ORIGINAL ARTICLE

Jörg Schaber · Franz-W. Badeck

Physiology-based phenology models for forest tree species in Germany

Abstract Models of phenology are needed for the projection of effects of a changing climate on, for example, forest production, species competition, vegetation–atmosphere feedback and public health. A new phenology model for deciduous tree bud burst is developed and parameters are determined for a wide geographical range (Germany) and several forest tree species. The new model is based on considerations of simple interactions between inhibitory and promotory agents that are assumed to control the developmental status of a plant. Several alternative model structures were formulated emphasizing different hypothetical physiological processes. The new models fitted the observations better than classical models. The bias of the classical models, i.e. overestimation of early observations and underestimation of late observations, could be reduced but not completely removed. Differences in the best-fitting model equations for each species indicated that, for the late spring phases (bud burst of *Fagus sylvatica* and *Quercus robur*), the photoperiod played a more dominant role than for early spring phases (bud burst of *Betula pendula* and *Aesculus hippocastanum*). Chilling only plays a subordinate role for spring bud burst compared to temperatures preceding this event in our data. The presented modeling approach allowed for a species-specific weighting of the dominant processes. The model results are in accordance with experimental findings that indicate an important role of day length in late spring BB. Potentials for model improvement are discussed.

J. Schaber (✉) · F.-W. Badeck
Potsdam Institute for Climate Impact Research,
P.O. Box 601203, 14473 Potsdam, Germany,
e-mail: schaber@pik-potsdam.de
Tel.: +34-96-354-3666
Fax: +34-96-354-3670

Introduction

The length of the vegetation period timed by the phenological events bud burst and leaf coloring determines, to a large part, the annual carbon balance of deciduous trees (Kramer 1994; Goulden et al. 1996; Chen et al. 1999; Kramer et al. 2000). Goulden et al. (1996) and Chen et al. (1999) showed that it is the day of leaf emergence rather than the day of leaf coloring that controls annual carbon sequestration in forests. This is mainly because, at the time of leaf emergence, days are usually fairly long and temperatures high, so that a one-day advancement in spring adds more to the annual carbon balance than an extra day in autumn where days are short and temperatures low. Moreover, models for leaf senescence that offer a better prediction than some fixed average date do not exist, yet. This is mainly because the factors influencing leaf senescence have not yet been unequivocally detected. Temperature, for example has no clear effect (Schaber 2002). For these reasons phenological models concentrate on simulating the day of bud burst rather than leaf coloring or leaf fall.

Several phenological bud burst models have been published for tree species (Cannell and Smith 1983; Hänninen 1994; Kramer 1994, 1995; Linkosalo et al. 2000). All these models, however, have so far been applied either to data sets of limited size for single species (Cannell and Smith 1983; Hannerz 1999), small data sets from a geographically limited region (Kramer 1994, 1995; Murray et al. 1989) or to clones (Diekmann 1996; Menzel 1997). When the phenology of a wider geographical region, including several species and provenances, is to be modeled to simulate the sensitivity of regional carbon budgets to climate change or validate satellite observations, the limitations of these models with their parameterizations need to be tested.

The purpose of this study was to find a model that is able to simulate bud burst of several deciduous tree species over a wide geographical region where several provenances for each species can be expected, i.e. to model the phenology that is actually being observed either on the ground or from satellites.

Models, data and methods

Phenological models

There have been attempts to base phenological modeling on physiological aspects (Hänninen 1994) rather than empirical temperature sums (Hunter and Lechowicz 1992; Worrall 1993; Diekmann 1996; Hannerz 1999) or observed relationships between chilling and temperature sums (Cannel and Smith 1983; Murray et al. 1989; Hunter and Lechowicz 1992). Hänninen (1994) concentrates on the distinct phases of ontogenetic development of tree buds that have been identified by many authors (Samish 1954; Smith and Kefford 1964; Vegis 1964) and their relation to external factors, mainly temperature and photoperiod.

The nature of these different phases and their physiological and morphological status is not yet fully understood. However, most investigators agree that the regulation of these phases and their transitions is controlled by the actions and interactions of growth-promoting compounds or hormones and growth-inhibiting compounds (Samish 1954; Wareing and Saunders 1971; Hanover 1980; Powell 1987). Thus, it is the hypothesized balance or ratio between promotory and inhibitory agents that determines the physiological state of development of the plant and its reaction to external driving forces. Several candidates for such substances have been isolated. However, the nature of their interaction remains unclear and experimental evidence is often contradictory (Lavender and Silim 1987). Despite the lack of knowledge about the details of hormonal control of tree phenology, the first attempt was made to put the agreed first principles of the hormonal aspects of phenology into a simple conceptual framework that can be easily translated into a mathematical model: the abundance or concentration of enzymes is determined by the rates of synthesis and breakdown. It is known that temperature and photoperiod play a prominent role in controlling these processes. Temperature, for instance, can act through purely physical mechanisms, as when it influences viscosity and diffusion. Moreover, synthesis of proteins is maximal in an optimal temperature range and declines at higher and lower temperatures (Vegis 1973; Johnson and Thornley 1985). Photoperiod has been observed to be the driving force of a trigger acting through the phytochrome system (Wareing 1956; Nitsch 1957; Perry 1971; Heide 1993a, b). From these simple but basic principles a model for the abundance or concentration of an inhibitory compound I and a promotory compound P is hypothesized and formulated as a system of two simple coupled difference equations:

$$\Delta I = a_1 f_1(T) g_1(L) - a_2 f_2(T) g_2(L) I$$
$$\Delta P = a_3 f_3(T) g_3(L)(1 - I) - a_4 f_4(T) g_4(L) P \tag{1}$$

where the a_i are scaling parameters and the f_i and g_i, $i = 1, …, 4$, are functions of temperature T and photoperiod (day length) L, respectively. Temperature T and photoperiod L are themselves functions of time, in our case of the day of the year. Breakdown of the compounds P and I, indicated by the negative terms in Eq. 1, is assumed to be a first-order reaction, whereas the synthesis terms, indicated by the positive terms in Eq. 1, are formulated as simple forcing terms. The synthesis term of the promoter P is damped by the presence of the inhibitor I. P and I are normalized arbitrarily to vary between one and zero.

To obtain an applicable model, assumptions about the parameters, functions, and boundary and initial conditions have to be made: we let the model start at t_0, the observed date of leaf coloring of the previous year. For reasons of simplicity, on this date rest is assumed to be deepest. That means we assume the maximal concentration of inhibitory substances and minimal abundance of promotory compounds, setting $I(t_0) = 1$ and $P(t_0) = 0$, i.e. with the current family of models the initial deepening of rest is not simulated. Bud burst takes place when P reaches the threshold $P = 1$.

Following the observation that almost all physiological processes have a minimum, optimal and maximum temperature, the temperature dependence of the functions f_i is modeled as a triangular function (Hänninen 1994) where

$$f_i(T) = \begin{cases} \frac{T - T_{\min}}{T_{opt} - T_{\min}} & \text{for } T_{\min} \leq T \leq T_{opt} \\ \frac{T_{\max} - T}{T_{\max} - T_{opt}} & \text{for } T_{opt} \leq T \leq T_{\max} \\ 0 & \text{for } T \leq T_{\min} \text{ or } T \geq T_{\max} \end{cases}$$

for $i = 1, …, 4$ (Eq. 1), T being the daily mean temperature and $T_{\min}$, T_{opt}, $T_{\max}$ fixed temperature thresholds that are different for each f_i.

The factors can be modified by multiplication by a function dependent on day length, L, i.e. either $L/24$ or $(24 - L)/24$, depending on whether the process is assumed to be promoted by long days or long nights. Inhibitor breakdown is assumed to be accelerated by long days, whereas build-up of an inhibitor is assumed to be supported by long nights. The contrary is assumed for the promoter. Synthesis of the promoter is favored by long days and its breakdown is accelerated by long nights (Wareing 1956; Nitsch 1957; Perry 1971).

To keep the number of parameters as low as possible, some restrictions were introduced in the general framework of Eq. 1. The conditions $f_1(T) = f_4(T) = 1$ were imposed, i.e. the forcing term for the inhibitor and the breakdown term for the promoter only depend on photoperiod. Further, either a_1 or a_4 had to be zero. This resulted in 12 possible models each one having nine parameters. They are listed in Table 1 and are called promoter–inhibitor models (PIM) in the following.

Table 1 Promoter–inhibitor models derived from the general framework in Eq. 1 and the imposed restrictions described in the paragraph above. ΔI, ΔP rates of change of the inhibitor I and promoter P. f_I, f_P temperature-dependent triangular functions according to Eq.1 for the inhibitor I and promoter P, respectively, each having three parameters $T_{\min}$, T_{opt}, $T_{\max}$. a_i, $i = 1, …, 4$ are scaling parameters. $I(t_0) = 1$, $P(t_0) = 0$, time step: 1 day. $t_0 =$ day of leaf coloring the previous year. Bud burst: $P > 1$

Model no.		Model specification
1	$\Delta I = -a_2 f_2(T) I$	$\Delta P = a_3 f_3(T)(1 - I) - a_4 P$
2	$\Delta I = -a_2 f_2(T) I$	$\Delta P = a_3 f_3(T)(1 - I) - a_4 P(24 - L)/24$
3	$\Delta I = a_2(24 - l)/24 - a_2 f_2(T) I$	$\Delta P = a_3 f_3(T)(1 - I)$
4	$\Delta I = -a_2 f_2(T) I$	$\Delta P = a_3 f_3(T)(1 - I) L/24 - a_4 P$
5	$\Delta I = -a_2 f_2(T) I$	$\Delta P = a_3 f_3(T)(1 - I) L/24 - a_4 P(24 - L)/24$
6	$\Delta I = a_1(24 - l)/24 - a_2 f_2(T) I$	$\Delta P = a_3 f_3(T)(1 - I) L/24$
7	$\Delta I = -a_2 f_2(T) I L/24$	$\Delta P = a_3 f_3(T)(1 - I) - a_4 P$
8	$\Delta I = -a_2 f_2(T) I L/24$	$\Delta P = a_3 f_3(T)(1 - I) - a_4 P(24 - L)/24$
9	$\Delta I = a_1(24 - l)/24 - a_2 f_2(T) I L/24$	$\Delta P = a_3 f_3(T)(1 - I)$
10	$\Delta I = -a_2 f_2(T) I L/24$	$\Delta P = a_3 f_3(T)(1 - I) L/24 - a_4 P$
11	$\Delta I = -a_2 f_2(T) I L/24$	$\Delta P = a_3 f_3(T)(1 - I) L/24 - a_4 P(24 - L)/24$
12	$\Delta I = a_1(24 - l)/24 - a_2 f_2(T) I L/24$	$\Delta P = a_3 f_3(T)(1 - I) L/24$

For each species the model formulation was determined that best suited the observed dates of bud burst and the corresponding course of temperature and day length.

Three models from the literature that have been successfully applied to different data sets were also applied, i.e. the sequential model (SM) (Hänninen 1994; Kramer 1994), the Cannell and Smith model (CSM) (Cannell and Smith 1983) with the modification introduced by Menzel (1997) and the classical linear temperature-sum model (TSM) (Wang 1960; Robertson 1968; Kramer 1994; Menzel 1997). The same model equations as in the above publications were used. The sequential model is also based on physiological considerations (Hänninen 1994) and was found to offer the best fit among several other models of beech bud burst (Kramer 1994; Kramer et al. 2000). The SM has seven parameters. The CSM model empirically models the observation that increased chilling in winter decreases the required temperature sum for bud burst in spring. The CSM has four parameters. The linear temperature-sum model (TSM) simply integrates temperatures above a certain threshold from a defined date up to a fixed critical value, which defines the TSM as a three-parameter model. For a full documentation of the models see the original papers cited or Schaber (2002).

All models are also compared to the null model, which is simply the mean date of bud burst.

Data

The phenological data used for this study come from the phenological network of the German Weather Service. For the purpose of modeling tree leaf emergence in relation to weather, phenological stations were selected where corresponding weather measurements were available. Weather data were also provided by the German Weather Service.

Combined phenological time series (Häkkinen et al. 1995; Linkosalo et al. 1996; Schaber and Badeck 2002) were calculated for all phenological stations available in the vicinity of a weather station (dots in circles in Fig. 1). Assuming that the weather station measurements are characteristic for a certain region, it is assumed that the corresponding characteristic phenology is obtained when the time series from at least five phenological stations are combined, the stations not being further than 10 km away and not differing by more than 50 m in elevation from the weather station. Phenological observations at stations with the latter characteristics were found to be highly correlated (Hense et al. 2002).

With the above criteria a data set of 81 weather stations and 495 phenological stations was obtained (Fig. 1). The combined time series were obtained with the procedure recommended in Schaber and Badeck (2002) with outlier removal according to the 30-day rule (Schaber and Badeck 2002).

The model fits were done for those deciduous tree species for which a high number of observations of bud burst were available. Only those observations were taken into account that had a corresponding observation of leaf coloring in the previous year because the calculations for the promoter–inhibitor model were assumed to start on this date. The exception was *Larix decidua*, for which calculations were allowed to start on 1 November because there were not many leaf coloring dates available. The resulting combined time series had a total number of observations over all years and stations of

– 2,475 data points for horse chestnut (*Aesculus hippocastanum* L.),
– 1,937 data points for birch (*Betula pendula* Roth.)
– 1,536 data points for beech (*Fagus sylvatica* L.)
– 1,811 data points for oak (*Quercus robur* L.)
– 1,441 data points for larch (*Larix decidua* Mill.)

Fig. 1 Phenological and weather stations in Germany that were used for this study. *Dots* The locations of the phenological stations. Climate stations are situated in the center of the circles. The circles have a radius of 10 km

Fitting phenological models

The phenological models (except the null model) were fitted to the observations by minimizing the sum of squared residuals

$$f(x) = \sum_i r_i^2(x) \qquad (2)$$

in the parameter space $x = x_1, \ldots, x_n$ (n parameters), where the residuals $r_i(x) = m_i(x) - o_i$, $m_i(x)$ is the day of bud burst according to the model with its respective parameter x, and o_i is the observed day of bud burst in year i.

The phenological models (except the null model) were fitted to the combined time series by a simulated annealing algorithm (Metropolis et al. 1953; Ingber 1989). This algorithm has previously been successfully applied to phenological data (Chuine et al. 1988). Here we used the implementation of Ingber (1989), which is available on the internet (http://www.ingber.com), and parallelized the code with a message-passing interface (http://www.mpi-forum.org ; Gropp et al. 1994). The available data were split into two sets of equal size for each species. This was done by splitting each station series into even and uneven years. One data set was used to fit the model, the other set was used for independent testing. Convergence of the algorithm with respect to the parameters was characterized by their standard error over ten runs with different initial parameters.

The difference of the averages over the ten optimization runs was tested against the critical value of least significant difference (Sokal and Rohlf 1995; Sachs 1997) to see which model results did significantly ($P < 0.05$) differ.

The quality of the model fits was determined by four different measures. The first is the sum of squared residuals (SSR), i.e. $f(x)$ in Eq. 2, which is used as the optimization criterion. A more descriptive version of the SSR is the root mean square error

(RMSE). The second is the mean absolute error (MAE) and the third is the r^2 value of the linear regression of the simulated against the observed values. The fourth measure of model performance is a function of the slope of the regression of the simulated against the observed values. We defined *bias = 1 – slope of the regression* as a measure of the trend in the residuals of the model fit.

Results

Parallelization significantly speeded up the optimization process depending on the number of processors used. For the optimization of the PIM for ten runs with different initial parameters for, *A. hippocastanum*, for example, the computation time could be decreased from more than 2 weeks to 3 days using 1 and 40 processors, respectively.

In Table 2 the minimum of the SSR over the ten optimization runs for the different models is displayed.

The variation of the SSR of the different repetitions was small (less than 2%, results not shown, for details see Schaber 2002). The CSM was most sensitive to initial conditions.

Some model fits were not significantly different with respect to their mean SSR over ten runs. In Table 3 the models where the average SSR over ten optimization runs was not significantly different and which, at the same time, were in the group with the lowest SSR are shown per species. For all species, the models PIM 4, 5, 8, 10, and 11 were always in the group of best-fitting models. For each species the model with the lowest average SSR is not necessarily the model with the lowest overall SSR (results not shown, for details see Schaber 2002). For *F. sylvatica*, for example, the model with the lowest mean SSR is PIM 4; the lowest overall SSR, however, was found among the ten optimization runs of PIM 11. For *L.*

Table 2 Optimization results. After the species name the number of observations used for model fitting is given in parenthese. An equal number were used for model testing. For the PIM the model with the lowest sum of squared residuals in a single optimization run (*Min SSR*) is given. r^2 and the bias of a regression of observed against simulated values are given. The root-mean-square error (*RMSE*) and mean absolute error (*MAE*) of the fits are rounded to days

Species	Model	Fitting					Testing				
		Min SSR	r^2	Bias	RMSE	MAE	Min SSR	r^2	Bias	RMSE	MAE
A. hippocastanum (1238)	PIM 8	36,477	0.695	0.3	5	4	35,312	0.708	0.278	5	4
	SM	70,320	0.416	0.558	8	6	68,440	0.434	0.548	7	6
	CSM	41,654	0.654	0.313	6	4	40,275	0.671	0.29	6	4
	TSM	42,586	0.648	0.305	6	5	43,396	0.648	0.304	6	5
	NULL	119,645	–	–	10	8	12,0762	–	–	10	8
B. pendula (987)	PIM 2	24,036	0.722	0.27	5	4	31,726	0.657	0.326	6	4
	SM	58,960	0.335	0.583	8	6	66,701	0.298	0.621	8	7
	CSM	25,906	0.7	0.32	5	4	30,580	0.668	0.356	6	4
	TSM	30,543	0.646	0.345	7	6	36,229	0.605	0.389	6	5
	NULL	86,093	–	–	9	7	91,813	–	–	10	8
F. sylvatica (786)	PIM 11	21,583	0.486	0.504	5	4	20,485	0.514	0.477	5	4
	SM	36,836	0.149	0.792	7	5	35,457	0.172	0.781	7	5
	CSM	26,063	0.386	0.562	6	4	25,315	0.407	0.54	6	4
	TSM	26,101	0.379	0.630	6	4	24,050	0.433	0.599	6	4
	NULL	41,956	–	–	7	6	42,155	–	–	7	6
Q. robur (906)	PIM 12	27,824	0.613	0.346	6	4	27,304	0.618	0.35	5	4
	SM	51,025	0.313	0.59	8	6	50,392	0.323	0.576	7	6
	CSM	28,994	0.593	0.426	6	4	28,231	0.603	0.414	6	4
	TSM	30,291	0.576	0.401	6	4	30,600	0.569	0.412	6	4
	NULL	71,105	–	–	9	7	70,934	–	–	9	7
L. decidua (721)	PIM 2	41,204	0.585	0.403	8	5	37,044	0.60	0.343	7	5
	SM	61,633	0.381	0.628	9	7	56,611	0.381	0.601	9	7
	CSM	41,541	0.583	0.405	8	5	37,420	0.594	0.36	7	5
	TSM	42,259	0.575	0.403	8	5	36,661	0.604	0.349	7	5
	NULL	99,176	–	–	12	8	91,384	–	–	11	9

Table 3 Models with lowest SSR but non-significant (*P <* 0.05) difference among SSR after ten optimizations with randomly selected initial parameter values. Ascending order of minimum SSR from the top

A. hippo-castanum	*B. pendula*	*F. sylvatica*	*Q. robur*	*L. decidua*
PIM 8	PIM 2	PIM 11	PIM 11	PIM 2
PIM 2	PIM 8	PIM 5	PIM 12	PIM 5
PIM 5	PIM 5	PIM 4	PIM 10	PIM 4
PIM 10	PIM 11	PIM 10	PIM 5	PIM 1
PIM 11	PIM 4	PIM 8	PIM 6	CSM
PIM 4	PIM 10	PIM 2	PIM 4	PIM 11
			PIM 8	PIM 8
				PIM 10

Table 4 Parameter values for the PIM models. For a description of the parameters see Eq. 1 and Table 1. *min*, *max* Allowed ranges for the parameters during model fits, *Min SSR* parameters for the PIM that performed best in terms of SSR for model fitting, *SE* standard error of the average parameters after ten optimization runs

| Model		f_I					f_P				
		T_{min}	T_{opt}	T_{max}	a_1	a_2	T_{min}	T_{opt}	T_{max}	a_3	a_4
Allowed range	Min	–25	–15	0	0	0	–20	0	5	0	0
	Max	10	20	35	1	1	15	40	45	1	1
A. hipp.	Min SSR	–24.2	–15	23.13		0.095405	–10.94	23.55	38.57	0.053592	0.044243
PIM 8	SE	1.01	0.49	0.36		0.002	0.3	1.83	1.71	0.002563	0.000837
B. pendula	Min SSR	–24.96	–10.0	15.05		0.030619	–7.03	21.8	25.35	0.064803	0.045432
PIM 2	SE	0.79	0.16	0.29		0.002	0.48	1.72	2.06	0.004	0.001
F. sylvatica	Min SSR	–10.34	–0.89	18.11		0.058326	–10.03	28.61	44.49	0.109494	0.039178
PIM 11	SE	0.38	0.32	0.97		0.002	0.25	1.24	1.77	0.004	0.000
Q. robur	Min SSR	–23.05	–0.3	16.91	0.010379	0.055149	3.46	34.55	34.55	0.331253	
PIM 12	SE	0.22	0.22	0.28	0.001	0.003	0.06	1.87	1.49	0.02	
L. decidua	Min SSR	–16.71	–13.60	34.98		0.005598	–3.24	18.38	37.02	0.112225	0.017149
PIM 2	SE	0.60	0.22	1.08		0.001	0.18	2.10	2.23	0.066568	0.001485

decidua the CSM is within the group with the smallest SSR, for the other species this group only consists of PIM. For the PIM, the minimum SSR and r^2 could result from different models for fitting and testing respectively. For *B. pendula* and *Q. robur* the model with the lowest SSR and lowest r^2 when tested against independent data was PIM 5, for *F. sylvatica* it was PIM 10, whereas when fitted to the data it was PIM 2, PIM 12 and PIM 11 for *B. pendula*, *Q. robur* and *F. sylvatica* respectively (Table 2). In these cases, however, the average SSR over ten optimization runs of these models was not significantly different (Table 3) and the r^2 values differed by 0.01. For *L. decidua* and *A. hippocastanum* it was the same PIM that found the lowest SSR for fitting and testing respectively. For the fitting procedure the PIM always converged to the lowest SSR and highest r^2 value compared to the classical models. When the models were tested with independent data, it was the TSM and the CSM for *L. decidua* and *B. pendula*, respectively, that had the lowest SSR and highest r^2 values. The PIM had the lowest bias for all species fitted and tested, except for *L. decidua*, where it was the TSM. Except for *F. sylvatica* and *L. decidua* the CSM performed better than the TSM. In terms of RMSE and MAE the PIM performed best for *A. hippocastanum, F. sylvatica* and *Q. robur*. For *B. pendula* and *L. decidua* the CSM and TSM performed as well as the PIM as far as RMSE and MAE were concerned. The SM always had the lowest performance for all measures but still improved on the null model.

In general, the PIM improved on the classical models (SM, CSM, TSM). In terms of r^2 for model testing, the maximal improvement – an increase of 11% explained variance – was achieved for *F. sylvatica* (Table 2). For the other species this increase ranged from 0 to 3% in the amount of explained variance for *B. pendula* and *A. hippocastanum* respectively. For model fitting, improvement of model performance for the PIM in terms of r^2 ranged from 1% for *B. pendula* to 10% for *F. sylvatica* relative to the best classical model (Table 2). In Fig. 2, plots of simulated against observed dates of bud burst are shown for the five species and best PIM models tested.

For the model results shown in Fig. 2 the respective parameter values and their allowed ranges during the optimization process are tabulated in Table 3. For the SM, CSM, TSM and the other PIM, the parameters and ranges can be obtained from the authors. In general, the standard error of the parameters over the ten optimization runs is small (Table 4, more parameter sets can be downloaded from Schaber 2002, where Table 5-4 provides more results for PIM and Tables A5-1 to A5-3 contain the results for the other models). For *Q. robur* the variation of the parameters of the triangular temperature dependence function f_I is higher than that for the other species. The limits of the parameters of f_I and f_P were adjusted after several test optimization runs such that the resulting optimal parameter set would not reach the limits of the allowed interval. Only for T_{opt} of f_I for *A. hippocastanum* do the parameters still reach the allowed limits after several readjustments. The resulting temperature levels were lower for f_I than for f_P for the allowed ranges as well as for the final values.

A sensitivity analysis for PIM that best fitted the data was conducted for the parameters with respect to the SSR for each species. The parameters of the temperature-dependent breakdown function f_I showed a low sensitivity for all models and species. The scaling parameter a_2 of f_I was slightly more sensitive as far as the resulting fit in terms of SSR is concerned, but still low. The parameters of the temperature-dependent synthesis function f_P showed a substantially higher sensitivity than the parameters of f_I. T_{opt} of f_P especially played a prominent role. T_{max} showed no marked influence on SSR in both f_I and f_P. The most sensitive were the f_P scaling parameters a_3 and a_4. However, sensitivity did not only depend on the model structure but also on the other parameters. This showed the substantial difference in sensitivity of a_3 and a_4 for PIM 2 between *A. hippocastanum* and *B. pendula* (for details see Schaber 2002).

Fig. 2 Simulated day of bud burst (*BB*) in Julian day of year (*DOY*) against observed values. *PIM* Promoter–inhibitor model. Results for the models that best reproduced an independent data set. Linear regression of simulated BB (*y*) as a function of observed BB (*x*). 1:1 line, indicating the perfect fit
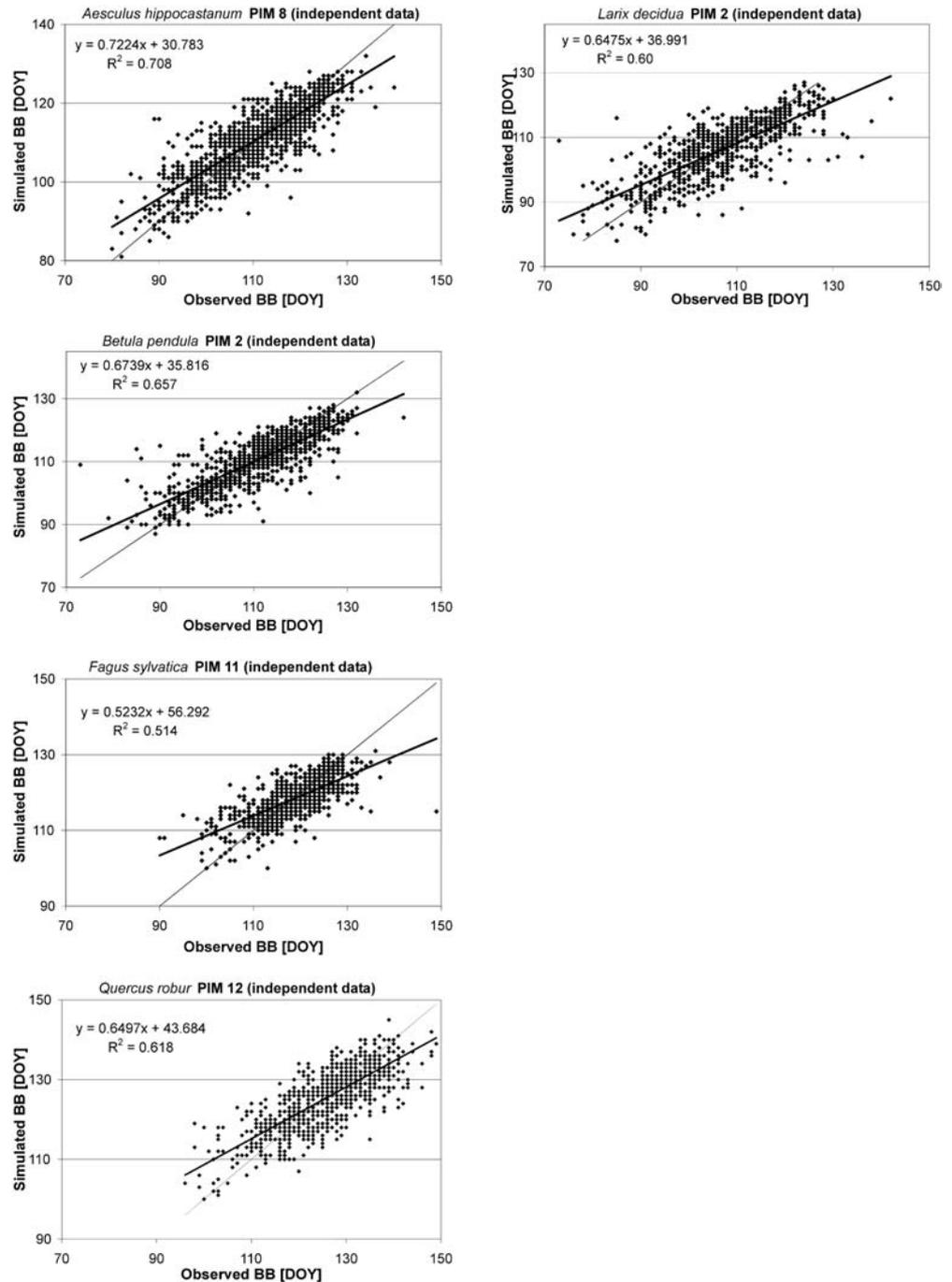
Discussion

The promoter–inhibitor model (PIM) has similarities to other physiology-based models that have been proposed before, specifically the so-called parallel and deepening-rest model (Hänninen 1994; Kramer 1994). Landsberg (1974) stated that, for dormancy release, it is essential that the response to forcing temperatures is taken into account, even when the critical state of chilling has not yet been attained. The potential of the buds to respond to forcing temperatures increases with the time spent under chilling conditions. Thus, chilling and the accumulation of heat units parallel each other, in contrast to the successive completion of chilling and forcing as prescribed in the sequential model (SM). The concept of decreased accumulation of forcing temperatures through incomplete chilling is also incorporated in the PIM by the dampening factor $(1 - I)$ in the forcing term for the promoter concentration P (Eq. 1, Table 1). The deepening-rest model (Kobayashi et al. 1982) stipulates that the state of chilling must increase before it can loosen its block on assimilation of heat units again. This notion is also allowed for in the PIM through the forcing term for the inhibitor concentration I, but it depends on the special

model structure whether increasing *I* is allowed and, if so, then increasing *I* depends on the day length. For the other synthesis and breakdown terms the PIM also consequently features the influence of day length, which has not been done in this form before. In particular, the breakdown of the promoter is a new feature in the PIM. This relates to a decrease of state of forcing that has not been allowed before in the classical models. However, it is interesting to note that, even the PIM were developed from another point of view, as in other phenological models, into involve hormonal aspects, this would lead to a similar mathematical description for certain processes which is only interpreted differently in the different approaches.

Hänninen (1994) has combined several mathematical formulations of classical processes that are agreed to play a role in bud burst phenology; he incorporated them into a common framework and came up with 96 different models. Two of them, the SM and the parallel model, have also been used in this study. There has been no attempt so far to test all these models thoroughly with a consistent data set. Hänninen (1994) has tested all of his proposed methods but used the parameter sets for the submodels as described in the literature. Kramer (1994) and Menzel (1997) fitted a limited number of these models to a limited data set. With the simulated annealing optimization algorithm in combination with parallel computing, as featured in this study, it now becomes possible not only to test but also to fit all available phenology models to extensive data sets with a reasonable effort in modeling and computation time. A complete testing of all available models promises new insights into the important processes determining the phenology of trees.

The fact that several model structures yielded similar or not significantly different results when fitted to the data gives rise to the notion that the model might be overparameterized or that some model elements or processes do not add to the explanatory power of the model. However, those models that were always among the best, i.e. PIM 4, 5, 8, 10, 11 and mostly also 2, had some interesting features in common (Table 1). Concerning the inhibitor part, those models exhibited only breakdown of the inhibitor, i.e. allowed only a continuous release of growth-arresting internal conditions. Consequently, the promoter part showed a contrasting reaction and included breakdown of promotory compounds. Because the models could only include either synthesis of inhibitor or breakdown of promoter, a direct comparison of the two processes was not possible. However, because breakdown of promoter was favored, the deepening of rest after partial rest release did not seem to play a dominant role. This suggests that the state of forcing can indeed decrease in spring. This means that, under unfavorable conditions in spring time, a once-attained physiological state of development is not only retained but set back and has to resume before it can develop further. PIM 1 and 7 included the promoter breakdown term but were not among the best-fitted models. These two models, however, did not include any day length modification

function, showing another important result: day length seems to be an important factor for spring phenology.

Even though this has been known for a long time and often proved by experiments, day length has only rarely been used in phenological models (Hänninen 1994; Menzel 1997). Partly because former modeling studies did not find any relevant effect of day length, temperature was thought to be sufficient. However, in this study the inclusion of day length did improve the models.

The temperature response functions adjusted to the expected levels. This was also partly due to the predefined model structure, but the fact that the breakdown function of inhibitor f_I had optimal temperatures below or around zero supports the chilling concept, meaning that low temperatures do enhance subsequent development.

There was a distinction of model structure among species. For the best-fitted models only the late-spring deciduous species, *F. sylvatica* and *Q. robur*, included the day length factor in their promoter synthesis term (PIM 11 and PIM 12). The early-spring and the needle-leafed species, *A. hippocastanum*, *B. pendula* and *L. decidua*, did not include the day length factor in the promoter synthesis term of the model that fitted the data best (PIM 8 and PIM 2). On the basis of these findings it can be speculated that the models support experimental evidence for late spring species day length playing a more pronounced role than it does for early species. The importance of the photoperiod for bud burst has also been shown in experiments with *F. sylvatica* (Wareing 1956; Heide 1993a, b). However, the distinction of model structures among species is weak because some model results were not significantly different.

One of the possible model refinements allowing a better discrimination between the different model structures can be derived from the results shown in Table 4. It has been mentioned above that the model might be overparameterized. The optimal temperature of the triangular response functions are, for most species, at a level that is close to the maximal temperatures observed, meaning that one half of the triangular functions becomes obsolete. For f_I the temperature allowing for the maximal rate of inhibitor breakdown T_{opt} is below -10 °C, except for *F. sylvatica* and *Q. robur* where it is 0.89 °C and -0.3 °C respectively. In the data used here, daily mean temperatures only seldom reach levels below -10 °C. The same is the case for T_{opt} of f_P. Here optimal temperature is always above 18 °C, which is a level that is only seldom reached in spring time in Germany. The sensitivity analysis supports the notion that the lower (upper) part of the triangular function f_I (f_P) is obsolete because T_{min} of f_I and T_{max} of f_P have no marked influence on model performance in the case of SSR. Thus, the parameters T_{min} of f_I and T_{max} of f_P can be dropped from the model without substantial impact on model performance. Thus, a two-parametric saw-tooth function or just a two-parametric linear function could have also been used. To obtain an optimal parameter set for only 7 parameters the models would have to be fitted again, which is an exercise for further studies.

Other functions could also be used. A triangular function might not be optimal. The triangular functions seem to adjust themselves to wide ranges in order to accomodate a broad range of optimal temperatures. A function with a broad range of optimal temperatures and a faster (nonlinear) decrease to zero outside the optimal range might be more appropriate here.

As mentioned before, the starting dates of the models are rather arbitrary. Other starting points, e.g. 1 January, could also improve the models. The starting date of the model can be hypothesized to have a minor effect on model performance because the chilling term is less sensitive than the forcing terms.

Even though the SM was found to be best in earlier simulation studies (Kramer 1994), it was outperformed by the classical models TS and CSM in this study. The convergence criteria for the SM were as good as for the other models, implying that this was not due to an optimization problem. However, Kramer (1994) allowed much broader ranges for the temperature functions. It also shows that a large number of parameters does not necessarily improve model results. Substantial improvement of bud burst simulation by the PIM compared to the classical models was achieved for *F. sylvatica*, implying that the TSM and the CSM lack one or more important processes for this species, e.g. inclusion of day length (Heide 1993a). Among the best-fitting models, only for *L. decidua* could a classical model (CSM) not be distinguished from other best-fitting models (Table 3). For the other modeled species both the TSM and the CSM perform satisfyingly well and, for the sake of parsimony, they seem to be sufficient for some applications.

For model simulations over longer periods, such as several decades, possibly in connection with tree growth models, there is one factor that might become important and favors the PIM again, i.e. the bias. The general notion of phenological models is that they tend to overestimate early observations and underestimate late observations. The models are not flexible enough to map extreme conditions and tend to be suited to simulated average observations. A measure for this systematic mistake is the introduced bias. For *B. pendula* the CSM was better than the PIM as far as SSR and r^2 are concerned. However, the CSM underestimated the observed values on day (DOY) 130 by 4 days and observed values on day 140 by 9 days on average whereas the PIM underestimated only by 5 days. For *F. sylvatica*, improvement of the bias also had significant effects. The day of bud burst in late spring has a substantial impact on the annual carbon balance of trees because of the long days and favorable temperatures. When incorporating models of phenology into forest growth models, a reduction in the bias of bud burst for late or early years by 2 or more days will have significant long term-effects on tree growth simulations. The fact that the bias could not be entirely removed, even with the better PIM, indicates that the models are still not flexible enough with respect to extreme weather conditions. The growth-promoting influence of high temperatures, for instance, might increase more strongly in time than is at present modeled by the influence of day length. Introduction of more non linearities into the model could further improve the bias.

References

Cannell MGR, Smith R (1983) Thermal time, chill days and prediction of budburst in *Picea sitchensis*. J Appl Ecol 20:951–963

Chen WJ, Black TA, Yang PC, Barr AG, Neumann HH, Nesic Z, Blanken PD, Novak MD, Eley J, Ketler RJ (1999) Effects of climatic variability on the annual carbon sequestration by a boreal aspen forest. Global Change Biol 4:41–53

Chuine I, Cour P, Rousseau DD (1988) Fitting models predicting dates of flowering of temperate-zone trees using simulated annealing. Plant Cell Environ 21:455–466

Diekmann M (1996) Relationship between phenology of perennial herbs and meteorological data in deciduous forests of Sweden. Can J Bot 74:528–537

Goulden ML, Munger JW, Fan S-M, Daube BC (1996) Exchange of carbon dioxide by a deciduous forest response to interannual climate variability. Science 271:1576–1578

Gropp W, Lusk E, Skjellum A (1994) Using MPI: portable parallel programming with the message-passing interface. MIT Press, Cambridge, Mass

Häkkinen R, Linkosalo T, Hari P (1995) Methods for combining phenological time series: application to bud burst in birch (*Betula pendula*) in Central Finland for the period 1896–1955. Tree Physiol 15:721–726

Häninnen H (1994) Effects of climatic change on trees from cool and temperate regions: an ecophysiological approach to modeling of bud burst phenology. Can J Bot 73:183–199

Hannerz M (1999) Evaluation of temperature models for predicting bud burst in Norway spruce. Can J For Res 29:9–19

Hanover JW (1980) Control of tree growth. Bioscience 30:756–762

Heide OM (1993a) Daylength and thermal time responses of budburst during dormancy release in some northern deciduous trees. Physiol Plant 88:531–540

Heide OM (1993b) Dormancy release in beech buds (*Fagus sylvatica* L.) requires both chilling and long days. Physiol Plant 89:187–191

Hense A, Glowienka-Hense R, Muller M, Braun P (2002) Spatial modelling of phenological observations to analyse their interannual variations in Germany. Agric For Meteorol 112:161–178

Hunter AF, Lechowicz MJ (1992) Predicting the timing of bud burst in temperate trees. J Appl Ecol 29:597–604

Ingber L (1989) Very fast simulated re-annealing. Math Comput Model 12:967–973

Johnson IR, Thornley JHM (1985) Temperature dependance of plant and crop processes. Ann Bot 55:7–24

Kobayashi K, Fuchigami L, English MJ (1982) Modeling temperature requirements for rest development in *Cornus sericea*. J Am Soc Hortic Sci 107:914–918

Kramer K (1994) Selecting a model to predict the onset of growth of *Fagus sylvatica*. J Appl Ecol 31:172–181

Kramer K (1995) Modelling comparison to evaluate the importance of phenology and spring frost damage for the effects of climate change on growth of mixed temperate-zone deciduous forsts. Clim Res 7:31–41

Kramer K, Leinonen I, Loustau D (2000) The importance of phenology for the evaluation of impact of climate change on growth of boreal, temperate and Mediterranean forests ecosystems: an overview. Int J Biometeorol 44:67–75

Landsberg JJ (1974) Apple fruit bud development and growth; analysis and an empirical model. Ann Bot 38:1013–1023

Lavender DP, Silim SN (1987) The role of plant growth regulators in dormancy in forest trees. Plant Growth Regul 6:171–191

Linkosalo T, Häkkinen R, Hari P (1996) Improving the reliability of a combined phenological time series by analysing observation quality. Tree Physiol 16:661–664

Linkosalo T, Carter TR, Häkkinen R, Hari P (2000) Predicting spring phenology and frost damage risk of *Betula* spp. under climatic warming: a comparison of two models. Tree Physiol 20:1175–1182

Menzel A (1997) Phänologie von Waldbäumen unter sich ändernden Klimabedingungen – Auswertung der Beobachtungen in den Internationalen Phänologischen Gärten und Möglichkeiten der Modellierung von Phänodaten. Thesis Forstwissenschaftliche Fakultät der Uni München. München, Universität München

Metropolis N, Rosenbluth AW, Rosenbluth M, Teller AH (1953) Equation of state calculations by fast computing machines. J Chem Phys 21:1087–1092

Murray MB, Cannel MGR, Smith RI (1989) Date of bud burst of fifteen tree species in Britain following climatic warming. J Appl Ecol 26:693–700

Nitsch JP (1957) Photoperiodism in woody plants. Proc Am Soc Hortic Sci 79:526–544

Perry TO (1971) Dormancy of trees in winter. Science 171:29–36

Powell LE (1987) Hormonal aspects of bud and seed dormancy in temperate-zone woody plants. Hortscience 22:845–850

Robertson GW (1968) A biometeorological time scale for a cereal crop involving day and night temperatures and photoperiod. Int J Biometeorol 12:191–223

Sachs L (1997) Angewandte Statistik. Springer, Berlin Heidelberg New York

Samish RM (1954) Dormancy in woody plants. Annu Rev Plant Physiol 5:183–204

Schaber J (2002) Phenology in Germany in the 20th century: methods, analyses and models. Dissertation, University of Potsdam, published as PIK-Report no. 78, Potsdam Institute of Climate Impact Research. (Can be downloaded from http://pub.ub.uni-potsdam.de/2002/0022/schaber.pdf)

Schaber J, Badeck F (2002) Evaluation of methods for the combination of phenological time series and outlier detection. Tree Physiol 22:973–982

Smith H, Kefford NP (1964) The chemical regulation of the dormancy phases of bud development. Am J Bot 51:1002–1012

Sokal RR, Rohlf FJ (1995) Biometry, 3rd edn. Freemann, New York

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

Vegis A (1973) Dependence of growth processes on temperature. In: Precht H, Christophersen J, Hensel H, Larcher W (eds) Temperature and life. Springer, Berlin Heidelberg New York, pp 145–169

Wang JY (1960) A critique of the heat unit approach to plant response studies. Ecology 41:785–790

Wareing PF (1956) Photoperiodism in woody plants. Annu Rev Plant Physiol 7:191–214

Wareing PF, Saunders PF (1971) Hormones and dormancy. Annu Rev Plant Physiol:261–288

Worrall J (1993) Temperature effects on bud-burst and leaf-fall in subalpine larch. J Sustain For 1:1–18