Process-based models outcompete correlative models in projecting spring phenology of trees in a future warmer climate

Daphné Asse, Christophe Randin, Marc Bonhomme, Anne Delestrade, Isabelle Chuine

Process-based models outcompete correlative models in projecting spring phenology of trees in a future warmer climate

Running title: Correlative *vs* process-based phenology models

Authors: Daphné Asse[1,2,3*], Christophe F. Randin[2†], Marc Bonhomme[4], Anne Delestrade[1], Isabelle Chuine[3†]

Adresses:

[1] Centre de Recherches sur les Ecosystèmes d'Altitude, Chamonix Mont-Blanc, France

[2] Department of ecology and Evolution, University of Lausanne, Switzerland

[3] Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175, CNRS- Université de Montpellier - Université Paul-Valéry Montpellier – EPHE. 1919 route de Mende, Montpellier F-34293 cedex 05, France

[4] Université Clermont Auvergne, INRA, UMR 547 PIAF, Clermont-Ferrand F-63100, France

*First and corresponding author: daphne.asse@gmail.com

[†]Lead author

Corresponding authors:

Isabelle Chuine, tel. +33 467 613 279, fax +33 467 613 336, e-mail: isabelle.chuine@cefe.cnrs.fr, postal address: CEFE, 1919 route de Mende, F-34293 Montpellier cedex 05

Christophe Randin, tel. +41 21 316 99 91, e-mail: christophe.randin@unil.ch, postal address: DEE UNIL, Biophore, CH-1015 Lausanne

Abstract: 330 words

Main text: 7323 words

Number of figures: 7

Number of tables: 1

Abstract

Many phenology models have been developed to explain historical trends in plant phenology and to forecast future ones. Two main types of model can be distinguished: correlative models, that statistically relate descriptors of climate to the date of occurrence of a phenological event, and process-based models that build upon explicit causal relationships determined experimentally. While process-based models are believed to provide more robust projections in novel conditions, it is still unclear whether this assertion always holds true and why. In addition, the efficiency and robustness of the two model categories have rarely been compared.

Here we aimed at comparing the efficiency and the robustness of correlative and process-based phenology models with contrasting levels of complexity in both historical and future climatic conditions. Models were calibrated, validated and compared using budburst dates of five tree species across the French Alps collected during 8 years by a citizen-science program.

Process-based models were less efficient, yet more robust than correlative models, even when their parameter estimates relied entirely on inverse modeling, i.e. parameter values estimated using observed budburst dates and optimization algorithms. Their robustness further slightly increased when their parameter estimates relied on forward estimation, i.e. parameter values measured experimentally. Our results therefore suggest that the robustness of process-based models comes both from the fact that they describe causal relationships and the fact that their parameters can be directly measured.

Process-based models projected a reduction in the phenological cline along the elevation gradient for all species by the end of the 21st century. This was due to a delaying

effect of winter warming at low elevation where conditions will move away from optimal chilling conditions that break bud dormancy *vs* an advancing effect of spring warming at high elevation where optimal chilling conditions for dormancy release will persist even under the most pessimistic emissions scenario RCP 8.5.

These results advocate for increasing efforts in developing process-based phenology models as well as forward modelling, in order to provide accurate projections in future climatic conditions.

Key words: *Budburst, elevation gradients, Alps, citizen science, endodormancy release, climate change impact*

1. Introduction

Phenology is a key aspect of plant and animal life strategies because it determines the timing of growth and reproduction. Life cycle of species must be adapted to the local weather conditions and resources. As a consequence, phenology is one of the top controls of crop yield (Olesen *et al.,* 2011; Nissanka *et al.,* 2015), population dynamics (Anderson et al., 2013), species distribution (Chuine, 2010) and evolutionary dynamics (Duputié *et al.,* 2015; Burghardt *et al.,* 2015). Phenology also ultimately regulates many functions of ecosystems such as productivity (Richardson et al., 2012), ecosystem carbon cycling (Delpierre et al., 2009), water (Hogg et al., 2000) and nutriment cycling (Cooke and Weih, 2005).

Since the 1970s, spring phenology has been reported to advance in response to warming (Walther *et al.* 2002; Parmesan and Yohe, 2003; Menzel *et al.* 2006; Fu *et al.* 2014). For instance, it has been shown that the apparent response of leaf unfolding to temperature was -3.4 days per °C between 1980 and 2013 in temperate Europe (Fu et al., 2015). This advance in spring phenology events is due to the warming of springs as bud growth rate is positively and strongly related to temperature (see for review Chuine and Régnière, 2017). However, this trend has been slowing down by about 40% after 2000 (Fu et al., 2015). One of the most likely hypotheses to explain this slowdown is the warming of winters (Asse et al., 2018). Indeed, most of temperate and boreal trees have developed a key adaptation to winter cold: the inability to resume growth despite transient favorable growing conditions in terms of temperature (Howe et al., 2000). This particular physiological state, called endodormancy (Lang et al., 1987), establishes in fall and disappears in early to late winter after a certain exposure to cold temperatures. Therefore, the warming of winters is suspected to delay endodormancy release and be responsible for the apparent decrease in the response of leaf

unfolding to warming after 2000. Ultimately, this lack of chilling temperatures might compromise budburst itself at some point if warming continues (Chuine et al., 2016). Such a situation is more likely to occur in populations inhabiting the warm edge of a species range and/or lower elevations in mountain regions, where species are already in suboptimal chilling conditions (Benmoussa et al., 2018; Guo et al., 2015). More than ever there is a need for more accurate projections of tree phenology for the upcoming decades to remove the large uncertainties that still remains.

Two main categories of predictive phenology models exist although there can be a continuum in-between: correlative and process-based models (for review see (Chuine *et al.,* 2013; Chuine and Régnière, 2017). Correlative models statistically relate descriptors of climate to phenological variables (i.e. usually the occurrence dates of a phenological phase such as bud break or flowering). In correlative models, parameters have no *a priori* defined ecological meaning and processes can be implicit (process-implicit) (Lebourgeois et al., 2010). In contrast, process-based models are built around explicitly stated mechanisms and parameters have a clear ecological interpretation that is defined *a priori*. In this category of model, response curves are often obtained directly from experiments, contrasting with empirical relationships of correlative models. Consequently, process-based models often require a larger number of parameters to be estimated or measured, with the consequence of a higher level of complexity than correlative models. However, they provide greater insights into how precisely each driver affects the trait, and they are expected to provide more robust projections in new climatic conditions corresponding either to other geographical areas or other time periods (Chuine et al., 2016).

Among the most widely used process-based phenology models, are the so-called 1-phase models that describe solely the ecodormancy phase, which follows the endodormancy phase. During the ecodormancy phase, bud cell elongation can take place whenever temperatures are appropriate, and the higher the temperature is, the higher is the rate of cell elongation during this phase. This category of models has been shown to be efficient in predicting accurately budburst date under historical climate (Vitasse *et al.,* 2011; Basler, 2016). Another category of models, called 2-phase models, describes additionally the endodormancy phase, and take into account the possible negative effect of winter warming on endodormancy release. This category of models is thus considered to provide more accurate projections in future climatic conditions (Chuine, 2010; Vitasse *et al.,* 2011). However, it has been shown recently that this second type of models might suffer from flawed parameter estimation when dates of endodormancy release have not been used for their calibration (Chuine et al., 2016). Unfortunately, observations of endodormancy release date are very rare because they are very difficult to determine (Jones et al., 2013; Chuine et al., 2016). Models are thus usually calibrated using solely bud break or flowering dates (Chuine, 2000; Caffarra *et al.,* 2011; Luedeling *et al.,* 2009; but see Chuine *et al.,* 2016). Some other models go further in the description of the processes by integrating a photoperiod cue (Schaber and Badeck, 2003; Gaüzere *et al.,* 2017). Some studies indeed support the hypothesis that in photosensitive species, which might represent about 30% of temperate tree species (Zohner et al., 2016), long photoperiod might compensate for insufficient chilling (Caffarra *et al.,* 2011a; Gaüzere *et al.* 2017).

There is now a large number of phenology models that differ in their level of complexity and in the types of response function to environmental cues they use (see for review Chuine *et al.,* 2013; Basler, 2016; Chuine and Régnière, 2017). However, very few

studies aimed at comparing their efficiency and robustness so far (Basler, 2016), especially in future climatic conditions (but see Vitasse *et al.,* 2011; Chuine *et al.,* 2016; Gaüzere *et al.,* 2017), while this has been done multiple times for species distribution models (e.g. Cheaib *et al.,* 2012; Higgins *et al.,* 2012; Morin and Thuiller, 2009; Kearney *et al.,* 2010) and crop models (Lobell and Burke, 2010) for example. By efficiency, we mean here the ability of the model to provide accurate predictions in conditions that have been used to calibrate the model (Janssen and Heuberger, 1995); and by robustness, we mean here the ability of the model to provide accurate predictions in external conditions (Janssen and Heuberger, 1995), i.e. other conditions than those used to calibrate the model. Model's robustness determines its transferability in time and space.

Process-based models are usually expected to provide more accurate projections for the future than correlative models because they describe causal relationships. The effect of each driver identified as affecting a particular trait value can be described by a causal relationship, sometimes involving other drivers as well (interaction between drivers). For this reason, process-based models have also an expected greater potential to deal with non-analog situations. However, the putative higher robustness of process-based models could also come from the fact that parameter values describing the causal relationships, or at least some of them, can be measured directly (forward estimation of parameter values) instead of being inferred statistically through inverse modelling techniques and data assimilation (backward estimation of parameter values). Yet, there has been no attempt so far to validate this widely accepted expectation.

Here, we aimed at comparing the efficiency and robustness of correlative *vs* process-based phenology models with contrasting levels of complexity, both in space and time. More

158 precisely, we aimed at answering the following questions: (1) Are process-based phenology
159 models more robust than correlative models? (2) If so, is it because they describe causal
160 relationships or because they can be less dependent on statistical inference (i.e. back
161 estimation of parameter values) and rely more on experimental measurements (i.e. forward
162 estimation of parameter values)? (3) How do projections of both types of model differ in
163 future climatic conditions?

164 Using observations of budburst dates collected over the Western Alps by a citizen
165 science program during 8 years, and experimental data, we calibrated correlative and process-
166 based phenology models with three levels of complexity for five major tree species: *Corylus*
167 *avellana* (L.), *Fraxinus excelsior* (L.), *Betula pendula* (Roth), *Larix decidua* (Mill.) and *Picea*
168 *abies* (L.). We then compared their predictions and projections over the Western Alps in
169 historical climate and in future climate respectively.

170 The Western Alps are particularly interesting to evaluate phenology models because
171 the elevation gradient provides a wide temperature range on a very short distance. In addition,
172 the southern part of the Western Alps is nearly located at the warmest edge of the geographic
173 range of the five studied species, where it has been shown that winter warming is already
174 affecting endodormancy release and budburst (Asse et al., 2018). Finally, temperature has
175 already increased in the Western Alps twice as fast as in the northern hemisphere during the
176 20th century (Rebetez and Reinhard, 2008) and recent evidence indicates that the current
177 warming rate increases with increasing elevation (Mountain Research Initiative EDW
178 Working Group, 2015). Consequently, mountain summits might warm faster than lower
179 elevation sites, so that the response of mountain ecosystems to climate change might be non-

linear along elevation gradients. Therefore, ultimately, we aimed at answering a fourth question: (4) How will climate change alter the budburst date of alpine species?

2. Methods

2.1. Phenological and meteorological data

We used observations of the budburst date, defined as the first day when 10% of vegetative buds of a given individual tree are opened (BBCH 07), of five common tree species: ash (*Fraxinus excelsior* L.), birch (*Betula pendula* Roth), hazel (*Corylus avellana* L.), larch (*Larix decidua* Mill.), and spruce (*Picea abies* L.). These species show different elevation ranges (from 150 - 1300 m a.s.l. for *Corylus* to 700 – 2100 m a.s.l for *Larix*), which allowed us to compare the two types of model over a large climatic gradient. The data were extracted from the Phenoclim database of CREA (Centre de Recherches sur les Ecosystèmes d'Altitude, Chamonix, France) (www.phenoclim.org) which covers the entire French Alps (for further details of the Phenoclim protocol see Appendix A) (Fig. 1). In total, and irrespective of species, 242 sites were surveyed for budburst between 2007 and 2014.

Sixty of the observation sites (ranging from 372 to 1919 m a.s.l,) were equipped with meteorological stations that recorded air temperature at 2-m height every 15 min with a DA8B20 digital thermometer placed in a white ventilated plastic shelter (Dallas Semiconductor MAXIM, "http://www.maxime-ic.com", operating range of −55 to 125 °C and an accuracy of 0.5 °C over the range of −10 to 85 °C) (Fig. 1). These data were used to calculate the daily mean air temperature at the sixty sites and were also interpolated at observation sites without meteorological stations using a 25-m digital elevation model (DEM) as follows. We first constructed a linear model of daily temperature at 2-m height as a function of elevation. Residuals of this linear model were then interpolated using the calibrations points over the entire French Alps according to the inverse distance weighted algorithm (IDW; see also: Kollas *et al.,* 2014; Cianfrani *et al.,* 2015, for further details).

In addition to Phenoclim data, we used data from an experiment on *Larix decidua*

207 realized at INRA UMR PIAF (Clermont-Ferrand, France) during winter 2010-2011. The aim
208 of this experiment was to determine the date of endodormancy release, the chilling required
209 to break endodormancy and the response of bud growth to temperature during the
210 ecodormancy phase of *Larix decidua* Mill. In a first experiment, 60 twigs of 30 to 40 cm,
211 were sampled on 4 sites (Saulzet, Orcival, Lamartine, Laqueuille) on February 2nd and
212 immediately taken to the laboratory. Twigs were recut under water and put in flasks filled up
213 with water. Water was subsequently changed twice a week. Twigs were divided into six sub-
214 sets that were placed in growth chambers (Sanyo MLR 351H) which were set to: 25, 20, 15,
215 10, 5, 2°C with 16h photoperiod and a light intensity of 50 µE/m2/s, relative humidity (RH)
216 of 70%, except for the 5°C and 2°C chambers which received 20 µE/m2/s and RH 50%. This
217 experiment was designed to investigate the response to temperature of bud growth during the
218 ecodormancy phase. In a second experiment, 5 twigs were sampled on 5 trees of *Larix*
219 *decidua* Mill. from the end of September to early February on the same 4 sites at ten dates: 27
220 Sep; 18 Oct; 17 Nov; 14 Dec; 29 Dec; 04 Jan; 11 Jan; 18 Jan; 24 Jan; 02 Feb. Twigs were
221 recut under water and put in flasks filled up with water. Water was subsequently changed
222 twice a week. Twigs were placed in growth chambers (Sanyo MLR 351H) at 20°C, 16h
223 photoperiod, 50 µE/m2/s, RH 70%. This experiment was designed to determine the
224 endodormancy release date and the chilling required to break endodormancy. In the two
225 experiments, each bud of each twig was monitored every second day to determine the date of
226 budburst (BBCH 07).

227

228 *2.2. Correlative phenology model*

229 We used mixed-effects models for each of the five species with budburst dates as
230 response variables and with growing degree-days and chilling as predicting variables (Asse *et*

al. 2018). We defined chilling for each species and each observation site as the frequency of days with a daily temperature < 5°C from September 1st to December 31st of the calendar year preceding budburst (Dantec et al., 2014). We then calculated for each species and each observation site growing degree-days (GDD) defined as the sum of positive daily mean temperature from January 1st to the median date of budburst observed over the 2007-2014 observation period. GDD was calculated with two different temperature thresholds: the more commonly-used 5°C (hereafter designated as GDD5) and 0°C (hereafter designated as GDD0) because daily mean temperatures between 0 and 5°C may contribute to trigger budburst events in plants growing in mountainous environments (Körner 1999; Vitasse et al. 2016). We chose Jan 1st as the starting date of accumulation because the ecodormancy phase can begin as early as January for some species and because the climatic conditions at the beginning of this phase vary a lot between species, years, and sites along elevation and latitudinal gradient (Asse et al., 2018).

2.3. Process-based phenology models

We used several process-based phenology models to simulate the budburst dates of each species.

First, we used 1-phase models, that describe the ecodormancy phase only and assume that endodormancy-release always occurs before temperature can trigger cell elongation and that there is no effect of chilling on forcing requirement (Chuine et al., 2013). Budburst occurs at t_f when the sum of the daily rates of development (R_f) reaches the critical value F^*:

$$\sum_{t_0}^{t_f} R_f(T_d) \geq F^* \quad (1)$$

with t_0 the starting date of forcing and T_d the daily mean temperature.

We used several versions of the model that differed by the response function to temperature (R_f), which determines the daily rate of bud development:

the Growing Degree Day function, similarly to the correlative models,

$$R_f(T_d) = Max(0;\ T_d - T_b) \quad (2)$$

with T_b the lower threshold temperature;

and two more complex functions: the sigmoid function,

$$R_f(T_d) = \frac{1}{1 + e^{-d_T(T_d - T_{50})}} \quad (3)$$

with d_T the steepness of the response and T_{50} the mid-response temperature;

and the Wang function (Wang and Engel, 1998)

$$R_f = Max\left[\left(2(T_d - T_{min})^{\alpha}\left(T_{opt} - T_{min}\right)^{\alpha} - \frac{(T_d - T_{min})^{2\alpha}}{\left(T_{opt} - T_{min}\right)^{2\alpha}}\right), 0\right] \quad (4)$$

with $\alpha = \ln(2) / \ln\left(\frac{T_{max} - T_{min}}{T_{opt} - T_{min}}\right)$, and T_{min}, T_{max} and T_{opt} the cardinal temperatures.

These models have from 3 to 5 parameters that were either calibrated using inverse modelling techniques or prescribed using experimental data (see 2.4).

Second, we used sequential 2-phase models (Hanninen, 1987) that take into account chilling requirements during the endodormancy phase (first phase) and forcing requirements during the ecodormancy phase (second phase; Chuine et al., 1999). The endodormancy phase ends at t_c when the accumulation of the daily rates of dormancy release (R_c) reaches the critical sum of chilling units C^*:

$$\sum_{t_0}^{t_c} R_c(T_d) \geq C^* \quad (5)$$

From t_c to t_f (date of budburst), forcing units are then accumulated as:

$$\sum_{t_c}^{t_f} R_f(T_d) \geq F^*. \quad (6)$$

We used several versions of the model that differed by the response function to temperature during the endodormancy phase (R_c) that determines the daily rate of endodormancy release:

the Wang function (Eq. 4) and a simple lower threshold function,

$$R_c(T_d) = \begin{cases} 1 & if\ T_d < T_h \\ 0 & if\ T_d \geq T_h \end{cases} \quad (7)$$

For the ecodormancy phase we used the best R_f function found with the 1-phase model for each species. These 2-phase models have 6 to 8 parameters that were either calibrated using inverse modelling techniques or prescribed using experimental data.

2.4 Parameter value estimation

The data set was divided into several data subsets: data subset 1 corresponded to data collected on sites equipped with a meteorological station; data subset 2 corresponded to data collected on sites not equipped with a meteorological station; data subset 3 corresponded to data from observation sites East of a species-specific longitudinal threshold; data subset 4 corresponded to data from observation sites West of this threshold. The species-specific longitude corresponded to the longitude on both sides of which the species was equally present (i.e. 50% of the observation sites West of the longitude and 50% East of the longitude). The data subsets contained a similar number of data ($\overline{\sigma}$=12).

First, all models were calibrated using the phenological data and corresponding meteorological data of data subset 1. The best models obtained were then additionally calibrated twice using data subset 3 and data subset 4 (Fig. 1). The three different calibrations aimed at evaluating the transferability of the models to conditions that differ from those used to calibrate the models (see section 2.6.1).

Correlative phenology models which corresponded to mixed effects models were generated in R (version 3.3.2; R Core Team, 2016) using the library nlme (Lindstrom & Bates 1990; Pinheiro & Bates 1996) with observation sites as random effect. Indeed, there may be some site-specific adaptations, which would blur an overall relationship between the temperature-based predicting variables (i.e chilling, GDD) and budburst. We considered models with all combinations of the two predicting variables (i.e. chilling and GDD), including also univariate models. GDD with two different thresholds were tested separately but in combination with chilling in multivariate models.

Process-based models were adjusted by minimizing the residual sum of squares using the simulated annealing algorithm of Metropolis (Chuine et al., 1998) using the Phenology Modelling Platform software (PMP5; http://www.cefe.cnrs.fr/fr/recherche/ef/forecast/phenology-modelling-platform) (Chuine et al., 2013). Adjustment was repeated 20 times to assure that the global optimum had been reached.

Finally, we also used the experimental observations on *Larix* to constrain the calibration of the best 2-phase model. Using the results of the two experiments, we were able to estimate the critical sum of chilling during the endodormancy phase (C^*), the response function to temperature (R_f) and the critical value F^* of the ecodormancy phase. We thus fixed C^*, the parameter values of the R_f function and F^* to those estimates. This model is

called hereafter “forward calibrated model”.

2.5. Model comparison

Models were compared using four performance indices: adjusted R-square, Akaike’s Information Criterion corrected for small sample size (AIC_C) (Burnham and Anderson, 2002); the efficiency (EFF) (Nash and Sutcliffe, 1970) ; the root mean square error for the calibration (RMSE) and validation ($RMSE_P$) data sets respectively, the systematic root mean square error (RMSEs) that estimates the model linear error (or bias).

$$\mathrm{AICc} = N \times ln\left(\frac{\sum_{i=1}^{n}(O_i - P_i)^2}{N}\right) + 2k + \left(\frac{2k(k+1)}{N-K-1}\right) \quad (8)$$

$$\mathrm{EFF} = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O}_i)^2} \quad (9)$$

$$\mathrm{RMSE_P\ or\ RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(O_i - P_i)^2}{n}} \quad (10)$$

$$\mathrm{RMSEs} = \sqrt{\frac{\sum_{i=1}^{n}\left(O_i - \hat{P}_i\right)^2}{n}} \quad (11)$$

where O_i represents observed dates, P_i represents projected dates, $\hat{P}_i$ represents regressed prediction dates and N or n is the number of observations and k is the number of parameters. We also used the ratio of performance to interquartile distance (RPIQ) that takes both the prediction error and the variation in observed values into account (Bellon-Maurel et al.,

2010). The greater the RPIQ, the better the model's predictive capacity.

$$\text{RPIQ} = \frac{(Q3-Q1)}{RMSE_p} \quad (12)$$

where *Q1* represents the value below which we can find 25% of the sample and *Q3* represents the value below which we find 75% of the sample.

Based on the AICc, EFF, RMSE, RMSEs, RPIQ and R^2 (adjusted R-squared), we first selected the best correlative model with GDD as variable, the best correlative model with GDD and chilling as variables, and the best 1-phase and 2-phase models. We additionally bootstrapped the best models (with 999 repetitions) to assess the effect of sampling bias on RMSE indices, a good overall measure of model performance. We then calibrated linear regressions of the model residuals (observed dates – predicted dates) as a function of elevation.

We additionally tested the performance of the two types of process-based model in usual *vs* unusual climatic conditions. We identified two climatically contrasted years over the observation period. First, a typical year with cold temperature during the fall and the early winter (called hereafter for the sake of simplicity winter period) followed by warm temperature during the late winter and the early spring (called hereafter for the sake of simplicity "preseason period"). Second, an unusually warm year with a warm winter period followed by a warm preseason period. Winter period corresponds to the development phase of endodormancy and preseason period corresponds to the development phase of ecodormancy. To identify the three climatically contrasted years, we compared quantile values of chilling and GDD between years for each species calculated as defined above.

2.6. Model projections

354 Following Bray & Storch (2009), we use the term “prediction” to refer to outputs of
355 models corresponding to the conditions used for the calibration; and the term ”projection” to
356 refer to outputs of models corresponding to conditions not used for the calibration. The term
357 “prediction” conveys a sense of certainty, while the term “projection” conveys a sense of
358 possibility.

359

360 *2.6.1. Projections across space*

361 We performed three spatial cross-validations to assess the transferability of the best
362 models to a wider geographic area. First, we used models calibrated on observation sites with
363 meteorological stations (data subset 1) to project the budburst date at observation sites
364 without meteorological stations (data subset2) using interpolated 2-m temperature for the
365 2007-2014 time period (see section 2.1). Second, we used models calibrated on observation
366 sites West of the species-specific longitude (data subset 3) to project the budburst date on
367 observation sites East of this longitude (data subset 4), and *vice versa*. The evaluation of the
368 models’ performance with data subsets 3 and 4 provides a more robust estimation of their
369 transferability compared to data subsets 1 and 2 because of the lower auto-correlation of the
370 meteorological conditions between them.

371 The accuracy of the projections was estimated using $RMSE_P$, $RMSE_S$, RPIQ and R^2.
372 We additionally bootstrapped models (with 999 repetitions) to assess the effect of sampling
373 bias on RMSE indices. We finally calibrated linear regressions of the model residuals
374 (observed dates – projected dates) as a function of elevation.

375

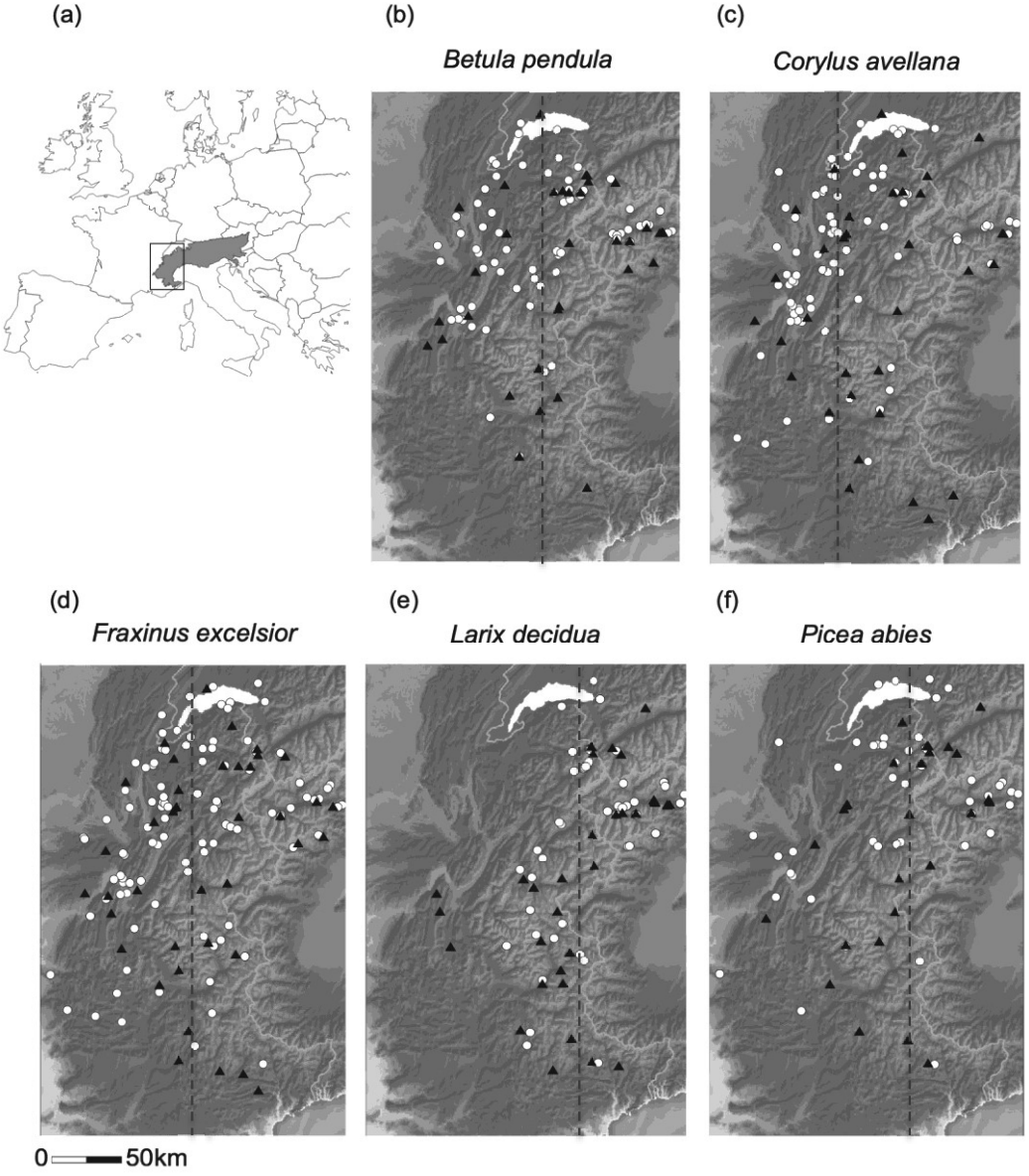
376 *2.6.2. Projections across time using climate scenarios*

We compared the projections of the best models calibrated on observation sites with meteorological stations, in historical (1950-2005) and future climatic conditions (2006 to 2100). We used the climatic data generated by the ALADIN-Climat v5 RCM model (CNRM) for the CMIP5 experiment at a 12-km resolution and downscaled at 8-km resolution using quantile-quantile method (http://drias-climat.fr/). For the future period (2006-2100), we used the scenario RCP8.5. We chose a single scenario because our objective here was to realize a sensitivity analysis of the different model types to climate change and not to provide impact projections. However, we chose the RCP 8.5 scenario because it is close to the current trajectory. Daily minimum and maximum temperatures were extracted for the grid covering the French Alps (43°48'N to 46°47'N, 5°59'E to 7°09'E). We compared the budburst dates projected by the best models for two different elevation ranges whose limits depended on the species distribution according to the CBNA (Conservatoire Botanique National Alpin) database.

The annual shift of projected budburst dates was calculated as the Sen's estimator of the slope of regressions (Sen, 1968) between projected budburst dates and years over the period 1950 to 2100. The shift in the budburst date over this period was also calculated.

Budburst dates projected by each model over the 1950-2100 time period were plotted with the associated confidence interval derived from the external validation (see section *2.5.1.*). The annual / decadal shift of projected budburst dates from Sen's estimator of slope was additionally provided for each model type and each of the 8-km cells in the French Alps for the same 1950-2100 period.

The different models that were compared and the different calibrations, validations and projections that were realized are summarized in Table 1.



401

402 **Fig. 1:** Locations of the data used for the study. Map showing the location of the study area
403 (western part of the Alps) at the scale of Europe (a). Maps showing the location of sites with
404 phenological records for the five species in the Western Alps (b-f): *Betula pendula* (birch) (b),
405 *Corylus avellana* (hazel) (c), *Fraxinus excelsior* (ash) (d), *Larix decidua* (larch) (e), *Picea abies*
406 (spruce) (f). Sites without meteorological station are shown with filled white circles and sites with
407 meteorological station are shown with filled black triangles. The dotted lines correspond to the
408 species-specific longitudes separating West and East sites that defined the sub-datasets for each
409 species (see section 2.4).

410

411

412

Models compared											
Model type	Correlative					Process-based					
Model family	single variable			bivariate		1-phase			2-phase		Forward calibrated 2-phase
Temperature response	GDD	GDD	TL	GDD +TL	GDD+ TL	GDD	sigmoid	Wang	TL + sigmoid	Wang + sigmoid	TL + sigmoid
Parameter values of the response	Tb=0	Tb=5	Tb=5	Tb=0; Th=5	Tb=5; Th=5	fitted					C*=58.1 ; F*=7.4
Model calibration: best correlative, 1-phase, 2-phase models + forward calibrated model											
Calibration datasets	meteo sites			east sites		west sites					
Time period	2007-2014										
Model validation: best correlative, 1-phase, 2-phase models + forward calibrated model											
Validation datasets (across space)	no meteo sites			west sites		east sites					
Time period	2007-2014										
Model projection: best correlative, 1-phase, 2-phase models + forward calibrated model											
Projection dataset (across time)	ALADIN-Climat v5 RCM model (CNRM) with RCP 8.5 at 8km resolution										
Time period	1950-2100										

413

414 **Table 1:** Summary of the different models that were compared and the different calibrations,
415 validations and projections that were realized. GDD is the abbreviation of growing degree
416 day function and TL of the threshold lower function. Meteo sites correspond to sites equipped
417 with a meteorological station, no meteo sites corresponds to sites not equipped with a
418 meteorological station, east sites correspond to observation sites East of a species-specific
419 longitudinal threshold, west sites correspond to observation sites West of this threshold.

420

3. Results

3.1. Selection of the best models

Mixed effects models which best explained budburst dates were generally models including GDD5, and chilling as predicting variables (Appendix B). However, budburst dates of *Corylus* and *Fraxinus* were best explained by GDD0 together with chilling. Variance partitioning also indicated an important joint contribution of GDD0 or GDD5 and chilling for all species (Appendix B). However, over all 8 years and all locations, chilling did not explain a significant part of the variance in spring phenology once forcing temperatures were taken into account except in *Picea* (Appendix B). For further analysis we only present the results of the correlative model with GDD0 or GDD5 (depending of the species) and chilling as variables.

The response function to temperature in process-based models that best explained the budburst date was the lower threshold function for *Fraxinus* and *Picea* and the Wang function for the three other species for the endodormancy phase; and the sigmoid function for all species for the ecodormancy phase (Appendix C). However, because the differences in efficiency and AICc were very small between models using the Wang function and the threshold function, we selected the lower threshold function for all species for the rest of the analyses, for the sake of parsimony.

Process-based models performed the best for *Picea* and *Betula* (Fig. 2, Appendix D, Appendix F and Appendix G). However, model error increased with the distance to the mean budburst date, irrespective of the model and the species, with a tendency towards exaggerating early and late budburst dates (predicted dates too early or too late, respectively) (Appendix F and Appendix G). We found no trend in model residuals (observed – predicted

444 dates) along elevation, except for *Larix* for which observed budburst dates tended to occur
445 earlier than predicted by the models at low elevation and reversely at high elevation
446 (Appendix H and Appendix I).

447 All performance indices indicated that correlative models predicted the budburst date
448 more accurately than process-based models whatever the calibration dataset for all species
449 (Appendix D). This result was confirmed by bootstrapped values of RMSE that were
450 significantly lower for correlative models than process-based models for all species (Fig. 2).
451 Performance indices were only slightly better for 2-phase models than 1-phase models (Fig.
452 2, Appendix D). However, comparing the performances for an unusually warm year, i.e. with
453 warm winter and warm preseason, we found that 2-phase process-based models performed
454 better for *Corylus* and *Picea* (Fig. 3). For the other species, model performance was similar
455 between the one- and 2-phase process-based models. For a typical year with a cold winter and
456 a warm preseason, performance indices were similar between the two types of process-based
457 model for all species (Fig. 3).

458

459 *3.2. Model transferability in space*

460 Whatever the validation data subset, process-based models provided more accurate
461 projections of the budburst date than correlative models, except for *Picea abies* for which
462 models' performed similarly (Fig. 2, Appendix E). Performances were similar between 2-
463 phase models and 1-phase models. Process-based models were thus more robust than
464 correlative models. Indeed, while RMSE increased from predictions to projections in
465 correlative models, they remained similar in process-based models (Fig. 2, Appendix E).
466 Therefore, model performance of process-based models was more stable (3.45 and 0.87 mean

467 difference in RMSE between calibration and validation for the correlative and process-based
468 models respectively).

469 Similarly to the predicted dates, error in projected dates increased with the distance to
470 the mean budburst date, irrespective of the model and the species (Appendix F and Appendix
471 G) with exaggerating early and late dates.

472 Model residuals were linearly related to elevation for *Corylus* and *Fraxinus*
473 (correlative models only), and *Larix* (all models), with dates projected slightly too late at low
474 elevation and too early at high elevation (Appendix H and I).

475

476 *3.3. Projections of the budburst date in future climatic conditions*

477 *Evolution across time*

478 We compared the projections of the budburst date of the five species by the best
479 correlative model and the best process-based 1-phase and 2-phase models over the historical
480 period and the 21st century in the French Alps using climatic data simulated by the ALADIN-
481 Climat RCM. Uncertainties in the climatic data (minimum temperatures, maximum
482 temperatures) were similar along the period and should not therefore add a bias across time in
483 the model projections (Appendix J).

484 Model projections differed substantially according to elevation and time period. At
485 low elevation, up to 2050, correlative models and 1-phase models projected a continuous
486 trend for earlier budburst dates -0.08 to -0.16 days/year and -0.09 to -0.12 days/year
487 respectively; Fig. 4 & 5, Appendix K) while 2-phase models showed a much weaker trend for
488 earlier budburst dates (-0.05 to -0.07 days/year; Fig. 4, Appendix K). Models projection
489 diverged substantially after 2050. While correlative models projected still earlier budburst

dates (-0.27 to -0.48 days/year; Fig. 4, Appendix K), trend weakened in 1-phase models projection (-0.03 to -0.20 days/year; Fig. 4, Appendix K); and vanished or even reverted in 2-phase models projection (-0.01 to +0.16 days/year; Fig. 4, Appendix K).

At high elevation, correlative models provided astonishing projections with very low interannual variability of budburst dates, a slight trend towards earlier date from 1950 to 2050 (-0.005 to -0.07 days/year; Fig. 4, Appendix K) that steepened after 2050 (-0.13 to -0.28 days/year; Fig. 4, Appendix K). Process-based models showed a trend for earlier budburst dates (-32 to -46 days) all along the 1950-2100 period. 1-phase models and 2-phase models both projected a remarkably similar trend towards earlier budburst dates all along the historical period and until 2050 (-0.16 to -0.21 days/year; Fig. 4, Appendix K). However, while after 2050, the trend towards earlier budburst dates remained similar between the two process-based models for *Larix* (-0.42 days/year) and *Picea* (-0.32 days/year), it weakened in the 2-phase model for the other species (Fig. 4, Appendix K).

For *Fraxinus*, *Betula* and *Picea,* 2-phase models episodically projected budburst failure due to endodormancy release failure because of a lack of chilling. This situation occurred especially at low elevation and with an increasing frequency along the 21th century (Fig. 5).

Variation across space

Correlative models projected earlier budburst dates over the French Alps at an 8-km resolution with a trend more pronounced in the outer and southern Alps than in the central Alps where the elevation is higher on average. The contrast between outer and central Alps was the most pronounced for *Betula*, *Larix* and *Picea* (Fig. 6). Reversely, 1-phase and 2-

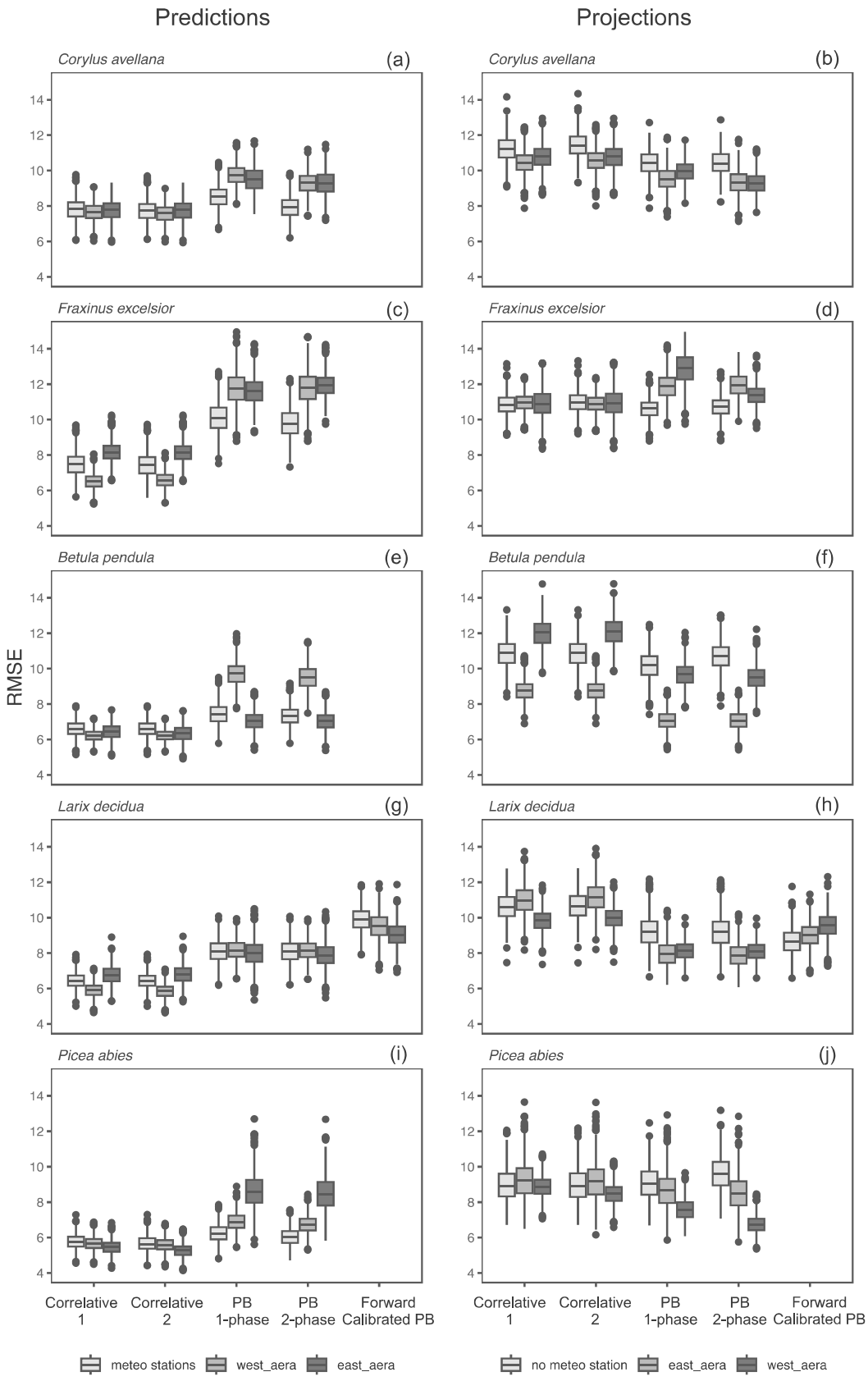
phase models projected a steeper trend towards earlier budburst dates in the central Alps than the outer Alps. However, contrary to 1-phase models, 2-phase models projected later dates in the southern Alps for all species, and also in the north of the outer Alps for *Larix* and *Picea* (Fig. 6). The contrasted projections between the outer and central Alps were more pronounced for *Larix* (Fig. 6).

3.4. Performance of the forward calibrated model

The forward calibrated 2-phase model for *Larix* performed the worst on average compared to other models in predicting the budburst date whatever the calibration data subset (Fig. 2, Appendix D), but nevertheless provided more accurate predictions of early dates than the 2-phase model calibrated with inverse modelling (Appendix L). The model also provided the most accurate projections for the sites without meteorological stations, but not for the other data subsets (Fig. 2, Appendix E). More interestingly, this model was the only one to provide a lower error on average with the validation data subsets than with the calibration data subsets (-0.39 *vs* +0.65 for 2-phase model and +0.84 for 1-phase model), and thus showed the reverse behavior we usually observe with models of an increased error in external (validation) conditions. Like other models (either correlative or process-based) for *Larix decidua*, residuals were linearly related to elevation, with projected dates tending to be overestimated at low elevation and underestimated at high elevation (Appendix M).

Over the historical period and the 21st century, budburst dates projected by the forward calibrated 2-phase model were very similar to those projected by the 2-phase model either at low or high elevation (Fig. 7). However, the forward calibrated model showed much more frequent endodormancy release failures than the 2-phase model which projected none for *Larix* (Fig. 5 g, h). For this reason, we could not represent the shift of the budburst date

projected by the forward calibrated model for *Larix* over the French Alps for each grid cell for the period 1950-2100 (Fig. 6).


Predictions
Projections
Corylus avellana
(a)
(b)
Fraxinus excelsior
(c)
(d)
Betula pendula
(e)
(f)
Larix decidua
(g)
(h)
Picea abies
(i)
(j)
RMSE
Correlative 1
Correlative 2
PB 1-phase
PB 2-phase
Forward Calibrated PB
meteo stations
west_aera
east_aera
no meteo station
east_aera
west_aera

539

Fig. 2: Boxplots of bootstrapped values of RMSE (days). Performance of the best models in predicting (a, c, e, g, i) and projecting the budburst date (b, d, f, h, j). Predictions correspond to the budburst dates predicted by the models for their respective calibration dataset: sites with meteorological stations (clear grey); sites West of the species-specific longitudinal threshold (grey); sites East of the species-specific longitudinal threshold (dark grey). Projections correspond to budburst dates predicted by the three different versions of calibrated models for respectively observation sites without meteorological stations (clear grey); observations sites East of the species-specific longitudinal threshold (grey); observation sites West of the species-specific longitudinal threshold (dark grey). Models on the X axis are the same calibrated models on right and left panels which differ only by the datasets used to evaluate the model performance distinguished by the different shades of grey. Correlative 1 refers to the linear mixed model with GDD only. Correlative 2 refers to the linear model with GDD and chilling as variables. PB 1-phase refers to the process-based 1-phase model. PB 2-phase refers to the process-based 2-phase model. Forward calibrated PB refers to the forward calibrated process-based 2-phase model.

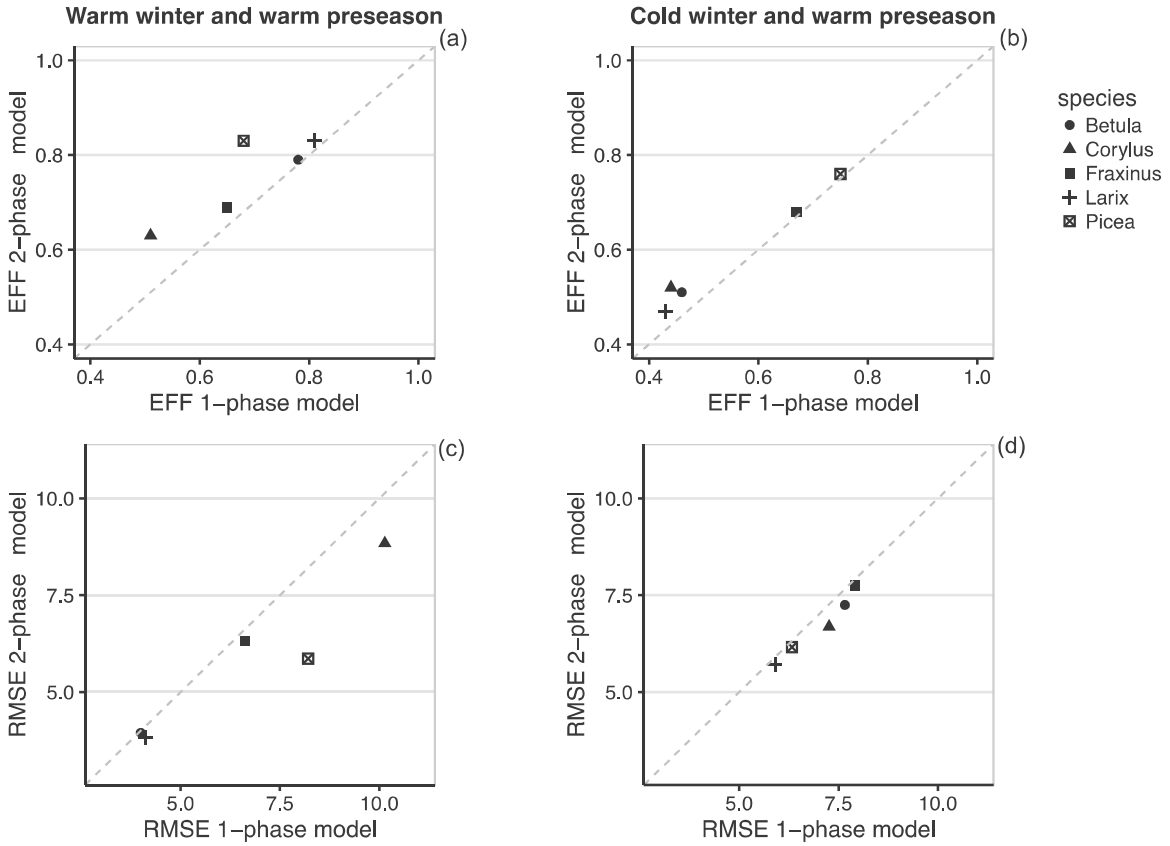


Fig. 3: Performance (Efficiency a, b; RMSE c, d) of the process-based models in predicting the budburst date of the calibration data subset 1 (sites with meteorological stations) for a year with a warm winter followed by a warm preseason (a, c) and a year with a cold winter followed by a warm preseason (b, d). Process-based 1-phase models use a sigmoid function of temperature for the ecodormancy phase. Process-based 2-phase models use a lower threshold function of temperature for the endodormancy phase and a sigmoid function of temperature for the ecodormancy phase. EFF, model efficiency; RMSE, roots-mean-squared error (days). Model labels as in Fig. 2.

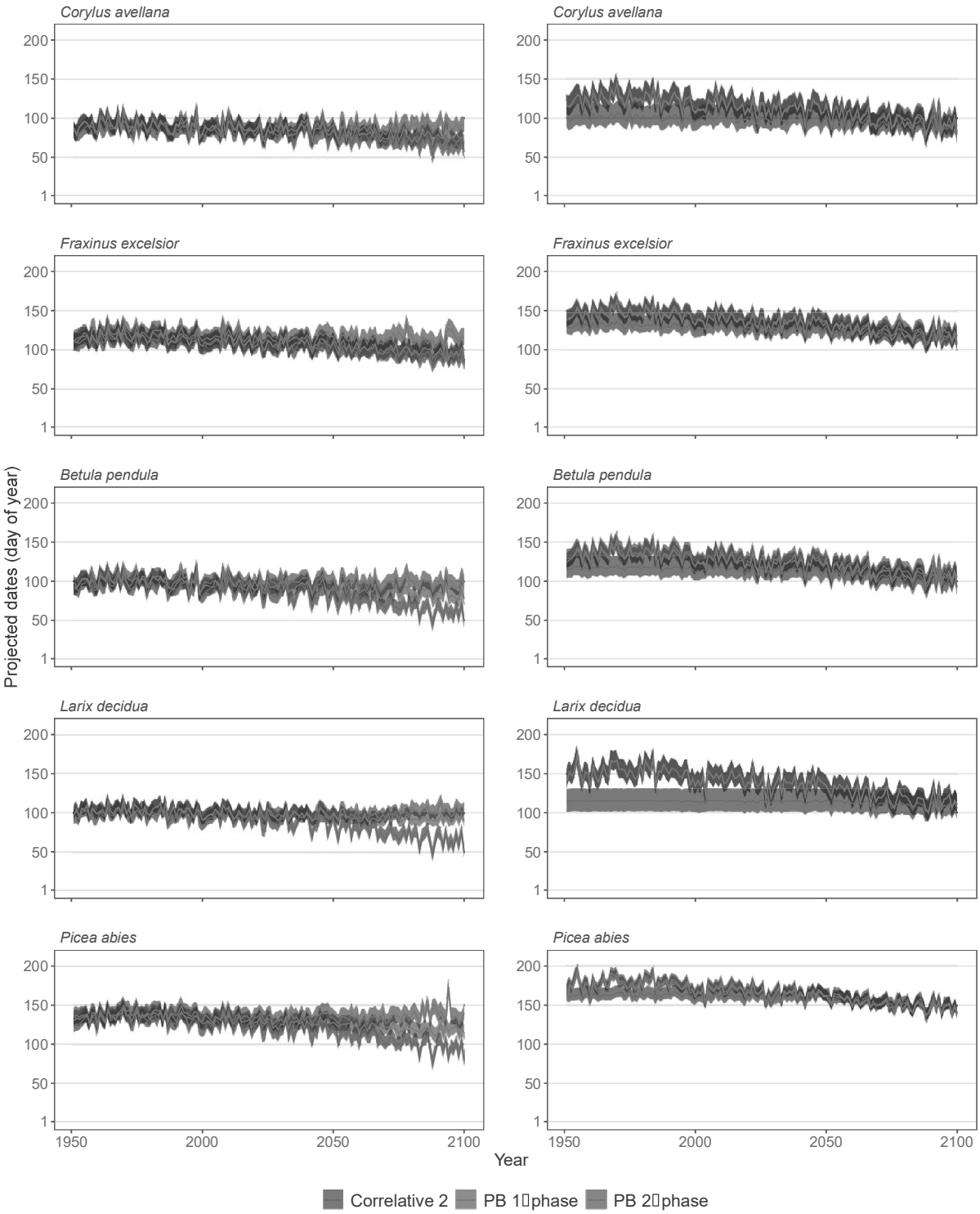
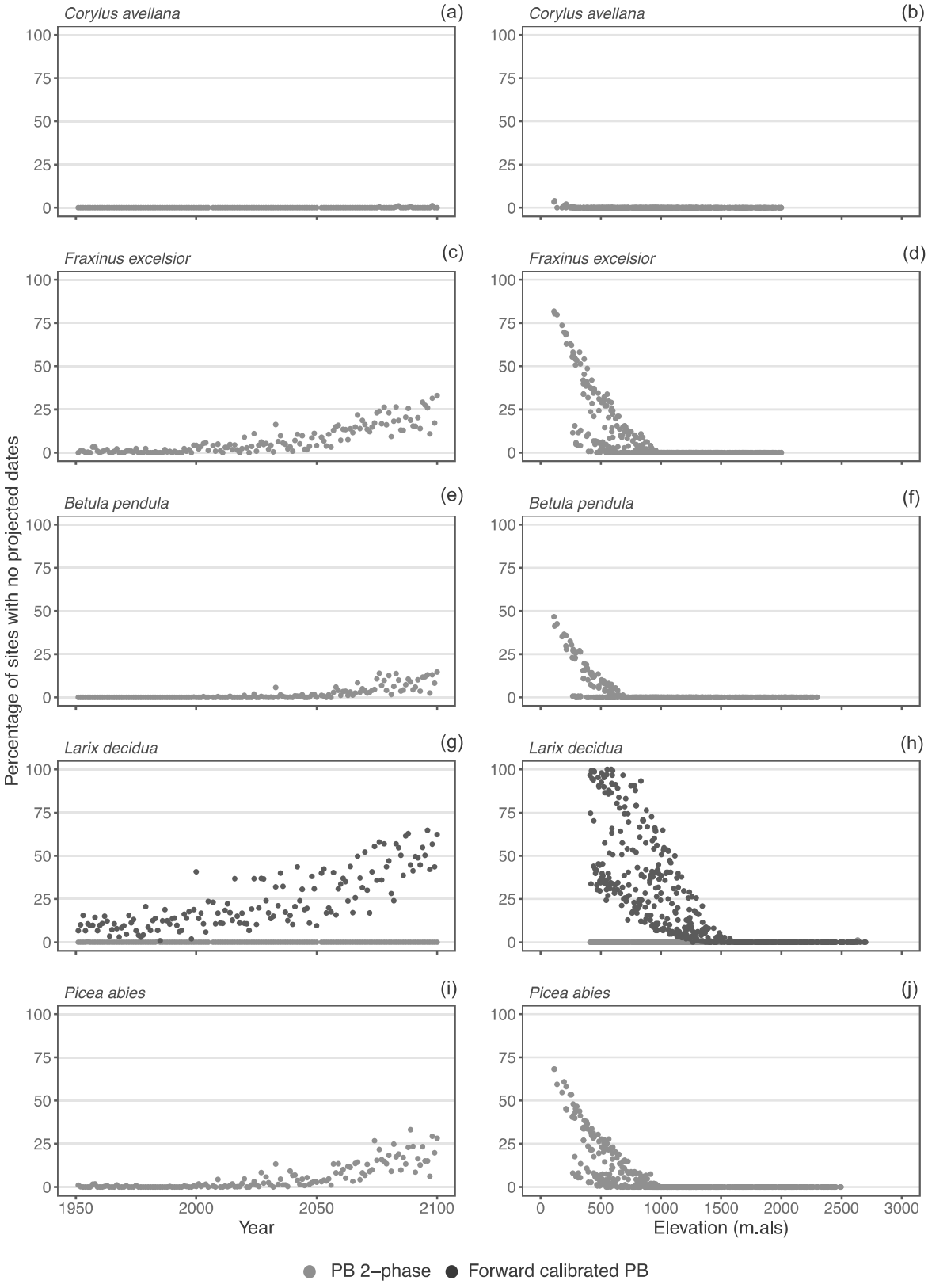


Fig. 4: Budburst date projected by the correlative 2 model (with GDD and chilling as variables) (blue), the process-based 1-phase (grey) and the process-based 2-phase model (red) in historical climate and future climate according to scenario RCP8.5, in the Alps at low

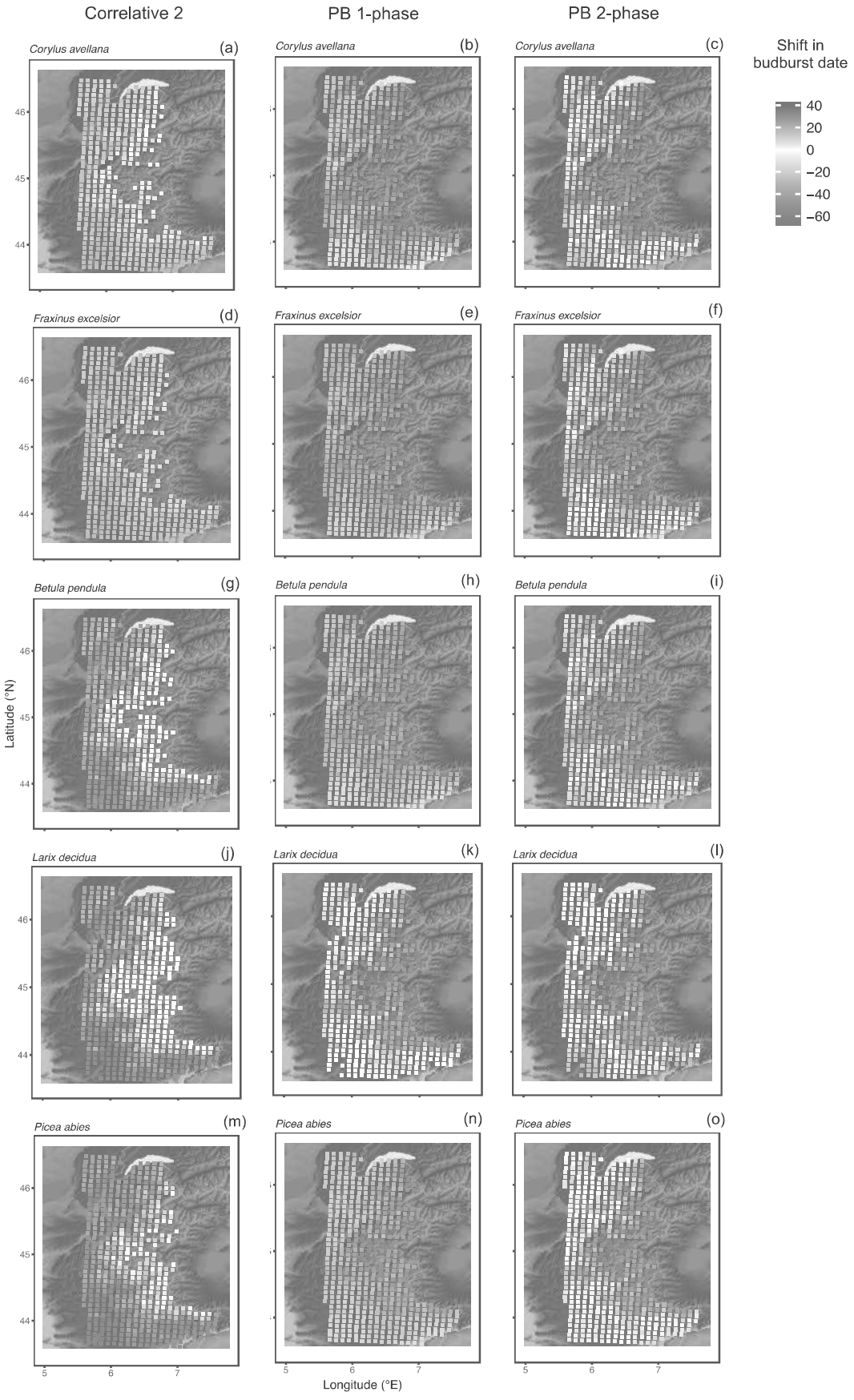
573 elevation (a, c, e, g, i) and high elevation (b, d, f, h, j). Elevation limits depend on the species.
574 Low elevation: *Corylus*= 100 to 1000m asl.; *Fraxinus*= 100 to 1000 m asl.; *Betula*= 100 to
575 1000 m asl.; *Larix*= 400 to 1500 m asl. and *Picea*= 100 to 1500 m asl.. For high elevation:
576 *Corylus*= 1001 to 2000m asl.; *Fraxinus*= 1001 to 2000 m asl.; *Betula*= 1001 to 2300 m asl.;
577 *Larix*= 1501 to 2700 m asl. and *Picea*= 1001 to 2500 m asl.. Bold lines indicate the projected
578 mean date and areas indicate the confidence interval of the model. Model labels as in Fig. 2.
579

580


Corylus avellana
(a)
Corylus avellana
(b)
Fraxinus excelsior
(c)
Fraxinus excelsior
(d)
Betula pendula
(e)
Betula pendula
(f)
Larix decidua
(g)
Larix decidua
(h)
Picea abies
(i)
Picea abies
(j)
Percentage of sites with no projected dates
0
25
50
75
100
1950
2000
2050
2100
Year
0
500
1000
1500
2000
2500
3000
Elevation (m.als)
PB 2–phase
Forward calibrated PB

581

582 **Fig. 5:** Percentage of sites showing dormancy breaking failure under climate scenario
583 RCP8.5 in the Alps along the years (a, c, e, g, i) and along the elevation gradient (b, d, f, h, j).
584 Grey dots indicate values for the process-based 2-phase model and black dots indicate values
585 for the forward calibrated process-based 2-phase model. Note that the 2-phase process-based
586 model did not project dormancy breaking failure for *Larix decidua*. Model labels as in Fig. 2.
587


Correlative 2
PB 1-phase
PB 2-phase
Corylus avellana
(a)
(b)
(c)
Fraxinus excelsior
(d)
(e)
(f)
Betula pendula
(g)
(h)
(i)
Larix decidua
(j)
(k)
(l)
Picea abies
(m)
(n)
(o)
Shift in budburst date
40
20
0
−20
−40
−60
Latitude (°N)
Longitude (°E)
46
45
44
5
6
7

588

589 **Fig. 6:** Shift in budburst date projected by the correlative (with GDD and chilling as variables)
590 (a, d, g, j, m), the process-based 1-phase (b, e, h, k, n) and 2-phase (c, f, i, l, o) models in the
591 French Alps (spatial resolution: 8 km) between 1950 and 2100 (scenario RCP8.5). Shift was
592 calculated as the Sen's estimator of the slope of regressions between projected budburst dates
593 and years over the period 1950 to 2100.
594

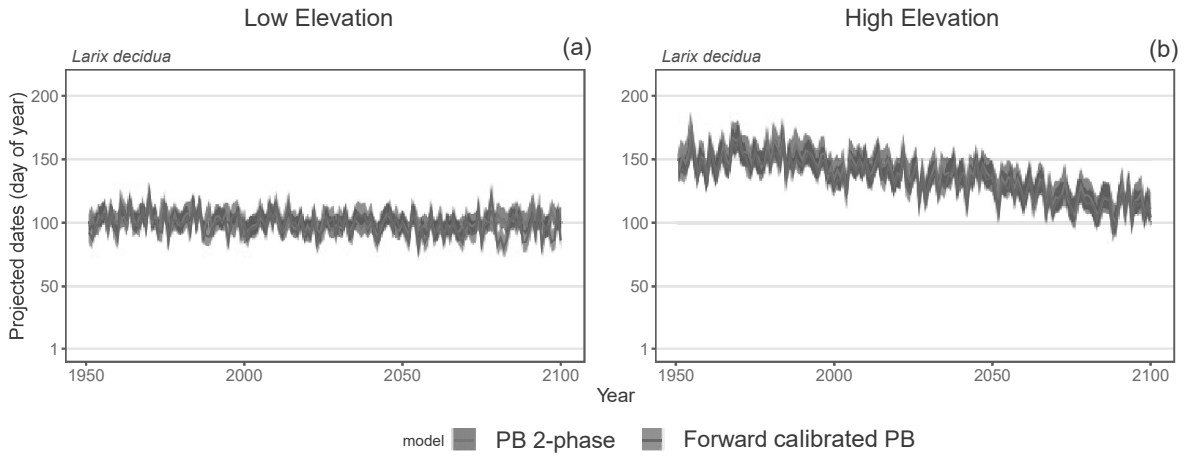


Fig. 7: Budburst date of *Larix decidua* projected by the 2-phase model (red) and the forward calibrated 2-phase model (yellow) in historical and future climate (scenario RCP8.5) in the Alps at low elevation (400 to 1500 m asl.) (a) and high elevation (1501 to 2700 m asl.) (b). Red curves indicate the mean date projected by the 2-phase model and red areas indicate the confidence interval of the model. Brown curves indicate the mean date projected by the forward calibrated model and orange areas indicate the confidence interval of the model. Because the forward calibrated model showed much more frequent budburst failures due to endodormancy release failures than the 2-phase model, we could not represent the deviation of budburst date projected by the calibrated model over the French Alps for the period 1950-2100. Model labels as in Fig. 2.

4. Discussion

4.1. Process-based phenology models provide more robust projections than correlative models

Linear mixed models using chilling days and growing degree days (GDD) were more efficient than process-based models in predicting the budburst date in calibration conditions (i.e. *sensu* Randin *et al.*, 2006). However, process-based models were more efficient than correlative models to project budburst date in external conditions (except for *Picea* for which the three model types revealed similar performance). Error in predicted dates and projected dates increased with the distance to the mean budburst date of the calibration datasets, irrespective of the model and the species. This error inflation at the edge of the calibration range leads to exaggerating early and late dates. However, errors were lower for process-based models compared to correlative models probably because the former formalize causal relationships between the dependent variable and the independent variables across the range of projection contrary to the latter. Although, these conclusions are dependent on the models we have used for this study the way we chose the models asserts a certain degree of generality to these conclusions because models differed essentially by their structure and not by their predicting variables.

Projections of the forward calibrated 2-phase model were not more accurate than those of the other 2-phase models except for sites without meteorological stations. However, it was the only model that did not show an increased error in external conditions. Increased accuracy in this model was mainly achieved through a more accurate prediction of early dates. Our results show that robustness of process-based models can come additionally from forward parameter estimation, but also that forward parameter estimation is not necessarily

633 the Holy Grail that we should seek. Another study conducted on rice (Nagano *et al.,* 2012;
634 but see Satake *et al.,* 2013) showed that phenology models calibrated by forward estimation
635 actually might perform worse *in situ* conditions than in controlled conditions and *vice versa*.
636 This study suggests that either the effect of some abiotic factors and their interactions or the
637 daily variation of these factors in nature might be poorly represented under controlled
638 conditions, and that epigenetic or acclimation effects might interfere. Besides this, results
639 from experiments can sometimes be biased by ontogenetic effects when young cohorts of
640 individuals are used (Vitasse, 2013). Thus, it seems that backward and forward parameter
641 estimations have both biases that come from the same reason, i.e. incomplete sampling of the
642 environmental conditions that affect the phenology, wherever natural or artificial.

643 Therefore, obtaining realistic and accurate response curves of plant responses to
644 environmental cues might actually require a combination of both the experimental approach
645 (forward estimation) and the statistical approach (backward estimation), especially when for
646 the latter, calibration datasets are small and span short time periods or small geographical
647 areas. However, forward estimation of model's parameters requires costly equipment such as
648 greenhouses or growth chambers for which temperature and light conditions can be
649 controlled. When forward estimation is impossible, backward estimation of parameter values
650 could nevertheless benefit from additional expert knowledge based on experimental grounds
651 already published in the literature (see for review Chuine and Régnière, 2017).

652

653 *4.2. Model transferability through time and space*

654 Divergence between projections of correlative and 2-phase models in future climatic
655 conditions across the Alps are striking. For all species at high elevation over the 1950-2050

656 period, correlative models projected earlier dates than 2-phase models, while after 2050,
657 projections were similar. Standard deviations of the projected dates before 2050 for
658 correlative models were also surprisingly small, particularly for *Larix.* This might be due to
659 the fact that the GDD variable was calculated over the same period of the year along the
660 elevation gradient. At high elevation, GDD remains low every year, yielding a low variation
661 in the projected budburst dates, which was particularly marked during the historical period.
662 Because *Larix* is the species that reaches the tree line, variations of GDD are even lower for
663 this species compared to the four other species. These results highlight the limitations of
664 correlative models to simulate budburst dates when transferred to another spatial or temporal
665 domain.

666 Performance indices were similar between 1-phase models and 2-phase models for
667 predictions and projections in historical conditions, suggesting that so far chilling had no
668 major effect on the budburst date because it was sufficient to fully release endodormancy.
669 Consequently, this means that 1-phase models are complex enough to predict and project
670 budburst dates accurately in historical climatic conditions (Chuine *et al.,* 1999 ; Linkosalo *et*
671 *al.,* 2008 ; Vitasse *et al.,* 2011). However, 2-phase models provided more accurate
672 predictions of the budburst date for particular recent years with warm winter. This was
673 especially marked for *Corylus* and *Picea*. Thus, a certain level of winter warming might have
674 been reached in the Alps, sometimes leading to non-optimal chilling conditions and later
675 dormancy release date compared to colder years. In such situation, the integration of an
676 endodormancy phase in the phenology model clearly increases the projections accuracy. The
677 negative effect of winter warming on endodormancy release in tree species has been shown in
678 a previous study (Asse et al. 2018) and has also been proposed as an explanation of the
679 dampening of the response of budburst to warming temperature over the past 25 years in

Europe (Fu et al., 2015). Although warm winter conditions are still rare in our past temperature records, they will become more frequent in the upcoming decades especially after 2050 according to climate projections. Such conditions are expected to delay dormancy release and ultimately delay or even prevent budburst (Fig. 2 & 4). In this context, it becomes essential and urgent that robust 2-phase models be available for as many tree species as possible.

Although chilling was taken into account in the correlative models, they did not project a negative effect of a lack of chilling on the budburst date like the 2-phase process-based model. Although the two types of models were calibrated with the same dataset which contained only one exceptionally warm winter (2006-2007) which had a delaying effect on the budburst date (Asse et al., 2018), it seems that only the process-based model was able to capture this important information.

4.3. Evolution of the budburst date in the Alps over the 21st century

At low elevation, projected budburst dates were similar until ca. 2050 for all models and for all species, with a weak linear trend towards earlier dates (Fig 5 & 6). This linear trend is due to the warming of spring which accelerates cell elongation while the warming of winter does not affect chilling accumulation until 2050 according to climate projections as previously highlighted by Vitasse *et al.,* (2011) ; Chuine *et al*., (2016) and Fig. 1 in Guo et al., (2015). However, after 2050, while the linear trend towards earlier dates continues and accelerates in correlative and 1-phase models (although less in the latter and especially for *Larix*), this trend progressively vanished in 2-phase models.

At high elevation, projected budburst dates were similar for all species over the 21st century between 1-phase and 2-phase models, suggesting that at high elevation, chilling requirements could be fulfilled until the end of the 21st century.

Thus, our results suggest that warmer winters might have opposite effects on spring phenology at high compared to low elevation, by advancing *vs* delaying dormancy release respectively, and consequently reducing the phenological cline across the elevation gradient. As a consequence, trees might become at increasing risk of late spring frost damage at high elevation compared to low elevation in the upcoming decades. A preview of such situations have already been reported in the Swiss Alps over the last decades (Vitasse et al., 2017, 2018). Considering that phenological traits have been shown to be primary determinants of species range (Chuine & Beaubien, 2001; Chuine *et al.*, 2013), a reduced phenological shift across the elevation gradient might in the medium term alter the altitudinal zonation of the vegetation.

Our results also support the hypothesis of Vitasse *et al.*, (2017) that winter temperatures are currently actually colder than the optimal chilling temperature at high elevation so that a warming of winter actually increases the number of chilling days and advance endodormancy release, and consequently budburst. Bud endodormancy indeed requires from a few weeks to several months of cold temperature, that can vary from 0°C to 12°C depending on the species, to be fully released (Lang et al., 1987; Caffarra et al., 2011a; Malagi et al., 2015).

From 1950 to 2100, 2-phase models projected a trend for earlier budburst date in the inner Alps (higher elevations), while they projected either very weak trend toward earlier dates or a trend toward later dates in the outer Alps (lower elevations) (particularly for *Larix*

and *Picea*), and a trend toward later dates whatever the species in the Southern Alps, which match or approach their southern range limit. In the short term, delaying effect of winter warming on budburst date might be beneficial to trees by reducing the risk of exposure to late spring frost. But in the longer term, it is expected to lead to partial dormancy release at low elevation and at species southern range limit leading to leaf and flower malformation and erratic bud break (Bonhomme *et al.*, 2005; Erez, 2000; Petri & Leite, 2004; Zguigal *et al.*, 2006; Caffarra *et al.*, 2011b; Laube *et al.*, 2014). Additionally to temperature, photoperiod might also co-control budburst in about 30% of temperate tree species (Caffarra et al., 2011; Basler and Körner, 2012; Laube et al., 2014; Zohner and Renner, 2015), with long photoperiod compensating for a lack of chilling. Integration of this compensatory effect in the models might change substantially the projections of budburst dates for the end of the 21st century for photosensitive species which might be the case for *Picea abies* (Blümel & Chmielewski, 2012; Gauzere *et al.*, 2017).

Conclusion

Our results showed that (1) process-based phenology models are more robust than correlative model even when they rely entirely on backward estimation (inverse modelling) of their parameter values. (2) They also demonstrated that the robustness of process-based models could be increased, though not substantially, when their calibration could rely on forward estimation. Therefore, the robustness of process-based models seems to come primarily from the explicit description of causal relationships rather than from forward estimation of model parameters, and we advise using, whenever possible, both backward and forward estimation of model parameters. (3) In opposite to correlative models, process-based

models projected a reduction in the phenological cline along the elevation gradient for all species by the end of the 21st century. This later result suggests that using linear relationships to provide projections for the second part of the 21st century will be vain, especially at low elevation and at species southern range limits.

Acknowledgements

We are grateful to all the volunteers and staffs of protected areas involved in the Phenoclim program for their help and support for collecting data on the study sites. We also thank Geoffrey Klein (Centre de Recherches sur les Ecosystèmes d'Altitude CREA, Chamonix-Mont-Blanc) for having prepared the temperature data used in this study. We thank the Conservatoire Botanique National Alpin for providing the species occurrence data, and two anonymous referees for their constructive comments on earlier versions of the manuscript. This research project was funded by the Association Nationale de la Recherche et de la Technologie (ANRT). This work is part of the project ''Phenoclim'' supported by the Rhône-Alpes and PACA Regions, French ministry of Transition écologique et solidaire. We thank the Société Académique Vaudoise (SAV) for having kindly provided matching funds.

766 **References**

767 Anderson, J.J., Gurarie, E., Bracis, C., Burke, B.J., Laidre, K.L., 2013. Modeling climate
768 change impacts on phenology and population dynamics of migratory marine species.
769 Ecol. Modell. 264, 83–97. doi:10.1016/j.ecolmodel.2013.03.009
770 Asse, D., Chuine, I., Vitasse, Y., Yoccoz, N.G., Delpierre, N., Badeau, V., Delestrade, A.,
771 Randin, C.F., 2018. Warmer winters reduce the advance of tree spring phenology
772 induced by warmer springs in the Alps. Agric. For. Meteorol. 252, 220–230.
773 doi:10.1016/j.agrformet.2018.01.030
774 Basler, D., 2016. Evaluating phenological models for the prediction of leaf-out dates in six
775 temperate tree species across central Europe. Agric. For. Meteorol. 217, 10–21.
776 doi:10.1016/j.agrformet.2015.11.007
777 Basler, D., Körner, C., 2012. Photoperiod sensitivity of bud burst in 14 temperate forest tree
778 species. Agric. For. Meteorol. 165, 73–81. doi:10.1016/j.agrformet.2012.06.001
779 Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.M., McBratney, A., 2010.
780 Critical review of chemometric indicators commonly used for assessing the quality of
781 the prediction of soil attributes by NIR spectroscopy. TrAC - Trends Anal. Chem.
782 doi:10.1016/j.trac.2010.05.006
783 Benmoussa, H., Ben Mimoun, M., Ghrab, M., Luedeling, E., 2018. Climate change threatens
784 central Tunisian nut orchards. Int. J. Biometeorol. 62, 2245–2255. doi:10.1007/s00484-
785 018-1628-x
786 Blümel, K., Chmielewski, F.M., 2012. Shortcomings of classical phenological forcing models
787 and a way to overcome them. Agric. For. Meteorol. 164, 10–19.
788 doi:10.1016/j.agrformet.2012.05.001
789 Bonhomme, M., Rageau, R., Lacointe, A., Gendraud, M., 2005. Influences of cold
790 deprivation during dormancy on carbohydrate contents of vegetative and floral
791 primordia and nearby structures of peach buds (Prunus persica L. Batch). Sci. Hortic.
792 (Amsterdam). 105, 223–240. doi:10.1016/j.scienta.2005.01.015
793 Bray, D., Storch, H. Von, 2009. "Prediction" or "Projection"? 534–543.
794 Burghardt, L.T., Metcalf, C.J.E., Wilczek, A.M., Schmitt, J., Donohue, K., 2015. Modeling
795 the Influence of Genetic and Environmental Variation on the Expression of Plant Life
796 Cycles across Landscapes. Am. Nat. 185, 212–227. doi:10.1086/679439
797 Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference: a practical
798 information-theoretic approach, Ecological Modelling.
799 doi:10.1016/j.ecolmodel.2003.11.004
800 Caffarra, A., Donnelly, A., Chuine, I., 2011a. Modelling the timing of Betula pubescens
801 budburst. II. Integrating complex effects of photoperiod into process-based models.
802 Clim. Res. 46, 159–170. doi:10.3354/cr00983
803 Caffarra, A., Donnelly, A., Chuine, I., Jones, M.B., 2011b. Modelling the timing of Betula
804 pubescens budburst. I. Temperature and photoperiod: A conceptual model. Clim. Res.
805 46, 147–157. doi:10.3354/cr00980
806 Cheaib, A., Badeau, V., Boe, J., Chuine, I., Delire, C., Dufrêne, E., François, C., Gritti, E.S.,
807 Legay, M., Pagé, C., Thuiller, W., Viovy, N., Leadley, P., 2012. Climate change impacts
808 on tree ranges: Model intercomparison facilitates understanding and quantification of
809 uncertainty. Ecol. Lett. 15, 533–544. doi:10.1111/j.1461-0248.2012.01764.x
810 Chuine, I., 2010. Why does phenology drive species distribution? Philos. Trans. R. Soc. B
811 Biol. Sci. 365, 3149–3160. doi:10.1098/rstb.2010.0142
812 Chuine, I., 2000. A Unified Model for Budburst of Trees. J. Theor. Biol. 207, 337–347.

doi:10.1006/jtbi.2000.2178
Chuine, I., Beaubien, E.G., 2001. Phenology is a major determinant of tree species range. Ecol. Lett. 4, 500–510. doi:10.1046/j.1461-0248.2001.00261.x
Chuine, I., Bonhomme, M., Legave, J.M., Garcia de Cortazar-Atauri, I., Charrier, G., Lacointe, A., Améglio, T., 2016. Can phenological models predict tree phenology accurately in the future? The unrevealed hurdle of endodormancy break. Glob. Chang. Biol. 22, 3444–3460. doi:10.1111/gcb.13383
Chuine, I., Cour, P., Rousseau, D.D., 1999. Selecting models to predict the timing of flowering of temperate trees: Implications for tree phenology modelling. Plant, Cell Environ. 22, 1–13. doi:10.1046/j.1365-3040.1999.00395.x
Chuine, I., Cour, P., Rousseau, D.D., 1998. Fitting models predicting dates of flowering of temperate-zone trees using simulated annealing. Plant, Cell Environ. 21, 455–466. doi:10.1046/j.1365-3040.1998.00299.x
Chuine, I., De Cortazar-Atauri, I.G., Kramer, K., Hänninen, H., 2013. Plant development models, in: Phenology: An Integrative Environmental Science. pp. 275–293. doi:10.1007/978-94-007-6925-0_15
Chuine, I., Régnière, J., 2017. Process-Based Models of Phenology for Plants and Animals. Annu. Rev. Ecol. Evol. Syst. 48, 159–182. doi:10.1146/annurev-ecolsys-110316-022706
Cianfrani, C., Satizábal, H.F., Randin, C., 2015. A spatial modelling framework for assessing climate change impacts on freshwater ecosystems : Response of brown trout (Salmo trutta L .) biomass to warming water temperature. Ecol. Modell. 313, 1–12. doi:10.1016/j.ecolmodel.2015.06.023
Cooke, J.E.K., Weih, M., 2005. Nitrogen storage and seasonal nitrogen cycling in Populus: Bridging molecular physiology and ecophysiology. New Phytol. doi:10.1111/j.1469-8137.2005.01451.x
Dantec, C.F., Vitasse, Y., Bonhomme, M., Louvet, J.M., Kremer, A., Delzon, S., 2014. Chilling and heat requirements for leaf unfolding in European beech and sessile oak populations at the southern limit of their distribution range. Int. J. Biometeorol. 58, 1853–1864. doi:10.1007/s00484-014-0787-7
Delpierre, N., Soudani, K., François, C., Köstner, B., Pontailler, J.Y., Nikinmaa, E., Misson, L., Aubinet, M., Bernhofer, C., Granier, A., Grünwald, T., Heinesch, B., Longdoz, B., Ourcival, J.M., Rambal, S., Vesala, T., Dufrêne, E., 2009. Exceptional carbon uptake in European forests during the warm spring of 2007: A data-model analysis. Glob. Chang. Biol. 15, 1455–1474. doi:10.1111/j.1365-2486.2008.01835.x
Duputié, A., Rutschmann, A., Ronce, O., Chuine, I., 2015. Phenological plasticity will not help all species adapt to climate change. Glob. Chang. Biol. 21, 3062–3073. doi:10.1111/gcb.12914
Erez, A., 2000. Bud Dormancy; Phenomenon, Problems and Solutions in the Tropics and Subtropics. Temp. Fruit Crop. Warm Clim. 17–48. doi:10.1007/978-94-017-3215-4_2
Fu, Y.H., Piao, S., Op de Beeck, M., Cong, N., Zhao, H., Zhang, Y., Menzel, A., Janssens, I.A., 2014. Recent spring phenology shifts in western Central Europe based on multiscale observations. Glob. Ecol. Biogeogr. 23, 1255–1263. doi:10.1111/geb.12210
Fu, Y.H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G., Ciais, P., Song, Y., Vitasse, Y., Zeng, Z., Janssens, I. a, Huang, M., Menzel, A., Pen, J., 2015. Declining global warming effects on the phenology of spring leaf unfolding. Nature 526, 104–107. doi:10.1038/nature15402
Gauzere, J., Delzon, S., Davi, H., Bonhomme, M., Garcia de Cortazar-Atauri, I., Chuine, I., 2017. Integrating interactive effects of chilling and photoperiod in phenological process-

based models. A case study with two European tree species: Fagus sylvatica and Quercus petraea. Agric. For. Meteorol. 244–245, 9–20. doi:10.1016/j.agrformet.2017.05.011
Guo, L., Dai, J., Wang, M., Xu, J., Luedeling, E., 2015. Responses of spring phenology in temperate zone trees to climate warming: A case study of apricot flowering in China. Agric. For. Meteorol. 201, 1–7. doi:10.1016/j.agrformet.2014.10.016
Hanninen, H., 1987. Effects of temperature on dormancy release in woody plants: implications of prevailing models. Silva Fenn.
Higgins, S.I., O'Hara, R.B., Römermann, C., 2012. A niche for biology in species distribution models. J. Biogeogr. 39, 2091–2095. doi:10.1111/jbi.12029
Hogg, E.H., Price, D.T., Black, T.A., 2000. Postulated feedbacks of deciduous forest phenology on seasonal climate patterns in the Western Canadian interior. J. Clim. 13, 4229–4243. doi:10.1175/1520-0442(2000)013<4229:PFODFP>2.0.CO;2
Howe, G.T., Saruul, P., Davis, J., Chen, T.H.H., 2000. Quantitative genetics of bud phenology, frost damage, and winter survival in an F 2 family of hybrid poplars. TAG Theor. Appl. Genet. 101, 632–642. doi:10.1007/s001220051525
Janssen, P.H.M., Heuberger, P.S.C., 1995. Calibration of process-oriented models. Ecol. Modell. 83, 55–66. doi:10.1016/0304-3800(95)00084-9
Jones, H.G., Hillis, R.M., Gordon, S.L., Brennan, R.M., 2013. An approach to the determination of winter chill requirements for different Ribes cultivars. Plant Biol. 15, 18–27. doi:10.1111/j.1438-8677.2012.00590.x
Kearney, M.R., Wintle, B.A., Porter, W.P., 2010. Correlative and mechanistic models of species distribution provide congruent forecasts under climate change. Conserv. Lett. 3, 203–213. doi:10.1111/j.1755-263X.2010.00097.x
Kollas, C., Randin, C.F., Vitasse, Y., Körner, C., 2014. Agricultural and Forest Meteorology How accurately can minimum temperatures at the cold limits of tree species be extrapolated from weather station data ? Agric. For. Meteorol. 184, 257–266. doi:10.1016/j.agrformet.2013.10.001
Körner, C., 1999. Alpine plant life: functional plant ecology of high mountain ecosystems, Alpine plant life functional plant ecology of high mountain ecosystems. doi:10.1659/0276-4741(2001)021[0202:APLFPE]2.0.CO;2
Lang, G., Early, J., Martin, G., Darnell, R., 1987. Endo-, para-, and ecodormancy: physiological terminology and classification for dormancy research. HortScience.
Laube, J., Sparks, T.H., Estrella, N., Höfler, J., Ankerst, D.P., Menzel, A., 2014. Chilling outweighs photoperiod in preventing precocious spring development. Glob. Chang. Biol. 20, 170–182. doi:10.1111/gcb.12360
Lebourgeois, F., Pierrat, J.C., Perez, V., Piedallu, C., Cecchini, S., Ulrich, E., 2010. Simulating phenological shifts in French temperate forests under two climatic change scenarios and four driving global circulation models. Int. J. Biometeorol. 54, 563–581. doi:10.1007/s00484-010-0305-5
Lindstrom, M., Bates, D., 1990. Nonlinear mixed effects models for repeated measures data. Biometrics. doi:10.2307/2532087
Linkosalo, T., Lappalainen, H.K., Hari, P., 2008. A comparison of phenological models of leaf bud burst and flowering of boreal trees using independent observations. Tree Physiol. 28, 1873–1882. doi:10.1093/treephys/28.12.1873
Lobell, D.B., Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate change. Agric. For. Meteorol. 150, 1443–1452. doi:10.1016/j.agrformet.2010.07.008

Luedeling, E., Zhang, M., McGranahan, G., Leslie, C., 2009. Validation of winter chill models using historic records of walnut phenology. Agric. For. Meteorol. 149, 1854–1864. doi:10.1016/j.agrformet.2009.06.013

Malagi, G., Sachet, M.R., Citadin, I., Herter, F.G., Bonhomme, M., Regnard, J.L., Legave, J.M., 2015. The comparison of dormancy dynamics in apple trees grown under temperate and mild winter climates imposes a renewal of classical approaches. Trees - Struct. Funct. 29, 1365–1380. doi:10.1007/s00468-015-1214-3

Menzel, A., Sparks, T.H., Estrella, N., Koch, E., Aaasa, A., Ahas, R., Alm-Kübler, K., Bissolli, P., Braslavská, O., Briede, A., Chmielewski, F.M., Crepinsek, Z., Curnel, Y., Dahl, Å., Defila, C., Donnelly, A., Filella, Y., Jatczak, K., Måge, F., Mestre, A., Nordli, Ø., Peñuelas, J., Pirinen, P., Remišová, V., Scheifinger, H., Striz, M., Susnik, A., Van Vliet, A.J.H., Wielgolaski, F.E., Zach, S., Zust, A., 2006. European phenological response to climate change matches the warming pattern. Glob. Chang. Biol. 12, 1969–1976. doi:10.1111/j.1365-2486.2006.01193.x

Morin, X., Thuiller, W., 2009. Comparing niche- and process-based models to reduce prediction uncertainty in species range shifts under climate change. Ecology 90, 1301–1313. doi:10.1890/08-0134.1

Mountain Research Initiative EDW Working Group, 2015. Elevation-dependent warming in mountain regions of the world. Nat. Clim. Chang. 5, 424–430. doi:10.1038/nclimate2563\rhttp://www.nature.com/nclimate/journal/v5/n5/abs/nclimate2563.html#supplementary-information

Nagano, A.J., Sato, Y., Mihara, M., Antonio, B.A., Motoyama, R., Itoh, H., Nagamura, Y., Izawa, T., 2012. Deciphering and prediction of transcriptome dynamics under fluctuating field conditions. Cell 151, 1358–1369. doi:10.1016/j.cell.2012.10.048

Nash, J.E., Sutcliffe, J. V., 1970. River flow forecasting through conceptual models part I - A discussion of principles. J. Hydrol. 10, 282–290. doi:10.1016/0022-1694(70)90255-6

Nissanka, S.P., Karunaratne, A.S., Perera, R., Weerakoon, W.M.W., Thorburn, P.J., Wallach, D., 2015. Calibration of the phenology sub-model of APSIM-Oryza: Going beyond goodness of fit. Environ. Model. Softw. 70, 128–137. doi:10.1016/j.envsoft.2015.04.007

Olesen, J.E., Trnka, M., Kersebaum, K.C., Skjelvåg, A.O., Seguin, B., Peltonen-Sainio, P., Rossi, F., Kozyra, J., Micale, F., 2011. Impacts and adaptation of European crop production systems to climate change. Eur. J. Agron. 34, 96–112. doi:10.1016/j.eja.2010.11.003

Parmesan, C., Yohe, G., 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421, 37–42. doi:10.1038/nature01286

Petri, J.L., Leite, G.B., 2004. Consequences of insufficient winter chilling on apple tree bud-break, in: Acta Horticulturae. pp. 53–60. doi:10.17660/ActaHortic.2004.662.4

Pinheiro, J.C., Bates, D.M., 1996. Unconstrained parametrizations for variance-covariance matrices. Stat. Comput. 6, 289–296. doi:10.1007/BF00140873

R Core Team, 2016. R: A language and environment for statistical computing. Version 3.3.2. R Found. Stat. Comput. Vienna, Austria. doi:10.1017/CBO9781107415324.004

Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M., Guisan, A., 2006. Are niche-based species distribution models transferable in space? J. Biogeogr. 33, 1689–1703. doi:10.1111/j.1365-2699.2006.01466.x

Rebetez, M., Reinhard, M., 2008. Monthly air temperature trends in Switzerland 1901-2000 and 1975-2004. Theor. Appl. Climatol. 91, 27–34. doi:10.1007/s00704-007-0296-2

Richardson, A.D., Anderson, R.S., Arain, M.A., Barr, A.G., Bohrer, G., Chen, G., Chen, J.M., Ciais, P., Davis, K.J., Desai, A.R., Dietze, M.C., Dragoni, D., Garrity, S.R.,

957 Gough, C.M., Grant, R., Hollinger, D.Y., Margolis, H.A., Mccaughey, H., Migliavacca,
958 M., Monson, R.K., Munger, J.W., Poulter, B., Raczka, B.M., Ricciuto, D.M., Sahoo,
959 A.K., Schaefer, K., Tian, H., Vargas, R., Verbeeck, H., Xiao, J., Xue, Y., 2012.
960 Terrestrial biosphere models need better representation of vegetation phenology: Results
961 from the North American Carbon Program Site Synthesis. Glob. Chang. Biol. 18, 566–
962 584. doi:10.1111/j.1365-2486.2011.02562.x
963 Satake, A., Kawagoe, T., Saburi, Y., Chiba, Y., Sakurai, G., Kudoh, H., 2013. Forecasting
964 flowering phenology under climate warming by modelling the regulatory dynamics of
965 flowering-time genes. Nat. Commun. 4. doi:10.1038/ncomms3303
966 Schaber, J., Badeck, F.W., 2003. Physiology-based phenology models for forest tree species
967 in Germany. Int. J. Biometeorol. 47, 193–201. doi:10.1007/s00484-003-0171-5
968 Sen, P.K., 1968. Estimates of the Regression Coefficient Based on Kendall's Tau. J. Am.
969 Stat. Assoc. 63, 1379–1389. doi:10.1080/01621459.1968.10480934
970 Vitasse, Y., 2013. Ontogenic changes rather than difference in temperature cause understory
971 trees to leaf out earlier. New Phytol. 198, 149–155. doi:10.1111/nph.12130
972 Vitasse, Y., François, C., Delpierre, N., Dufrêne, E., Kremer, A., Chuine, I., Delzon, S., 2011.
973 Assessing the effects of climate change on the phenology of European temperate trees.
974 Agric. For. Meteorol. 151, 969–980. doi:10.1016/j.agrformet.2011.03.003
975 Vitasse, Y., Rebetez, M., Filippa, G., Cremonese, E., Klein, G., Rixen, C., 2016. `Hearing'
976 alpine plants growing after snowmelt: ultrasonic snow sensors provide long-term series
977 of alpine plant phenology. Int. J. Biometeorol. 1–13. doi:10.1007/s00484-016-1216-x
978 Vitasse, Y., Signarbieux, C., Fu, Y.H., 2017. Global warming leads to more uniform spring
979 phenology across elevations. Proc. Natl. Acad. Sci. . doi:10.1073/pnas.1717342115
980 Walther, G.R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin,
981 J.M., Hoegh-Guldberg, O., Bairlein, F., 2002. Ecological responses to recent climate
982 change. Nature 416, 389–395. doi:10.1038/416389a
983 Wang, E., Engel, T., 1998. Simulation of phenological development of wheat crops. Agric.
984 Syst. 58, 1–24. doi:10.1016/S0308-521X(98)00028-6
985 Zguigal, A., Chahbar, A., Loudiyi, D.E.M.W., Crabbé, J., 2006. Caractéristiques de la
986 dormance des bourgeons du pommier dans les régions à hiver doux. Biotechnol. Agron.
987 Soc. Environ. 10, 131–137.
988 Zohner, C.M., Benito, B.M., Svenning, J.C., Renner, S.S., 2016. Day length unlikely to
989 constrain climate-driven shifts in leaf-out times of northern woody plants. Nat. Clim.
990 Chang. 6, 1120–1123. doi:10.1038/nclimate3138
991 Zohner, C.M., Renner, S.S., 2015. Perception of photoperiod in individual buds of mature
992 trees regulates leaf-out. New Phytol. 208, 1023–1030. doi:10.1111/nph.13510
993 Zohner, C.M., Renner, S.S., 2014. Common garden comparison of the leaf-out phenology of
994 woody species from different native climates, combined with herbarium records,
995 forecasts long-term change. Ecol. Lett. 17, 1016–1025. doi:10.1111/ele.12308
996

Appendix A

Phenoclim protocol

Phenoclim is a citizen science program initiated in 2004 (www.phenoclim.org) by CREA (Centre de Recherches sur les Ecosystèmes d'Altitude, Chamonix, France), which aims at assessing the long-term effects of climate change on plant phenology over the French Alps. The specificity of this program, compared to other existing citizen science initiatives, lies not only in its large geographic coverage in mountain environments, but also in its simultaneous acquisition of accurate temperature records in addition to phenological observations.

In the Phenoclim protocol, budburst is defined as the first day at which 10% of vegetative buds of the crown on a given individual are opened (BBCH 07). At each site, observers survey at least two individuals per species. These individuals are chosen to be adult and dominant trees taller than 7 m, sharing similar environmental conditions (i.e. soil, slope and aspect) and at a maximum horizontal distance of 100 m from each other. Phenological observations were recorded weekly. For each species, yearly budburst dates were calculated as the mean date of individuals observed at a given site. The five surveyed species do not occur simultaneously at all sites, therefore, the number of years surveyed varies for a given site (Fig.1). In total, and irrespective of species, 242 sites for budburst were surveyed between 2007 and 2014.

Appendix B

Linear mixed models of budburst dates as response variables and temperature-based factors as competing predicting variables. Models are calibrated both on observation sites equipped with temperature stations and on all observation sites. Adjusted R-squared (R^2), AIC, and *P*-value of ANOVA are provided for each model. Var-part= partitions of variance between growing degree-days > 0°C (GGD0), growing degree-days > 5°C (GGD5), and chilling. For models with both GDD and chilling, two P values are provided, one for each effect.

		Budburst												
		With temperature stations							All sites					
		N	AIC	R^2	*P-value (anova)*	*Var-part GDD0*	*Var-part chilling*	*Var-part - both*	N	AIC	R^2	*P-value (anova)*	*Var-part GDD0*	*Var-part chilling*	*Var-part - both*
C. avellana		126							319						
	GDD0		929	0.55	***					2377	0.62	***			
	GDD5		940	0.51	***					2404	0.59	***			
	Chilling		967	0.26	*					2502	0.38	***			
	GDD0+Chilling		**921**	0.54	***/NS	0.30	0.01	0.11		**2373**	0.62	***/NS	0.28	0.0	0.18
	GDD5+Chilling		932	0.52	***/NS	0.24	0.01	0.12		2392	0.60	***/**	0.23	0.0	0.18
F. excelsior		152							396						
	GDD0		1132	0.65	***					2949	0.70	***			
	GDD5		1133	0.65	***					2958	0.70	***			
	Chilling		1174	0.50	*					3106	0.53	***			
	GDD0+Chilling		**1126**	0.64	***/NS	0.24	0.0	0.22		**2944**	0.70	***/NS	0.23	0.0	0.33
	GDD5+Chilling		1127	0.65	***/NS	0.23	0.0	0.22		2953	0.70	***/NS	0.22	0.0	0.33
B. pendula		145							307						
	GDD0		1052	0.72	***					2224	0.76	***			
	GDD5		1049	0.73	***					2227	0.76	***			
	Chilling		1120	0.54	*					2393	0.55	***			
	GDD0+Chilling		1047	0.72	***/NS	0.34	0.0	0.2		**2220**	0.76	***/NS	0.32	0.0	0.23
	GDD5+Chilling		**1044**	0.73	***/NS	0.35	0.1	0.3		2222	0.76	***/NS	0.33	0.0	0.23
L. decidua		155							303						
	GDD0		1113	0.68	***					2174	0.74	***			
	GDD5		1109	0.69	***					2180	0.72	***			
	Chilling		1179	0.42	*					2367	0.43	**			
	GDD0+Chilling		1109	0.68	***/NS	0.29	0.0	0.12		**2171**	0.74	***/NS	0.36	0.0	0.13
	GDD5+Chilling		**1105**	0.69	***/NS	0.29	0.0	0.12		2176	0.72	***/NS	0.34	0.0	0.13
P. abies		107							248						
	GDD0		752	0.79	***					1757	0.81	***			
	GDD5		743	0.81	***					1738	0.84	***			
	Chilling		833	0.44	**					1949	0.43	***			
	GDD0+Chilling		745	0.78	***/NS	0.50	0.01	0.15		1752	0.81	***/NS	0.38	0.0	0.28
	GDD5+Chilling		**733**	0.81	***/*	0.54	0.02	0.14		**1725**	0.84	***/**	0.41	0.01	0.27

Appendix C

Parameter estimates obtained for each model when calibrated on observation sites with weather stations and performances indices. *N*: number of observations; *k*: number of parameters. N number of data; k number of parameters; SStot total variance (sum of squares), SSres residual variance (sum of squares); RMSE Root mean squared error; EFF efficiency; AICc Corrected Akaike's Information Criteria.

		1-phase model			**2-phase model**	
Species		GDD	Sigmoïd	Wang	Threshold lower/sigmoïd	Wang/sigmoïd
C. avellana	*Parameters*	$t_0 = 1$ $F^* = 21$ $T_b = 7.12$	$t_0 = 57$ $F^* = 8$ $d_T = -0.45$ $T_{50} = 8.44$	$t_0 = 51$ $F^* = 6$ $T_{opt} = 27.95$ $T_{min} = -2.94$ $T_{max} = 41.13$	$t_0 = -115$ $C^* = 145$ $F^* = 11$ $T_h = 16$ $d_T = -0.38$ $T_{50} = 8.61$	$t_0 = -119$ $C^* = 150$ $F^* = 4$ $T_{opt} = 5.07$ $T_{min} = -4.47$ $T_{max} = 22.07$ $d_T = -1.71$ $T_{50} = 7.91$
	Calibration performances					
	N	128	128	128	128	128
	k	3	4	5	6	8
	SStot	22262	19743	19743	19743	19743
	SSres	16559	9815	9796	8564	7470
	RMSE	10.91	8.76	8.75	8.18	7.64
	EFF	0.26	0.5	0.5	0.57	0.62
	AICc	626.51	563.80	565.72	550.72	537.74
F. excelsior	*Parameters*	$t_0 = 1$ $F^* = 121$ $T_b = 5.68$	$t_0 = 54.7$ $F^* = 17$ $d_T = -0.22$ $T_{50} = 12.25$	$t_0 = 55$ $F^* = 23$ $T_{opt} = 16.2$ $T_{min} = -38.8$ $T_{max} = 20.8$	$t_0 = -77$ $C^* = 125$ $F^* = 9$ $T_h = 10.6$ $d_T = -0.2$ $T_{50} = 16.2$	$t_0 = -78$ $C^* = 125$ $F^* = 9$ $T_{opt} = 0.1$ $T_{min} = -36.6$ $T_{max} = 21.3$ $d_T = -0.24$ $T_{50} = 10.41$
	Calibration performances					
	N	149	149	149	149	149
	k	3	4	5	6	8
	SStot	34545	31275	31275	31275	31275
	SSres	19787	12559	12650	11958	12430
	RMSE	10.98	9.22	9.21	8.96	9.13
	EFF	0.43	0.59	0.60	0.62	0.60
	AICc	732.52	670.37	672.20	665.98	676.19
B. pendula	*Parameters*	$t_0 = 1$ $F^* = 72$ $T_b = 4.9$	$t_0 = 60$ $F^* = 1$ $d_T = -0.2$ $T_{50} = 24.5$	$t_0 = 62$ $F^* = 9$ $T_{opt} = 17.4$ $T_{min} = -37.5$ $T_{max} = 21.5$	$t_0 = -101$ $C^* = 127$ $F^* = 5$ $T_h = 12.2$ $d_T = -0.2$ $T_{50} = 16.97$	$t_0 = -108$ $C^* = 157$ $F^* = 5$ $T_{opt} = 3.4$ $T_{min} = -9.9$ $T_{max} = 26.2$ $d_T = -0.2$ $T_{50} = 14.7$
	Calibration performances					
	N	146	146	146	146	146
	k	3	4	5	6	8
	SStot	30844	30313	30313	30313	30313
	SSres	14743	8425	8440	8298	7720
	RMSE	9.69	7.6	7.6	7.5	7.27
	EFF	0.52	0.72	0.72	0.73	0.75
	AICc	677.86	600.36	601.89	602.46	596.38

Species		1-phase model			2-phase model	
		GDD	Sigmoïd	Wang	Threshold lower/sigmoïd	Wang/sigmoïd
L. decidua	*Parameters*	t_0 = 1 F^* = 3 T_b = 9.6	t_0 = 83 F^* = 6 d_T = -37 T_{50} = 6.2	t_0 = 83 F^* = 6 T_{opt} = 10.1 T_{min} = 5.9 T_{max} = 37.7	t_0 = -100.8 C^* = 184 F^* = 6 T_h = 18.3 d_T = -37.4 T_{50} = 6.2	t_0 = 66 C^* = 134 F^* = 6 T_{opt} = 9 T_{min} = -38.9 T_{max} = 34.2 d_T = -7.1 T_{50} = 6.2
	Calibration performances					
	N	159	159	159	159	159
	k	3	4	5	6	8
	SStot	27728	26275	26275	26275	26275
	SSres	54044	11328	11240	11297	10911
	RMSE	17.83	8.44	8.41	8.43	8.28
	EFF	-0.95	0.57	0.57	0.57	0.58
	AICc	930.83	686.58	687.47	690.43	689.31
P. abies	*Parameters*	t_0 = 1 F^* = 290 T_b = 2.96	t_0 = 83 F^* = 5 d_T = -0.1 T_{50} = 24.2	t_0 = 84 F^* = 8 T_{opt} = 29.9 T_{min} = -36.3 T_{max} = 37.6	t_0 = -100 C^* = 174 F^* = 8 T_h = 13.9 d_T = -0.2 T_{50} = 18.1	t_0 = -107 C^* = 176 F^* = 3 T_{opt} = 0.01 T_{min} = -39.8 T_{max} = 24.9 d_T = -0.2 T_{50} = 27.6
	Calibration performances					
	N	112	112	112	112	112
	k	3	4	5	6	8
	SStot	22158	21453	21453	21453	21453
	SSres	10623	4989	4984	4660	4753
	RMSE	9.29	6.67	6.67	6.45	6.51
	EFF	0.52	0.77	0.77	0.78	0.78
	AICc	513.97	433.58	435.66	430.36	437.18

Appendix D

Model performance in predicting the budburst date in the calibration conditions. Meteo sites refers to observation sites with meteorological stations; West sites refers to observations sites West of the species-specific longitudinal threshold; East sites refers to observation sites East of the species-specific longitudinal threshold. Correlative 1 refers to the linear mixed model with GDD only. Correlative 2 refers to the linear model with GDD and chilling as variables. PB 1-phase uses a sigmoid function of temperature for the ecodormancy phase. PB 2-phase uses a lower threshold function of temperature for the endodormancy phase and a sigmoid function of temperature for the ecodormancy phase. PB calibrated refers to the 2-phase model which parameter estimates have been derived from experimental results. AICc, corrected Akaike criterion; EFF, model efficiency; RMSE, roots-mean-squared error (days); RMSEs, systematic root mean square error (days); RPIQ, ratio of performance to interquartile distance; N, number of data used to fit the model.

Calibration sites	Meteo sites						West sites						East sites					
	AICc	EFF	RMSE	RMSEs	RPIQ	N	AICc	EFF	RMSE	RMSEs	RPIQ	N	AICc	EFF	RMSE	RMSEs	RPIQ	N
Betula																		
Correlative 1	964.1	0.79	6.62	6.58	2.87	145	1036.9	0.87	6.23	6.19	3.61	159	974.4	0.72	6.45	6.37	2.48	148
Correlative 2	964.1	0.80	6.62	6.58	2.87	145	1037.0	0.87	6.23	6.19	3.61	159	969.8	0.73	6.35	6.27	2.52	148
PB 1-phase	999.7	0.74	7.46	7.44	2.55	145	1180.1	0.68	9.72	9.71	2.32	159	1004.2	0.66	7.07	7.05	2.26	148
PB 2-phase	995.4	0.75	7.36	7.33	2.58	145	1173.1	0.70	9.53	9.49	2.36	159	1003.8	0.66	7.06	7.04	2.27	148
Corylus																		
Correlative 1	882.5	0.59	7.89	7.83	1.94	126	1067.8	0.70	7.69	7.60	2.47	154	1149.1	0.68	7.81	7.72	2.18	165
Correlative 2	879.9	0.60	7.81	7.75	1.99	126	1064.0	0.71	7.60	7.51	2.50	154	1148.8	0.68	7.81	7.72	2.18	165
PB 1-phase	902.2	0.51	8.58	8.47	1.81	126	1141.0	0.52	9.80	9.64	1.94	154	1219.0	0.52	9.58	9.55	1.77	165
PB 2-phase	886	0.58	7.98	7.94	1.94	126	1129.8	0.57	9.31	9.29	2.04	154	1212.2	0.55	9.37	9.35	1.81	165
Fraxinus																		
Correlative 1	1047.7	0.73	7.5	7.44	2.56	152	1129.8	0.82	6.55	6.47	3.36	171	1300.2	0.75	8.22	8.14	2.80	184
Correlative 2	1045.9	0.73	7.46	7.4	2.57	152	1134.6	0.82	6.63	6.56	3.32	171	1299.2	0.75	8.20	8.12	2.81	184
PB 1-phase	1137.7	0.51	10.12	10	1.90	152	1281.1	0.42	11.83	10.06	1.86	171	1366.8	0.49	11.66	9.76	1.97	184
PB 2-phase	1130	0.54	9.82	9.76	1.96	152	1280.6	0.41	11.88	10.05	1.85	171	1379.5	0.46	11.98	10.10	1.92	184
Larix																		
Correlative 1	1016.3	0.74	6.48	6.43	2.78	154	923.2	0.77	5.89	5.84	2.93	144	1055.1	0.80	6.75	6.69	3.07	158
Correlative 2	1016.1	0.74	6.47	6.42	2.78	154	921.0	0.77	5.85	5.80	2.95	144	1056.6	0.80	6.78	6.72	3.06	158
PB 1-phase	1088.4	0.58	8.16	8.12	2.20	154	1019.3	0.56	8.16	8.16	2.11	144	1109.4	0.72	8.00	7.94	2.59	158
PB 2-phase	1088.1	0.59	8.15	8.12	2.21	154	1018.4	0.56	8.15	8.13	2.12	144	1105.5	0.73	7.90	7.85	2.63	158
PB calibrated	1127.5	0.38	9.95	9.22	1.81	154	1042.7	0.39	9.57	8.85	1.80	144	1149.3	0.65	9.04	9.01	2.30	158
Picea																		
Correlative 1	684.7	0.83	5.79	5.76	2.77	107	788.5	0.83	5.69	5.67	3.86	124	777.3	0.87	5.47	5.42	3.66	124
Correlative 2	681.4	0.83	5.7	5.68	2.81	107	785.3	0.86	5.62	5.60	3.92	124	768.5	0.88	5.27	5.23	3.79	124
PB 1-phase	702	0.8	6.29	6.25	2.54	107	837.6	0.78	6.93	6.91	3.17	124	894.6	0.68	8.71	8.70	2.29	124
PB 2-phase	695.6	0.82	6.08	6.06	2.63	107	831.2	0.79	6.76	6.74	3.25	124	890.7	0.69	8.57	8.57	2.33	124

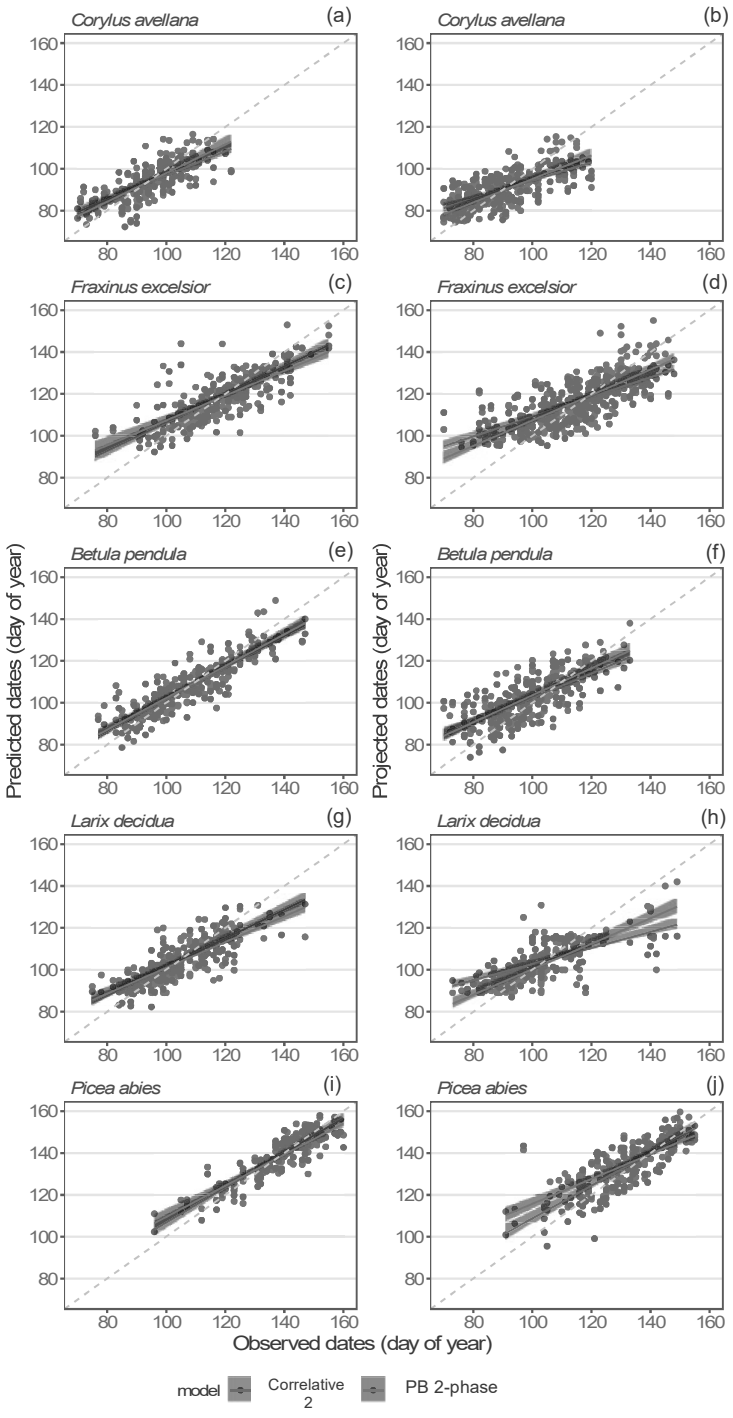
Appendix E

Model performance in projecting the budburst date in external conditions. Model and site labels as in appendix D. No meteo sites 1 refers to observation sites without meteorological stations. R^2, adjusted R-squared; RMSE, roots-mean-squared error (days); RMSEs, systematic root mean square error (days); Diff. RMSE, difference between RMSE of the calibration (shown in appendix D) and RMSE of the projection; RPIQ, ratio of performance to interquartile distance; N, number of data used to fit the model.

Calibration sites	Meteo stations						West sites						East sites					
Projection sites	No meteo station						East sites						West sites					
	R^2	RMSE	RMSEs	RPIQ	Diff RMSE	N	R^2	RMSE	RMSEs	RPIQ	Diff RMSE	N	R^2	RMSE	RMSEs	RPIQ	Diff RMSE	N
Betula																		
Correlative 1	0.46	11.23	10.70	1.73	4.24	194	0.47	10.49	10.03	1.82	2.55	165	0.44	10.82	10.57	1.86	5.63	154
Correlative 2	0.44	11.44	10.94	1.73	4.24	193	0.47	10.60	10.05	1.82	2.56	165	0.44	10.81	10.56	1.85	5.81	154
PB 1-phase	0.50	10.45	10.32	1.85	2.70	193	0.47	10.51	10.12	1.93	-1.42	165	0.52	9.97	9.81	2.32	2.65	154
PB 2-phase	0.51	10.46	10.25	1.75	3.33	193	0.42	11.09	10.48	1.86	-0.91	165	0.57	9.31	9.30	2.36	2.46	154
Corylus																		
Correlative 1	0.61	10.90	10.33	1.60	3.34	244	0.61	10.98	10.01	1.62	2.80	199	0.53	10.91	10.68	1.76	3.01	197
Correlative 2	0.60	11.03	10.42	1.57	3.63	244	0.60	10.91	10.08	1.60	3.00	199	0.53	10.93	10.71	1.76	3.00	197
PB 1-phase	0.58	10.66	10.62	1.72	1.87	244	0.64	10.37	9.54	1.62	0.71	199	0.53	12.89	10.73	1.91	0.39	197
PB 2-phase	0.59	10.77	10.59	1.72	2.48	244	0.60	11.78	10.10	1.53	1.78	199	0.53	11.36	10.75	2.04	-0.06	197
Fraxinus																		
Correlative 1	0.48	10.86	10.19	2.13	3.40	162	0.50	8.78	8.53	2.09	4.43	148	0.56	12.08	11.47	2.11	2.69	159
Correlative 2	0.48	10.86	10.19	2.11	3.57	162	0.50	8.79	8.53	2.11	4.28	148	0.56	12.16	11.50	2.10	2.73	159
PB 1-phase	0.53	10.16	9.70	2.18	0.54	162	0.62	8.30	7.44	2.22	-1.46	148	0.68	9.72	9.68	1.78	1.23	159
PB 2-phase	0.55	10.69	9.55	2.16	0.95	162	0.56	8.62	8.03	1.95	-0.10	148	0.70	9.52	9.48	2.03	-0.62	159
Larix																		
Correlative 1	0.55	10.66	10.05	1.74	4.18	148	0.54	11.07	10.31	1.87	5.18	158	0.40	9.85	9.49	1.75	3.10	144
Correlative 2	0.55	10.69	10.09	1.73	4.22	148	0.52	10.05	10.55	1.85	4.20	158	0.40	10.00	9.52	1.72	3.22	144
PB 1-phase	0.62	9.28	9.21	1.99	1.12	148	0.69	8.77	8.41	2.37	0.61	158	0.55	8.18	8.18	2.11	0.18	144
PB 2-phase	0.62	9.26	9.21	1.99	1.11	148	0.69	8.75	8.40	2.37	0.60	158	0.56	8.15	8.14	2.12	0.25	144
PB calibrated	0.67	8.68	8.64	2.13	-1.27	148	0.64	9.14	9.12	2.27	-0.43	158	0.48	9.57	8.85	1.80	0.53	144
Picea																		
Correlative 1	0.66	9.01	8.81	2.44	3.22	140	0.66	9.29	8.91	2.15	3.60	124	0.72	8.87	7.86	2.48	3.40	124
Correlative 2	0.66	9.01	8.80	2.44	3.31	140	0.67	9.23	8.85	2.17	3.61	124	0.73	8.47	7.72	2.60	3.20	124
PB 1-phase	0.65	9.10	8.98	2.41	2.81	140	0.68	9.16	8.69	2.18	2.23	124	0.77	7.61	7.05	2.89	-1.10	124
PB 2-phase	0.63	9.63	9.27	2.28	3.55	140	0.69	8.70	8.69	2.30	1.94	124	0.79	6.75	6.73	3.26	-1.82	124

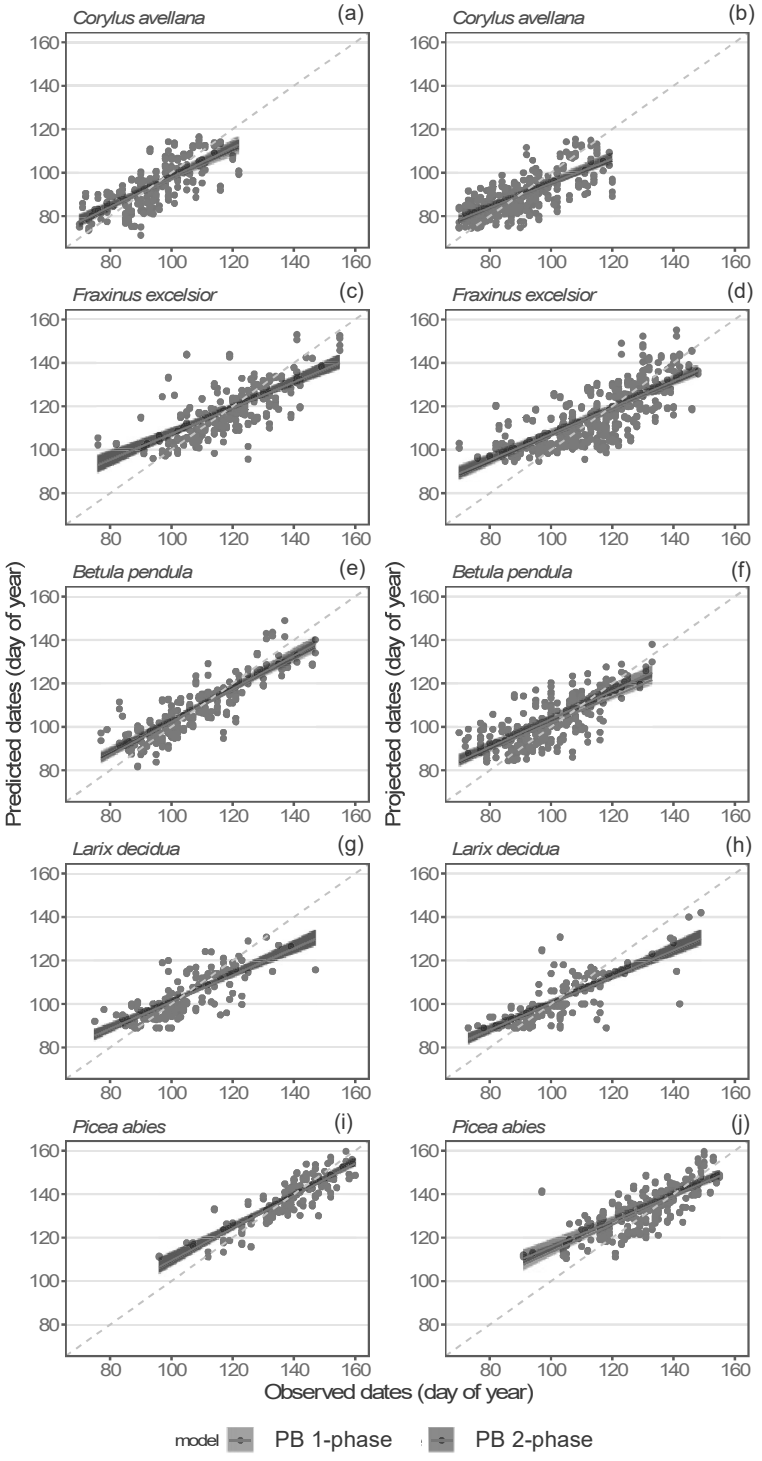
Appendix F

Comparison of predicted (a, c, e, g, i) and projected (b, d, f, h, j) budburst dates either by the correlative model (with GDD and chilling as variables) (blue dots) or the process-based 2-phase model (red dots) with observed budburst dates for each species. Linear regressions of predicted/projected dates on observed dates (superimposed lines) are shown for correlative model (blue) and process 2-phase model (red). Light grey area= ± .95 confidence interval of dates predicted/projected by the correlative model. Dark grey area= ± .95 confidence interval of dates predicted/projected by the 2-phase model.



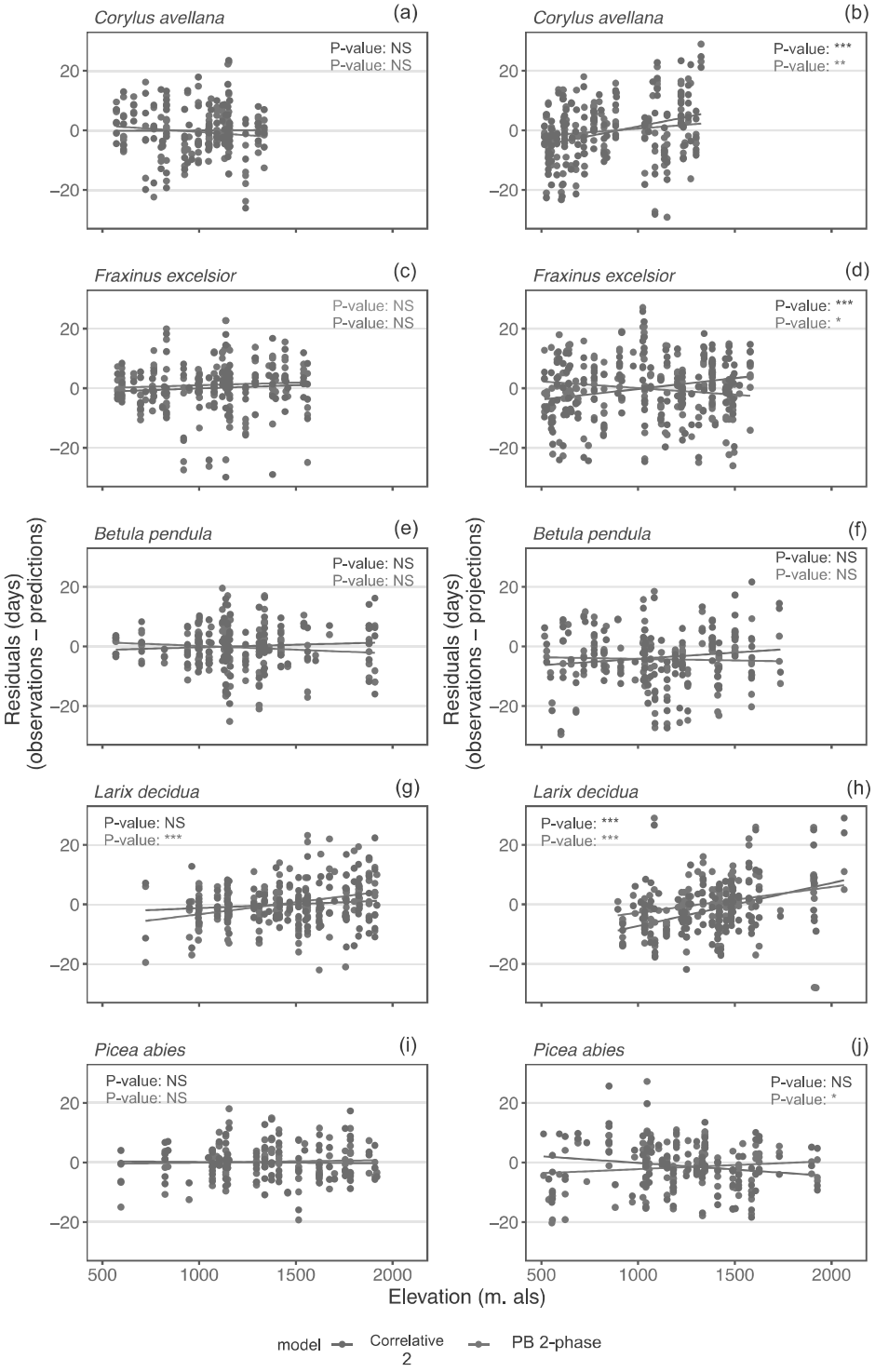
Appendix G

Comparison of predicted (a, c, e, g, i) and projected (b, d, f, h, j) budburst dates either by the process 1-phase model (grey dots) or the process 2-phase model (red dots) with observed budburst dates for each species. Linear regression of predicted/projected budburst dates on observed dates (superimposed lines) are shown for process 1-phase model (grey) and process 2-phase model (red). Light grey area= ± .95 confidence interval of dates predicted/projected by the 1-phase model. Dark grey area= ± .95 confidence interval of dates predicted/projected by the 2-phase model.



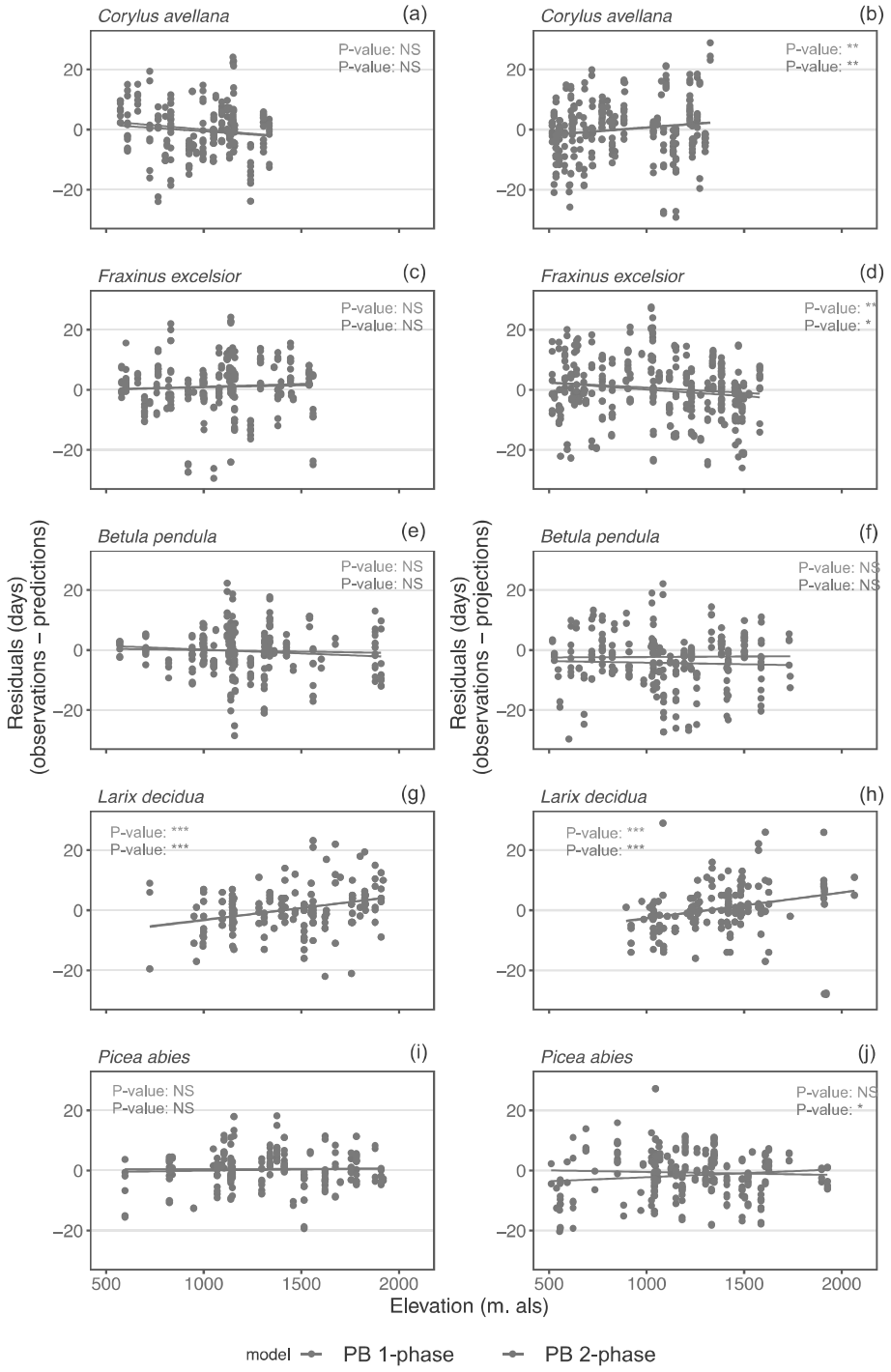
Appendix H

Variation of the residuals of correlative models (with GDD and chilling as variables) (blue dots) and process-based 2-phase models (red dots) calculated either with the calibration data subset1 (a, c, e, g, i) or the validation data subset 2 (b, d, f, h, j) along the elevation gradient. Superimposed lines are linear regression of residuals on elevation (same color as symbols).



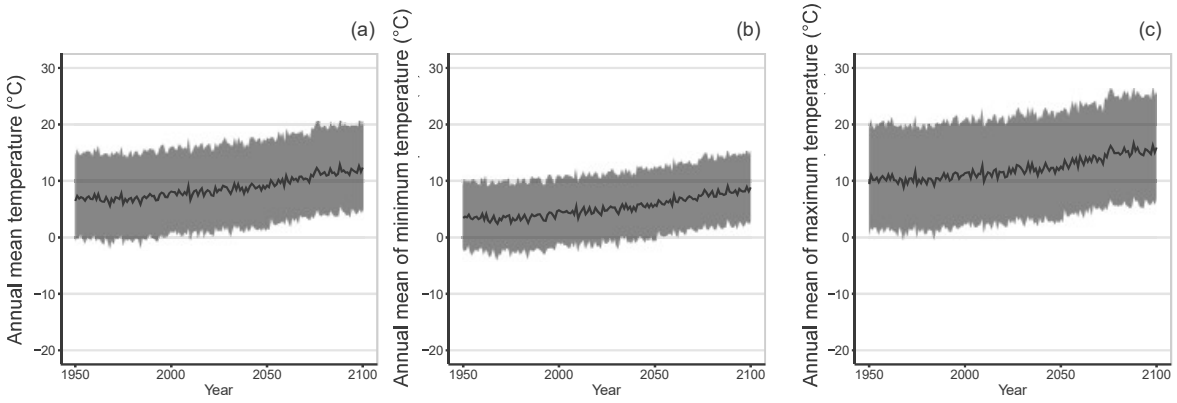
Appendix I

Variation of the residuals of process-based 1-phase models (grey dots) and process-based 2-phase models (red dots) calculated either with the calibration data subset1 (a, c, e, g, i) or the validation data subset 2 (b, d, f, h, j) along the elevation gradient. Superimposed lines are linear regression of residuals on elevation (same color as symbols).



Appendix J

Evolution and uncertainties of the climatic data generated by the ALADIN-Climat v5 RCM model (CNRM) for the CMIP5 experiment at a 12-km resolution, downscaled at 8-km resolution using quantile-quantile method (http://drias-climat.fr/). Daily minimum and maximum temperatures of RPC8.5 scenarios were extracted on the grid points of a rectangle area covering French Alps (43°48'N to 46°47'N, 5°59'E to 7°09'E). Black curve indicate the mean of annual mean temperature over the region (a), the mean of the annual average minimum temperature over the region (b), the mean of annual average maximum temperature over the region (d), and dark grey areas indicate the standard deviation.



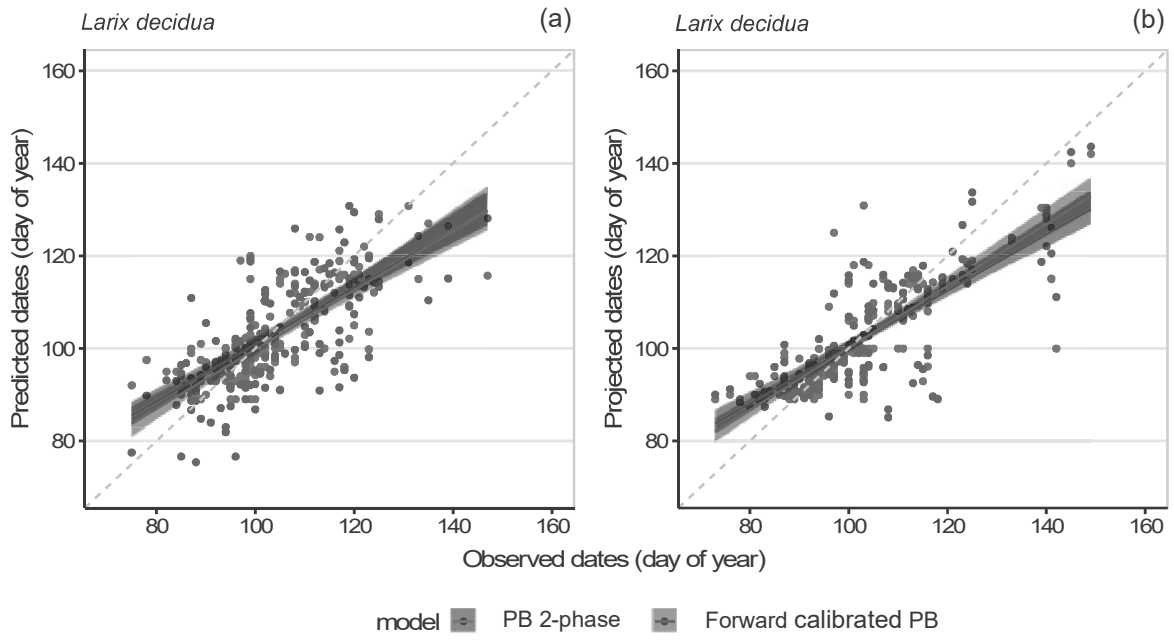
Appendix K

Mean annual shift of the budburst date projected over the periods 1950-2100, 1950-2050 and 2050-2100 in the French Alps at low elevation and high elevation. Trend values indicate the Sen's slope per year and trendp values indicate the Sen's slope over the time period. Future climate data are according to scenario RCP8.5, in the Alps. Model labels as in appendix D. The elevation limits depend on species (Fig. 1 for elevation ranges).

		Low elevation						High elevation					
Species	Model	1950-2100		1950-2050		2050-2100		1950-2100		1950-2050		2050-2100	
		trend	trendp	trend	trendp	trend	trendp	trend	trendp	trend	trendp	trend	trendp
C. avellana	*Correlative*	-0.15	-22.90	-0.08	-7.50	-0.27	-13.90	-0.08	-12.00	-0.03	-3.20	-0.20	-10.40
	Process-based 1-phase	-0.13	-19.30	-0.10	-9.30	-0.17	-8.70	-0.26	-37.80	-0.17	-16.80	-0.37	-18.50
	Process-based 2-phase	-0.02	-2.50	-0.05	-4.60	0.08	3.90	-0.23	-34.40	-0.17	-16.50	-0.29	-14.50
F. excelsior	*Correlative*	-0.16	-24.30	-0.10	-10.10	-0.27	-13.80	-0.12	-17.70	-0.07	-6.60	-0.25	-12.60
	Process-based 1-phase	-0.16	-23.10	-0.12	-11.30	-0.20	-10.00	-0.23	-34.60	-0.16	-15.80	-0.32	-16.20
	Process-based 2-phase	-0.05	-7.90	-0.07	-6.80	0.020	0.80	-0.21	-31.80	-0.16	-15.30	-0.27	-13.60
B. pendula	*Correlative*	-0.25	-37.70	-0.15	-14.60	-0.45	-23.10	-0.10	-15.20	-0.05	-4.50	-0.28	-14.50
	Process-based 1-phase	-0.14	-20.10	-0.10	-9.50	-0.19	-9.40	-0.23	-34.00	-0.16	-15.50	-0.33	-16.50
	Process-based 2-phase	-0.04	-5.70	-0.05	-4.60	-0.01	-0.50	-0.22	-32.60	-0.16	-15.60	-0.28	-13.80
L. decidua	*Correlative*	-0.24	-35.70	-0.13	-13.10	-0.47	-24.20	-0.02	-3.60	-0.01	-0.50	-0.13	-6.90
	Process-based 1-phase	-0.09	-13.00	-0.09	-9.30	-0.03	-1.30	-0.31	-46.20	-0.20	-19.70	-0.42	-20.90
	Process-based 2-phase	-0.02	-3.30	-0.07	-7.10	0.16	8.20	-0.31	-46.60	-0.21	-20.10	-0.41	-20.80
	Forward calibrated PB	-0.05	-8.10	-0.01	-1.10	-0.02	-1.10	-0.29	-43.40	-0.20	-19.60	-0.38	-19.10
P. abies	*Correlative*	-0.27	-40.60	-0.16	-16.30	-0.48	-24.40	-0.08	-12.70	-0.03	-3.00	-0.25	-12.80
	Process-based 1-phase	-0.14	-21.20	-0.11	-11.10	-0.19	-9.70	-0.25	-37.30	-0.17	-17.00	-0.32	-16.00
	Process-based 2-phase	-0.03	-3.80	-0.05	-5.30	0.10	5.10	-0.26	-38.80	-0.18	-18.00	-0.32	-15.80

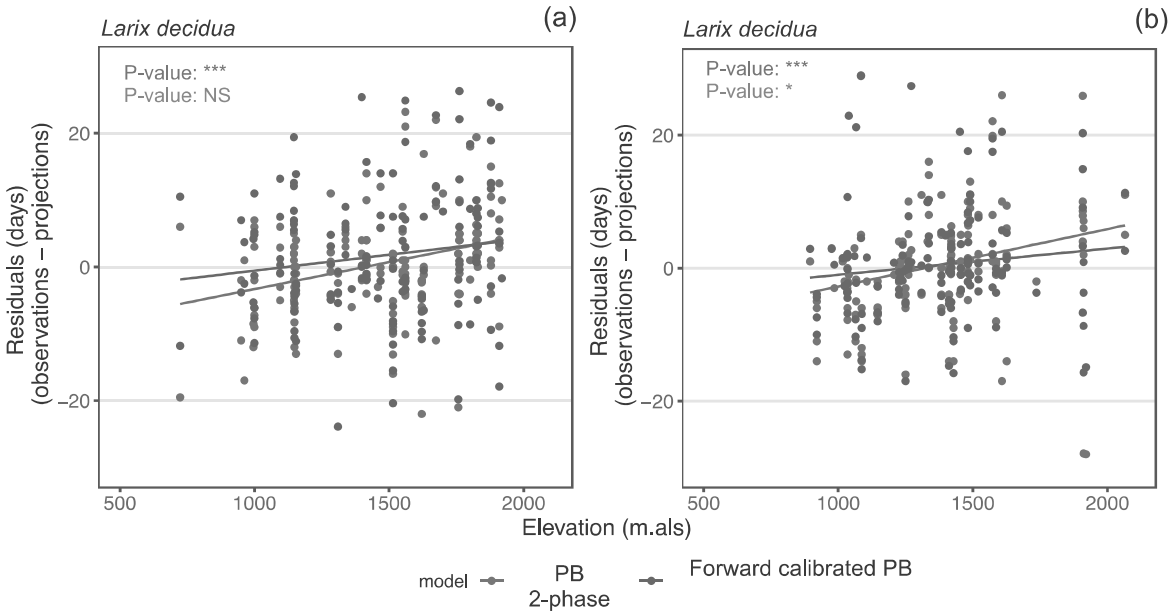
Appendix L

Comparison of the budburst dates of *Larix decidua* predicted (a) and projected (b) by the process-based 2-phase model (red dots) and the forward calibrated process-based 2-phase model (brown dots) with the observed budburst dates. Linear regression of predicted/projected budburst dates on observed dates (superimposed lines) are shown for the 2-phase model (red) and the 2-phase calibrated model (brown). Red area= ± .95 confidence interval of dates predicted/projected by the 2-phase model. Orange area= ± .95 confidence interval of dates predicted/projected by the calibrated 2-phase model.



Appendix M

Variation of the residuals of process-based 2-phase models (red dots) and forward calibrated process-based 2-phase models (brown dots) calculated either with the calibration data subset1 (a, c, e, g, i) or the validation data subset 2 (b, d, f, h, j) along the elevation gradient. Superimposed lines are linear regression of residuals on elevation (same color as symbols).



Declaration of Interest Statement

Declaration of interests

- ✓ ☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

~~☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:~~
