Global Change Biology (2004) **10,** 259–272, doi: 10.1111/j.1529-8817.2003.00731.x

Bayesian analysis of climate change impacts in phenology

VOLKER DOSE* and ANNETTE MENZEL†

*Centre for Interdisciplinary Plasma Science, Max-Planck-Institut für Plasmaphysik, EURATOM Association, Boltzmannstrasse 2, D-85748 Garching bei München, Germany, †Department of Ecology, Center of Life Science, TU Munich, Am Hochanger 13, D-85354 Freising, Germany

Abstract

The identification of changes in observational data relating to the climate change hypothesis remains a topic of paramount importance. In particular, scientifically sound and rigorous methods for detecting changes are urgently needed. In this paper, we develop a Bayesian approach to nonparametric function estimation. The method is applied to blossom time series of *Prunus avium* L., *Galanthus nivalis* L. and *Tilia platyphyllos* SCOP. The functional behavior of these series is represented by three different models: the constant model, the linear model and the one change point model. The one change point model turns out to be the preferred one in all three data sets with considerable discrimination of the other alternatives. In addition to the functional behavior, rates of change in terms of days per year were also calculated. We obtain also uncertainty margins for both function estimates and rates of change. Our results provide a quantitative representation of what was previously inferred from the same data by less involved methods.

Keywords: Bayesian theory, climate change, phenology, trend

Received 12 June 2003; revised version received and accepted15 September 2003

Introduction

The global average surface temperature has increased over the 20th century by about $0.6 \pm 0.2\,^{\circ}C$ and is projected to continue to rise at a rapid rate. However, the record shows a great deal of variability: Most of the warming has occurred during two periods (1910–1945 and 1976–2000) and it is very likely that the 1990s were the warmest decade (IPCC, 2001). Many studies have revealed evidence of ecological impacts of this recent climate change. In particular, shifts in plant and animal phenology for the boreal and temperate zones of the northern hemisphere have been reported (Menzel & Estrella, 2001; Sparks & Menzel, 2002; Walther *et al.*, 2002; Root *et al.*, 2003). Phenology is perhaps the simplest and most frequently used bio-indicator to track climate changes. Springtime phases are particularly sensitive to temperature. Thus, phenological observations can demonstrate a consistent temperature-related shift or 'fingerprint' and assess the amount of change. Reviews of phenological trend studies indicate that most of the data originate from the last four to five decades. These recent data predominantly reveal advancing of flowering and leaf unfolding in Europe and North America by 1.2–3.8 days decade^{-1} on average and a strong seasonal variation with highest advances in early spring.

However, there are several problems involved in the commonly used methods of searching for signals in phenological time series. The detection of shifts is mostly done by classical statistical methods, such as slopes of linear regression models (e.g. Bradley *et al.*, 1999; Menzel & Fabian, 1999; Jones & Davis, 2000; Schwartz & Reiter, 2000; Defila & Clot, 2001; Menzel *et al.*, 2001; Ahas *et al.*, 2002; Peñuelas *et al.*, 2002; Menzel, 2003), rarely by other curve fitting methods (e.g. Ahas, 1999; Sagarin & Micheli, 2001). Trends are then reported in days per year or decade, or days of change over the study period. The above methods are sensitive to extreme values (Schlittgen & Streitberg, 1999). It is apparent that using linear regression models the length of a time series and its start and end dates are critical in detecting changes and in determining their magnitude, especially when highly variable phenological time series of a few decades are analyzed. Thus,

Correspondence: Prof. Dr. Volker Dose, tel. + 49 89 3299 1265, fax + 49 89 3299 2591, e-mail: volker.dose@ipp.mpg.de

series that include the whole of the 1990s benefit from the decade being the warmest on record. However, the observation periods vary between phenological networks and among stations in networks because phenological observations mostly depend on volunteers and thus have often discontinuous, incomplete data series. Moreover, different individuals may apply different standards in their observations. Several studies have addressed this problem (e.g. Menzel & Estrella, 2001; Sparks & Menzel, 2002) and illustrate the variation of resulting changes with the period of interest (e.g. Scheifinger *et al*., 2002). Few studies use these linear regression models for even longer time series, some covering almost one century (Beaubien & Freeland, 2000; Kozlov & Berlina, 2002; Magnuson *et al*., 2000 as well as Sagarin & Micheli, 2001 for lake and river ice cover).

Reviews of phenological trend studies suggest that only about 40% of the reported trends have proved statistically significant. The significance is often tested by the *F*-test (Defila & Clot, 2001), and occasionally by the Mann–Kendall trend test, which does not require a Gaussian distribution of the data (e.g. Menzel, 2000). Few studies also report the standard error of the slope (e.g. Sagarin, 2001).

Studies analyzing long-term phenological records often reveal a heterogeneous pattern of temporal variability with sometimes alternating periods of advanced and delayed onset (e.g. Schnelle, 1950; Lauscher, 1978, 1983; Freitag, 1987; Sparks & Carey, 1995; Ahas, 1999). The advance of phenological events in the last decades is compared to the timing in preceding periods, mostly only by comparing averages in distinct periods (e.g. Fitter & Fitter, 2002). Another phenomenon in time domain described (Chmielewski & Rötzer, 2002; Scheifinger *et al*., 2002) is a discontinuity in time series behavior in the late 1980s, as in many areas almost no trend is observed before the discontinuous shift towards earlier occurrence dates after the late 1980s. However, change points have not been determined so far in phenological time series.

These limitations of the currently used methods render comparison and interpretation of the observed changes extremely difficult. They may partly account for the observed spatial variability among sites or the different response of species besides the inherent inhomogeneity caused by local microclimate conditions, natural variation, genetic differences or other nonclimatic factors. Thus, there is strong need to improve the recently applied method of change detection in phenological time series.

Although we are familiar with the ongoing battle between traditional statisticians and the (much older, Bayes, 1763) Bayesian school of thought, we shall not add to the many existing comments. Bayesian methods are presently undergoing an impressive renaissance since the computational problems involved in the application of Bayes theorem fade away with the ever more powerful personal computers available to every scientist. We identify ourselves as practising Bayesians (Dose, 2003) for the reason that this theory of probability requires to make explicit all assumptions that enter the analysis of a particular data set. Furthermore, the marginalization option of Bayesian probability theory, which has no satisfactory counterpart in traditional frequentist statistics, is of invaluable importance in the case of incomplete or poorly conditioned data. Such data sets may be associated with complicated, even multimodal likelihoods (see Discussion). Rigorous application of Bayesian methods allows also for a rigorous analysis of uncertainties of the results. The importance of uncertainty analysis in climate-related research has recently been vigorously pointed out by Katz (2002). We deem equally important the model comparison option of Bayesian probability theory. There is no such thing as a null hypothesis in Bayesian theory. Instead, the theory allows for a ranking of a bunch of models (at least two !) and provides numerical measures of their respective probabilities. The model comparison option of Bayesian theory rests on the built-in Ockham's razor (Garrett, 1991), which limits the complexity of a model to the amount necessary to explain the data, avoiding the fitting of noise. The analysis in the section on Model selection will compare three different models for the trend of flowering time series.

Bayesian statistical methods have been applied so far in climate change detection, analysis and attribution (e.g. Hobbs, 1997; Hasselmann 1998; Leroy, 1998; Tol & De Vos, 1998; Barnett *et al*., 1999; Berliner *et al*., 2000; Katz, 2002), and, for example, in climate reconstructions (Robertson *et al*., 1999). Previous work on model comparison addressed the test for changes in mean and/or variance in hydrological data (Perreault *et al*., 2000).

In this study, we will focus on long-term phenological observations (1896–2002) in Germany and analyze the variations of the onset of phenological phases in the 20th century. We chose time series from Geisenheim (49°59′N,7°58′E, 60 km west of Frankfurt am Main) where several phenological phases have been reported almost continuously since 1896. Three phases within the course of the phenological year were selected. These are: flowering of snowdrop (*Galanthus nivalis* L.) in earliest spring, sweet cherry (*Prunus avium* L.) in mid-spring, and lime tree (*Tilia platyphyllos* SCOP) in mid-summer. Data result from the Historical Phenological Database (1896–1935) and the Actual Phenological

Database (1951–2002) of the German Weather Service, as well as from Schnelle & Witterstein (1952) (1936–1944), and corresponding meteorological yearbooks (1945–1952).

In the next section, we shortly introduce the Bayesian concepts. In the subsequent section we perform a comparison of different models to describe the functional behavior of the blossom time series. The predictions based on the best model are presented in the penultimate section. Finally, an extention of the analysis to other species and the discussion of the results are presented.

Bayesian concepts

In this paper, we are concerned with the identification of trends in phenological data. Since there is no first-principles-based theory for such trends the method of choice is nonparametric function estimation within the Bayesian framework. We shall introduce the Bayesian concepts and terminology only to the extent that is necessary to fix the nomenclature in the rest of the article. Readers without knowledge of Bayesian probability theory are referred to the amply available literature for a more detailed introduction (Tribus, 1969, Box & Tiao, 1973, Sivia, 1996, Leonard & Hsu, 1999).

Bayesian probability theory is based on the application of two rules. The first is the product rule, which allows a probability or probability density function of two (or more) variables conditional on additional information I to be broken down into simpler functions:

$$p(\theta, D|I) = p(\theta|I)p(D|\theta, I). \tag{1}$$

The distributions $p(\theta|I)$ and $p(D|\theta, I)$ depend only on the single variables θ and D, respectively. We shall identify θ with 'parameter' and D with 'data' later on. $p(\theta|I)$ is conditional on I only while $p(D|\theta, I)$ is in addition conditional on θ. Because of the symmetry in θ, D of the left-hand side of Eqn (1) it may be expanded alternatively:

$$p(\theta, D|I) = p(D|I)p(\theta|D, I). \tag{2}$$

Equating the two equivalent expansions (1) and (2) we arrive at Bayes theorem:

$$p(\theta|D, I) = p(\theta|I)p(D|\theta, I)/p(D|I). \tag{3}$$

Bayes theorem tells us how to update knowledge about the parameter θ encoded in $p(\theta|I)$ by collecting appropriate data D. $p(\theta|I)$ is called the prior probability on θ, which we may also regard as expert knowledge. It can stem from earlier experiments, other observations or evolve from the experience of the scientist working in the respective field. $p(\theta|I)$ is usually very weak information. In fact, it is the large uncertainty about θ usually associated with $p(\theta|I)$, which motivates new experiments or observations. The prior $p(\theta|I)$is combined with $p(D|\theta, I)$, the sampling distribution of the data, to arrive at the posterior distribution of $\theta, p(\theta|D, I)$ given the new data D. $p(D|\theta, I)$ will in the following be regarded as a function of the parameter θ and is then called the likelihood function.

The last quantity to be explained in (3) is $p(D|I)$. It follows from the second rule governing Bayesian probability theory, the marginalization rule. This extremely important rule tells how to remove an unwanted 'nuisance' parameter from a Bayesian calculation:

$$p(D|I) = \int p(\theta, D|I)\,\mathrm{d}\theta = \int p(\theta|I)p(D|\theta, I)\,\mathrm{d}\theta. \tag{4}$$

Comparison with (3) shows that $p(D|I)$ is the normalization in Bayes theorem. Eqn (4) suggests, however, an even more extensive interpretation of $p(D|I)$. The meaning of the parameter θ is of course derived from a function or rather a class of functions chosen to model the data. $p(D|I)$ represents then the probability of the data given this class of functions regardless of which numerical value the parameter θ assumes. $p(D|I)$ is therefore also called the evidence of the data given a class of models and is the key quantity to decide how to rank a certain number of models when tried on the same set of data. The procedure is formally simple. Consider the models $\{M_j\}$. We then want to calculate $p(M_j|D, I)$, which is given by Bayes theorem as

$$p(M_j|D, I) = p(M_j|I)p(D|M_j, I)/p(D|I) \tag{5}$$

$p(D|I)$ follows as before from marginalization with the integral in (4) replaced by a sum.

Comparison of models M_j and M_k yields the odds ratio:

$$\frac{p(M_j|D, I)}{p(M_k|D, I)} = \frac{p(M_j|I)}{p(M_k|I)}\,\frac{p(D|M_j, I)}{p(D|M_k, I)} \tag{6}$$

It is composed of two factors. The first is $p(M_j|I)/p(M_k|I)$ and is called the prior odds. It summarizes the experts preference of model choice M_j over model choice M_k. This factor is frequently taken equal to unity since it is exactly the inability to prefer one model against another that enforces collection of new data. The second factor is called the Bayes factor. The probabilities entering the Bayes factor are obtained employing the marginalization rule:

$$\begin{aligned} p(D|M_j, I) &= \int p(\theta, D|M_j, I)\,\mathrm{d}\theta \\ &= \int p(\theta|M_j, I)p(D|\theta, M_j, I)\,\mathrm{d}\theta. \end{aligned} \tag{7}$$

We shall now approximate Eqn. (7) in order to perform a qualitative discussion of the Bayes factor (Gregory &

Loredo, 1992). We shall assume that the prior probability on θ, $p(\theta|M_j, I)$ is rather uninformative and wide in contrast to the likelihood, which we assume to be very informative on θ in which case $p(D|\theta, M_j, I)$ is a function that is concentrated in a narrow region $\Delta\theta$ around its maximum $\hat{\theta}$. The integral in (7) can then be written approximately as

$$p(D|M_j, I) = p(\hat{\theta}|M_j, I)p(D|\hat{\theta}, M_j, I)\Delta\theta. \tag{8}$$

Since the prior is a normalized function, we can define a prior range $\delta\theta$ by $1 = \delta\theta p(\hat{\theta}|M_j, I)$ and obtain

$$p(D|M_j, I) \approx p(D|\hat{\theta}, M_j, I)\frac{\Delta\theta}{\delta\theta}. \tag{9}$$

The arguments above hold of course also if θ is not a scalar, i.e. even if we assume that M_j has a j-dimensional parameter vector $\vec{\theta}_j$. The second factor in (9) must then be raised to the jth power. The approximation discussed so far allows an important conclusion for the case of nested models. We call M_j and M_k nested if all parameters entering M_j also enter M_k with the same meaning and prior probabilities and $k>j$. The Bayes factor for this case becomes then

$$\frac{p(D|M_j, I)}{p(D|M_k, I)} = \frac{p(D|\,\hat{\theta}_j, M_j, I)}{p(D|\,\hat{\theta}_k, M_k, I)}\left(\frac{\Delta\theta}{\delta\theta}\right)^j\left(\frac{\delta\theta}{\Delta\theta}\right)^k. \tag{10}$$

According to the assumption that the prior in (7) is rather diffuse while the likelihood is sharply peaked, we have that $\Delta\theta \ll \delta\theta$. Since by assumption $k>j$ the volume factor $(\delta\theta)^{k-j}/(\Delta\theta)^{k-j}$ is large compared to unity. On the other hand, since $k>j$ model M_k is expected to give a better fit to the data than model M_j and consequently $p(D|\,\hat{\theta}_j, M_j, I)/p(D|\,\hat{\theta}_k, M_k, I)$ is a number smaller than unity. Therefore, model M_k will only be preferred over model M_j if the fit is so much better that it overrides the volume factor. This important built-in feature of Bayesian probability theory is called Ockham's razor. This philosophical concept requires that one should take the simpler of two models, which both provide reasonable explanations of the data.

Finally we comment on the importance of priors in model comparison. The assumption above was that the prior distribution be rather diffuse so that it hardly influences the most probable parameter value $\hat{\theta}$. However, if the prior range $\delta\theta$ becomes too large or tends even to infinity, then it follows from (10) that always the simplest model would win the competition. As a consequence one should try hard to formulate as informative priors as possible for the purpose of model comparison. Improper priors are entirely useless in this context.

Having identified the appropriate model to explain the data we are left with the determination of the parameters, which specify the model. The full information on the parameters is of course contained in the posterior distribution (3). If it is sufficiently 'simple', meaning that $p(\theta|D, I)$ resembles a Gaussian function, then it may be summarized in terms of mean and variance

$$\begin{aligned}\langle\theta\rangle &= \int \theta p(\theta|D, I)\,\mathrm{d}\theta,\\ \langle\Delta\theta^2\rangle &= \int (\theta - \langle\theta\rangle)^2 p(\theta|D, I)\,\mathrm{d}\theta.\end{aligned} \tag{11}$$

This completes a Bayesian analysis if the problem was model selection and best estimate of parameters that specify the model. Frequently, however, the problem arises to make predictions on the basis of the available set of data. We shall defer the treatment of this problem to a separate section.

Model selection

The Bayesian tools will now be applied to the analysis of phenological data such as displayed in Fig. 1 as open circles. They represent the occurrence of cherry blossom (*P. avium* L.) in terms of days after the beginning of the year. The first impression is that they suffer from a considerable natural variability noise. A cursory view might deny any systematic trend at all. This is the first model that we shall consider and for which we shall calculate the evidence.

The constant model

The likelihood function for this model must incorporate the data $\vec{d}$, the years of observation $\vec{x}$, the scatter of the

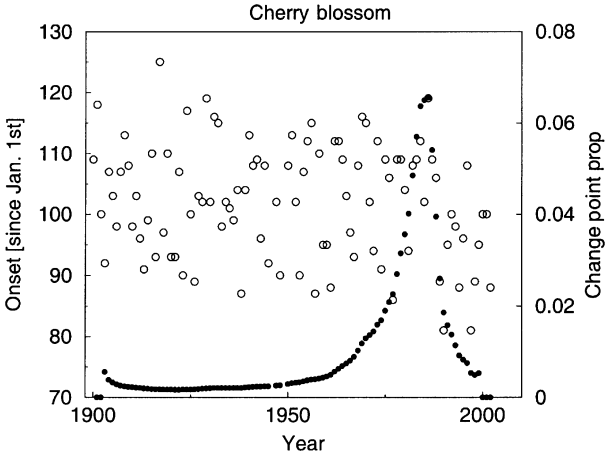

Fig. 1 The observations on cherry blossom at Geisenheim/Germany are displayed as open circles (left vertical scale). The probability of the change point for the one change point model is shown as the full dots (right vertical scale).

data that will be characterized by a variable σ and the constant f that we choose to define the 'no trend' blossom day. The model equation reads

$$d_i - f = \varepsilon_i, \quad \forall i. \tag{12}$$

If expectation value $\langle \varepsilon_i \rangle$ of the errors is zero and variance $\langle \varepsilon_i^2 \rangle$is assumed to be known as σ^2, then by the principle of maximum entropy (Jaynes, 1957a, b; Kapur & Kesavan, 1992) the explicit form of the likelihood becomes

$$p(\vec{d}|\vec{x},\sigma,f,I) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \exp\left\{-\frac{1}{2\sigma^2}\sum_i^N (d_i - f)^2\right\}. \tag{13}$$

The same expression follows if we abandon the principle of maximum entropy and make the much stronger assumption that the ε_i follow a Gaussian distribution with zero mean and variance σ^2. From (13) we must now calculate the evidence $p(\vec{d}|\vec{x},c,I)$. An additional c has been split off from the general conditional background I to make explicit that we treat the constant model here. The evidence is obtained from the marginalization theorem (4):

$$p(\vec{d}|\vec{x},c,I) = \int \mathrm{d}f\,\mathrm{d}\sigma p\,(\vec{d},f,\sigma|\vec{x},c,I). \tag{14}$$

This equation is an identity. The integrand in (14) will now be expanded using the product rule:

$$p(\vec{d}|\vec{x},c,I) = \int \mathrm{d}f\,\mathrm{d}\sigma p(f,\sigma|\vec{x},c,I)\,p(\vec{d}|\vec{x},f,\sigma,c,I). \tag{15}$$

The first distribution under the integral in (15) is logically independent of $\vec{x}$ and c and simplifies to

$$p(f,\sigma|I) = p(\sigma|I)p(f|\sigma,I). \tag{16}$$

Again logical independence applies to $p(f|\sigma,I)$ since the estimate of the average value f is not influenced by any knowledge about the variance σ^2. The prior distribution $p(f|I)$ on f is then chosen (weakly informative) to be constant over the range 2γ

$$p(f|I) = \frac{1}{2\gamma}. \tag{17}$$

The range γ can be estimated from the variance of the data in Fig. 1. A possible k-dimensional generalization of (17) would be $(1/2\gamma)^k$. A more parsimonious choice of the prior volume is, however, obtained if we replace the hypercube by a hypersphere $V_s(k,\gamma)$:

$$V_s(k,\gamma) = \gamma^k(\sqrt{\pi})^k/\Gamma\left(\frac{k+2}{2}\right) = p(\vec{f}|\gamma,k,I). \tag{18}$$

A similar uninformative choice is made for $p(\sigma|I)$. Since σ is a scale parameter for the difference $|d_i - f|$, we choose a normalized form of Jeffreys' prior:

$$p(\sigma|\beta,I) = \frac{1}{2\ln\beta}\frac{1}{\sigma}, \quad \frac{1}{\beta} < \sigma < \beta \tag{19}$$

The integrand in (15) is thereby fully specified. In order to perform the integration we rewrite the exponent:

$$\sum_i (d_i - f)^2 = N(f - \overline{d})^2 + N\overline{\Delta d^2}, \tag{20}$$

$$\overline{d} = \frac{1}{N}\sum_i^N d_i, \quad \overline{\Delta d^2} = \frac{1}{N}\sum_i^N (d_i - \overline{d})^2. \tag{21}$$

This leads to the marginal likelihood:

$$\begin{aligned} p(\vec{d}|\vec{x},c,I) = {} & \left(\frac{1}{2\pi}\right)^{N/2}\frac{1}{2\gamma}\,\frac{1}{2\ln\beta} \\ & \times \int \frac{\mathrm{d}\sigma}{\sigma}\,\frac{1}{\sigma^N}\exp\left\{-\frac{N\overline{\Delta d^2}}{2\sigma^2}\right\} \\ & \times \int \exp\left\{-\frac{N}{2\sigma^2}(f - \overline{d})^2\right\}\mathrm{d}f. \end{aligned} \tag{22}$$

Although the priors on f and σ have only finite range we assume that the ranges are sufficiently large such that the limits of integration can be extended to $(0, \infty)$ for σ and $(-\infty, \infty)$ for f, respectively, with negligible approximation error on the integrals. The inner integral is then of the standard Gaussian one-dimensional type and yields

$$\int_{-\infty}^{\infty} \mathrm{d}f \exp\left\{-\frac{N}{2\sigma^2}(f - \overline{d})^2\right\} = \sigma\sqrt{\frac{2\pi}{N}}. \tag{23}$$

The remaining integral can be converted to the integral representation of the gamma function by substitution $x = 1/\sigma^2$ and yields

$$\int_0^{\infty} \frac{\mathrm{d}\sigma}{\sigma}\,\frac{1}{\sigma^{N-1}}\exp\left\{-\frac{N\overline{\Delta d^2}}{2\sigma^2}\right\} = \frac{1}{2}\frac{\Gamma((N-1)/2)}{\{N\overline{\Delta d^2}/2\}^{(N-1)/2}} \tag{24}$$

Collecting terms the evidence for the constant model becomes

$$p(\vec{d}|\vec{x},c,I) = \frac{1}{2}\left(\frac{1}{\pi}\right)^{(N-1)/2}\frac{1}{2\gamma}\,\frac{1}{2\ln\beta}\,\frac{\Gamma((N-1)/2)}{\{N\overline{\Delta d^2}\}^{(N-1)/2}}\,\frac{1}{\sqrt{N}}. \tag{25}$$

The linear model

This model assumes a linear trend in the data leaving open the question whether we expect a rise or a fall as a function of observation year. The model equation for this case becomes

$$d_i - f_1\frac{x_N - x_i}{x_N - x_1} - f_N\frac{x_i - x_1}{x_N - x_1} = \varepsilon_i, \tag{26}$$

where x_1, x_N are the first and last year of observation and f_1, f_N design functional values that specify a linear trend between x_1 and x_N but are of course unknown. Eqn (26) suggests to go over to matrix notation:

$$\vec{d} - \mathbf{A}\vec{f} = \vec{\varepsilon}. \tag{27}$$

Keeping assumptions on ε_i as before the likelihood becomes

$$p(\vec{d}|\vec{x},\sigma,\vec{f},l,I) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \times \exp\left\{-\frac{1}{2\sigma^2}(\vec{d}-\mathbf{A}\vec{f})^{\mathrm{T}}(\vec{d}-\mathbf{A}\vec{f})\right\}. \tag{28}$$

The explicitly added l in the condition reminds that we are treating a linear trend model. The arguments that lead to priors on $\vec{f}$ and σ as well as considerations of logical independence remain unchanged. The evidence becomes then

$$p(\vec{d}|\vec{x},l,I) = \int p(\vec{f}|I)p(\sigma|I)p(\vec{d}|\vec{x},\vec{f},\sigma,I)\,\mathrm{d}\vec{f}\,\mathrm{d}\sigma. \tag{29}$$

The main difference compared to the previous model is that the $\vec{f}$-integration has become two dimensional. Accordingly the rewriting of the exponent ϕ is slightly more complicated now. We define vector $\vec{f}_0$, matrix $\mathbf{Q}$ and residue R by the equation

$$\begin{aligned}\phi &= (\vec{d}-\mathbf{A}\vec{f})^{\mathrm{T}}(\vec{d}-\mathbf{A}\vec{f}) \\ &= (\vec{f}-\vec{f}_0)^{\mathrm{T}}\mathbf{Q}(\vec{f}-\vec{f}_0)+R.\end{aligned} \tag{30}$$

The $\vec{f}$-integral is now straightforward and we obtain

$$p(\vec{d}|\vec{x},l,I) = \left(\frac{1}{2\pi}\right)^{N/2}\frac{1}{V_s(2,\gamma)}\,\frac{1}{2\ln\beta} \times \int_0^\infty \frac{\mathrm{d}\sigma}{\sigma}\,\frac{1}{\sigma^2}\exp\left\{-\frac{R}{2\sigma^2}\right\}\frac{2\pi\sigma^2}{\sqrt{\det\mathbf{Q}}}. \tag{31}$$

The remaining σ-integral is of the same type as before and the evidence becomes

$$p(\vec{d}|\vec{x},l,I) = \frac{1}{2}\frac{1}{V_s(2,\gamma)}\,\frac{1}{2\ln\beta}\left(\frac{1}{\pi}\right)^{(N-2)/2} \times \frac{1}{\sqrt{\det\mathbf{Q}}}\,\frac{\Gamma((N-2)/2)}{\{R\}^{(N-2)/2}}. \tag{32}$$

Matrix $\mathbf{Q}$ and residue R are so far unknown. These quantities follow from a comparison of coefficients in $\vec{f}$ of the two equivalent forms of ϕ in (30):

$$\begin{aligned}&\mathbf{Q} = \mathbf{A}^{\mathrm{T}}\mathbf{A}, \\ &\mathbf{Q}\vec{f}_0 = \mathbf{A}^{\mathrm{T}}\vec{d} \rightarrow \vec{f}_0 = \mathbf{Q}^{-1}\mathbf{A}^{\mathrm{T}}\vec{d}, \\ &R = \vec{d}^{\mathrm{T}}\vec{d} - \vec{d}^{\mathrm{T}}\mathbf{A}\mathbf{A}^{-1}\mathbf{A}^{\mathrm{T}}\vec{d}.\end{aligned} \tag{33}$$

The expression for the residue R can be simplified considerably if we employ singular value decomposition on the matrix $\mathbf{A}$. Assume

$$\mathbf{A} = \sum_i \lambda_i \vec{U}_i \vec{V}_i^{\mathrm{T}}. \tag{34}$$

This yields for matrix $\mathbf{Q}$:

$$\begin{aligned}\mathbf{Q} = \mathbf{A}^{\mathrm{T}}\mathbf{A} &= \sum_{i,k}\lambda_i\lambda_k\,\vec{V}_i\,\vec{U}_i^{\mathrm{T}}\,\vec{U}_k\,\vec{V}_k^{\mathrm{T}} \\ &= \sum_k \lambda_k^2\,\vec{V}_k\,\vec{V}_k^{\mathrm{T}}.\end{aligned} \tag{35}$$

The last equality follows from the fact that the $\{\vec{U}_i\}$and the $\{\vec{V}_i\}$each form an orthogonal normalized vector system. From (34) and (35) we get immediately

$$\det\mathbf{Q} = \prod_k \lambda_k^2, \tag{36}$$

$$\mathbf{A}\mathbf{Q}^{-1}\mathbf{A}^{\mathrm{T}} = \sum_k \vec{U}_k\,\vec{U}_k^{\mathrm{T}}, \tag{37}$$

$$R = \vec{d}^{\mathrm{T}}\left\{\mathbf{I} - \sum_k \vec{U}_k\,\vec{U}_k^{\mathrm{T}}\right\}\vec{d}. \tag{38}$$

This completes the analysis of the linear model.

The change point model

A change point model is a natural refinement of the linear model. With this model we represent the trend by piecewise linear sections. The simplest case would be one change point separating two linear sections. This model contains four parameters: the design values at the boundaries and at the change point, and in addition, the parameter that specifies the position of the change point. The model equation for the general case is very similar to (25), namely,

$$\begin{aligned}d_i - f_k\frac{x_{k+1}-x_i}{x_{k+1}-x_k} - f_{k+1}\frac{x_i-x_k}{x_{k+1}-x_k} \\ = \varepsilon_i,\quad x_k \le x_i \le x_{k+1}.\end{aligned} \tag{39}$$

The likelihood is identical to (28) expect for the additional parameter $\vec{E}$, which specifies n change point positions. Note that the $\vec{f}$-space is $(n+2)$ dimensional in this model and that

$$p(\vec{d}|\vec{x},\sigma,\vec{f},\vec{E},I) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \times \exp\left\{-\frac{1}{2\sigma^2}(\vec{d}-\mathbf{A}(\vec{E})\vec{f})^{\mathrm{T}}(\vec{d}-\mathbf{A}(\vec{E})\vec{f})\right\}. \tag{40}$$

The matrix $\mathbf{A}$ depends now on the change point positions $\vec{E}$. The formal similarity of the change point model and the linear model allows duplication of most of the previous analysis. The evidence that we need now is $p(\vec{d}|\vec{x},n,I)$ where n is the number of change points. It is given by

$$p(\vec{d}|\vec{x},n,I) = \sum_{\vec{E}} p(\vec{d},\vec{E}|\vec{x},I) = \sum_{\vec{E}} p(\vec{E}|I)p(\vec{d}|\vec{E},\vec{x},I). \tag{41}$$

The marginal likelihood $p(\vec{d}|\vec{E},\vec{x},I)$ can be taken from the linear model case if we remember that determinant of $\mathbf{Q}$ and residue R will now depend on $\vec{E}$.

Hence

$$p(\vec{d}|\vec{x},n,I) = \frac{1}{2}\left(\frac{1}{\pi}\right)^{(N-n-2)/2} \frac{1}{V_s(n+2,\gamma)} \frac{1}{2\ln\beta} \times \sum_{\vec{E}} p(\vec{E}|I) \frac{1}{\sqrt{\det \mathbf{Q}(\vec{E})}} \frac{\Gamma((N-n-2)/2)}{\{R(\vec{E})\}^{(N-n-2)/2}}. \quad (42)$$

In order to proceed, we need to specify the number of change points n and the prior probability of their distribution $p(\vec{E}|I)$. We shall assume a flat uninformative prior on $\vec{E}$ meaning that $p(\vec{E}|I) \equiv p(n|I)$. Let m be the number of data points by which we define a regression problem of a straight-line segment. Of course $m_{\min} = 3$. If we have N data at hand, then the number of ways to choose one change point out of $N-2m$ is

$$Z(n=1) = N - 2m. \quad (43)$$

For each additional change point we must add a fit region of size m. Hence

$$Z(n) = \frac{(N-(n+1)m)!}{(N-(n+1)m-n)!n!} = 1/p(E|I). \quad (44)$$

Combining (43) and (45) we arrive at the evidence for an n-change point model:

$$p(\vec{d}|\vec{x},n,I) = \frac{1}{2}\left(\frac{1}{\pi}\right)^{(N-n-2)/2} \frac{1}{V_s(n+2,\gamma)} \frac{1}{2\ln\beta} \times \sum_{\vec{E}} \frac{n!(N-(n+1)m-n)!}{(N-(n+1)m)!} \frac{1}{\sqrt{\det \mathbf{Q}(\vec{E})}} \times \frac{\Gamma((N-n-2)/2)}{\{R(\vec{E})\}^{(N-n-2)/2}}. \quad (45)$$

The evaluation of (45) poses computational problems as may be seen by reference to (44). For large N and small to moderate n $Z(n)$ rises approximately proportional to N^n and the computational effort becomes prohibitive.

A possible way out of this difficulty is to replace the exact summation by an approximate one employing Monte Carlo techniques. The simplest approach would consist of replacing the exact sum over all possible distributions $\vec{E}$ by a uniform random sample $\vec{E}_{\mathrm{MC}}$ of size Z_{MC}, and the prior $1/Z(n)$ accordingly by $1/Z_{\mathrm{MC}}$. We shall not elaborate this problem further since data analysis in this paper will employ only the one change point model.

Model selection results

Recalling the results for the evidence in the three cases we see that the factor $1/2\ln\beta$ shows up in all cases and, therefore, does not influence the ranking of the evidences regardless of what numerical value we assign to β. In this case, we could have used the improper prior $1/\sigma$ from the very beginning since it enters the same way in all results. This is not true for the value of γ that specifies the prior range for the design ordinates $\vec{f}$. Our previous choice of γ does, therefore, affect the model probabilities. Everything is then uniquely specified. Given the marginal likelihoods $p(\vec{d}|\vec{x},M_j,I)$ and assuming equal prior probability for each model, an uninformative and not necessarily the most educated choice, we arrive at the probability for the individual model by

$$p(M_j|\vec{d},\vec{x},I) = \frac{p(\vec{d}|\vec{x},M_j,I)}{\sum_k p(\vec{d}|\vec{x},M_k,I)}. \quad (46)$$

Numerical values for the cherry blossom data set are presented in Table 1.

The numbers in the column 'residue' are a measure of the goodness of fit. In the case of the constant model this is given by $N\overline{\Delta d^2}$(25), in the case of the linear model by R (30). For the change point model, the case is more complicated and will be explained later. First of all, we see that the one change point model is singled out of the three with high selectivity. It is the only one that we shall consider further. A very interesting pattern arises when comparing model probabilities and goodness of fit. The last column in Table 1 indicates the degrees of freedom, e.g. the number of parameters that enter the corresponding model. We notice that the residue is a monotonically decreasing function of the number of parameters. The model probability rises monotonically and favors clearly the one change point model.

Table 2 extends the model comparison results to snow drop blossom data (see Fig. 4) and lime tree blossom data (see Fig. 6). Both data sets support the one change point model with high significance. However, while the lime tree results are associated with a small residual sum of squares of 5045, which is appreciably lower than in the cherry case, the snow drop data analysis yields the large residual sum of squares of 21202 for the one change point model. This indicates

Table 1 Model comparison results

	$p(M_j\|\vec{d},\vec{x},I)$	Residue	DOF
Constant model	0.074	8874	1
Linear model	0.104	8710	2
One change point	0.822	8118	4

The goodness of fit represented by the residual sum of squares increases monotonously with increasing model complexity. Although the built-in Ockham factor of Bayesian probability theory counter balances this effect a strong preference for the one change point model obtained. DOF, degrees of freedom.

Table 2 Model comparison results extended to lime tree and snow drop blossom data

	Limetree	Cherry	Snow drops
Constant model	0.018	0.074	0.0003
Linear model	0.028	0.104	0.021
One change point	0.954	0.822	0.978

The entries are the respective model probabilities.

that the class of models considered here is very appropriate for the lime tree and cherry data but probably poor for the snow drop data. We recognize the necessity of future investigation of higher-order change point models.

Predictions from the one change point model

So far we have evaluated which one out of the chosen group of models is best suited to describe the trend that is hidden in the data. We now proceed to calculate the trend itself for the one change point model. To this end it is useful to derive first the probability distribution for the single change point. This is given in terms of the marginal likelihood $p(\vec{d}|\vec{x},E,I)$ and the prior $p(E|I)$ by Bayes theorem:

$$p(E|\vec{d},\vec{x},I)=\frac{p(E|I)p(\vec{d}|\vec{x},E,I)}{p(\vec{d}|\vec{x},I)}. \tag{47}$$

Since $p(E|I)$ has previously been chosen flat, all the E-dependence is concentrated in the marginal likelihood. The marginal likelihood for the change point problems has been evaluated in the last section and is given by (45) with $n=1$. The properly normalized probability distribution of E is shown in Fig. 1 as full dots. Numerical values can be derived from the right-hand scale. The maximum probability is reached near 1985 and has a value of about 0.07. But the distribution $p(E|\vec{d},\vec{x},I)$ is very much smeared. The message is that every change point position will have to be taken into account for evaluating the trend because all of them have a non-negligible probability.

We now proceed to the evaluation of some function $T=\varphi(\vec{f},E,z)$ from the information contained in the data. For this goal we require the distribution $p(T|\vec{d},\vec{x},z,I)$ where z is the year for which we want to predict the value of T. The required distribution is again obtained via the marginalization rule:

$$\begin{aligned}p(T|\vec{d},\vec{x},z,I)&=\sum_E\int \mathrm{d}\vec{f}\,\mathrm{d}\sigma p(T,E,\vec{f},\sigma|\vec{d},\vec{x},I)\\&=\sum_E\int \mathrm{d}\vec{f}\,\mathrm{d}\sigma p(E,\vec{f},\sigma|\vec{d},\vec{x},z,I)\\&\quad\times p(T|E,\vec{f},\sigma,\vec{d},\vec{x},z,I).\end{aligned} \tag{48}$$

The first probability density under the integral is logically independent of z since the posterior distribution of the parameters $E,\vec{f},\sigma$ depends of course on $\vec{d}$ and $\vec{x}$ but is independent of some future estimation independent variable z. Similarly, the second factor is logically independent of $\vec{d},\vec{x},\sigma$ since T is a unique function of $\vec{f}$, E and z. In fact

$$p(T|\vec{f},E,z)=\delta(T-\varphi(\vec{f},E,z)). \tag{49}$$

This introduces a nasty structure into (48). But since we shall be content with calculating moments from (48), things simplify again by appropriate change of the order of integration:

$$\begin{aligned}\langle T^k\rangle&=\sum_E\int \mathrm{d}\sigma\,\mathrm{d}\vec{f}p(E,\vec{f},\sigma|\vec{d},\vec{x},I)\int \mathrm{d}T\,\delta(T-\varphi)T^k\\&=\sum_E\int \mathrm{d}\sigma\,\mathrm{d}\vec{f}p(E,\vec{f},\sigma|\vec{d},\vec{x},I)\{\varphi(\vec{f},E,z)\}^k.\end{aligned} \tag{50}$$

Further expansion of the posterior distribution of the parameters and repeated use of Bayes theorem until $p(E,\vec{f},\sigma|\vec{d},\vec{x},I)$ is expressed in terms of the full likelihood $p(\vec{d}|\vec{x},\vec{f},\sigma,E,I)$ and the priors $p(\vec{f}|\gamma,I)$ and $p(\sigma|I)$ finally yields:

$$\begin{aligned}\langle T^k\rangle=&\sum_E\frac{p(E|I)}{p(\vec{d}|\vec{x},I)}\frac{1}{2\ln\beta}\frac{1}{V_s(3,\gamma)}\\&\times\int\frac{\mathrm{d}\sigma}{\sigma}\int \mathrm{d}\vec{f}\varphi^k(\vec{f},E,z)p(\vec{d}|\vec{x},\vec{f},\sigma,E,I).\end{aligned} \tag{51}$$

We shall now consider in detail the cases that are important for the present problem. Note that the integral with prefactors in (51) is equal to the previously calculated (42) marginal likelihood $p(\vec{d}|\vec{x},E,I)$ for $k=0$. We shall show in the following that if φ^k is either a linear or a quadratic function in $\vec{f}$ then the integration in (51) leads to a function $\varphi_k(E,z)$ times the basic integral for $k=0$. Hence the expectation values of a piecewise linear function and its square, which we need to determine mean and variance, are given as the sum of all possible $\varphi_k(E,z)$ weighted with their respective posterior probability $p(E|\vec{d},\vec{x},I)$. This is indeed a transparent and intuitively plausible result.

Let

$$\varphi(z|\vec{f},E)=\begin{cases}f_1\frac{x_2-z}{x_2-x_1}+f_2\frac{z-x_1}{x_2-x_1}, & z\le x_2,\\ f_2\frac{x_3-z}{x_3-x_2}+f_3\frac{z-x_2}{x_3-x_2}, & z>x_2,\end{cases} \tag{52}$$

where x_1 and x_3 denote the initial and final data point abscissa and x_2 denotes the change point position. This function is linear in $\vec{f}$ and can be written as

$$\varphi(z|\vec{f},E)=\vec{b}^{\mathrm{T}}(z,E)\vec{f}. \tag{53}$$

An even simpler case arises if we consider the derivative of φ which, when inserted in (51) yields the rate of change of the trend for example in days per year for the present data. Since the derivative is

constant in $x_1 \le z \le x_2$ and also in $x_2 < z \le x_3$, we have

$$\varphi'(z|\vec{f},E) = \vec{c}^{\,\mathrm{T}}(E)\vec{f}. \tag{54}$$

Hence the only z-dependence arises from whether $z \le x_2$ or $z > x_2$ in this case. Both cases have the same $\vec{f}$ dependence and the inner, $\vec{f}$-integral I_f becomes (see (30))

$$I_f = \int \mathrm{d}\vec{f}\vec{f} \exp\left\{-\frac{1}{2\sigma^2}(\vec{f}-\vec{f}_0)^{\mathrm{T}}\mathbf{Q}(\vec{f}-\vec{f}_0) - \frac{R}{2\sigma^2}\right\}. \tag{55}$$

By substituting $\vec{g} = \vec{f} - \vec{f}_0$this can be shown to be equal to

$$I_f = \vec{f}_0 \int \mathrm{d}\vec{f} \exp\left\{-\frac{1}{2\sigma^2}(\vec{f}-\vec{f}_0)^{\mathrm{T}}\mathbf{Q}(\vec{f}-\vec{f}_0) - \frac{R}{2\sigma^2}\right\}. \tag{56}$$

This completes the proof for a linear $\vec{f}$-dependence of φ^k and leads to the explicit expression for the trend φ and its rate of change φ':

$$\varphi(z) = \langle \varphi(z|\vec{f},E)\rangle = \sum_E p(E|\vec{d},\vec{x},I)\ \vec{b}^{\,\mathrm{T}}(z,E)\vec{f}_0(E), \tag{57}$$

$$\varphi'(z) = \langle \varphi'(z|\vec{f},E)\rangle = \sum_E p(E|\vec{d},\vec{x},I)\ \vec{c}^{\,\mathrm{T}}(E)\vec{f}_0(E). \tag{58}$$

In order to evaluate also the uncertainties associated with these estimates we must finally consider the squares of (53) and (54). We treat the case (53) explicitly since it is slightly more complicated due to the $\vec{z}$-dependence of $\vec{b}$:

$$(\vec{b}^{\,\mathrm{T}}(E,z)\vec{f})^2 = (b_1f_1 + b_2f_2 + b_3f_3)^2 = \vec{f}^{\,\mathrm{T}}\,\mathbf{B}\vec{f}. \tag{59}$$

The elements B_{ik} of matrix $\mathbf{B}$ are given by $B_{ik} = b_ib_k$. In order to perform the $\vec{f}$-integration in the inner integral (51) we substitute $\vec{g} = (\vec{f} - \vec{f}_0)/\sigma$. This yields

$$\begin{aligned} I_{\vec{f}} =& \sigma^3 \int \mathrm{d}\vec{g}(\sigma\vec{g}+\vec{f}_0)^{\mathrm{T}}\mathbf{B}(\sigma\vec{g}+\vec{f}_0) \\ &\times \exp\left\{-\frac{1}{2}\vec{g}^{\mathrm{T}}\,\mathbf{Q}\vec{g}\right\} \\ =& \left\{\sigma^5 \mathrm{Trace}(\mathbf{Q}^{-1}\mathbf{B}) + \sigma^3 \vec{f}_0^{\,\mathrm{T}}\,\mathbf{B}\vec{f}_0\right\} \\ &\times \int \mathrm{d}\vec{g} \exp\left\{-\frac{1}{2}\vec{g}^{\mathrm{T}}\,\mathbf{Q}\vec{g}\right\}. \end{aligned} \tag{60}$$

By reference to Eqn. (24) we derive that the remaining two σ-integrations are related by

$$\begin{aligned} &\int_0^\infty \frac{\mathrm{d}\sigma}{\sigma}\ \frac{1}{\sigma^{N-3}}\ \exp\left\{-\frac{R}{2\sigma^2}\right\} \\ &\quad = \frac{N-5}{R}\int_0^\infty \frac{\mathrm{d}\sigma}{\sigma}\ \frac{1}{\sigma^{N-5}}\ \exp\left\{-\frac{R}{2\sigma^2}\right\}. \end{aligned} \tag{61}$$

Collecting terms we obtain

$$\begin{aligned} \langle(\vec{b}^{\,\mathrm{T}}(E,z)\vec{f})^2\rangle =& \sum_E p(E|\vec{d},\vec{x},I) \\ &\times \left\{\mathrm{Trace}(\mathbf{Q}^{-1}\mathbf{B})\,\frac{R}{N-5} + \vec{f}_0^{\,\mathrm{T}}\,\mathbf{B}\,\vec{f}_0\right\}. \end{aligned} \tag{62}$$

The evaluation of (62) is complicated by the fact that matrix $\mathbf{B}$ depends on E and z while $\mathbf{Q}$ and R depend on E only. The problem can be simplified by making the z-dependence of $\mathbf{B}$ explicit. We rewrite (52) as

$$\begin{aligned} &\varphi(z|\vec{f},E) \\ &= \begin{cases} \frac{x_2}{x_2-x_1}f_1 - \frac{x_1}{x_2-x_1}f_2 + z\left(-\frac{1}{x_2-x_1}f_1 + \frac{1}{x_2-x_1}f_2\right), & z \le x_2, \\ \frac{x_3}{x_3-x_2}f_2 - \frac{x_2}{x_3-x_2}f_3 + z\left(-\frac{1}{x_3-x_2}f_1 + \frac{1}{x_3-x_2}f_2\right), & z > x_2. \end{cases} \end{aligned} \tag{63}$$

By appropriate obvious definitions of vectors $\vec{\alpha}$ and $\vec{a}$, φ^2 can be expressed as

$$\begin{aligned} \varphi^2(z|\vec{f},E) =& (\alpha_1f_1 + \alpha_2f_2 + \alpha_3f_3)^2 \\ &+ z^2(a_1f_1 + a_2f_2 + a_3f_3)^2 \\ &+ 2z(\alpha_1f_1 + \alpha_2f_2 + \alpha_3f_3) \\ &\times (a_1f_1 + a_2f_2 + a_3f_3). \end{aligned} \tag{64}$$

From (64) we derive the matrix representation

$$\varphi^2(z|\vec{f},E) = \vec{f}^{\,\mathrm{T}}\ \mathbf{B}_0\vec{f} + z\vec{f}^{\,\mathrm{T}}\ \mathbf{B}_z\vec{f} + z^2\vec{f}^{\,\mathrm{T}}\ \mathbf{B}_{zz}\vec{f}, \tag{65}$$

$$(B_0)_{ik} = \alpha_i\alpha_k, (B_z)_{ik} = \alpha_ia_k + \alpha_ka_i, (B_{zz})_{ik} = a_ia_k. \tag{66}$$

Matrices $\mathbf{B}_0$, $\mathbf{B}_z$ and $\mathbf{B}_{zz}$ depend now only on E. Hence, if all quantities are calculated as a function of E and are stored, evaluation of (62) for different z is computationally cheap. This completes our analysis.

The result for the average functional behavior with the associated uncertainty range is displayed in Fig. 2. The reader might wonder at this stage why we end up with relatively narrow confidence ranges compared to the strong scatter of the data. The answer is simply that we estimate essentially only four parameters $(\vec{f},E)$ from a data set of about one hundred points. This yields obviously a rather stiff function.

The rate of change is displayed in Fig. 3. There is essentially zero change over most of the century with the deviation starting about 1985 and a current rate of change of $-0.6\,\mathrm{days\,yr^{-1}}$. This rate is associated with an uncertainty of $\pm 0.5\,\mathrm{days\,yr^{-1}}$. A final word about extrapolation into the future is in order. The predictive function is a superposition of straight-line segments weighted with respective probabilities $p(E|\vec{d},\vec{x},I)$. Hence, it is itself a straight line between successive years of observation. The function in Fig. 2 extrapolates, therefore, linearly and the rate correspondingly at a constant level. Extrapolation in Fig. 2 into the region

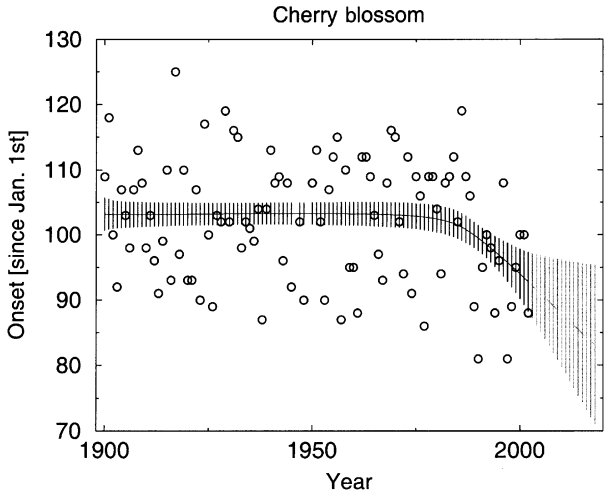

Fig. 2 The data and the average functional behaviour estimated from the one change point model. Note the rapidly widening confidence range in the extrapolation region.

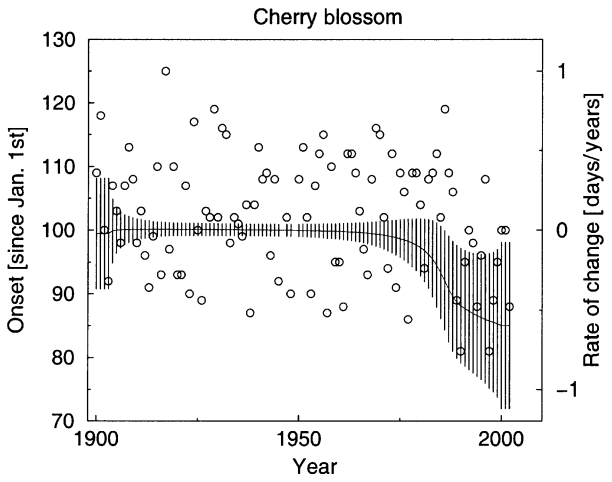

Fig. 3 The rate of change (trend) in days per year (right-hand scale).

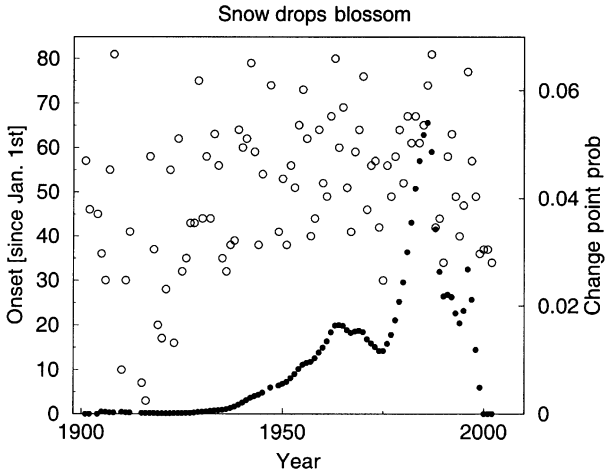

Fig. 4 Observations on snow drop blossom (open circles) and the change point probability for the one change point model (full dots).

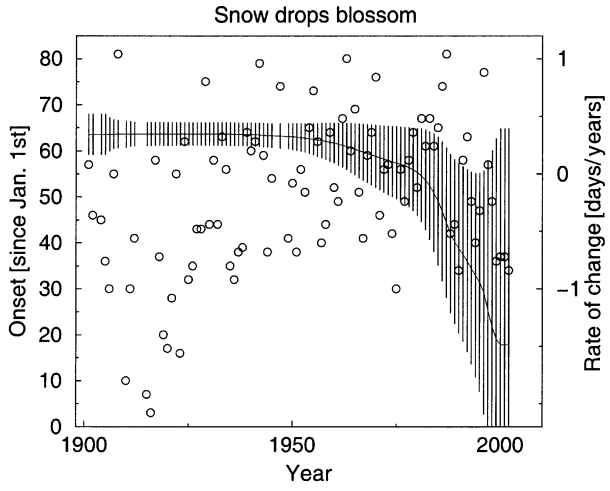

Fig. 5 The rate of change of the onset of snow drop blossom in days per year. Data as in Fig. 4.

that is no longer supported by observational data exhibits a rapid widening of the confidence range. This is the price for extrapolation and supports the common notion that predictions are difficult, in particular, when the future is concerned. Bayesian probability theory provides a beautiful quantification of this statement.

Discussion

The model comparison option of the Bayesian approach as described above was used to compare three different types of models (constant, linear, and one change point) for the analysis of the Cherry blossom (results in Figs 1–3). Figures 4–7 provide analogous results of the one change point model for the other species' time series, blossom of snowdrop and lime tree.

Comparison of the results for the three phases analyzed

We know that spring phenology strongly depends on the temperature of the previous months, e.g. blossom of *G. nivalis* correlates best with the mean temperature of Jan–Mar ($T_{\text{Jan–Mar}}$)($r = -0.747$, $n = 66$, $P < 0.0001$), blossom of *P. avium* with ($T_{\text{Feb–Apr}}$)($r = -0.807$, $n = 66$, $P < 0.0001$), and of *T. platyphyllos* with ($T_{\text{Mar–May}}$)($r = -0.798$, $n = 65$, $P < 0.0001$) (data from the climate station at Geisenheim). As most of the global temperature rise in the 20th century has occurred during two periods (1910–1945 and 1976–2000), it is of interest for the study of climate change impacts in

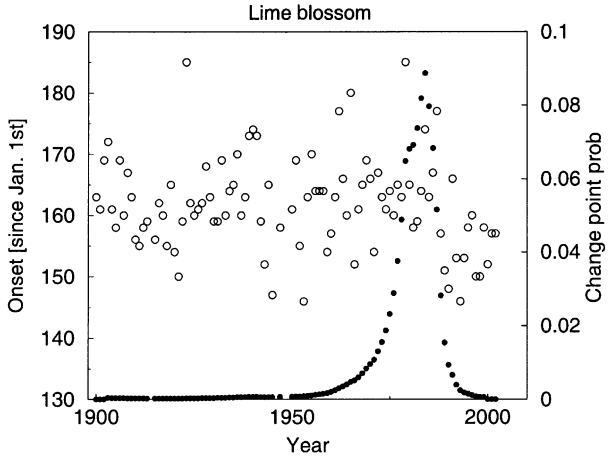

Fig. 6 Observations on lime tree blossom (open circles) and the change point probability for the one change point model (full dots).

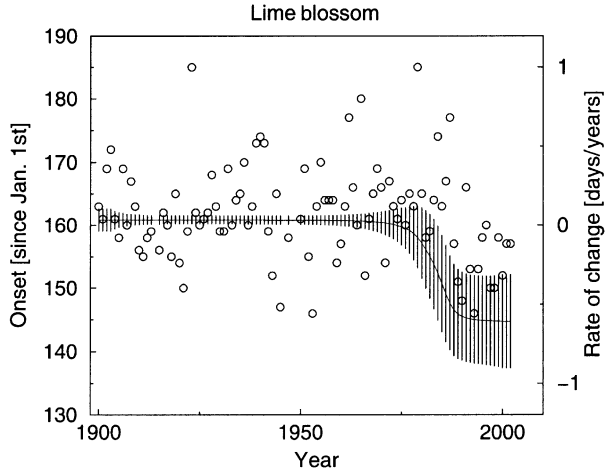

Fig. 7 The rate of change of the onset of lime tree blossom in days per year. Data as in Fig. 6.

phenology, if phenological time series have experienced changes in comparable time frames. The model comparison option of the Bayesian theory here allows the evaluation of which one out of the chosen group of models is best suited to describe the trend that is hidden in the data.

For the three time series at Geisenheim, it turns out that the most likely model is the one change point model. The resulting change point probability has its maximum near 1985 (~ 0.07) for cherry blossom (Fig. 1) and, with a quite similar distribution, a maximum near 1984 (0.09) for lime tree blossom (Fig. 6). The change point probability distribution of snowdrop blossom is much broader, even multi-modal, with a maximum near 1986 (0.055) and a secondary maximum (0.02) near 1964 (Fig. 4).

The rate of change determined by the one change point model reaches $-1.5\,\text{days}\,\text{yr}^{-1}$ in 2002 for snowdrop blossom (Fig. 5), and around $-0.6\,\text{days}\,\text{yr}^{-1}$ both for cherry (Fig. 3) and lime tree blossom (Fig. 7).

Several studies report time of season differences with highest advances often in early spring, and notable advances of succeeding phenophases in full spring and early summer (e.g. Bradley *et al.*, 1999; Defila & Clot, 2001, Menzel *et al.*, 2001). Sparks & Smithers (2002) suggest higher temperature changes in early season as the reason for this differentiation. The relative order of changes revealed for these three species' blossom series at Geisenheim is well in accordance with other results: Regression coefficients from mean anomaly curves of Germany (1951–2000, Menzel, 2003) ranged from $-0.25\,\text{days}\,\text{yr}^{-1}$ for snowdrop, to $-0.14\,\text{days}\,\text{yr}^{-1}$ (lime tree) and $-0.09\,\text{days}\,\text{yr}^{-1}$ (cherry). It is evident that their absolute magnitude for the 1951–2000 period is much smaller than rates of change determined for the year 2002 with the one change point model. A recalculation of the regressions in Menzel (2003) for the 1985–2000 period revealed advances of $-1.42\,\text{days}\,\text{yr}^{-1}$ for snowdrop blossom, $-0.67\,\text{days}\,\text{yr}^{-1}$ for cherry blossom and $-0.86\,\text{days}\,\text{yr}^{-1}$ for lime tree blossom. These average rates of change for Germany mirror the results of our new method based on Bayesian theory. Over most of the century, there is essentially zero change; from mid-1980s onwards, the rate of change is negative for cherry (Fig. 3) and lime tree blossom (Fig. 7). This finding is in accordance with results of other studies (e.g. Scheifinger *et al.*, 2002; Chmielewski & Rötzer, 2002) that describe a discontinuous shift towards earlier occurrence dates after the late 1980s and almost no trends before that date. However, the rate of change in snowdrop blossom (Fig. 5) indicates advancing onset not before 1992; although the rates of change are relatively high in the 1990s they are associated with a considerable uncertainty range.

Comparison with the traditional statistical approach

The proposed new method of selecting models to describe the trend in phenological time series has major advantages compared to the traditional statistical approach of linear regression. Fig. 8 displays the slopes of the linear regression and the corresponding significance by the Mann–Kendall trend test for all possible combinations of starting (x-axis) and ending year (y-axis) with 10 or more years of observation of cherry blossom at Geisenheim. This example clearly demonstrates that the resulting trends start to be representative in the temporal scale when more than 30 years are included. The main obstacle of this traditional approach is evident: The rate of change

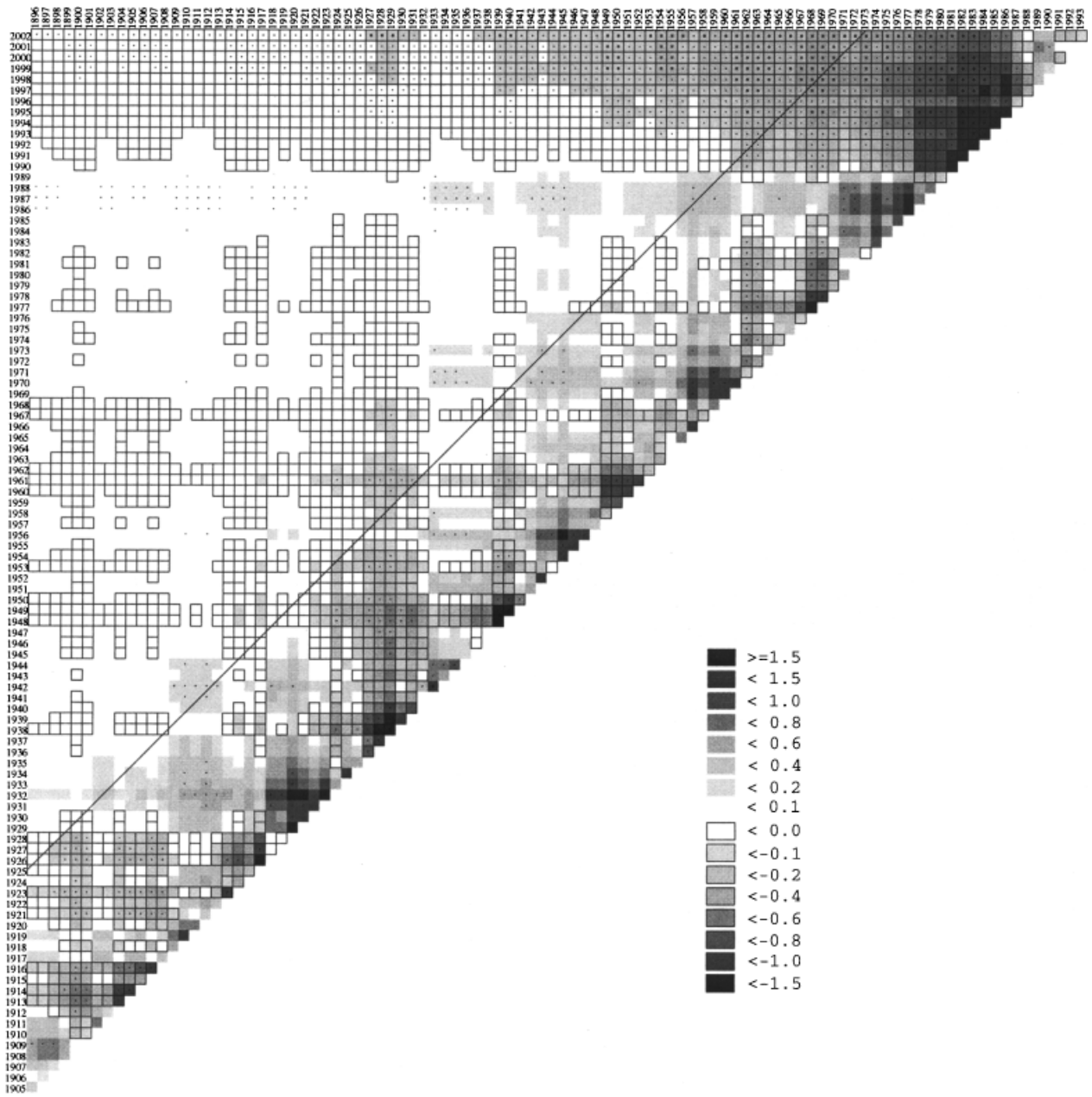

Fig. 8 Trend matrix: Linear regression coefficients and their significance by the Mann–Kendall trend test for *Prunus avium* blossom time series at Geisenheim/Germany. (.1 Q1 ≥ 1…2, 1 Q1 > 2)

strongly depends on the time period and no distinct rate of change for single years can be given. However, all time series ending in 1989 and later reveal advancing trends, especially if they start in the second half of the 20th century. The corresponding results for lime tree (trend matrix not displayed) are very similar to Fig. 8. The trend matrix for snowdrop (not shown) is even more heterogeneous because snowdrop blossom was observed extremely early in several years between 1910 and 1923, and thus nearly all time series starting before 1920 reveal delayed blossom.

One of the traditional ways to determine possible change points is by a modified version of the Mann–Kendall test (e.g. Böhm *et al.*, 2001) based on the normalized values of Kendall's Q (see Rapp, 2002):

$$Q = \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \operatorname{sgn}(y_j - y_i)}{\sqrt{\frac{1}{18}[N(N-1)(2N+5) - \sum_{l} b_l(b_l-1)(2b_l-1)(b_l+5)]}}. \quad (67)$$

N is the length of the original time series. Summation is extended over all $1/2N\ (N-1)$ pairs of data (y_j, y_i) with $i<j$. b_l is the number of identical observational data y_l. The Q values are determined progressively for the $n-1$ subseries of the original time series as well as for the backward original time series (regressive). Intersections of the progressive and regressive Kendall's Q series approximate the change points of the original series. Their significance is ranked according to the excess over suitably chosen significance thresholds. With this traditional approach, change points were determined around 1929/1930 for blossom of snowdrop, around 1989/1990/1993 for blossom of cherry, and even three possible change points 1905–1909, 1923–1926, and 1989/1990/1992 for blossom of lime tree. The results for cherry blossom match almost the maximum change point probability around 1985.

Conclusions

Regional studies of plant phenology, often using phenological network data, are extremely important for assessing the impacts of global change as they can shed light on regional peculiarities. However, this information is only revealed if the spatial variability can be separated from inherent problems in the temporal significance due to different underlying time periods. This latter point is the crucial limitation of the currently used linear regression approach. In contrast, Bayesian concepts allow the comparison of different models to describe the functional behavior of phenological time series and provide even annual predictions for the change point probability and rates of changes. Thus, this new approach will allow an intensified comparison of regional changes in phenology. Our next step will consist of a combined analysis of phenological and temperature time series in order to investigate whether temperature changes support phenological change points. If so, the pooling of temperature and phenological data is expected to provide times series descriptions of higher precision for both.

Acknowledgements

We thank the German Weather Service (Deutscher Wetterdienst) for providing the phenological data analyzed in this work.

References

Ahas R (1999) Long-term phyto-, ornitho- and ichthyophenological time-series analyses in Estonia. *International Journal of Biometeorology*, **42**, 119–123.

Ahas R, Aasa A, Menzel A *et al*. (2002) Changes in European spring phenology. *International Journal of Climatology*, **22**, 1727–1738.

Barnett TP, Hasselmann K, Chelliah M *et al*. (1999) Detection and attribution of recent climate change: a status report. *Bulletin of the American Meteorological Society*, **80**, 2631–2659.

Bayes T (1763) An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions*, **53**, 370.

Beaubien EG, Freeland HJ (2000) Spring phenology trends in Alberta, Canada: links to ocean temperature. *International Journal of Biometeorology*, **44**, 53–59.

Berliner LM, Levine RA, Shea DJ (2000) Bayesian climate change assessment. *Journal of Climate*, **13**, 3805–3820.

Böhm R, Auer I, Brunetti M *et al*. (2001) Regional temperature variability in the European Alps: 1760–1998 from homogenized instrumental time series. *International Journal of Climatology*, **21**, 1779–1801.

Box GEP, Tiao GC (1973) *Bayesian Inference in Statistical Analysis*. Addison-Wesley, Reading, MA.

Bradley NL, Leopold AC, Ross J *et al*. (1999) Phenological changes reflect climate change in Wisconsin. *Proceedings of the National Academy of Sciences USA Ecology*, **96**, 9701–9704.

Chmielewski FM, 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.

Defila C, Clot B (2001) Phytophenological trends in Switzerland. *International Journal of Biometeorology*, **45**, 203–207.

Dose V (2003) Bayesian inference in physics: case studies. *Reports on Progress in Physics*, **66**, 1421–1461.

Fitter AH, Fitter RSR (2002) Rapid changes in flowering time in British plants. *Science*, **296**, 1689–1691.

Freitag E (1987) *Auswirkungen von Klimaänderungen auf den Entwicklungsrhythmus der Pflanzen für historische Zeiträume*. Schlussbericht zum BMFT-Förderungsvorhaben KF2008, Deutscher Wetterdienst, Offenbach am Main.

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

Gregory PC, Loredo TJ (1992) A new method for the detection of a periodic signal of unknown shape and period. *Astronomical Journal*, **398**, 146–168.

Hasselmann K (1998) Conventional and Bayesian approach to climate-change detection and attribution. *Quarterly Journal of the Royal Meteorological Society*, **124**, 2541–2565.

Hobbs BF (1997) Bayesian methods for analysing climate change and water resource uncertainties. *Journal of Environmental Management*, **49**, 53–72.

IPCC (2001) Climate change 2001: the scientific basis. In: *Third Assessment report of the Working Group I* (eds Houghton JT *et al*.), Cambridge University Press, Cambridge.

Jaynes ET (1957a) Information theory and statistical mechanics. *Phys. Rev.*, **106**, 620–630.

Jaynes ET (1957b) Information theory and statistical mechanics II. *Physical Review*, **108**, 171–190.

Jones GV, Davis RE (2000) Climate influences on grapevine phenology, grape composition, and wine production and quality for Bordeaux, France. *American Journal of Enology and Viticulture*, **51**, 249–261.

Kapur JN, Kesavan HK (1992) *Entropy Optimization Principles with Applications*. Academic, Boston.

Katz RW (2002) Techniques for estimating uncertainty in climate change scenarios and impact studies. *Climate Research*, **20**, 167–185.

Kozlov M, Berlina N (2002) Decline in the length of the summer season on the Kola peninsula, Russia. *Climatic Change*, **54**, 387–398.

Lauscher F (1978) Neue Analysen ältester und neuerer phänologischer Reihen. *Archiv fur Meteorologie Geophsik und Bioklimatologie Series B*, **26**, 373–385.

Lauscher F (1983) Weinlese in Frankreich und Jahrestemperatur in Paris seit 1453. *Wetter und Leben*, **35**, 39–42.

Leonard T, Hsu JSJ (1999) *Bayesian Methods: An Analysis for Statisticians and Interdisciplinary Researchers*. Cambridge University Press, Cambridge.

Leroy SS (1998) Detecting climate signals: some Bayesian aspects. *Journal of Climate*, **11**, 640–651.

Magnuson JJ, Robertson DM, Benson BJ *et al*. (2000) Historical trends in lake and river ice cover in the northern hemisphere. *Science*, **289**, 1743–1746.

Menzel A (2000) Trends in phenological phases in Europe between 1951 and 1996. *International Journal of Biometeorology*, **44**, 76–81.

Menzel A (2003) Plant phenological anomalies in Germany and their relation to air temperature and NAO. *Climatic Change*, **57**, 243–263.

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

Menzel A, Estrella N, Fabian P (2001) Spatial and temporal variability of the phenological seasons in Germany from 1951–1996. *Global Change Biology*, **7**, 657–666.

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

Peñuelas J, Filella I, Comas P (2002) Changed plant and animal life cycles from 1952 to 2000 in the Mediterranean region. *Global Change Biology*, **8**, 531–544.

Perreault L, Bernier J, Bobée B *et al*. (2000) Bayesian change-point analysis in hydrometeorological time series. Part 2. Comparison of change-point models and forecasting. *Journal of Hydrology*, **235**, 242–263.

Rapp J (2002) *Konzeption, Problematik und Ergebnisse klimatologischer Trendanalysen für Europa und Deutschland*, Berichte des Deutschen Wetterdienstes 212.

Robertson I, Lucy D, Baxter L *et al*. (1999) A kernel-based Bayesian approach to climatic reconstruction. *Holocene*, **9**, 495–500.

Root TL, Price JT, Hall KR *et al*. (2003) Fingerprints of global warming on wild animals and plants. *Nature*, **421**, 57–60.

Sagarin R (2001) False estimates of the advance of spring. *Nature*, **414**, 600.

Sagarin R, Micheli F (2001) Climate change in nontraditional data sets. *Science*, **294**, 811.

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.

Schlittgen R, Streitberg BHJ (1999) *Zeitreihenanalyse*. R. Oldenburg Verlag, München.

Schnelle F (1950) Hundert Jahre phänologische Beobachtungen im Rhein–Main–Gebiet, 1841–1859, 1867–1947. *Meteorologische Rundschau*, **3**, 150–156.

Schnelle F, Witterstein F (1952) *Beiträge zur Phänologie Deutschlands, II. Tabellen Phänologischer Einzelwerte von etwa 500 Stationen der Jahre 1936–1944* (Berichte des Deutschen Wetterdienstes in der US-Zone Nr. 41.

Schwartz MD, Reiter BE (2000) Changes in North American Spring. *International Journal of Climatology*, **20**, 929–932.

Sivia DS (1996) *Data Analysis – A Bayesian Tutorial*. Clarendon, Oxford.

Sparks TH, Carey PD (1995) The responses of species to climate over two centuries: an analysis of the Marsham phenological record. *Journal of Ecology*, **83**, 321–329.

Sparks TH, Menzel A (2002) Observed changes in seasons: an overview. *Journal of Climatology*, **22**, 1715–1725.

Sparks TH, Smithers RJ (2002) Is spring getting earlier? *Weather*, **57**, 157–166.

Tol RSJ, De Vos AF (1998) A Baysian statistical analysis of the enhanced greenhouse effect. *Climatic Change*, **38**, 87–112.

Tribus M (1969) *Rational Descriptions, Decisions and Designs*. Pergamon, Elmsford.

Walther GR, Post E, Convey P *et al*. (2002) Ecological responses to recent climate change. *Nature*, **416**, 389–395.