Christoph Schleip – Volker Dose– Annette Menzel

Norway spruce (*Picea abies*): Bayesian analysis of the relationship between temperature and bud burst.

Corresponding author:
Dipl.-Forstw. Christoph Schleip
Lehrstuhl für Ökoklimatologie
Technische Universität München
Am Hochanger 13
D- 85354 Freising - Weihenstephan

Email: Christoph.Schleip@web.de
Tel.: (+49) 8161 714744
Fax.: (+49) 8161 714753

Prof. Dr. Annette Menzel
Lehrstuhl für Ökoklimatologie
Technische Universität München
Am Hochanger 13
D- 85354 Freising - Weihenstephan

Email: menzel@forst.tu-muenchen.de
Tel.: (+49) 8161 714740

Prof. Dr. Volker Dose
Max-Planck-Institut für Plasmaphysik
Boltzmannstraße 2
D- 85748 Garching bei München

Tel.: (+49) 8932991265
Fax.: (+49) 8932992591

Keywords

Bayesian analysis, simulated annealing, Norway spruce, bud burst, time series, nonlinearity, temperature response, prompt response.

Abstract

Climate change has already affected the phenology of several species. To be able to assess the impacts of climate change under various climate scenarios, we need superior models of the phenology of different species. Linear regression methods alone are of limited value for the analyses of natural indicators or phenological data because most time series of naturally-occurring events in ecosystems do change. In this paper, we applied a Bayesian probability approach to investigate time series of the phenological phase of bud burst in Norway spruce (*Picea abies* (L.) Karst.) and mean monthly/weekly temperatures of corresponding climate stations in Germany. Our aim was to detect, in these temperature and Norway spruce bud burst time series, years with the highest probability for discontinuities. We analysed rates of change and the relationship between temperature changes and bud burst of Norway spruce in the 51 year period 1953-2003.

We used a Bayesian method for a coherence analysis between phenological onset dates and an effective temperature generated as a weighted average of monthly and weekly means from January to May. Weight coefficients were obtained from an optimization of the coherence factor by simulated annealing.

In all investigated cases we found coherence factors that suggested a relationship between temperature and phenological time series. Norway spruce bud burst and mean temperature times series of April and May exhibited abrupt changes, particularly at the beginning of the 1980s. April and May temperature time series revealed an increased warming until 2003, and bud burst events advanced. Norway spruce bud burst, in particular, exhibited responses to temperatures of the previous (April) and current month (May). We suggest that forcing temperatures in phenology models should include beside commonly used sums of daily mean temperatures also solutions where weighted effective temperatures in a sensitive time span are considered.

1 Introduction

In the Northern Hemisphere, spring events now occur earlier in the year than in previous decades (Menzel and Fabian 1999). For Norway spruce (*Picea abies* (L.) Karst.), and other evergreen and deciduous trees in mid and higher latitudes, an optimal temperature response of bud burst is one of the most essential factors in natural selection. The ability to avoid spring frost damage and, at the same time, the effective exploitation of the growing season by an early initiation of growth, will greatly improve the tree's fitness and ability to compete. In this study we concentrate on Norway spruce (*Picea abies* (L.) Karst.) which is an economically important timber species. Several studies have used linear regression methods to investigate the relationship between spring phenology and air temperatures (e.g. Sparks and Carey 1995, Beaubien and Freeland 2000, Sparks et al. 2000, Menzel 2003, Menzel et al. 2006). Menzel (2003) used a subset regression technique to describe the correlation of phenological anomalies with air temperature. Although linear regression models have been widely used they are of limited value for the analysis of natural indicators or phenological data. Most time series in ecosystems exhibit various kinds of trends, cycles and seasonal patterns. Phenological records frequently reveal a heterogeneous pattern of temporal variability with sometimes alternating periods of advanced and delayed onset (e.g. Sparks and Carey 1995; Ahas 1999; Schleip et al. 2006). The functional behaviour of phenological time series often exhibits a discontinuity in the late 1980s (Chmielewski and Rötzer, 2002; Scheifinger et al., 2002, Dose and Menzel 2004; Schleip et al. 2006).

Pearson correlations show the strength of a linear relationship between two variables. However, if an obvious nonlinear relationship between two variables can be observed, the Pearson correlation coefficient is suboptimal (Anscombe, 1973).

Recently, new approaches on a physiological basis use promoting and inhibiting variables which are related to temperature (Schaber and Badeck 2003). Mostly models of bud-burst

timing were based on the concepts of stage of bud dormancy and stage of bud ontogenesis (Hari, 1972; Sarvas, 1974; Häkkinen et al., 1998). The simplest phenological model considers that only forcing temperatures cumulated from a fixed date to a given sum explain the dates of bud burst [‘Thermal Time model’] (Cannell and Smith 1983). Rates of forcing are either growing degree-days (Murray et al., 1989) or a sigmoid function of the temperature (Kramer, 1994a, Hänninen, 1990a). Häkkinen (1999) also refers to the theory that bud burst takes place when the stage of ontogenesis exceeds a defined threshold value. Phenological models are often driven by three main assumptions: the type of response of bud growth to temperature, the dependency of chilling and forcing temperature effects, and the time windows when buds are assumed to be receptive to chilling and forcing temperatures (Chuine, 2000).

The present paper looks for supporting results particularly for the first assumption: the type of the response of bud burst to temperature. We are looking especially for alternative explanation of the rates of forcing. For an improved understanding of ecological responses to climate change we seek methods which are equally applicable with nonlinear changes in time series and with linear and nonlinear dependences.

Our study addresses three main questions:

1) Do phenological time series of Norway spruce bud burst and temperature time series contain discontinuities and if so, when do these changes occur?
2) Do temperature and bud burst time series behave independently or do they exhibit coherence? Which monthly and weekly mean temperatures time series exhibit the highest coherence with Norway spruce bud burst times series at different stations in Germany?
3) How do the rates of change of those months with the highest coherence develop between 1951-2003?

To answer the first question we tested the functional behaviour of phenological and temperature time series for changes. With a Bayesian approach developed by Dose and

Menzel (2004) we analysed three different models/functions; a constant model, a linear model and a change-point model.

To answer the second question we used the conceptually new Bayesian correlation approach, recently proposed by Dose and Menzel (2006), and advanced it methodologically. Dose and Menzel (2006) introduced the concept of a coherence factor as the odds ratio of the probability that the trajectory of temperature and phenological events occurs coherently or independently. Dose and Menzel (2006) applied their approach to blossom onset of three different species at Geisenheim in Germany using average temperatures over a three month period. We expanded the Bayesian correlation approach to eighteen meteorological stations throughout Germany using a higher resolution of monthly and even weekly mean temperatures. In contrast to the approach of Dose and Menzel (2006), we implemented a simulated annealing optimization algorithm to generate the coherence factor and temperature weights. The method of simulated annealing is a technique that has attracted substantial attention as being suitable for optimization problems at large scales, especially for those where a global maximum may be present among many, poorer, local maxima. High temperature weights signify a high coherence of the monthly temperature change-point curves with the phenological change-point curve.

Finally, to answer the third question, the rates of change were obtained by a overlay of a constant, linear and a change-point rate of change weighted by their respective model probabilities.

2 Material and methods

Climatic data

Daily temperature data from 18 meteorological stations in Germany (Fig. 1), collected by the German Meteorological Service (DWD) within the period 1951-2003 were used in our analysis. We concentrated on those met stations which had at least eight phenological stations within 25 km. The altitude of the met stations ranged from 5 m to 516 m above sea level. The temperature observations were condensed into monthly and weekly average temperatures. Weekly averages were calculated with the help of a SAS WEEK function. The WEEK function returns the week-number value of the current date as a decimal number in the range 0-53. Monday was considered the first day of the week.

Phenological data

The phenological data of bud burst of Norway spruce were also provided by the DWD. In the DWD phenological network, volunteers at around 1,600 stations observe defined plant development stages (DWD 1991). The phenological data were obtained by averaging time series from at least eight phenological stations within 25 km of the met station (Fig. 1), which did not differ by more than 50 m in elevation from that of the met station. The combination of several phenological stations within an area reduces any influence of local microclimate. In this study, we focused on the phenophases of bud burst of Norway spruce (*Picea abies* (L.) Karst.) between 1951 and 2003. Overall the mean bud burst date was May 8; the earliest bud burst was April 16 1961 at Trier, and the latest was May 30 1984 at Hof (Fig. 2). The DWD observer manual (DWD 1991) defines a bud as opened when the fresh green needles are clearly visible and separated.

Methods of Analysis

We used a new Bayesian approach for the description of climatic and phenological time series that was introduced by Dose and Menzel (2004). That paper details the computational and mathematical formulae we use. Here we briefly summarize the main features of the Bayesian approach. An important feature of Bayesian probability theory is that it allows calculation of the probability of different competing models. It often happens that no single model is clearly superior. In such cases, a marginalization of all three models weighted by their respective model probability is used to achieve the most probable functional description and annual rate of change.

We selected three models to describe the temperature and phenological data: constant, linear and change-point models. The constant model assumes a functional behaviour constant in time with an associated zero rate of change. The linear model assumes a linear change in time of the observed phenomenon, i.e. with an associated constant rate of change. The change-point model offers a time varying change. The change-point model is a triangular function which is supported at the beginning of the time series in year x_1 and assumes there a function value f_1, and the endpoint of the time series in year x_N and assumes there the functional value f_N. Although the endpoints of the time series remain fixed in the subsequent calculations, the intermediate point x_E with associated functional value f_E can be any year such that $x_2 < x_E < x_{N-1}$. The functional values, as well as the matching point (change point) of the two linear sections making up the triangular function, are variables of the calculation. The change point model is not simply another arbitrary functional form which is likely to provide a better fit to the data due to its extra parameter. The assessment of the quality of a model is determined by the so-called odds ratio which is described in detail in Dose and Menzel (2006). The odds ratio assumes same prior probabilities for competing models and is equal to the so called "Bayes Factor". The Bayesian approach provides a powerful way of assessing competing models at the forefront of science by automatically quantifying Occam`s razor (Garret 1991;

Gregory 2005). Occam`s razor is a principle attributed to the medieval philosopher William of Occam (or Ockham). The principle states that one should not make more assumptions than necessary. It chooses the simplest from a set of otherwise equivalent models of a given phenomenon. In any given model, Occam`s razor helps us to "shave off" those variables that are not really needed to explain the phenomenon (Garret 1991; Gregory 2005). In Fig. 3 an example of the three model fits are given for the phenological phase bud burst of Norway spruce in Hof. At this station the Bayesian model comparison reveals a change-point model probability of 100% and underlines the presence of one major change-point within the period 1951-2003. The one change-point model is sufficient enough to represent the major change in a 50 year long phenological and temperature time series. Especially when investigating long-term (>100 years) temperature or phenology changes (see Schleip et al. in review) a multiple change-point model will be of course capable of modelling a more detailed structure in a time series and therefore would mirror more adequately several temporal changes within long-term time series. But each added change-point adds two more variables to the likelihood that may be punished by the Ockham`s razor because of overestimation.

If N is the number of entries on the time scale, there are N-2 possibilities (excluding the endpoints) for the change-point position. The Bayesian marginalization rule is employed to remove the change-point variable by marginalization. This extremely important rule removes 'nuisance' parameter from a Bayesian calculation (Dose and Menzel 2004). The change-point variable is such a 'nuisance' parameter because we do not consider the best solution to be that which minimises the root mean square error (RMSE), or the two or three best triangular functions, but all of them. The advantage is that the marginalisation rule overlays all possible triangular functions and then weights them with their respective change-point probability. By using the Bayesian marginalization rule the support functional values can be eliminated resulting in a probability, p(E), for a particular change-point choice. If the data exhibited an abrupt change, then this change-point probability would be zero except for that particular E.

In cases of more gradual change, appreciable change-point probabilities are also observed for E+/-1, E+/-2. The associated probability of a change-point position can be rigorously calculated. An example is shown in Fig. 4a and 4b. It exhibits the change-point probabilities as a function of time for a temperature (thick dashed line) and a phenological time series (continuous line) as well as their overlap renormalized to unit area (thin dashed line). The upper panel shows a case of small overlap, characteristic of a small coherence factor and the lower panel a case of large overlap characteristic of a high coherence factor. The calculation of the coherence factor relates to the change-point distributions in the following way:

The variable "change-point position" (E) is eliminated using the Bayesian marginalisation rule. With a flat prior distribution for the change-point position this amounts to averaging over all N-2 change-point positions per series. N does not need to be identical to the number of observations, because the algorithm tolerates missing data. The calculation of the probability p(x) that temperature and phenology observations evolve either independently or coherently (= synchronously) reduces to performing this average independently for the change-point positions in the temperature (ET) and in the phenology series (EP) or for ET = EP only. The ratio of probabilities p(coherent)/p(independent) is equivalent to a Bayes Factor. In the absence of qualified prior information the Bayes Factor equals the posterior odds (Dose and Menzel, 2006). The Bayes Factor will be called coherence factor henceforth. A coherence factor above one signifies that the two time series are more probably synchronous than independent.

In the recent work of Dose and Menzel (2006) the phenology time series of snowdrops, cherry and lime tree at Geisenheim were related to the three-month mean temperatures January – March, February – April and March –May. In the current paper we generalize the temperature averaging and choose as the average effective temperature $T(y_i)$ in year y_i

$$T(y_i) = \sum_{k=1}^{k_{\max}} w_k \cdot T_k(y_i), \qquad \sum w_k = 1\,,\; w_k > 0\,, \tag{1}$$

where $T_k(y_i)$ are the average temperatures in month or week k of year y_i, w_k are positive weight coefficients that add up to unity and k=1 is either associated with January or with the first week of the year, while k_{max} is the last month or week in which the phenological event occurs. The unknown weight coefficients are determined by maximizing the coherence between temperature and phenology time series. In the first cycle of the calculation (n=1, where n is the index of the current cycle of the calculation) we start with an assumption of equal weights $w_k=1/k_{max}$ for all k. These weights are then used to calculate the coherence factor C using the procedure of Dose and Menzel (2006).

For cycle n=2, a random new set of unnormalized weights is generated according to

$$w_{new} = w_{old}\left(1 + r * \frac{N^2}{N^2 + 4\,n^2}\right) \qquad (2)$$

where r is a uniform random number $-0.25 < r < 0.25$, n is the index of the current cycle of the calculation and N the predetermined number of cycles chosen to find an optimum set of weights. w_{new}, w_{old} are components of vectors with the dimension= k_{max}. The factor multiplying the random number r is near unity at the beginning of the calculation, it drops to one half at $n= N/2$, and converges to 0.2 for $n=N$. w_{new} must, of course, finally be renormalized to sum to unity.

In the present analysis, two routes were found to improve the initial choice of weights once the new coherence factor C_{new} was known. The simple uphill search algorithm accepts the new set of weights only if they lead to an improved coherence factor. If not, a new proposal set is generated with the previous set of w_{old}. On the other hand, if a higher coherence factor results from the calculation, the associated weights become w_{old} for the next calculation cycle.

The relationship between weights and coherence factor is nonlinear and complex. In fact there is no guarantee that the function "coherence factor" exhibits a unique maximum as a function of the k_{max} weights. In such a situation the simple "uphill search" algorithm may converge to a local maximum and miss the global maximum. This multimodal possibility can be resolved

by using a simulated annealing approach, which accounts for the multimodal possibility and finds the global maximum in the presence of one or several lower satellite maxima. For this purpose it is necessary to accept not only uphill steps but, conditional on a certain probability p, where

$$p = Min(1, \exp\{ (C_{new} - C_{old})/T \}) \tag{3}$$

to allow also for downhill moves. The latter can cross a valley and find another possibly higher maximum. T is the annealing temperature and scales the difference between the new (C_{new}) and the old (C_{old}) coherence factors. The approach for one step of the calculation is then

1.) accept uphill moves with probability one, i.e. always
2.) if $C_{new} < C_{old}$ choose a random number R from a uniform (0,1) distribution and accept the downhill move if

$$\frac{C_{new} - C_{old}}{T(n)} > \ln R \tag{4}$$

There is no general rule for the choice of $T(n)$.

The present calculations were done with $T(n+1)=T(n)/1.01$ e.g. a one percent decrease of temperature per step. Fig. 5 shows the random walks of coherence factor and weights using the simulated annealing approach. It leads to the same approximate optimum as the simple "uphill search" discussed before. This is not necessarily always the case but, so far, we have not observed a difference between the two approaches in calculations on more than fifty data sets. However, since the computational effort is practically the same for the two approaches, we have chosen to use the simulated annealing route as our tool for the current and future calculations. For our application we chose five months with mean monthly and weekly temperatures. We selected the months January to May, as May is the last month in which

Norway spruce bud burst occurs in Germany. For the weekly resolution we chose 20 weeks with mean weekly temperatures. The choice of January or alternatively the first week of the year as the initial month or week is somewhat arbitrary. In our example that illustrates the method (Fig. 5), it is evident that the weights from January, February and March temperatures show no systematic pattern when compared between different data stations. In Hof, April and May temperature weights of 0.58 and 0.37 exhibited the highest values. Note that the coherence factor converges to 3.3 in this case.

3 Results

Model probabilities

To answer the first question of the introduction, we tested the model preferences of 18 mean phenological time series of Norway spruce bud burst and of the corresponding monthly temperature time series from January to May. The Bayesian model comparison revealed that at 17 phenological stations the time series were described best with a change-point model (Fig. 6a). The only exception is found at Wuerzburg where the linear model exhibited a model probability of 51%. Overall the change-point model was the best model for phenological data description (87% average model probability). The linear model was the second best (12%) and the constant model was the least preferred model (1%).

The change-point model also provided the best data description for April and May temperature time series. The average change-point model probability of the monthly temperature time series increased from 39% in January to a maximum probability of 61% in May (Fig. 6b). The linear and constant models exhibited a considerably large probability in the months January to March. The mean probability of the constant model had its maximum for March temperatures (47%).

Change-point probability distribution

As demonstrated above the change-point model was the preferred model to describe phenological and temperature time series in April and May. In Fig. 7 the corresponding change-point probability distributions at all 18 stations are summarized with the help of box plots, including Norway spruce bud burst temperature time series of April and May and additionally the joint change-point probability of all temperature and bud burst time series (Fig 7d).

At the beginning of the 1980s the change-point probabilities of Norway spruce bud burst encompassed values from 2% to 13%. During the decades of the 1950s, 1960s and 1990s the change-point probabilities only ranged from 0% to 2% (Fig. 7a). In the 1970s and especially in the 1980s the upper and lower change-point probability limits, which are symbolized by the vertical line within each box plot, were higher than in other decades (Fig. 7a).

At the beginning of the 1980s the change-point distributions of April and May temperatures revealed the highest upper change-point probability limits with values exceeding 6% in Figs. 7b and 7c. The joint (temperatures and phenological) change-point probability distribution exhibited a clear single peak at the beginning of the 1980s (Fig 7d).

Coherence factors

To answer the second question of whether temperature and Norway spruce time series evolve independently or coherently we calculated coherence factors for each station in Germany. At nine of 18 stations, the coherence factors had a value of two or higher (Fig. 8a). In the remaining seven cases the factor remained below two. For Frankfurt and Hof the coherent option was superior to the incoherent treatment by a factor greater than three. In all presented cases, the coherence factor was never less than 1.1. This important result signifies that temperature and bud burst time series are more probably synchronous than independent. The average coherence factor of all climate stations in the monthly resolution amounted to 2.07 and, for the weekly resolution, 2.4 (not shown in Figures).

Temperature weights

The temperature weights reveal that, at nearly all investigated stations, the change-point distributions of April and May temperatures correlated with the change-point distributions of bud burst of Norway spruce (Fig. 8a). At the monthly resolution April and May revealed the highest temperature weights (of 0.48 and 0.28 respectively). Some exceptionally high

temperature weights were also found at single stations e.g. Wuerzburg for January and February (Fig. 8a).

Similar results can be seen in the weekly resolution. Within April (week number 14 to 17) and May (week number 18 and 20) the temperature weights had their highest values (Fig. 8b). The maximum temperature weights were seen in week 18 and had values above 0.2. A smaller accumulation of high temperature weights was seen at the end of February (week 9) with temperature weights of 0.15. For both the monthly and the weekly resolution, March exhibited the lowest weights.

Model averaged rates of change

To answer question three, we compared the model averaged rates of change of Norway spruce bud burst and of the corresponding temperature time series of April and May. Over the period 1951-2003, the rates of change of Norway spruce bud burst exhibited a discontinuity at the beginning of the 1980s (Fig. 9a), where the upper limit of the box plots drops to negative rates of change; equivalent to advancing bud burst. In 2003, the rates of change of all 18 phenological stations ranged from -0.25 to -0.75 days year^{-1} (Fig 9a). At the beginning of the time series in 1951 the rates of change ranged from 0.35 to -0.02 days year^{-1} which meant that bud burst was delayed at the majority of the investigated stations.

For the months of April and May we detected cooling and warming periods from 1951 to 2003. In 1951, change rates of April temperatures were between -0.07 and 0.00 °C year^{-1}; equivalent to cooling. At the end of the time series, April temperatures warmed at 18 stations (0.01 to 0.07 °C year^{-1}). In comparison May temperatures showed a stronger warming and a larger variability. At the beginning of the time series the rates of change of May temperatures ranged from -0.03 to 0.03 °C year^{-1}, at the end (2003) from 0.03 to 0.17 °C year^{-1}.

4 Discussion

This study does not only deliver quantitative results on the correlations between temperature and Norway spruce bud burst in Germany, but it also offers new insights on model improvement and to methods for the understanding of ecological responses to climate change.

Most time series of naturally-occurring events in ecosystems do change their slopes. In the present work we showed, using the approach of Dose and Menzel (2004) that linear regression models alone are of limited value for the analyses of temperature or phenological time series. Norway spruce bud burst time series of 17 out of 18 German stations revealed an abrupt change at the beginning of the 1980s. The change-point model proved to be the preferred model with an average model probability of 87% to describe this observed discontinuity. We show that temperature time series also exhibited this discontinuity at the beginning of the 1980s. April and May revealed the highest change-point probabilities. The advantage of the Bayesian probability method is that it allows an accurate analysis of the relationship between phenology and temperature observations. In all cases investigated here, the results clearly suggest a coherent development of temperature and phenological time series, with some coherence factors as large as three. Therefore, we expect that matching point probabilities derived from the two data sets (joint change-point distribution) will be more informative (e.g. better localized in time than that obtained from a single series of data). Norway spruce bud burst represents a phenological phase which shows a prompt response to temperatures of the previous (April) and current month (May) with average temperature weights of 0.48 and 0.28, respectively. A high coherence factor signifies that the change-point distribution curves of phenological and temperature time series are largely synchronous, e.g. exhibit large overlaps (as shown in Fig. 4b). A change-point distribution curve with a clear peak signifies a higher probability of an abrupt change. The higher the estimated temperature

weights for a certain month, the more overlapping can be expected in the change-point distributions of temperatures and the phenological event. It is important to note that no conclusion can be made regarding the existence or the direction of a cause and effect relationship; only that Norway spruce bud burst change-point distributions are correlated with the temperature change-point distributions of a certain month or week. Phenophases are responding to many meteorological and environmental factors such as light, photoperiod, temperature, precipitation, humidity, wind, soil conditions etc. (Schnelle, 1955; Menzel, 2002). Despite the many influencing factors, the timing of leaf unfolding of trees is very likely triggered mainly by temperature. Specifically, chilling temperatures break winter dormancy and subsequent warming temperatures induce budburst (Dose and Menzel 2006). Determining whether there is an actual cause and effect relationship requires further investigation. The fact that spring phenology is very likely primarily driven by temperatures suggests that we can attribute the observed biological rates of change to the effects of climate variation.

The comparison of rates of change of Norway spruce bud burst time series with those of monthly temperatures that exhibit the highest temperature weights gives us further insights into the relationship. We calculated model averaged rates of change, using the Bayesian probability approach of Dose and Menzel (2004). Model averaged rates of change are obtained by the superposition of the constant, the linear and the change-point model rates of change, weighted by the respective model probabilities. It is worth noting that the model averaging process does not alter the shape of the rate of change derived from the change-point model. The model averaging procedure adds a counterbalance due to the constant distribution from the linear model and a reduction of the amplitude by the amount of the model probability obtained for the change-point model. The model averaged rates of change of April and May temperatures have increased from 1951 to 2003, which is equivalent to increased

warming. In 2003, May temperature change rates of 18 climate stations ranged from 0.03 to 0.17 °C year^{-1} whereas in 1951 May temperature change rates ranged from -0.03 to 0.03 °C year^{-1}. In contrast, change in bud burst of Norway spruce in 2003 was estimated from -0.25 to - 0.75 days year^{-1} but had showed a delay in 1951. Over most of the investigated period, there was essentially a zero rate of change; but from the 1980s onwards the rate of change was negative for Norway spruce bud burst. This finding is consistent with results of other studies (e.g. Scheifinger et al., 2002; Chmielewski and Rötzer, 2002; Dose and Menzel, 2004, Schleip et al., 2006) that describe an abrupt change towards earlier occurrence dates after the late 1980s and almost no rates of change before that date. Thus the results of our paper clearly reveal that the phenological phase has a discontinuity in the 1980s. We confirmed for several climate stations in Germany that temperatures in April and May had a very similar discontinuity in the 1980s. The reason for this specific timing of change-points in the 1980s is most likely linked to altered atmospheric circulation patterns, such as the North Atlantic Oscillation (NAO) (e.g. Menzel, 2003).

The results of Menzel et al. (2006) and Menzel (2003) underline our findings. Menzel (2003) found that the anomaly curve of Norway spruce revealed notable phenological advances of 0.13 days year^{-1} during the previous 5 decades (1951-2000). Menzel (2003) detected that, in general, later spring phases (including Norway spruce) responded to March to May temperatures. Menzel (2003) calculated the subset regression between phenological anomalies of bud burst of Norway spruce and the three-monthly running mean temperatures of March, April and May. Her results showed a R^2 of 0.79 and a slope of -4.7 days year^{-1}. Menzel (2003) also applied a two-variable model where the month preceding bud burst (April) was chosen as the first variable and the mean temperatures of March-May as the second variable; and explained most of the variability (R^2 =0.85). In the work of Menzel et al. (2006), most phases correlated significantly with mean monthly temperatures of the month of onset and the

two preceding months. For 19% of the phenophases the highest correlation was seen with the month of onset, 63% with the preceding month and 18% with that 2 months earlier.

The enhancement of resolution of our approach by weekly or even shorter temperature intervals has pros and cons. On one hand such an enhancement of resolution inherits a loss in the achievable precision and very likely causes unwanted noise. In other words, if we conducted our analysis with a daily resolution, we might get high temperature weights of a certain day which is more likely accidentally and not because of a biological dependence. But, on the other hand, the results of our weekly analysis reveal more specific information about further systematic biological dependences. Beside April and May, the end of February exhibited a systematic accumulation of higher temperature weights.

The state of forcing is often described as a sum of daily rates of forcing (Chuine, 2000). Our results suggest that bud burst does not simply react to a rate of forcing with fixed temperature sum or defined threshold value as used by Cannell and Smith (1983), Murray et al. (1989) and Häkkinen (1999) and others. Forcing temperatures rather exhibit a periodic pattern with a smaller first signal at the end of February and a greater temperature prompt in April and May. However, forcing temperatures have changed in recent decades in a nonlinear way.

Our approach of first analysing the properties of the time series, such as model preferences, change-point probabilities and rates of change, and then, secondly, investigating the coherence of the temperature and phenological time series gives more detailed insights into the nature of the interdependences than just analysing directly the effects of air temperature on the phenological timing. We demonstrated and emphasized how well nonlinear temperature change patterns are mirrored by the phenological event. As support for more ecophysiological approaches, one could say that they should incorporate specific forcing temperature change patterns rather than temperature threshold sums of previous and current years. Linkosalo (2000) concluded that the formulation of commonly used phenological models seems to be general enough to suit several different plant species and various

phenological phenomena. But Linkosalo (2000) also mentioned that it is also possible that the nature of the control mechanism is not straightforward triggering as stated in the same models.

Häkkinen (1999) has already discussed the disadvantages of standard statistical methods because of the dynamic nature of the models of bud development theories. He suggested an alternative approach of a bootstrap and cross validation method for the evaluation of theories based on the numerical comparison of the model mean square errors only.

Chuine et al. (1998) tested four commonly used models to predict the dates of flowering of temperate-zone trees, the spring warming (Hunter and Lechowicz 1992), sequential (Sarvas 1974; Hänninen 1987, 1990b; Kramer 1994b), parallel (Landsberg 1974; Hänninen 1987, 1990b; Kramer 1994a) and alternating models (Cannell and Smith 1983; Murray et al 1989; Kramer 1994a, 1994b). The main disadvantages of these models are that they are unable to make accurate predictions based on external data (Kramer 1994a). Chuine et al. (1998) stated that the external validity is still not existent for the majority of the species. They suggested that a wrong estimation of the starting date of the forcing phase and a wrong estimation of the critical state of forcing may be the reason. The comparison of the accuracy of different models for different species shows that there is no consensus model even if some models seem consistently more accurate than others (Chuine et al. 1998).

Our study indicates that the method of Bayesian analysis combined with the method of simulated annealing may bring a non-negligible contribution to the estimation of forcing temperatures and model selection. The great advantage of Bayesian analysis is that it considers the inability to prefer one model against another that enforces the collection of new data. The description of the data in terms of only one model is often unsatisfactory (Dose and Menzel 2004, Schleip et al. 2006). The Bayesian model comparison analysis allows us to estimate a reliable combined model averaged rate of change. Compared to the commonly used linear regression approach, we are able to provide model averaged rates of change at an

annual resolution. This helps us to describe discontinuities and to quantify the direction and speed of the changes. Further more the implemented simulated annealing method allows determining temperature weight coefficients that show us which temperature changes support phenological change-points.

Conclusions

We have shown that Norway spruce bud burst and temperature time series both reveal nonlinear changes at the beginning of the 1980s. For nearly all phenological data examined, the change-point model was the preferred model to describe the time series. Change-point distributions of Norway spruce bud burst exhibited the highest Bayesian correlations with temperatures at the end of February, and in April and May. The annual resolution of the rates of change of the Norway spruce time series and April and May temperature time series gives further insight into the coherence of these time series. Since the beginning of the 1980s, April and May temperature rates of change of all 18 investigated stations increased to positive values (warming) and Norway spruce bud burst time series started to reveal an enhanced advancing of the phenological phase. With the help of our method we suggest for phenology models to incorporate specific forcing temperature change patterns for each phase. The influence of forcing temperatures may be defined beside daily temperature sums also by weighted effective temperatures in a sensitive time span. Thus it would be possible to detect different time spans of relevant forcing temperatures for one phenological phase. Further more the method allows to determine for each species the individually temperature respond pattern.

The model comparison option of Bayesian probability theory enables us to test for further bud burst theories in the future. The theory allows a ranking of a number of different models and provides numerical measures of their respective probabilities. The model comparison option of Bayesian theory rests on the built in Occam's razor (Garret 1991), which limits the complexity of a model to the amount necessary to explain the data, avoiding the fitting of noise. Bayesian analysis provides a powerful way of analysing competing models.

References

Anscombe, F. J., 1973. Graphs in statistical analysis. American Statistician, **27**, 17–21.

Beaubien, E.G., Freeland, H.J., 2000. Spring phenology trends in Alberta, Canada: links to ocean temperature. *Int. J. Biometeorol.*, **44**, 53-59.

Cannell, M.G.R. and R.I. Smith., 1983. Thermal time, chill days and prediction of budburst in Picea sitchensis. J. Appl. Ecol. **20**, 951-- 963

Cannell, M.G.R., Smith, R.I., 1986. Climatic warming, spring budburst and forest damage on *Trees Journal of Applied Ecology*, **23**, 177-191.

Chmielewski, F.M., Rötzer, T., 2002. Annual and spatial variability of the beginning of growing season in Europe in relation to air temperature changes. Climate Research, **19**, 257–264.

Chuine, I., Cour, P. and D.D., Rousseau, 1998. Fitting models predicting dates of flowering of temperatue-zone trees using simulated annealing. *Plant, Cell and Environment,* **21**, 455-466.

Chuine, I. 2000. A Unified Model for Budburst of Trees. J. theor. Biol., **207**, 337-347.

Dose, V., Menzel, A., 2006. Bayesian correlation between temperature and blossom onset data. *Global Change Biology*, **12**, 1451-1459.

Dose, V., Menzel, A., 2004. Bayesian analysis of climate change impacts in phenology. *Global Change Biology*, **10**, 259-272.

DWD: Deutscher Wetterdienst (Hrsg.), 1991. *Anleitung für die phänologischen Beobachter des Deutschen Wetterdienstes. 3. Aufl., Selbstverlag des DWD, Offenbach am Main*.

Garrett, A.J.M., 1991. In: Ockham's Razor in: Maximum Entropy and Bayesian Methods (eds Grandy WT, Schick LH), pp. 357–364. Kluwer, Dordrecht.

Gregory P.C., 2005. Bayesian logical data analysis for the physical sciences; A Comparative approach with Mathematica Support. Cambridge University Press, Boston.

Hänninen, H., 1987. Effects of temperature on dormancy release in woody plants: implications of prevailing models. Silva Fennica **21**, 279-299.

Hänninen, H., 1990a. Modelling bud dormancy release in trees from cool and temperate regions. Acta For. Fenn. **213**, 1-47.

Hänninen, H., 1990b. Modelling dormancy release in trees from cool and temperate regions. In Process Modeling of Forest Growth Responses to Environmental Stress (ed. R.K. Dixon, R.S. Meldahl, G.H. Ruark and W.G. Warren), pp. 159-165. Timber Press, Portland.

Häkkinen, R., T. Linkosalo and P. Hari, 1998. Effects of dormancy and environmental factors on timing of bud burst in *Betula pendula*. Tree Physiol. **18**, 707-712.

Häkkinen, R., 1999. Statistical evaluation of bud development theories: application to bud burst of *Betula pendula* leaves. Tree Physiology **19**, 613-618.

Hari, P., 1972. Physiological stage of development in biological models of growth and maturation. Ann. Bot. Fenn. **9,** 107-115.

Hunter, A.F. and M.J. Lechowicz, 1992. Predicting the timing of budburst in temperate trees. J. Appl. Ecol. **29**, 597-604.

Kramer, K. 1994a. Selecting a model to predict the onset of growth of *Fagus sylvatica*. J. Appl. Ecol. **31**, 172-181.

Kramer, K. 1994b. A modelling analysis of the effects of climatic warming on the probability of spring frost damage to tree species in The Netherlands and Germany. Plant, Cell and Environment **17**, 367-377.

Landsberg JJ., 1974. Apple fruit bud development and growth; analysis and an empirical model. Annals of Botany **38**, 1013-1023.

Linkosalo, T., 2000. Mutual regularity of spring phenology of some boreal tree species: predicting with other species and phenological models. Canadian Journal of Forest Research **30**, 667-673.

Menzel, A., Fabian, P., 1999. Growing season extended in Europe. *Nature*, **397**, 659.

Menzel, A., 2000. Trends in phenological phases in Europe between 1951 and 1996. *Int. J. Biometeorol.*, **44**(2), 76-81.

Menzel, A., Estrella, N., 2001. 'Plant Phenological Changes', *in Walther. G.R., Burga, C.A., and Edwards, P.J., (eds.) Fingerprints of Climate Change – Adapted Behaviour and Shifting Species Ranges, Kluwer Academic/Plenum Publishers, New York and London, pp. 123–137.*

Menzel A., 2002. Phenology: its importance to the global change community. *Climatic Change*, **54**, 379–385.

Menzel A., 2003. Phenological anomalies in Germany and their relation to air temperature and NAO. *Climatic Change*, **57**, 243-263.

Menzel A., Sparks T., Estrella N., Koch E., Aasa A., Ahas R., Alm-Kubler K., Bissolli P., et al. 2006. European phenological response to climate change matches the warming pattern. *Global Change Biology*, **12** (10), 1969-1976, doi:10.1111/j.1365-2486.2006.01193.x.

Murray M.B., Cannel M.G.R., Smith R.I., 1989. Date of bud burst of fifteen tree species in Britain following climatic warming. *Journal of Applied Ecology*, **26**, 693–700

Parmesan C., Yohe G., 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**, 37–42.

Root, T.L., Price, J.T., Hall, K.R., et al., 2003. Fingerprints of global warming on wild animals and plants. *Nature*, **421**, 57–60.

Rötzer, T., Grote, R., Pretzsch, H., 2004. The timing of bud burst and its effect on tree growth *International Journal of Biometeorology*, **48**, 09–118, doi:10.1007/s00484-003-0191-1.

Rötzer T, Chmielewski FM (2001) Phenological maps of Europe. *Climate Research*, **18**, 249–257.

Sarvas, R., 1974. Investigations on the annual cycle of development of forest trees. II. Autumn dormancy and winter dormancy. Commun. Inst. For. Fenn. **84**(1), 1--101.

Scheifinger, H., Menzel, A., Koch, E. et al., 2002. Atmospheric mechanisms governing the spatial and temporal variability of phenological observations in central Europe. *International Journal of Climatology*, **22**, 1739–1755.

Schnelle, F., 1955. Pflanzenphänologie. Akademische Verlagsgesellschaft Geest und Portig, Leipzig.

Schleip, C., Menzel, A., Estrella, N., Dose, V., 2006. The use of Bayesian analysis to detect recent changes in phenological events throughout the year. *Agricultural and Forest Meteorology*, **141**, 179-191, doi:10.1016/j.agrformet.2006.09.013.

Schleip, C., Rutishauser, T., Luterbacher, J., and A., Menzel. Time series modelling and central European temperature impact assessment of phenological records in the last 250 years. in review.

Sparks, T.H., Carey, P.D., 1995. The responses of species to climate over two centuries: an analysis of the Marsham phenological record. *Journal of Ecology*, **83**, 321–329.

Sparks, T.H., Jeffree, E.P., Jeffree, C.E., 2000. An examination of the relationship between flowering times and temperature at the national scale using long-term phenological records from the UK. *International Journal of Biometeorology*, **44**(2), 82-87.

Sparks. T.H., Menzel, A., 2002. Observed changes in the seasons: an overview. *International Journal on Climatology*, **22**, 1715-1725.

Zavaleta, E.S., Shaw, M.R., Chiariello, N.R., Mooney, H.A., Field, C.B., 2003. Additive effects of simulated climate changes, elevated CO_2, and nitrogen deposition on grassland diversity. *Proceedings of the National Academy of Science*, **100**(13), 7650-4.

Figure legends

Fig. 1: Distribution and altitude of the climate stations in Germany (black dots) and corresponding phenological stations (small dots). The radius of the circles around each climate station is 25 km.

Fig. 2: Horizontal boxplots of the onset date of bud burst at all 18 climate stations. The 25th percentile is found at the left end and the 75th percentile is found at the right end of the box. The range is marked as black horizontal line, the median as black vertical line in the boxes. The mean is marked as circle with cross.

Fig. 3: Bayesian change-point, linear and constant model estimation of the onset of bud burst Norway spruce (*Picea abies* L.) in Hof. In this example the change-point model exhibits a probability of 100%.

Fig. 4: Distributions of temperature, bud burst and joint (temperature and bud burst) change-point probability of Norway spruce bud burst (*Picea abies* L.) in Schleswig (a) and in Hof (b). In the upper panel the coherence factor has a value of 1.2 and in the lower panel a value of 3.3. Note that the y-axes have different scales. The thick dashed line symbolises the averaged change-point probability distribution of the weighted temperatures for the months January to May. The continuous line represents the probability distribution of the phenological data. The thin dashed line stands for the joint change-point probability.

Fig. 5: Random walks of coherence factor and monthly mean temperature weights using the simulated annealing approach for Norway spruce (*Picea abies* L.) in Hof, Germany. w[1] to w[5] are weights of January to May mean temperatures respectively, co_fac= Coherence factor. Note that the x-axis shows the number of random steps and the left y-axis describes the

values of the coherence factor, the right y-axis represents the proportions of the temperature weights.

Fig. 6: Bayesian model probabilities of the change-point, linear and constant model of a) Norway spruce bud burst at 18 phenological stations in Germany and of b) mean temperatures from January to May at 18 corresponding climate station.

Fig. 7: Box plots of change-point probability distributions of a) Norway spruce bud burst at 18 phenological stations and of b) April mean temperature time series and of c) May mean temperature time series d) joint (temperature and phenological) change-point probability at the corresponding 18 climate stations. Change-point model probability distributions were calculated for the period 1951-2003. The median is represented by the horizontal line within each box plot. The top of each box is the third quartile (Q3) - 75% of the data values are less than or equal to this value. The bottom of the box is the first quartile (Q1) - 25% of the data values are less than or equal to this value. The lower whisker extends to this adjacent value - the lowest value within the lower limit. The upper whisker extends to this adjacent value - the highest data value within the upper limit.

Fig. 8: Coherence factors and a) monthly and b) weekly temperature weights of bud burst Norway spruce in Germany. In a) the coherence factors are in brackets following the names of the climate stations. The bars represent the temperature weights for a) the months January to May and for b) the weeks since the beginning of the year. Temperature weights were obtained by the simulated annealing optimization.

Fig. 9: Box plots of Bayesian model averaged rates of change of a) Norway spruce bud burst at 18 phenological stations in days year^{-1} and of b) April mean temperature time series and of

c) May mean temperature time series in °C year^{-1} at the corresponding 18 climate stations. Model averaged rates of change were calculated for the period 1951-2003.