warming is known to accelerate the heat-sensitive forcing phase, thus contributing to bringing forward budburst timing, the inverse effect is generally expected on rest completion, as more time is needed to accumulate chilling units in milder Autumns/Winters.

This response was found on both cultivars and both models when they were calibrated on cuttings. The largest shifts of the rest release date, 19.7 days on average, were estimated in A1B, while the lowest ones (14.6 days on average) were observed again in B1 (Fig. 5).

Field-calibrated models predicted a strongly different response between cultivars. In Montepulciano rest exit was delayed as in cutting-calibrated models, but the variation took place about 80 days later, in the first to the third week of February. In Sangiovese there was almost no change in the rest exit date, which remained in the second week of January irrespective of model and projected time period.

Delays in rest exit are known to counterbalance the advancing trend of Spring phenology (Yu et al., 2010). In our results however, this effect is noticeable only in field-based predictions for Montepulciano. In cutting-based estimates the rest exit delay was mostly restricted to Autumn even in the latest climate projections, that is, largely before heat accumulation starts.

4. Discussion

4.1. Can growth room experiments surrogate field observations to calibrate/validate budburst models?

When used as data source for model calibration, growth room experiments estimated budburst with good accuracy. As a
validation dataset they provided a stringent test platform that evaluated the models across a large set of conditions, well beyond the variation interval of the field dataset used for calibration.

Observing budburst on cuttings kept under controlled conditions proved therefore to represent a viable alternative to the lack of field data to support model-based applications.

4.2. Do cutting-calibrated models provide consistent parameterizations?

Despite strong differences in the time required to complete chilling and forcing, as well as in how these processes respond to temperature (Figs. 2 and 3), cutting- and field-calibrated models showed comparable performances over a wide range of cold exposure durations, provided these stay above a critical value (Fig. 1). Estimation accuracy was therefore unrelated to how chilling and forcing dynamics were described, and this raises questions about whether the parameters of these models have a real physiological meaning for grapevine. A comparison of our results to literature available on the subject was carried out to find out which parameterization method gives the more consistent and informative results.

4.2.1. Chilling-related parameterization

Growth-room experiments showed that the grapevine cultivars under study had low chilling requirements, as budburst was reached even by the earliest samplings in November.

The optimization process had to select from iteratively generated solutions that accounted for both early and late budburst realizations. This led optimization to converge towards low $C_{crit}$ values, whose principal effect is to favour early exits from Winter rest. Another outcome of training models on a very variable dataset is the obtaining of a narrow interval of optimal chilling temperatures, in this case between 0 and 10.5°C or less. This makes models more sensitive to interannual variability: the time taken to accumulate the required amount of chilling units rises in milder Autumn/Winters, since fewer days are active on dormancy and vice versa in colder years. These results are fully consistent with previous studies reporting that chilling exposures between 50 and 400 h at temperatures <7°C are sufficient for normal bud sprouting (Dokoozlian, 1999; Botelho et al., 2007).

As for the optimal temperature interval for chilling, those derived from cutting-calibration confirm what is known from the available experimental knowledge. The temperature of 10°C as upper limit to optimal chilling temperatures is a long established notion for grapevine (Pouget, 1972; Bernstein, 1984) and for other tree species (Perry, 1971; Hänninen and Backman, 1994; Sarvas, 1974). Calò et al. (1976) working on 22 cultivars found that a cold pre-treatment at 1°C was more effective than one at −6°C for accelerating sprouting. Dokoozlian (1999) observed little or no difference in the effect on budburst when chilling temperatures varied between 0 and 10°C.

These previous findings render the longer chilling requirements obtained from field-calibration unrealistic, most likely artefacts
generated by the low variability of the field dataset. In the field most of the forcing process takes place after January–February, long after actual rest release. Under these conditions the exact assessment of the date of rest break is not critical, and the lack of constraints allows $C_{crit}$ and the chilling response function to temperature to assume nonsense values.

This especially applies to parameters obtained for Sangiovese. No previous report or common practical experience supports chilling effects of temperatures close to 40 °C. With such a wide optimal temperature interval, almost every day in Autumn/Winter can satisfy chilling requirements and contribute to rest completion. In other words, Winter variability is not accounted for by this parameterization, which can ultimately be simplified to a Thermal Time model, which considers only heat summation for predicting budburst.

Although field-derived $C_{crit}$ in Sangiovese is very high, the wide chilling temperature interval allowed to fulfil it at the beginning of January, in a period where normally forcing units accumulation has yet to start. Therefore, differently from what was observed in Montepulciano, $F_{crit}$ was not affected by the high $C_{crit}$ value (Table 6).

This reasoning may help interpretation of similar findings recently reported by Caffarra and Eccel (2010) who parameterized a modified version of the Unified model against field data from a Chardonnay vineyard in Conegliano, North-Eastern Italy, located only 6 km from where part of our field data come from. These authors obtained a wide range of optimal chilling temperatures, and a $C_{crit}$ of 79 days, and recognized the connection between these two results. They remarked that neither such a high chilling requirement, neither a chilling effect of temperatures above 10 °C are supported by existing experimental knowledge. They therefore suggested that “chilling accumulation”, as described by the model, in fact may have been influenced by an effect of Autumn mild temperatures in anticipating dormancy onset, and by a delaying effect of warm temperatures, above 20 °C, on dormancy break.

Whilst confirming the possibility that field-based calibration can result in parameterizations contrasting with experimental evidence, our results also proved that this does not prevent models from giving accurate estimates, at least under present climatic conditions.

Further evidence was reported by Parker et al. (2011), who showed that Unichill and Thermal Time models, tested on a very wide dataset, gave comparable performances.

4.2.2. Forcing-related parameterization

Both cutting- and field-calibrated forcing unit curves corroborated literature data reported in Fig. 3. A comparison with these data requires some caution, particularly with those of Pouget’s, since they were taken on cuttings pre-treated with a hot-water bath to break rest. This treatment is known to accelerate budburst
(Orffer and Goussard, 1980; Halaly et al., 2008) and it could possibly have altered the relation between temperature and forcing rate.

A noteworthy side-result is the small standard deviation of the \( F_{\text{crit}} \) predicted by the Unified model, since higher variations were expected on the basis of previous experience, showing that the time taken to budburst varies sensitively on the total amount of chilling received (Dokoozlian, 1999). In fact, in our results the optimized Unified model tends to assume fixed values for \( F_{\text{crit}} \), hence cancelling the difference with Unichill, which at this point is to be considered the preferable model, giving equivalent performances with a more simple structure. The apparent contradiction with experimental data may be explained by the difficulty of separating the effect of early heat unit accumulation from that of eventual reduction of \( F_{\text{crit}} \) due to increased chilling. As a matter of fact, accumulation of forcing units can take place even at temperatures <10 °C, as our forcing functions and previous studies show (de Cortázaro-Atauri et al., 2009; Nendel, 2010), so this may contribute to quiescence completion during cold periods following rest exit, thus being confused with an \( F_{\text{crit}} \) reduction effect. A similar interpretation was advanced by Chuine and Cour (1999), who even questioned the existence of an \( F_{\text{crit}} \) sensitivity to chilling and considered the possibility a modelling artefact.

The low chilling requirements found in cutting-based models complements a longer forcing time, between 20 and 40 days which is not incompatible with the available experimental knowledge. Calò et al. (1976) observed durations of 13.5–35.5 days at 23 °C on 22 cultivars. Hellman et al. (2006) recorded 28 days at 21 °C for ‘Sangiovese’. Dokoozlian (1999) found that at 22 °C it may take any duration from about 18 to more than 40 days, depending on the length and temperature of the chilling treatments (cv. ‘Perlette’). Weaver et al. (1975) reported that cuttings from the middle of the cane of ‘Carignane’ taken from October to February, took from 20 to about 40 days at 25 °C.

### 4.3. Long term budburst estimation

Under present climatic conditions the two parameterization types showed similar budburst estimation capacity, provided that cold Winter conditions last long enough to prevent heat accumulation from starting before mid-February. In the study regions it is very unlikely for this to occur, but ongoing climate changes could gradually increase the probability of approaching this critical time. It is therefore reasonable to wonder whether and to what extent the calibration data source may affect long-term budburst estimation. According to Figs. 4 and 5, the parameterization type had an impact on model response when they were applied to scenario analysis.

The increasing divergence between cutting- and field-based models which was observed in Montepulciano, highlighted the better capacity of the cutting-based models to simulate grapevine behaviour in very short Winters, as already seen in cutting chamber experiments.

This finding suggests that caution should be observed when using field-calibrated models in applications involving high extrapolation degrees, since the magnitude of change may be underestimated.

Further caution is advisable even when field-based models appear to provide results similar to those from cutting-based ones, such as those observed on Sangiovese, as this may not be due to model accuracy, but to possible irrelevance of a dormancy dynamics description when cool conditions over-fulfil chilling requirements, which does not necessarily always hold true.

### 5. Conclusions

A series of experiments under semi-controlled conditions was designed for the purpose of inducing budburst in cuttings exposed to different chilling treatments. The observations were assembled into a dataset which was used to calibrate and validate two grapevine budburst models. From results obtained we draw the following conclusions:

- Datasets built on growth-room experiments can effectively replace decade long field observations for calibrating and/or validating grapevine budburst models. When used for model calibration they allow obtaining more consistent parameterizations than with field data.
- Although calibrations carried out on field observations or growth room experiments yielded contradictory parameterizations, they can both perform well under present climatic conditions. Discrepancies between the two types of calibration emerged, however, when they were applied to climate change scenario analysis.
- When using field data for model calibration, the size of the dataset does not guarantee against inconsistent parameterization if its variability is low. Some caution is therefore recommended when field-calibrated models are used in applications involving a high degree of extrapolation.

### References


