TECHNISCHE UNIVERSITÄT MÜNCHEN
Fachgebiet für Ökoklimatologie

Climate change detection in natural systems by Bayesian methods

Christoph Schleip

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender : Univ.-Prof. Dr. A. Fischer

Prüfer der Dissertation:

1. Univ.-Prof. Dr. A. Menzel
2. apl. Prof. Dr. K. F. Auerswald
3. Visiting Prof. T. H. Sparks, Ph. D.
   University of Liverpool / UK (schriftliche Beurteilung)

Die Dissertation wurde am 20.01.2009 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 15.06.2009 angenommen.

Abstract

Aims

The present PhD thesis focuses on climate change detection in natural systems by Bayesian analysis. In particular it seeks to detect changes in temperature and biological systems (vegetation; phenology of plants) and intends to improve the understanding of responses to climate change with the help of the Bayesian analysis. This PhD is segmented into three leading questions: (1) What are the advantages and disadvantages of the Bayesian approach compared to conventional statistical methods when analysing climate change impacts on natural systems? (2) Which potentials of the Bayesian approach (such as model probabilities, functional behaviours, model averaged rates of change, confidence intervals and time spans of elevated change point probability) contribute to an accurate assessment of climate change impacts on natural systems? (3) What kind of biological insights into the triggering climate change factor temperature and its influence on phenology can be gained by the Bayesian concept?

Material and Methods

Long-term (>30 years) plant phenological time series of different species within Europe and temperature time series adjacent to the respective phenological station are analysed. In one study also global atmospheric temperature time series from the surface up to the stratosphere are used. We use a Bayesian approach and employ three different models to describe the time series. The constant model represents the hypothesis of no change with a functional behaviour constant in time and an associated zero rate of change. The linear model assumes a linear change of the observed phenomenon with an associated constant rate of change. The change point model allows for a time-varying trend and thus allows the identification of nonlinear changes. Its development starts from triangular functions, hence two linear segments, which match at particular change point choices. Although the endpoints of the time series remain fixed in the subsequent calculations, the intermediate change point can be at any year. If N is the number of entries on the time scale, there are N-2 possibilities (excluding the endpoints) for the change point position. Change point probability distributions exhibit the change point probabilities as a function of time for a temperature or a phenological time series. Since the change point probability distribution is extended over several years, it does not make sense to select the maximum-likelihood triangular function for the time series model. Instead we employed the Bayesian marginalization rule to integrate out the change point variable from the model function. This extremely important rule removes 'nuisance' parameters from a Bayesian calculation.

The result of the Bayesian marginalization rule is a superposition of all possible triangular functions for the present data weighted by their respective change point probability that leads to the change point function. An analogous procedure is applied to the model averaged trend estimation. The rigorous application of Bayesian probability theory describes that the proper functional behaviour and the proper trend are obtained by superposition of a constant, a linear function and the one change point model function again weighted with their respective model probabilities.

Results and Discussion

The great advantage of Bayesian analysis is that it considers the inability to prefer one model against another. Compared to the commonly used linear regression approach, we are able to provide change point probabilities and 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. Thus Bayesian model averaged results are more informative than results based on single model approaches.

With the help of the Bayesian approach we detect an earlier start of spring plant phenology in the last five decades and more heterogeneous changes in autumn. The change point model provides the best description of the data for all seasons of the year. High probabilities for this specific model reveal Europe-wide nonlinear changes in the examined phenology. The dominance of the change point model is most pronounced for phases in summer to late autumn.

Change point distributions of Norway spruce bud burst exhibit the highest coherence with change point distributions of temperatures at the end of February and in April and May. Since the beginning of the 1980s, April and May temperature rates of change increase to positive values (warming) and Norway spruce bud burst time series reveal an enhanced advancing of the phenological phase.

In the context of the last 250 years the end of the 20th century represents a period with unique major increases in temperatures of all seasons and earlier grape harvest phenology as derived from model averaged trends. Furthermore a study of atmospheric temperature data from the surface up to the stratosphere verifies with the Bayesian approach predominant nonlinear temperature changes in nearly all pressure levels and underlines the importance of alternatives to the often used linear models.

Conclusion

The Bayesian approach offers new possibilities including robust model selection for time series description, assessment of functional behaviour and rates of change with uncertainty margins as well as evaluation of coherent or independent treatment of time series of triggering parameters and affected systems.

With our practical employment of the Bayesian concept we enhance the richness of biological insights. The diffentiation of temporal and spatial changes in phenology and temperature time series as well as the potential to judge and incorporate outputs of competing mathematical models are an attractive contribution to the studies of climate change and of its multiple impacts.

Zusammenfassung

Zielsetzung

Die vorliegende Dissertation befasst sich mit der Detektion des Klimawandels in natürlichen Systemen mit Hilfe der Bayes'schen Analyse. Es sollen die Änderungen in Temperaturzeitreihen und phänologischen Zeitreihen analysiert werden, um das Verständnis für die Reaktionen von natürlichen Systemen auf den Klimawandel zu vertiefen. Die Promotionsarbeit wird durch folgende drei Fragen untergliedert:

(1) Was sind die Vor- und Nachteile der Bayes'schen Statistik im Vergleich zu herkömmlichen statistischen Methoden bei der hier vorgestellten Anwendung?

(2) Welche Resultate der Bayes'schen Analyse (wie zum Beispiel Modellwahrscheinlichkeiten, Funktionsschätzungen, modellgemittelte Änderungsraten, Vertrauensintervalle und Zeitspannen erhöhter „Change point“-Wahrscheinlichkeit) tragen in besonderem Maße zu der Detektion des Klimawandels bei?

(3) Welche Einsichten über ausschlaggebende Temperaturen und ihren Einfluss auf die Phänologie können mit dem Bayes'schen Ansatz gewonnen werden?

Material und Methoden

In der vorliegenden Arbeit wurden langfristige (>30jährige) phänologische Zeitreihen verschiedener Pflanzenarten sowie dazugehörige Temperaturzeitreihen in Europa analysiert. Außerdem wurden globale atmosphärische Temperaturen von der Troposphäre bis zur Stratosphäre analysiert.

Die vorliegende Bayes'sche Analyse berücksichtigt drei verschiedene Modelle zur Beschreibung der Zeitreihen. Das konstante Modell verkörpert die Hypothese, dass über den untersuchten Zeitraum keine Veränderung des beobachteten Phänomens eingetreten ist. Dies wird durch eine Funktion repräsentiert, die in der Zeit konstant bleibt und eine Änderungsrate mit dem Wert Null besitzt. Das lineare Modell beschreibt eine lineare Änderung des beobachteten Phänomens mit einer konstanten Änderungsrate. Das so genannte „Change point“-Modell ermöglicht eine Identifizierung von zeitlich veränderlichen Änderungsraten und kann somit nichtlineare Veränderungen beschreiben. Das „Change point“-Modell wird aus triangulären Funktionen entwickelt, die jeweils aus zwei linearen Segmenten bestehen. Die beiden linearen Segmente jeder Funktion sind durch die Endpunkte des Datensatzes fixiert und berühren sich an den jeweiligen Änderungspunkten (Change point). Wenn N Zeitunterteilungen betrachtet werden, gibt es (mit Ausnahme der Endpunkte) für die Lage des Änderungspunktes N-2 Möglichkeiten. „Change point“-Wahrscheinlichkeitsverteilungen zeigen für Temperatur- oder phänologische Zeitreihen die Wahrscheinlichkeit für jede mögliche Lage eines Änderungspunktes als Funktion der Zeit. Oftmals kann nicht eine einzelne trianguläre Funktion sämtliche Daten alleingültig und umfassend erklären. In diesen Fällen erlaubt es die Bayes’sche Analyse, die wahrscheinlichste mathematische Beschreibung zu erzielen. Die Marginalisierung ist dabei der entscheidende Schritt innerhalb des Bayes’schen Theorems.

Durch die Marginalisierung werden "störende" Parameter wie zum Beispiel. der Änderungspunkt aus der Bayes'schen Berechnung entfernt. Das Ergebnis der Marginalisierung ist eine Überlagerung aller möglichen dreieckigen Funktionen gewichtet mit ihrer jeweiligen Änderungspunktwahrscheinlichkeit. Ein analoges Verfahren wird zur

Schätzung der modellgemittelten Änderungsraten verwendet. Mit der Anwendung des Bayes'schen Theorems kann schließlich durch eine Überlagerung von allen konstanten, linearen und „Change point“-Funktionen, gewichtet mit ihren jeweiligen Modellwahrscheinlichkeiten, die wahrscheinlichste Funktion für die untersuchte Zeitreihe dargestellt werden.

Ergebnisse und Diskussion

Der große Vorteil der Bayes'schen Analyse liegt in dem Angebot eines Modellvergleichs sowie der integrativen Verwertung aller drei aufgeführten Lösungsmodelle. Im Vergleich zu dem üblichen linearen Regressionansatz sind wir in der Lage, „Change point“-Wahrscheinlichkeitverteilungen und modellgemittelte Änderungsraten in jährlicher Auflösung zu liefern. Dies hilft uns dabei, Diskontinuitäten zu beschreiben und die Ergebnisse zu quantifizieren sowie die Richtung und Geschwindigkeit der Veränderungen zu bestimmen. Dadurch sind Bayes'sche Modellergebnisse gegenüber einem einzelnen Modellansatz korrekter und informativer.

Mit Hilfe der Bayes'schen Analyse stellen wir in fast allen hier vorgestellten Studien einen früherern Beginn der Phänophasen im Frühling und heterogenere Veränderungen der Herbstphasen in den letzten fünf Jahrzehnten fest. Das „Change point“-Modell lieferte die beste Beschreibung der Daten für alle Jahreszeiten. Dieser Vorteil des „Change point“-Modells ist am stärksten für Sommer- und Spätherbstphasen ausgeprägt.

Die „Change point“-Verteilungskurven des Fichtenknospenaufbruchs und der Temperaturen von Ende Februar, April und Mai zeigten eine hohe Kohärenz. Seit Anfang der 1980er Jahre beschleunigt sich der Temperaturanstieg im April und Mai und verfrühen sich die jährlichen Termine für den Fichtenknospenaufbruch. Eine weitere, 250 Jahre zurückreichende Studie zeigte für das Ende des 20. Jahrhunderts einen einzigartigen Anstieg der Temperaturen sämtlicher Jahreszeiten und eine zunehmende Vorverlegung der Weinlese. Diese Erkenntnis wurde aus gemittelten Änderungsraten des Bayes'schen Modells abgeleitet. Außerdem gelingt es uns mit Hilfe der Bayes'schen Analyse dominierende nichtlineare Temperaturveränderungen global und in fast allen atmosphärischen Druckschichten von der Troposphäre bis zu der Stratosphäre nachzuweisen.

Schlussfolgerung

Die Bayes'sche Analyse bietet mit einer Modellauswahl für Zeitreihenbeschreibungen und der Ermittlung der Funktionsverläufe und Änderungsraten mit entsprechenden Vertrauensgrenzen wertvolle Analysemöglichkeiten. Sie eignet sich für die Analyse kohärenter oder unabhängiger Zeitreihen der Schlüsselfaktoren betroffener Systeme. Die Möglichkeit einer differenzierten Untersuchung zeitlicher und räumlicher Veränderungen von Temperaturentwicklungen und phänologischem Geschehen und das Angebot eines bewertenden Vergleichs unterschiedlicher mathematischer Lösungsmodelle sind ein attraktiver Beitrag zur Erforschung der Formen des Klimawandels und seiner vielfältigen Einflüsse.

Contents

List of Figures

List of Tables

1 General Introduction

1.1 Motivation and Problem Description

Growing general concern about Global Change and its impacts on ecosystems and society has increased the awareness of the need of accurate climate information in the past, present and the future and its impacts on diverse systems. Observed changes have not been globally uniform, in fact changes varied over regions (Walther et al., 2002; Root et al., 2003; Parmesan and Yohe, 2003). Global surface temperature has increased by an estimated 0.74°C over the past century, a change that is widely believed to result primarily from the effects of anthropogenic emissions of carbon dioxide and other greenhouse gases (Trenberth et al. 2007). Recent regional climate changes, particularly temperature increases, have already affected many physical and biological systems (Lucht et al., 2002; Rosenzweig et al., 2007, 2008). Many physical changes have been attributed to this warming, including sea level rise, melting of glaciers and ice sheets, decreased snow and ice cover, increased depth to permafrost and changes in patterns of wind, temperature, and precipitation (summarized in Rosenzweig et al., 2007). Numerous studies have sought evidence of such biological effects in nature. Several recent papers summarize the results of these studies and conclude that biological effects are already evident and have affected numerous taxa in different geographical areas (Walther et al., 2002; Parmesan and Yohe, 2003; Menzel et al., 2006; Parmesan, 2006).

Concerning possible impacts on various systems, it is important to keep in mind that climate change is probably linked to higher variability, in the end altering the duration, location, frequency and intensity of events (e.g. Easterling et al., 2000). Prominent examples in recent years include heat waves (e. g. the summer 2003 in Europe, Luterbacher et al., 2004; Schär et al., 2004), extended winter 2006/07 (Luterbacher et al., 2007; Rutishauser et al., 2008; Maignan et al., 2008), droughts, floods, heavy precipitation events (e.g. Kunkel, 2003; Groisman et al., 2005), storms, tornadoes and tropical cyclones (e.g. Emanuel, 2005; Webster et al., 2005). A further increase in climate variability is predicted and could trigger both short and long-term abrupt and nonlinear, changes in many ecosystems (Gutschick and BassiriRad, 2003; Peters et al., 2004). The nonlinearity of the climate system may lead to abrupt climate change, sometimes called rapid climate change, abrupt events or even surprises. The term abrupt often refers to time scales faster than the typical time scale of the responsible forcing (Trenberth et al. 2007). In the context of temperature and phenolgogical change, 'abrupt' designates regional events of large amplitude, typically a few degrees Celsius or days year $^{-1}$, occurring within several decades. The understanding of possible effects of such changes on ecosystems is associated with many problems and is part of the present work.

From the biological point of view, the assessment of impacts of climate change on natural systems requires change detection and the attribution to climate change. One of the most prominent and clearest bio-indicator of climate change is phenology (e.g. Menzel and Fabian, 1999; Walther et al., 2002; Root et al., 2003; Parmesan and Yohe, 2003). Phenological studies in ecological systems focus on the timing and magnitude of recurring biological phases (phenophases), the influence of biotic and abiotic forces on timing, and the interrelation between phases of the same or different species (Schwartz, 2003; Betancourt et al., 2005; Post et al., 2008).

In this PhD we mainly focus on the analysis of this bio indicator. The recording of phenological observations has a long history, nowhere more evident than in the several centuries of records of cherry blossoming in Japan (Menzel and Dose, 2005; Aono and Kazui, 2008). The sensitivity of spring phenophases to temperature is identifiable with the observation that heat sums for the late winter or spring months often are accurate predictors of phenophase's timing (Sparks and Carey, 1995; Diekmann, 1996; Kai et al., 1996; Heikinheimo and Lappalainen, 1997; Thórhallsdóttir, 1998; Schwartz, 1999; Spano et al., 1999; Van Vliet et al., 2002; Galán et al., 2005). Menzel and Fabian (1999) found that 70% of interannual variation in bud burst in a group of European species was explained by daily temperature patterns. Average February and March temperatures explained 75% of the variation in flowering time of Japanese cherries (*Cerasus* spp., Miller-Rushing et al., 2008).

Phenological changes over the past few decades (usually starting in the 1970s or 1980s, depending on location) are much greater than those from the previous several decades (Sparks et al., 1997, Peñuelas et al., 2002). An examination of 17 phenophases at 6500 stations in central Europe revealed "almost no trend" prior to changes after the late 1980s in most areas (Scheifinger et al., 2002). A clear spatial and temporal variability of spring and summer onset dates and their changes can be mainly attributed to regional and local temperature. Menzel et al. (2008) give a compact general overview of the impacts of climate variability and recent climate change on the European plant phenology across the 20th century.

Advances of springtime phenological events were detected by broad scale studies using satellite imagery (Schwartz et al., 2006) and by several meta-analyses (Parmesan and Yohe 2003; Root et al., 2003; Menzel et al., 2006). Average advances in spring phenophases have been 1-3 days per decade during the last several decades in temperate regions of the Northern Hemisphere (Menzel, 2000; Walther et al., 2002; Parmesan and Yohe, 2003; Wolfe et al., 2005; Menzel et al., 2006; Schwartz et al., 2006, Parmesan, 2007), though studies of particular species or particular regions give much more variable results (Scheifinger et al., 2002, Menzel et al., 2006). Changes in summer and autumn phenophases are less consistent in direction and magnitude than changes in spring phenophases (Walther et al., 2002), though the most typical response of autumn phenophases is a slight delay. In a large series of observations covering the late 1950s through the 1990s from the International Phenological Gardens in Europe, spring events advanced on average by 6 days, while fall events were delayed on average by 4-5 days (Menzel and Fabian, 1999; Menzel, 2000). Ahas and Aasa (2006) found in a sample of 753 series that most phenophases exhibited a delay of autumn events. Menzel et al. (2006) revealed advances in 78% of flowering and leafing phenophases but only in 48% of leaf coloring (autumn) phenophases.

Here, temperature is selected as the major climate variable because it represents a strong and widespread documented signal of climate warming in recent decades and has an important direct influence on many physical and biological processes. Physical and biological responses to changing temperatures are often better understood than responses to other climate parameters. Certainly other variables influence timing of at least some phenophases. Timing of snow melt can be an important variable for early spring phenophases in northern and alpine climates (Saavedra et al., 2003; Molau et al., 2005). While snow melt is strongly influenced by temperature, it is also influenced by amount of precipitation and other factors. Flowering of many plant species is responsive to photoperiod (Raven et al., 2005) and precipitation influences the timing of various plant phenophases, especially in dry or seasonally dry habitats (Keatley et al., 2002; Kramer et al., 2000, Peñuelas et al., 2004).

Temperature rise has resulted in marked changes in the timing of life cycle events of plants and animals (e. g. Schwartz, 1999; Menzel and Estrella, 2001). Phenological responses are particularly significant when individual plant-level responses are intense enough to turn into whole-ecosystem responses (Post et al., 2008). Perhaps most notably, the timing and degree of ‘‘green-up’’ are key ecosystem responses that reflect fundamental climate–vegetation couplings (Chapin et al., 2000; Clark et al., 2001; Walther et al., 2002). Although the emphasis of phenological research has often been on interannual variability (Schwartz, 2003), the largest changes in phenological patterns are likely to be linked with major ecosystem disturbances associated with extreme climate or weather events (Peñuelas and Filella, 2002).

Temperature data (typically, most easily and widespread monthly means) are often recorded from adjacent weather stations. Numerous studies have examined the relationship between phenological events and temperature over several seasons to derive predictive relationships between temperature and the timing of a phenophase (e.g. Root et al., 2003; Parmesan and Yohe, 2003; Menzel et al., 2006; Menzel, 2000). Such functions are typically used as the basis for predicting phenological changes likely to be associated with future temperature changes, with a linear relationship generally assumed. Sparks et al. (2000) note that the plant response to temperature, even if linear over a certain range, must gradually decrease, at a certain temperature.

Methods analysing phenology and its relationship to temperature underlie various sources of error (reviewed by Dose and Menzel, 2004) and the extent of these potential errors can not be precisely estimated. Typical sources of errors are that criteria for recording a particular phenological event (e.g. cherry flowering) can be interpreted differently by different observers. Genetic variation is unvoidable in studies of natural and most cultivated species (Kriebel and Wang, 1962; Defila, 1991). Plant age often varies among sites, with unknown effects on phenology (Baumgartner, 1952). Environmental and cultural condition such as soil type, soil moisture, aspect, and exposure are mostly not fully considered. Several variables, such as precipitation, or an urban heat island effect, sometimes have impacts on temperature trends. Influences of urbanization to flowering advancement have been estimated at 4 days over 30 years in central Europe (Roetzer et al., 2000) and 4-6 days in the past century in China and Japan (Yoshino and Ono, 1996). Comparisons of temperature changes in rural and urban sites in Massachusetts revealed that urban effects accounted for half of the total change in greater Boston (Primack et al., 2004).

A further potential bias in phenological studies is the possibility that studies representing significant changes are more likely to be published than those showing no change (Kozlov and Berlina, 2002; Menzel et al., 2006). On one hand such bias could result from the greater possibility of a paper's acceptance if it reveals significant patterns of change. While the interpretation of a significant result can be straightforward, a non-significant result could mean either that no trend exists or that the method used was insufficient to resolve the pattern, and uncertainty over which of these is correct may reduce the likelihood of a paper being submitted. This problem is eliminated by analyzing one or more entire sets of records with all results reported, as done by Menzel et al. (2006) for numerous European records and in meta-analyses by Parmesan and Yohe (2003) and Parmesan (2007). Despite these limitations of data and analysis, attempts at comprehensive and standardized analyses have been made (Menzel et al., 2006; Schwartz et al., 2006) and robust patterns have emerged.

Where long data series exist, the detection of trends or changes in system properties that are beyond natural variability has most commonly been made with regression and correlation analyses. Regression and correlation methods are frequently used in the detection of a relationship of the observed trend with climate variables (e. g. Bradley et al., 1999; Menzel and Fabian, 1999; Jones and Davis, 2000; Schwartz and Reiter, 2000; Defila and Clot, 2001; Menzel et al., 2001; Ahas et al., 2002; Peñuelas et al., 2002; Menzel, 2003; Rutishauser, 2009), rarely by other curve fitting methods (e.g. Ahas, 1999; Sagarin and Micheli, 2001). Linear regression methods are connected with certain limitations such as inadequate separation of temporal and spatial variability, inadequate assessment of uncertainty of functional behaviour, trends and interdependence. For example the linear regression procedure allows calculation of a rate of change in the phenological event over time though, the result depends on whether the series covers just a period of relatively rapid change (e.g., post-1970s or post-1980s) or includes a period of stable or declining temperatures (e.g., 1940-1970). Different studies have different starting dates, ending dates, durations, and frequencies of observation, and temperature change has not been constant over the past few centuries. Calculated rates of change vary depending on what time period is included in the particular set of records (Menzel, 2000; Roetzer et al., 2000; Sparks and Menzel, 2002; Badeck et al., 2004; Dose and Menzel, 2004). Menzel et al. (2006) reduced this problem in their comprehensive analysis of European phenological records by standardizing the time period over all sites.

Several papers have used other methods, including dynamic factor analysis (Gordo and Sanz, 2005), chronological clustering (Doi, 2007), and Bayesian methods (Dose and Menzel, 2004, 2006) to investigate phenological changes, and these have been helpful in detecting different parts of a single time series that show different patterns. In order to identify (abrupt) changes in climatic and natural systems accurately, this dissertation focuses on the application of a more suitable and precise method for the climate change detection based on the Bayesian concepts. Climate change detection employing nonparametric Bayesian function estimation is especially useful for studies of climate change impacts in natural systems where conditions are prescribed to change. Studies analysing long-term phenological records often reveal a heterogeneous pattern of temporal variability with alternating periods of advanced and delayed onset (e.g. Schnelle, 1950; Lauscher, 1978, 1983; Sparks and Carey, 1995; Ahas, 1999). Alternative methods to analyse and quantify changes in phenological time series are highly needed as rigorous treatment of uncertainty, separation of spatial and temporal variability can not be achieved with the traditional methods.

Only a few studies already exist, which applied Bayesian statistical methods in climate change detection, analysis and attribution (e.g. Hobbs, 1997; Hasselmann, 1998; Leroy, 1998; Katz, 2002; Berliner et al., 2000; Menzel and Dose 2004, 2006). Several recent Bayesian detection analyses have used this approach for the assessment of evidence of anthropogenic influence on climate (e.g., Min et al., 2004; Schnur and Hasselmann, 2005; Lee et al., 2005; Min and Hense, 2006, 2007). Leroy (1998) was among the first to explore the Bayesian approach in climate change detection. He also pointed out the need for the estimation of model uncertainties. Berliner et al. (2000) used a robust Bayesian approach to investigate the uncertainties in assessing anthropogenic impacts on climate change resulting from uncertainties described by the priors.

Significant progress has been made since the TAR (Third Assessment Report, 2001) in exploring ensemble approaches to provide uncertainty ranges and probabilities for global and regional climate change. Different methods show consistency in some aspects of their results, but differ significantly in others because they depend to varying degrees on the nature and use of observational constraints, the nature and design of model ensembles and the specification of prior distributions for uncertain inputs (Solomon et al., 2007). Thus a preferred method cannot yet be recommended, but the assumptions and limitations underlying the various approaches, and the sensitivity of the results to them, should be communicated to users. A good example concerns the treatment of model error in Bayesian methods, the uncertainty in which affects the calculation of the likelihood of different model versions, but is difficult to specify (Rougier, 2007). Awareness of this issue is growing in the field of climate prediction (Annan et al., 2005; Knutti et al., 2006) and probabilistic descriptions, particularly at the regional level, are new to climate change science. Bayesian analysis looks to bee a practical and theoretically appropriate tool for making inferences about climate change.

1.2 Bayes's theorem – an Introduction

Bayesian data analysis is based on two rules. The first is the conventional product rule for manipulating conditional probabilities. It allows breaking down a probability density function depending on two (or more) variables $p\left(\vec{\theta},\vec{d}\,\middle|\,M,I\right)$ conditional on the model M which specifies the meaning of the parameters $\vec{\theta}$ and additional information I into simpler functions

$$p\left(\vec{\theta},\vec{d}\,\middle|\,M,I\right)=p\left(\vec{\theta}\,\middle|\,M,I\right)*p\left(\vec{d}\,\middle|\,\vec{\theta},M,I\right) \tag{1}$$

where $p\left(\vec{\theta}\,\middle|\,M,I\right)$ and $p\left(\vec{\theta},\vec{d}\,\middle|\,M,I\right)$ depend only on the single (vector-) variables $\vec{\theta}$ and $\vec{d}$ respectively. Equ. (1) may be expanded in an alternative way due to symmetry in the variables $\vec{\theta},\vec{d}$.

$$p\left(\vec{\theta},\vec{d}\,\middle|\,M,I\right)=p\left(\vec{d}\,\middle|\,M,I\right)*p\left(\vec{\theta}\,\middle|\,\vec{d},M,I\right) \tag{2}$$

Equating the right hand sides of (1) and (2) yields Bayes' theorem.

$$p\left(\vec{\theta}\,\middle|\,\vec{d},M,I\right)=p\left(\vec{\theta}\,\middle|\,M,I\right)*p\left(\vec{d}\,\middle|\,\vec{\theta},M,I\right)\Big/p\left(\vec{d}\,\middle|\,M,I\right) \tag{3}$$

The function on the left hand side is called the posterior density of the parameters $\vec{\theta}$ given data $\vec{d}$ and model M. It is equal to the prior density of the parameters $\vec{\theta}$, $p\left(\vec{\theta}\,\middle|\,M,I\right)$ which encodes our information on $\vec{\theta}$ prior to considering the data $\vec{d}$ times the likelihood $p\left(\vec{d}\,\middle|\,\vec{\theta},M,I\right)$.

$p\left(\vec{d}\,\middle|\,M,I\right)$ is formally the normalisation for the posterior density

$$p\left(\vec{d}\,\middle|\,M,I\right)=\int d\vec{\theta}\;\,p\left(\vec{\theta}\,\middle|\,M,I\right)*p\left(\vec{d}\,\middle|\,\vec{\theta},M,I\right) \tag{4}$$

By inverse application of the product rule we arrive at the Bayesian marginalization rule which completes Bayes' theory and has no counterpart in traditional statistics

$$p\left(\vec{d}\,\middle|\,M,I\right)=\int d\vec{\theta}*p\left(\vec{d},\vec{\theta}\,\middle|\,M,I\right) \tag{5}$$

Equ. (5) allows for an important interpretation. It is obviously the likelihood of the data $\vec{d}$ given the model M regardless of the numerical values of the parameters $\vec{\theta}$.
Employing Bayes' theorem to invert (5) we obtain

$$p\left(M\middle|\vec{d},I\right)=p\left(M\middle|I\right)*p\left(\vec{d}\middle|M,I\right)\Big/p\left(\vec{d}\middle|I\right) \tag{6}$$

Equ. (6) is then the probability of a model M out of a possible variety given the data $\vec{d}$.
We shall now adapt these abstract concepts to the problems of the following chapters. The data $\vec{d}$ are then phenology or temperature time series. They are modelled by either a constant implying time independence or by a linear function in time which associated constant rate of change or by a function consisting of two linear segments matching at a given time t_E. Apparently the latter model, which we call the (one-) change point model, is not only the most complicated but reduces also to the other two by deleting variables. The likelihood for the change point model reads

$$p(\vec{d}|\vec{f},t_E,M,I) \tag{7}$$

where $\vec{f}$ is a three component vector of the support functional values at the beginning of the time series (f_1), the change point (f_2) and the end of the time series (f_3). The likelihood for the linear model evolves from (7) by deleting f_2 and t_E, and for the constant model be deleting f_2, f_3 and t_E. Our first task is to find the change point probability distribution $P\left(t_E\middle|\vec{d},M,I\right)$. By Bayes' theorem it is given by

$$p\left(t_E\middle|\vec{d},M,I\right)=p\left(t_E\middle|M,I\right)*p\left(\vec{d}\middle|t_E,M,I\right)\Big/p\left(\vec{d}\middle|M,I\right) \tag{8}$$

The marginal likelihood $p\left(\vec{d}\middle|t_E,M,I\right)$ is derived from (7).

$$p\left(\vec{d}\middle|t_E,M,I\right)=\int d\vec{f}\,p\left(\vec{f},\vec{d}\middle|t_E,M,I\right)=\int d\vec{f}\,p\left(\vec{f}\middle|t_E,M,I\right)*p\left(\vec{d}\middle|\vec{f},t_E,M,I\right) \tag{9}$$

Application of (8) requires the specification of $p\left(t_E\middle|\vec{d},M,I\right)$, which was taken flat, independent of t_E in all subsequent applications.
Note that (8) contains also the marginal likelihood $p\left(\vec{d}\middle|M,I\right)$ which is needed to infer the probability of the model M given the data $\vec{d}$ using (6).
One final point needs to be mentioned. Having obtained the posterior distribution of the parameters using (3) and (7) we can calculate expectation values of the parameters given the data. For example the expectation value of θ_k is given by

$$\langle\theta_k\rangle=\int d\vec{\theta}\theta_k * p(\vec{\theta}|\vec{d},M,I) \tag{10}$$

It can be shown that (10) holds also for any function $\phi(\vec{\theta}|\vec{d},M,I)$ and can be used in particular to derive estimates of the moments μ_1 and μ_2 of our model functions at any given time t. t is not restricted to the time interval covered by the data but can also lie in extrapolation regions. Defining μ_N as

$$\mu_N=\sum\nolimits_E\int d\vec{f}p(\vec{f},E|\vec{d},M,I)*\left\{\phi(\vec{f},E|t,\vec{d},M,I)\right\}^N \tag{11}$$

we find for the mean of the model functions μ_1 and for the standard deviation

$$\left\langle\Delta\phi^2\right\rangle^{1/2}=\left\{\mu_2-\mu_1{}^2\right\}^{1/2} \tag{12}$$

This completes the formal calculations referred to in the following chapters. Specialities of the Bayes' theorem are for example the inclusion of prior knowledge in the process of parameter estimation and the model comparison option. Prior knowledge on parameter estimates can be obtained by earlier experiments or information or just belief. In case of totally uninformative data the prior estimates of the parameters will not be changed, while in the case of highly informative data the posterior estimates of the parameters will not be influenced by the prior knowledge anymore (Dose, 2007).

In addition to this Introduction I recommend the method description of Chapter 2 (Section 2.2.2), where the reader will find an explanation of the Bayes' application illustrated with several figures. Furthermore I recommend the Appendix of Chapter 5 (Section 5.6). For further details of the algebra the reader is referred to Dose and Menzel (2006). This chapter and the following chapters can only touch upon a few aspects of the rich theory that underlies Bayesian analysis. Thorough introductions to Bayesian analysis and its applications are available in specialist literature (e.g. Gregory, 2007 and Sivia, 2005).

1.3 Objectives of Research

The topic of the present PhD thesis focuses on climate change detection in natural systems by Bayesian analysis. This work seeks to detect changes in temperature and biological systems (vegetation by the use of the phenology of plants) and intends to improve the understanding of responses to climate change with the help of the Bayesian analysis. According to the Intergovernmental Panel on Climate Change (IPCC) detection of climate change is the process of demonstrating that the climate has changed in some defined statistical sense (Le Treut et al., 2007). Detection relies on observational data and model output. Using knowledge of past climates to qualify the nature of ongoing changes has become a concern of growing importance during the last decades, as reflected in the successive IPCC reports.

This work analyses the correlation of long-term (> 30 years) plant phenological time series to rising temperature in Europe. Temperature time series nearby the respective phenological stations are analysed. Beside the phenological studies a further study is concerned with temperature changes in the atmosphere, differences in these changes at various levels in the atmosphere, and the understanding of the causes of these changes and differences. In this study our Bayesian analysis tool that has been originally designed for phenological analysis is firstly applied and tested to detect atmospheric temperature changes. Rates of change and variations in atmospheric temperatures are an integral part of the changes occurring in the Earth's climate system (Santer et al., 1996; Tett et al., 1996; Thorne et al., 2005).

On one hand I attempt to confirm detected changes in plant phenology and temperature time series with the Bayesian analysis. On the other hand I test if the Bayesian approach provides new biological insights that can not be gained by the conventional methods. In this PhD thesis the following research questions will be addressed and finally discussed in the "General and Summarizing Discussion" (Chapter 6):

1. What are the advantages and disadvantages of the Bayesian approach compared to conventional statistical methods when analysing climate change impacts on natural systems?

2. Which potentials of the Bayesian approach (such as model probabilities, functional behaviours, model averaged rates of change, confidence intervals and time spans of elevated change point probability) contribute to an accurate assessment of climate change impacts on natural systems?

3. What kind of biological insights into the triggering climate change factor temperature and its influence on phenology can be gained by the Bayesian concept?

1.4 Outline of Thesis

The substantial core of this PhD consists of four published first author peer-reviewed scientific papers of the candidate. These four publications are shown in detail in Chapter 2, 3, 4 and 5. The four first author publications are arranged chronologically and additionally by two further criteria. The first criterion is the degree of the complexity of the applied analysis. The second criterion is the complexity of the respective research question and type of used time series.

The first publication in **Chapter 2** "*The use of Bayesian analysis to detect recent changes in phenology events throughout the year"* focuses methodically on the model comparison option of the Bayesian approach that is used to compare three different types of models (constant, linear, and one change point). In addition to change point probability curves, rates of change in terms of days per year are also presented. The Bayesian concept is illustrated as an alternative to the classical statistic methods. In this chapter we analyse the variations of the onset of different long-term phenological phases and illustrate phenological changes of different seasons in Germany and Switzerland in the 20th century.

In the second publication in **Chapter 3** *"Norway spruce (Picea abies): Bayesian analysis of the relationship between temperature and bud burst"* the Bayesian analysis is expanded to investigate the relationship of phenology and temperature. Methodically we use 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 are obtained from an optimization of the coherence factor by simulated annealing.
We 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. In these temperature and Norway spruce bud burst time series we study years with the highest probability for discontinuities. We analyse rates of change and the relationship between temperature changes and bud burst of Norway spruce in the 51-year period 1953–2003.

The third publication with the title *"Time series modelling and central European temperature impact assessment of phenological records in the last 250 years"* is presented in **Chapter 4**. Within this chapter we compare Pearson correlation coefficients and linear moving window trends of two different lengths with a Bayesian correlation and model comparison approach. The latter is applied to calculate model probabilities, change point probabilities, functional descriptions and rates of change of three selected models with increasing complexity and temperature weights of single months. We assess the linear dependence of phenological variability by a linear Pearson correlation approach. In addition we apply the Bayesian correlation to account for nonlinearities within the time series. Long-term spring and autumn phenological observations from Switzerland and Burgundy (eastern France) as well as long-term Swiss monthly and seasonal temperature measurements are analysed to evaluate plant phenological variability and temperature impacts over the last 250 years.

The fourth publication in **Chapter 5** with the title:*"Bayesian analysis of changes in Radiosonde Atmospheric Temperature"* applies the Bayesian approach in the field of atmospheric change detection. We examine and compare model probabilities, change points and rates of change of global annual mean temperature anomalies in the period 1979–2004. The Bayesian analysis is firstly tested as an alternative approach to assess more of the potential of the current adjusted radiosonde data. We provide results of global temperature data for 13 atmospheric pressure levels from the surface up to the lower stratosphere. Such data at discrete vertical levels provide unique information for assessing changes in the structure of the atmosphere.

The core content of this PhD is enbedded by the "*General Introduction*" (**Chapter 1**) and the "*General and Summarizing Discussion*" (**Chapter 6**). The introductory chapter summarizes the features, motivation and the aims of using the Bayesian analysis to detect and attribute climate change in natural systems.

In **Chapter 6** the *"General and Summarizing Discussion"*connects all ideas, respective research questions and results of the above presented chapters to derive a comprehensive conclusion. Particularly I discuss, summarize and structure the main conclusions in the context of three leading questions presented in Section 1.4. First I consider the advantages and disadvantages of the Bayesian approach compared to conventional statistical methods. Secondly I discuss and summarize the importance of the Bayesian features to contribute to an accurate assessment of climate change impacts. And third I discuss results regarding temperature as a triggering factor and its influence on phenology. Finally in Section 6.4 I provide a compact "*Summary and Conclusion*".

An overview of all candidates' publications is listed in **Chapter 7**. In **Chapter 8** a short appraisal of the candidate's contribution is illustrated which is followed by the Acknolegements.

Finally the Appendix provides Abstracts of all candidate's peer-reviewed scientific papers and books which are not included in the present PhD as complete single chapters. Not included publications are mostly co-author papers where the candidate has not contributed more than 25% or first author publications that are still in review progress. In the Appendix the reader will also find one Abstract of an accepted first author book chapter with the title: Bayesian methods in phenology (Schleip et al., forthcoming). This book chapter in substance consists of the single chapters of the publications already presented in this PhD.

1.5 References

Annan, J.D., Hargreaves J.C., Ohgaito R., Abe-Ouchi A., Emori S., (2005) Efficiently constraining climate sensitivity with paleoclimate simulations. *SOLA* **1**, 181–184.

Aono, Y, Kazui K., (2008) Phenological data series of cherry tree flowering in Kyoto, Japan, and its application to reconstruction of springtime temperatures since the 9th century. *Int. J. Climatol.* **28**, 7, 905-914.

Ahas R., (1999) Long-term phyto-, ornitho- and ichthyophenological time-series analyses in Estonia. *Int. J. Biometeorol.* **42**, 119-123.

Ahas R., Aasa A., Menzel A., Fedotova V.G., Scheifinger H., (2002) Changes in European spring phenology. *Int. J. Climatol.* **22**, 1727-1738.

Ahas R., Aasa A., (2006) The effect of climate change on the phenology of selected Estonian plant, bird and fish populations. *Int. J. Biometeorol.* **51**, 17-26.

Badeck F. W., Bondeau A., Bottcher K., Doktor D., Lucht W., Schaber J., Sitch S., (2004) Responses of spring phenology to climate change. *New Phytol.* **162**, 295–309.

Baumgartner A. (1995) Zur Phänologie von Laubhölzern und Ihre Anwendung bei lokalklimatischen Untersuchungen. *Berichte des DWD in der US-Zone* **42**, 69.73.

Berliner L.M., Levine R.A., Shea D.J., (2000) Bayesian climate change assessment. *J. Climate* **13**, 3805-3820.

Betancourt J.L., Schwartz M.D., Breshears D.D., Cayan D.R., Dettinger M.D., Inouye D.W., Post E., Reed B.C. (2005) Implementing a U.S.A. National Phenology Network. Eos, Transactions, *American Geophysical Union* **86**, 539–540.

Bradley N.L., Leopold A.C., Ross J., Huffaker W., (1999) Phenological changes reflect climate change in Wisconsin. *Proceedings of the National Academy of Sciences USA Ecology* **96**, 9701–9704.

Chapin F.S., Zavaleta E.S., Eviner V.T., Naylor R.L., Vitousek P.M., Reynolds H.L., Hooper U.D., Lavorel S., Sala O. E., Hobbie S. E., Mack M. C., Dıaz. S., (2000) Consequences of changing biodiversity. *Nature* **405**, 234–242.

Clark J.S., Carpenter S.R., Barber M., Collins S., Dobson A., Foley J.A., Lodge D.M., Pascual M., Pielke R.J., Pizer W., Pringle C., Reid W.V., Rose K.A., Sala O., Schlesinger W.H., Wall D.H., Wear D, (2001) Ecological forecasts: an emerging imperative. *Science* **293**, 657–660.

Defila C., (1991) Planzenphänologie der Schweiz.. *Veröffentlichungen der Schweizerischen Meteorologischen Anstalt* **1**, 235.

Defila C., Clot B., (2001) Phytophenological trends in Switzerland. *Int. J. Biometeorol.* **45**, 203-207.

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

Doi H., (2007) Winter flowering phenology of Japanese apricot *Prunus mume* reflects climate change across Japan. *Climate Research.* **34**, 99-104.

Dose V., Menzel A., (2004) Bayesian Analysis of Climate Change Impacts in Phenology. *Global Change Biol.* **10**, 259-272.

Dose V., Menzel A., (2006) Bayesian correlation between temperature and blossom onset data. *Global Change Biol.* **12**, 1451–1459.

Dose V., (2007) EPL goes Bayesian. *Europhysics Letters (EPL)* **79**, 30000 (2pp).

Easterling, D.R., Meehl G.A., Parmesan C., Changnon S.A., Karl T.R., Mearns L.O., (2000) Climate extremes: observations, modeling, and impact. *Science* **289**, 2068-2074.

Emanuel K., (2005) Increasing destructiveness of tropical cyclones over the past 30 years. *Nature* **436**, 686-688.

Galán C., Garoa-Mozo H., Vazquez L., Ruiz L., Díaz de la Guardia C., Trigo M.M., (2005) Heat requirement for the onset of the Olea europaea L. pollen season in several sites in Andalusia and the effect of the expected future climate change. *Int. J. Biometeorol.* **49**, 184-188.

Gordo O., Sanz J.J., (2005) Phenology and climate change: a long-term study in a Mediterranean locality. *Oecologia* **146**, 484-495.

Gutschick V.P., BassiriRad H., (2003) Extreme events as shaping physiology, ecology, and evolution of plants: toward a unified definition and evaluation of their consequences. *New Phytologist* **160**, 21–42.

Gregory P.C., (2007) Bayesian logical data analysis for the physical sciences. In: A Comparative approach with Mathematica Support, Cambridge University Press, Boston.

Groisman P.Y., Knight R.W., Easterling D.R., Karl T.R., Hegerl G.C., Razuvaev V.A.N., (2005) Trends in intense precipitation in the climate record. *J. Climate* **18**, 9, 1326-1350.

Hasselmann K., (1998) Conventional and Bayesian approach to climate change detection and attribution. *Quart. J. R. Meteor. Soc.* **124**, 2541-2565.

Heikinheimo M., Lappalainen H., (1997) Dependence of the flower bud burst of some plant taxa in Finland on effective temperature sum: implications for climate warming. *Ann. Bot. Fenn.* **34**, 229-243.

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

Jones G.V., Davis, R. E., (2000). Using A Synoptic Climatological Approach to Understand Climate/Viticulture Relationships. *Int. J. Climatol.* **20**, 813-837.

Kai K., Kainuma M., Murakoshi N., (1996) Effects of global warming on the phenological observation in Japan, p. 85-92. In K. Omasa, K. Kai, H. Taoda, Z. Uchijima, and M. Yoshino [eds.], Climate change and plants in East Asia. Springer-Verlag, Tokyo.

Katz R.W., (2002) Stochastic modeling of hurricane damage. *J. Appl. Meteorol.* **41**, 754-76.

Keatley M.R., Fletcher T.D., Hudson I.L., Ades P.K., (2002) Phenological studies in Australia: potential application in historical and future climate analysis. *Int. J. Climatol.* **22**, 1769-1780.

Knutti R., Meehl G.A., Allen M.R., Stainforth D.A., (2006) Constraining climate sensitivity from the seasonal cycle in surface temperature, *J. Climate* **19,** 17, 4224-4233.

Kozlov M.V., Berlina N. G., (2002) Decline in length of the summer season on the Kola Peninsula, Russia. *Clim. Change* **54**, 387-398.

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

Kriebel K.I., Wang C.W., (1962) The inteaction between povenance and degree of chilling in bud-break of sugar mable. *Silvae Genet.* **11**, 125-130

Kunkel K.E., (2003) North American trends in extreme precipitation. *Natural Hazards* **29,** 291-305

Lauscher F., (1978) Neue Analysen ältester und neuerer phänologischer Reihen. *Archiv für Meteorologie Geophysik 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.

Lee T., Zwiers F., Hegerl G., Zhang X., Tsao M., (2005) A Bayesian approach to climate change detection and attribution assessment. *J. Climate* **18**, 2429–2440.

Le Treut H., Somerville R., Cubasch U., Ding Y., Mauritzen C., Mokssit A., Peterson T., Prather M., (2007) Historical Overview of Climate Change. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Leroy S.S., (1998) Detecting cimate signals: some Bayesian aspects. *J. Climate* **11**, 640-651.

Lucht W., Prentice I. C., Myneni R. B., Sitch S., Friedlingstein P., Cramer W., Bousquet P., Buermann W., Smith B., (2002) Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. *Science* **296**, 1687–1689.

Luterbacher J., Dietrich D., Xoplaki E., Grosjean M., Wanner H., (2004) European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* **303,** 5663, 1499-1503.

Luterbacher J., Liniger M., Menzel A., Estrella N., Della-Marta P.M., Pfister C., Rutishauser, T., Xoplaki E., (2007) Exceptional European warmth of autumn 2006 and winter 2007: Historical context, the underlying dynamics, and its phenological impacts *Geophysical Research Letters* **34**, L12704.

Maignan F., Breon F., Vermote E., Ciais P., Viovy N., (2008) Mild winter and spring 2007 over western Europe led to a widespread early vegetation onset. *Geophysical Research Letters* **35** , L02404.

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**, 76-81.

Menzel A., Estrella N., (2001) Plant phenological changes. In: Walther GR, Burga CA and Edwards PJ (Eds.) “Fingerprints” of Climate Change – Adapted behaviour and shifting species ranges. Kluwer Academic/Plenum Publishers, New York and London, 123-137.

Menzel A., Estrella N., Fabian P. (2001) Spatial and temporal variability of the phenological seasons in Germany from 1951 to 1996. *Global Change Biol.* **7,** 6, 657-666.

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

Menzel A., Dose V., (2005) Analysis of long-term time-series of beginning of flowering by Bayesian function estimation. *Meteorologische Zeitschrift* **14**, 3, 429-434.

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 Biol.* **12,** 10, 1969–1976.

Menzel A., Estrella N., Schleip C., (2008) Impacts of Climate Variability, Trends and NAO on 20th Century European Plant Phenology. In: Climate variability and extremes during the past 100 years. Series: Advances in Global Change Research , Vol. 33 Brönnimann, S.; Luterbacher, J.; Ewen, T.; Diaz, H.F.; Stolarski, R.S.; Neu, U. (Eds.), XVI, 364 p. 146 illus., 50 in color., Hardcover ISBN: 978-1-4020-6765-5. Springer, 221-233.

Miller-Rushing A.J., Primack R., (2008) Global warming and flowering times in Thoreau's Concord; A community perspective. *Ecology* **89**, 332-341.

Min S-K., Hense A., Paeth H., Kwon W-T., (2004) A Bayesian decision method for climate change signal analysis. *Meteorol Z.* **13**, 421–436.

Min S-K., Hense A., (2006) A Bayesian assessment of climate change using multimodel ensembles. Part I Global mean surface temperature. *J. Climate* **19**, 3237–3256.

Min S-K., Hense A., (2007) A Bayesian assessment of climate change using multimodel ensembles. Part II regional and seasonal mean temperatures. *J. Climate* **20**, 2769–2790.

Molau U., Nordenhäll U., Eriksen B., (2005) Onset of flowering and climate variability in an alpine landscape: a 10-year study from Swedish Lapland. *Am. J. Bot.* **92**, 422-431.

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

Parmesan C., (2006) Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.* **37**, 637-669.

Parmesan C., (2007) Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Global Change Biol.* **13**, 1860-1872.

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

Peñuelas J., Filella I.; Zhang X., Llorens L., Ogaya R., Lloret F., Comas P., Estiarte M., Terradas J., (2004) Complex spatiotemporal phenological shifts as a response to rainfall changes. *New Phytologist* **161**, 837-846.

Peters D.P.C., Pielke R. A., Bestelmeyer B.T., Allen C.D., Munson-McGee S., Havstad K.M., (2004) Cross-scale interactions, nonlinearities, and forecasting catastrophic events. *Proceedings of the National Academy of Sciences* (USA) **101**, 15130–15135.

Post E.S., Pedersen C., Wilmers C.C., Forchhammer M.C., (2008) Phenological sequences reveal aggregate life history response to climatic warming. *Ecology* **89**, 363–370.

Primack D., Imbres C., Primack R.B., Miller-Rushing A.J., Del Tredici P., (2004) Herbarium specimens demonstrate earlier flowering times in response to warming in Boston. *American Journal of Botany* **91**, 1260-1264.

Raven P. H., Evert R. F, Eichhorn S. E., (2005) Biology of plants, 7th edition. Freeman, New York, NY.

Roetzer T., Wittenzeller M., Haeckel H. Nekovar J., (2000) Phenology in central Europe: differences and trends of spring phenophases in urban and rural areas. *Int. J. Biometeorol.* **44**, 60-66.

Root T.L., Price J.T., Hall K.R., Schneider S.H., Rosenzweig C., Pounds J.A., (2003) Fingerprints of global warming on wild animals and plants. *Nature* **421**, 57–60.

Rosenzweig C., Casassa G., Karoly D.J., Imeson A., Liu C., Menzel A., Rawlins S., Root T.L., Seguin B., Tryjanowski P., (2007) Assessment of observed changes and responses in natural and managed systems. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Parry M.L., Canziani O.F., Palutikof J.P., van der Linden P.J., Hanson C.E., Eds., Cambridge University Press, Cambridge, UK, 79-131.

Rosenzweig C., Karoly D., Vicarelli M., Neofotis P., Wu Q., Casassa G., Menzel A., Root T.L., Estrella N., Seguin B., Tryjanowski P., Liu C., Rawlins S., Imeson A., (2008) Attributing physical and biological impacts to anthropogenic climate change. *Nature* **453**, 353-357.

Rougier J.C., (2007) Probabilistic Inference for Future Climate Using an Ensemble of Climate Model Evaluations. *Climatic Change* **81**, 247-264.

Rutishauser T., Luterbacher J., Defila C., Frank, D., Wanner H., (2008) Swiss spring plant phenology 2007: Extremes, a multi-century perspective and changes in temperature sensitivity. *Geophysical Research Letters* **35**, L05703.

Rutishauser T., (2009) Historical Phenology in Central Europe. Seasonality and Climate During the Past 500 Years. *Geographica Bernensia*, G82.

Saavedra F., Inouye D. W., Price M. V., Harte J., (2003) Changes in flowering and abundance of *Delphinium nuttallianum* (Ranunculaceae) in response to a subalpine warming experiment. *Global Change Biol.* **9**, 885-894.

Santer B.D., Taylor K.E., Wigley T.M.L., Johns T.C., Jones P.D., Karoly D.J. , Mitchell J.F.B., Oort A.H., Penner J.E., Ramaswamy V., Schwarzkopf M.D., Stouffer R.J., Tett S., (1996) A search for human influences on the thermal structure of the atmosphere. *Nature* **382**, 39–46.

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

Schleip C., Menzel A., Dose V., (forthcoming) Bayesian methods in phenology in Hudson IL & Keatley M.R., (Eds) Phenological Research: Methods for environmental and climate change analysis. Springer. Berlin, Heidelberg, New York.

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

Schnur R., Hasselmann K., (2005) Optimal filtering for Bayesian detection and attribution of climate change. *Clim Dyn* **24**, 45–55.

Schwartz M.D., (1999) Advancing to full bloom: planning phenological research for the 21st century. *Int. J. Biometeorol.* **42**, 113-118.

Schwartz M.D., Reiter B.E., (2000) Changes in North American spring. *Int. J. Climatol.* **20**, 929-932.

Schwartz M. D., (2003) Phenology: an integrative environmental science. Kluwer Academic Publishing, Dordrecht, The Netherlands.

Schwartz M.D., Ahas R., Aasa A., (2006) Onset of Spring Starting Earlier Across the Northern Hemisphere. *Global Change Biol.* **12**, 343-351.

Scheifinger H., A., Menzel A., Koch I, Peter C., Ahas R., (2002) Atmospheric mechanisms governing the spatial and temporal variability of phenological observations in central Europe. *Int. J. Climatol.* **22**, 1739-1755.

Schär C., Vidale P.L., Luthi D., Frei C., Haberli C., Liniger M.A., Appenzeller C., (2004) The role of increasing temperature variability in European summer heatwaves. *Nature* **427,** 6972, 332-336.

Sivia D.S., (2005) Data Analysis—A Bayesian Tutorial. Clarendon, Oxford.

Solomon S., Qin D., Manning M., Marquis M., Averyt K., Tignor M., Le Roy Miller H., Chen Z., (ed.): Climate Change (2007) The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change IPCC Cambridge University Press.

Spano D., Cesaraccio C., Duce P., Snyder R. L., (1999) Phenological stages of natural species and their use as climate indicators. *Int. J. Biometeorol.* **42**, 124-133.

Sparks T.H., Carey P.D., Combes J., (1997) First leafing dates of trees in Surrey between 1947 and 1996. *London Nature*. **76**: 15-20.

Sparks T.H., Carey P.D., (1995) The response of species to climate over two centuries: and analysis of the Marsham phenological record, 1736-1947. *J. Ecol.* **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. *Int. J. Biometeorol.* **44**, 82-87.

Sparks T.H., Menzel A., (2002) Observed changes in seasons: an overview. *Int. J. Climatol.* **22**, 1715-1725.

Tett S.F.B., Mitchell J.F.B., Parker D.E., Allen M.R., (1996) Human influence on the atmospheric vertical temperature structure: Detection and observations. *Science* **274**, 1170–1173.

Trenberth, K.E., Jones P.D., Ambenje P., Bojariu R., Easterling D., Klein Tank A., Parker D., Rahimzadeh F., Renwick J.A., Rusticucci M., Soden B., Zhai P., (2007) Observations: Surface and Atmospheric Climate Change. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., Qin D., Manning M., Chen Z., Marquis M., Averyt K.B., Tignor M., Miller H.L. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Thorne P.W., Parker D.E., Tett S.F.B., Jones P.D., McCarthy M., Coleman H., Brohan P., (2005) Revisiting radiosonde upper air temperatures from 1958 to 2002. *Journal of Geophysical Research*. **110**, D18105.

Thórhallsdóttir T.E., (1989) Flowering phenology in the central highland of Iceland and implications for climatic warming in the Arctic. *Oecologia* **114**, 43-49.

Van Vliet A.J., Overeem H., A., De Groot R.S., Jacobs A.F.G., Spieksma F.T.M., (2002) The influence of temperature and climate change on the timing of pollen release in the Netherlands. *Int. J. Climatol.* **22**, 1757-1767.

Walther G.R., Post E., Convey P., Menzel A., Parmesan C., Beebee T.J.C., Fromentin J.M., Hoegh-Guldberg O., Bairlein F., (2002) Ecological responses to recent climate change. *Nature* **416**, 389-395.

Webster P.J., Holland G.J., Curry J.A., Chang H.R., (2005) Changes in tropical cyclone number, duration, and intensity in a warming environment, *Science* **309**, 1844-1846.

Wolfe D.W., Schwarz M.D., Lakso A.N., Otsuki Y., Pool R.M., Shaulis N.J., (2005) Climate change and shifts in spring phenology of three horticultural woody perennials in northeastern USA. *Int. J. Biometeorol.* **49**, 303-309.

Yoshimo M., Ono. H.-S. P., (1996) Variations in the plant phenology affected by global warming, p. 93-107. In K. Omasa, K. Kai, H. Taoda, Z. Uchijima, and M. Yoshino [eds.], Climate change and plants in East Asia. Springer-Verlag, Tokyo, Japan.

2 The use of Bayesian analysis to detect recent changes in phenological events throughout the year

Christoph **Schleip** [a], Annette **Menzel** [a], Nicole **Estrella** [a], Volker **Dose** [b],

[a] *Fachgebiet für Ökoklimatologie, Technische Universität München, Am Hochanger 13, D-85354 Freising-Weihenstephan, Germany*

[b] *Centre of Interdisciplinary Plasma Science, Max-Planck-Institut für, Plasmaphysik, Boltzmannstraße 2, D-85748 Garching bei München, Germany*

Agricultural and Forest Meteorology (2006), Volume 141, Issues 2-4, 20, Pages 179-191, doi:10.1016/j.agrformet.2006.09.013.

Abstract

The most recent assessment report of the Intergovernmental Panel on Climate Change (IPCC, 2001) predicted further increases in global mean temperature as well as in climate variability and extreme events. The latter changes imply an increased risk of more abrupt and non-linear changes in many ecosystems. Phenology is one of the main bio-indicators of climate change impacts on ecosystems. In order to analyse observed phenological changes accurately we used the Bayesian approach for phenological time series analysis developed by Dose and Menzel (2004) including the model comparison option. This option offers new opportunities to analyse and quantify changes in phenological time series.

Our comprehensive phenological data set consisted of long-term observational records from the 1951-2000 period across central Europe. We analysed the data as constant (mean onset date), linear (constant trend over time) and one change point models. The one change point model involves the selection of two linear segments which match at a particular time ("broken stick"). The break is estimated by an examination of all possible breaks weighted by their respective change point probability.

The change point model provided the best description of the data from all seasons of the year. High probabilities for this specific model reveal Europe-wide non-linear changes in phenology. The dominance of the one change point model was most pronounced for phases in summer to late autumn. Between forest trees, fruit trees and herbaceous plants there were no significant differences in the preferred model. For different phenological stages of *Avena sativa* and *Aesculus hippocastanum* we observed an increasing probability of the one change point model through the year. Increasing model probabilities, especially at the end of the growing season, indicate abrupt and fast changes in phenology. A detailed analysis for eleven stations in Switzerland (1959-1999) revealed that the maximum change point probabilities of leaf unfolding records were concentrated in the mid 1980s whereas leaf colouring displayed a more heterogeneous pattern. The strength of the assessed trends in Switzerland differed by altitude. Only at a few places did the phenological time series exhibit a constant or linear course. The frequency of maximum change point probabilities in the 1980s indicate that most changes in our analysed data sets occurred simultaneously. The advantages of the Bayesian approach are discussed.

Keywords: Bayesian analysis; Change point model; Phenological seasons; Phenology; Central Europe; Climate change.

2.1 Introduction

In many studies ecological, often phenological, impacts of recent climate change have already been analysed. As summarised in various reviews, the main findings of these studies, (e.g. IPCC, 2001; Walther et al., 2002; Parmesan and Yohe, 2003), include an earlier onset of spring events in mid and higher latitudes, which is described as globally consistent. Mean changes, depending on the method used in meta-analyses, range between –2.3 and –5.1 days / decade (Parmesan and Yohe, 2003; Root et al., 2003). The published literature also suggests seasonal variations with stronger advances in earlier spring and less pronounced ones in later spring and summer (e.g. Sparks and Menzel, 2002).

The detection of shifts is traditionally performed by classical statistical methods, such as slopes of linear regression models or a comparison of mean values of distinct time periods (e.g. Bradley et al., 1999; Menzel and Fabian, 1999; Jones and Davis, 2000; Schwartz and Reiter, 2000; Defila and Clot, 2001; Menzel et al., 2001; Ahas et al., 2002; Peñuelas et al., 2002; Menzel, 2003) and rarely by curve fitting methods (e.g. Ahas, 1999; Sagarin and Micheli, 2001). Trends are then reported in days per year or decade, or days of change over the study period. Reviews of phenological trend studies suggest that only about 40% of the reported trends have proved to be statistically significant (Menzel et al., 2001; Defila and Clot, 2001). Studies analysing long-term phenological records often reveal a heterogeneous pattern of temporal variability with alternating periods of advanced and delayed onset (e.g. Schnelle, 1950; Lauscher, 1978, 1983; Freitag, 1987; Sparks and Carey, 1995; Ahas, 1999).

Most naturally-occurring time series in ecosystems are not at all stationary; instead they exhibit various kinds of trends, cycles and seasonal patterns. Although linear trend models have their uses, they are often inappropriate for natural indicators or phenological data. Most time series of naturally occurring events do change their slopes and/or their intercepts over time. However, the linear regression fits the slope and intercept of a regression line that gives the best average fit to all the past data, and unfortunately its deviation from the data is often greatest near the end of the time series.

Dose and Menzel (2004) introduced Bayesian analysis for describing climate change impacts on phenology. The model comparison option of Bayesian theory limits the complexity of a model to the amount necessary to explain the data and improves the description of the functional behaviour of many phenological time series. Dose and Menzel (2004) applied this method to three long-term flowering records (1896-2002), which were observed at the Research Station of the German Meteorological Service at Geisenheim (49°59'N, 7°58'E). Their data set comprised first flowering of snowdrops (*Galanthus nivalis L.*), sweet cherry (*Prunus avium L.*) and lime (*Tilia platyphyllos L.*). In their work, they evaluated the constant, linear, and one change point models (see method section) and provided annual predictions for the change point probability and rates of changes. The one change point model was a function consisting of two linear segments defined by the endpoints of the series and a change point in-between.

One of the major findings of Dose and Menzel (2004) was that flowering of cherry and lime significantly advanced from the mid 1980s onwards and that among the models, the one change point model was superior in describing the functional behaviour of the flowering records. Bayesian model averaging offers new possibilities to analyse and quantify changes in phenological time series in more detail. We utilize this new approach with a very comprehensive phenological data set, covering the major parts of Europe.

The special emphasis of this paper is to look at the impact of climate change on different phenological seasons across the year. Our study addresses three main questions:
(1) Are there differences across seasons in the preferred model (constant onset dates, linear change over time or one change point model with two linear segments and a change point) to describe the phenological time series and their changes? (2) Are greater changes, e.g. as known from literature for developmental stages in early spring, connected to a stronger preference for discontinuous changes (one change point model) within selected species? (3) Are there differences concerning the model preference or trends across altitudes and stations?

2.2 Material and Methods

2.2.1 Material

We analysed phenological data from the common database of the 5 FP-EU Project POSITIVE - Phenological Observation and Satellite Data (NDVI): Trends in the Vegetation Cycle in Europe (Contract No. EVK2-CT-1999-00012). The data comprised phenological records on common species including herbs, agricultural crops, fruit and forest trees. All 7087 observation stations together delivered more than 2600 time series of a clearly defined seasonal event. In total, 119 different phenological events, such as leaf unfolding of oak, were identified in this comprehensive data set. Not all stations recorded the same species and phenophases. The data set encompassed the whole seasonal cycle and also included phases in plant production which are mainly determined by farmers' decisions. Most of the data consisted of spring and summer phases such as leaf unfolding or flowering, but also include autumn phases such as leaf colouring and leaf fall. The phenological observations provided a good geographical coverage for Central, Eastern and part of Western Europe (Austria, Switzerland, Germany, Slovenia, Poland, Estonia and former Soviet Union) within the period 1951-1999. The longitudes and latitudes of the observational stations ranged from 5° to 65° E and 45° to 65°N respectively and the altitudes from 0 and 1400 metres above sea level.

Table 2.1: Indicator species of the four phenological seasons (very early spring, early spring, mid summer/early autumn and late autumn) with their phenological seasons, identification codes and numbers of investigated stations

Season	Species	Common name	Phenological stage	Code	Number of stations
Very early spring	*Anemone nemorosa* L.	Wood anemone	Beginning of flowering	1a	5
			General flowering	1b	17
	Corylus avellana L.	Hazel	Beginning of flowering	2a	102
			General flowering	2b	26
	Galanthus nivalis L.	Common snowdrop	Beginning of flowering	3a	37
			General flowering	3b	19
	Tussilago farfara L.	Colt's foot	Beginning of flowering	4a	106
			General flowering	4b	20
Early spring	*Aesculus hippocastanum* L.	Horse chestnut	Beginning of flowering	1	52
	Betula pendula Roth	Silver birch	Leaf unfolding	2	130
	Fagus sylvatica L.	Common beech	Leaf unfolding	3	49
	Prunus avium L.	Wild cherry	Beginning of flowering	4	49
	Prunus spinosa L.	Blackthorn	Beginning of flowering	5	18
Mid summer and early autumn	*Aesculus hippocastanum* L.	Horse chestnut	Ripe fruit	1	57
	Colchicum autumnale L.	Autumn crocus	General flowering	2	14
	Malus domestica Borkh.	Apple	Ripe fruit	3	16
	Prunus americana Marsh.	American plum	Ripe fruit	4	13
	Tilia cordata Mill.	Small leaved lime	General flowering	5	8
	Sambucus nigra L.	Elderberry	Ripe fruit	6	40
Late autumn	*Aesculus hippocastanum* L.	Horse chestnut	Autumn colouring	1a	30
			Leaf dropping	1b	16
	Betula pendula Roth	Silver birch	Autumn colouring	2	74
	Fagus sylvatica L.	Common beech	Autumn colouring	3	43
	Malus domestica Borkh.	Apple	Autumn colouring	4	28

A careful quality check following the method of Scheifinger et al. (2002) was conducted to exclude outliers and single observations which had been possibly incorrectly recorded.

Our Bayesian analysis is tolerant of missing values. In order to detect trends, the minimum number of observational years required at a single station was set to 30. 90 of the 119 phases met these selection criteria. Thus our corrected European phenological data set was composed of 2600 quality checked phenological time series of 30+ years comprising 90 phenological phases and covering a geographical range of 6°17'-63°7'E and 45°98'-61°81'N. Substantial differences in observed species and the recorded years existed among the surveyed countries. Consequently, the spatial density of all available series was different between phases and countries: there were less data available for Eastern Europe compared to Western Europe because only a few long-term (30+ years) observational records were available in countries such as the former Soviet Union, Poland, Estonia and the Czech Republic.

In order to analyse differences across seasons (first question, see introduction), we assigned 19 so-called indicator phases to four specific periods; "very early spring", "early spring", "mid-summer combined with early autumn", and "late autumn" (see Table 1). Schnelle (1955) defined 10 "phenological seasons" based on physiological and biological constraints. We restricted our analysis to four "seasons" due to the absence of indicator species for other periods.

The season "very early spring" begins at the end of February or during March when snowdrop and hazel flowers. Shortly afterwards, farmers begin work in the fields as soon as the soil is dry enough. The emergence of fruit blossom indicates the beginning of "early spring". Blackthorn and maple flower at the same time as sweet cherry, while the other deciduous trees have not yet stirred. Horse chestnut and birch begin to unfold their leaves. The season "mid summer combined with early autumn" is characterised by the harvesting of winter cereals, which starts with rye and ends with oats. The course of the phenological year nears its end when the horse chestnuts, acorns, beech-nuts and walnuts ripen. The second half of autumn ("late autumn") is characterised by the beginning of leaf colouring (e.g. horse chestnut, beech, birch, oak and ash), and the subsequent leaf fall from these and the fruit trees. The beets are then harvested and the winter sown cereals soon emerge. The phenological year ends with severe frosts, which force field work to cease altogether.

A second data set of different phenological stages of common oat *(Avena sativa)* and horse chestnut (*Aesculus hippocastanum*) were analysed in a sequence across the year in order to answer the second question about differential changes (within a species) across seasons.

A third data set was assembled in order to compare the changes in early spring and autumn across observation stations at different altitudes (see Figure 2.1). This smaller unified subset consisted of two species and two phases (*Aesculus hippocastanum* leaf unfolding/ autumn colouring and *Fagus sylvatica* leaf unfolding/ autumn colouring); all of which had complete records (1959-1999) at the selected eleven stations in Switzerland.

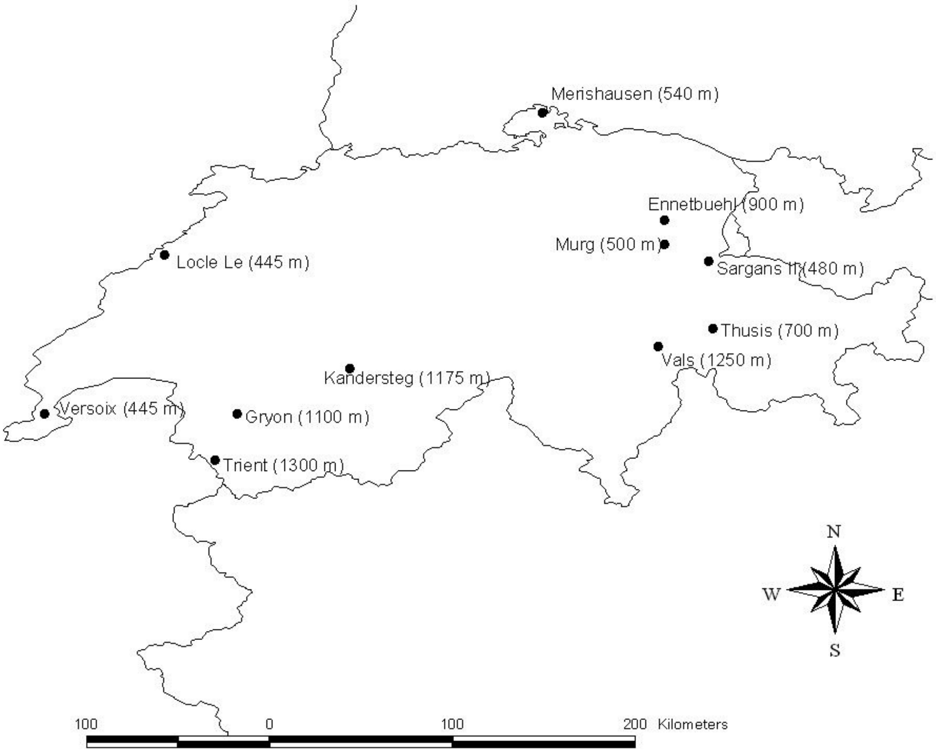


Figure 2.1: Geographical location of the stations in the Swiss subset.

2.2.2 Methods

The Bayesian approach to analyse the functional behaviour of phenological time series and their trends followed the methodology introduced by Dose and Menzel (2004). We refer to this publication for further insight in mathematical formulae. Here, the main features of the Bayesian approach are explained in 6 graphs (Figure 2.2 a-f).

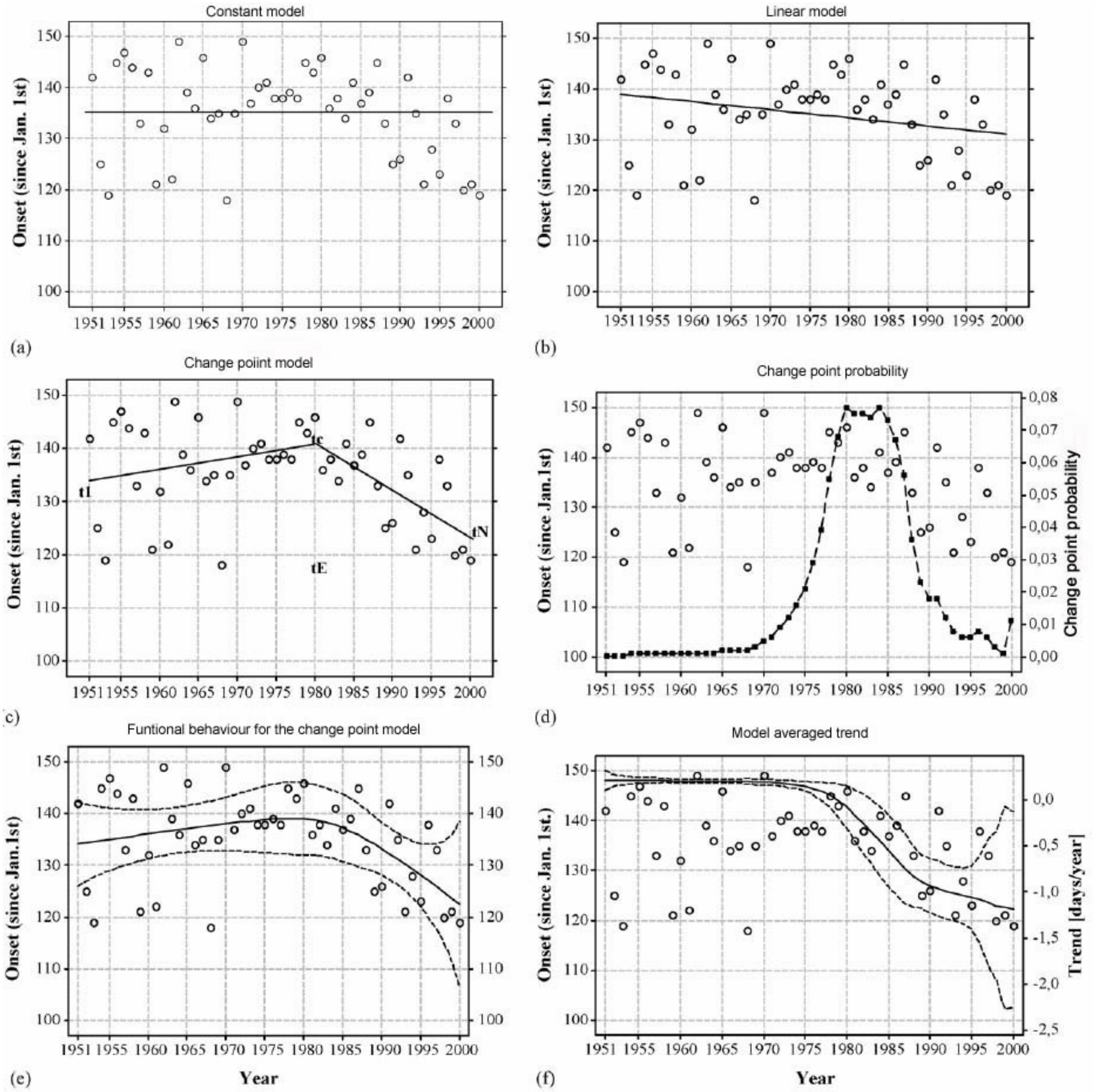


Figure 2.2: The Bayesian approach to phenological time series analysis (an example using the beginning of flowering of Syringa vulgaris at Grünenplan, Germany). (a) Constant model, (b) linear model, (c) one change point model, (d) change point probability distribution for the one change point model, (e) the functional behaviour of the time series (continuous line) with confidence intervals (dashed lines) for the change point model and (f) the derivative of the time series, the trend, with dashed lines representing the upper and lower confidence interval.

The phenological time series plotted in Figure 2.2 can be described by different models/ functions. We selected three of them: (1) a constant model with a fixed date of onset (Figure 2.2 a), (2) a model with linear change of onset dates in time (Figure 2.2 b) and (3) a one change point model, which is the choice of a function that consists of two linear segments which match at a particular time t_E (Figure 2.2 c). We call the point t_E the change point. The constant model represents the hypothesis of no change at all. It assumes a functional behaviour constant in time with associated zero rates of change/ trend. The forecasting equation for the linear trend model is: $y(t) = \alpha + \beta t$ where t is the time index. The parameters alpha and beta (the "intercept" and "slope" of the trend line) are usually estimated via a simple regression in which y is the dependent variable and the time index t is the independent variable.

The one change point model with two linear segments (Figure 2.2 c) fits our data with a residual sum of squares of 2800 compared to 3593 for the linear regression model and 3865 for the constant model. The obvious improvement in fit is penalized by a further complication as the model needs four parameters (t_E, the functional values at t_E (t_c) and a functional value for each segment). Further complications of the model (i.e. more change points) would achieve better fits – ending with an ultimate model containing as many parameters as data points. Bayesian probability theory generally calculates the probability of a given model compared to one or more alternatives. A built-in mechanism (so called Ockham's razor) balances the quality of the fit against the complexity of the model.

Figure 2.2 c displays only the maximum likelihood fit of a continuous function with two linear segments. In Bayesian probability, however, not only the single triangular function shown in Figure 2.2 c, but the overlay of all possible triangular functions (their number is N-2 for N data points) weighted with their associated probability should be considered. Readers who want to know more detail about the Bayesian ideas are referred to the excellent tutorial by Sivia (2005). We skip the problem of how these matching probabilities are calculated but display the result as a curve in Figure 2.2 d. In Figure 2.2 e the functional behaviour of the data, following the above procedure, is displayed, including the uncertainty of the function estimate which should not be confused with the variability of the data.

Even more interesting than the functional behaviour of a time series is the functional behaviour of its derivative, the trend. The derivative of a linear function is a constant, the derivative of a triangular function is a constant c1 in the range $t_1<t<t_c$ and another constant c2 in the range $t_c<t<t_N$. The Bayesian result for the overall trend is, as for the function estimation, the overlay of all possible step functions weighted by their respective change point probabilities. The resulting trend estimates including pointwise uncertainties are shown in Figure 2.2 f. What happens if all three model alternatives exhibit similar probabilities?

Bayesian probability theory calculates the probability of a given model, i.e in our case the one change point model compared to one or more alternatives, such as the constant and linear models. Unless one of the models sticks out with a very high probability, the description of the data in terms of only one model is again unsatisfactory. Instead, the rigorous application of Bayesian probability theory describes that the proper functional behaviour and the proper trend are obtained by superposition of a constant, a linear function and the one change point model function again weighted with their respective model probabilities.

Depending on the investigated season and phase, the entire corrected data set was used to study the model probabilities of the one change point, linear and constant models. Higher model probabilities indicate a preference for the respective model. All model probabilities are displayed as boxplots, which show the quartiles, the range, the mean, 95% confidence interval of the median and the median (Figure 2.3).

The second data set focuses on the change point probabilities of the one change point model for two selected species with several phenological stages across the year (Figure 2.4).

For 11 Swiss stations of the third data set only, the leaf unfolding and the colouring stages of two deciduous tree species was analysed. We demonstrate the change point probabilities and trends (Figure 2.5, Figure 2.6). A particular advantage of the Bayesian approach is that it also provides us with confidence intervals for actual trends. We discuss these trends and their associated confidence intervals for leaf unfolding and colouring of the two deciduous tree species at two interesting stations with rather high change point probability curves, at Versoix (6°17’ E, 46°28’ N, 445 m) and at Ennetbuehl (9°22’ E, 47°23’ N, 900 m) (Figure 2.7, Figure 2.8). Since the change point probability distributions are normalised to unit area, the height of their maximum is a measure of how abrupt the changes occur.

2.3 Results

Analysing all 2600 observational time series in Europe (Figure 2.3), we found that in all seasons (very early spring, early spring, mid summer and early autumn, late autumn) and for all phases the one change point model possessed, on average, the highest model probabilities. In particular, the phases at the middle and at the end of the year were characterised by higher one change point model probabilities (Figure 2.3 c, d); almost all time series of mid summer, early autumn and late autumn season had median values of over 60%. The highest one change point model probability was reached in late autumn by *Aesculus hippocastanum* leaf colouring (Figure 2.3, Figure 2.3 d). No significant differences concerning model probabilities existed between forest trees, fruit trees and other plants. The probabilities of the linear model mostly ranged between 20 and 30%; but in mid summer to late autumn the constant model had very low probabilities (Figure 2.3 c, 10% on average).

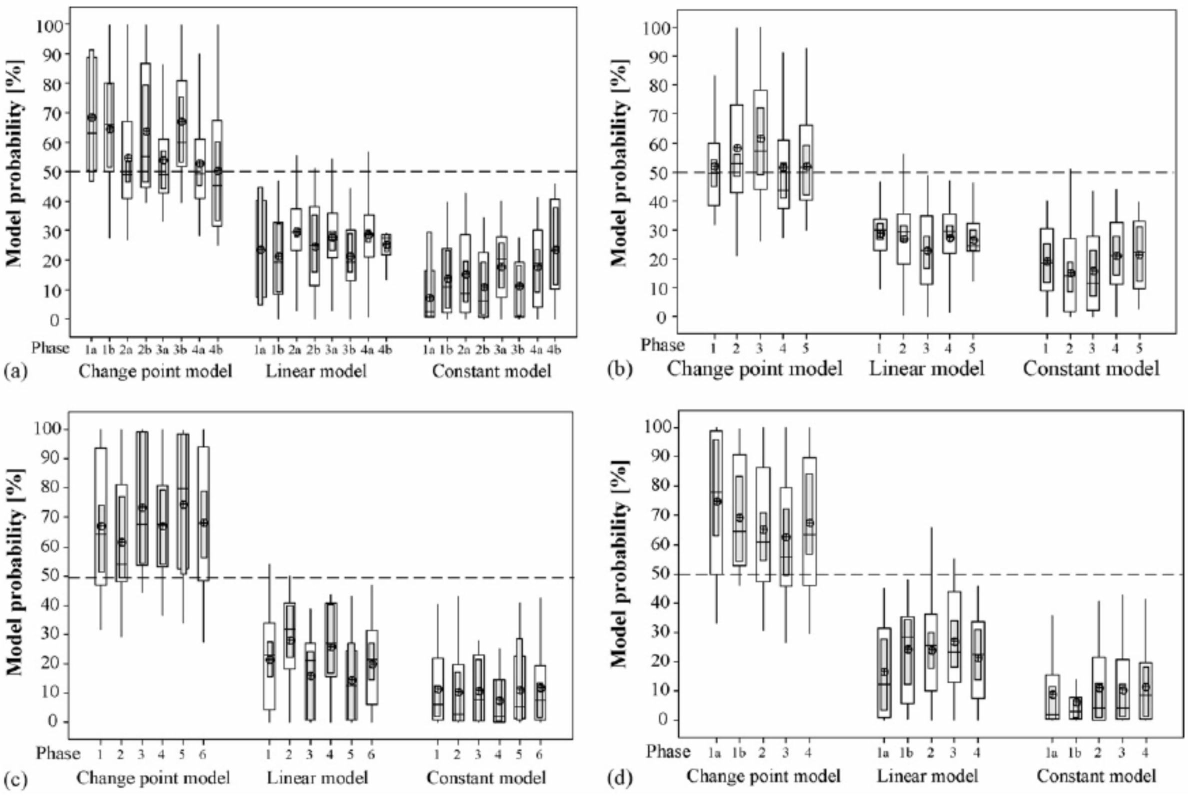


Figure 2.3: Boxplots of the one change point, linear and constant model probabilities for the four seasons: (a) very early spring, (b) early spring, (c) mid summer/early autumn and (d) late autumn (code numbers of the indicator species in Table 1). 95% confidence interval for the median is marked as the inner grey box, the 25th percentile is found at the lower end and the 75th percentile is found at the upper end of the box. The range is marked as black vertical line, the median as black horizontal line in the boxes. The mean is marked as circle with cross. The horizontal dashed line marks the 50% change point probability line.

In Figure 2.4 the respective results of analysis of phenological stages across the year of the two selected species, horse chestnut (*Aesculus hippocastanum*) and common oat (*Avena sativa*), are displayed. Both species revealed a nearly continuous increase of mean and median probabilities of the one change point model from the beginning of the growing period (leaf unfolding, flowering) until its end (full ripeness, autumn colouring).

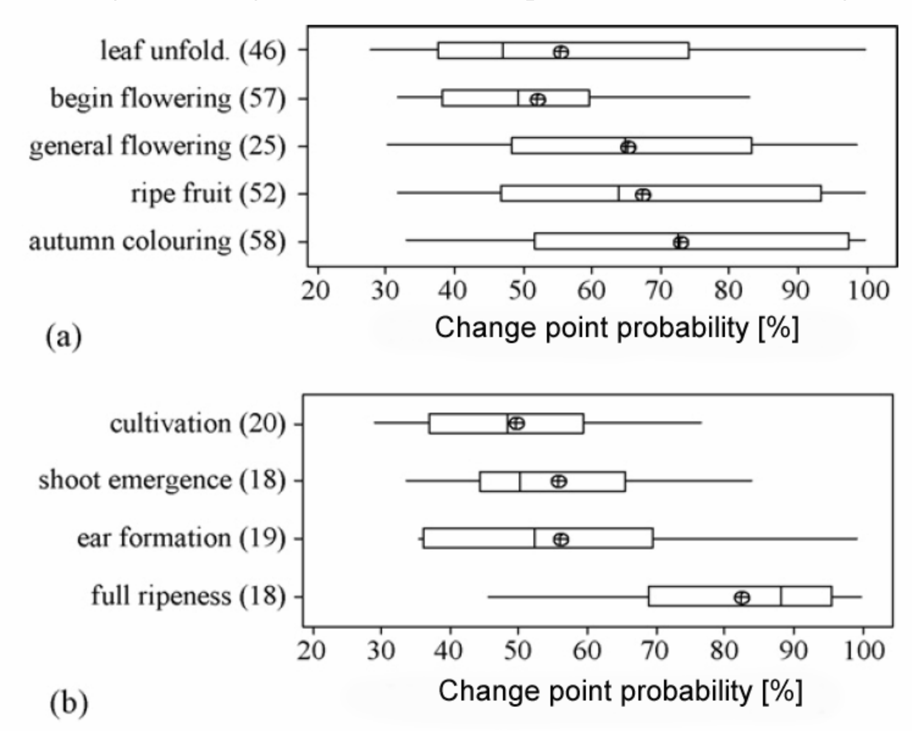


Figure 2.4: Horizontal boxplots of change point probabilities for phenological stages across the year (for description see Figure 2.3, numbers of stations in brackets). The definition of the phenological stage labelled "cultivation" includes all processes which involve a tilling and manipulation of the soil such as ploughing, disk harrowing and seed bed preparation.

The results of both Figure 2.3 and Figure 2.4 indicated that the probabilities of the three different models to describe the observational records clearly changed with season. In general, observed changes at the end of the growing season were more abrupt; thus, the one change point model was noticeably superior for description of these time series. The independent results for single species in Figure 2.4 clearly confirm the results of Figure 2.3 and prove that they are not an artefact of studying different species' phases in different seasons.

The Bayesian approach was then used for the description of phenological records within the third data set from Switzerland. The graphs in Figure 2.5 a-d illustrate the probabilities of annual change points for leaf unfolding and colouring of *Aesculus hippocastanum* and *Fagus sylvatica* at eleven stations over 40 years. There was a wide range of years where maximum change point probabilities occurred. In principle, maxima of the change point probability could occur from the very beginning of the analysed period in 1959 until 1999. However, the majority of change point probability maxima concentrated somewhere in the mid 1980s for the spring phases (Figure 2.5 a, c). Autumn phases exhibited a more dispersed pattern, with, in particular, few high change point probability maxima in the 1990s (Figure 2.5 b, d).

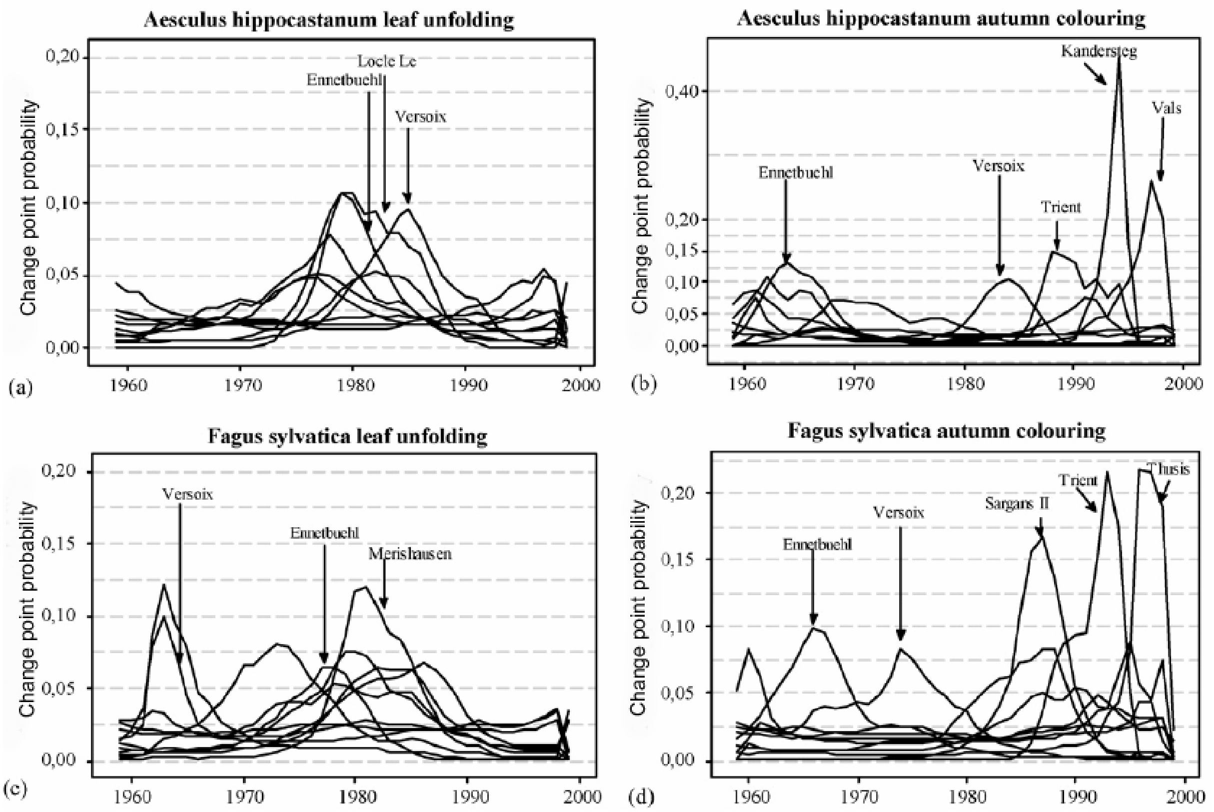


Figure 2.5: Change point probability distributions at 11 stations in Switzerland for Aesculus hippocastanum (a) leaf unfolding, (b) autumn colouring and for Fagus sylvatica, (c) leaf unfolding and (d) autumn colouring. Note that not all distribution curves are labelled.

High change point probabilities in the mid 1980s do not necessarily signify similar trend courses. They simply indicate that in this time period, temporal trends in these phenological records changed drastically. Examples are encountered within the third data set comprising complete records for the Swiss stations. During the four decades from 1959 to 1999, leaf unfolding data of the two deciduous tree species were first delayed and then advanced. The temporal trends for *Aesculus hippocastanum* (Figure 2.6 a) and *Fagus sylvatica* leaf unfolding (Figure 2.6 c) evolved, on average, from 0.22 to - 0.14 and from 0.37 to – 0.34 days/year (average data calculations are not shown in Figure 2.6). The trends of the respective leaf colouring phases in late autumn were positive in 1959 (on average 0.11 and 0.10 day/year), and thus delayed at the beginning of the time series (Figure 2.6 b, d). For autumn colouring of *Aesculus hippocastanum* (Figure 2.6 b), seven of the eleven Swiss stations displayed rather similar trends in the following decades and ended in 1999 with an average rate around – 0.35 days/year. For autumn colouring of *Fagus sylvatica* (Figure 2.6 d), eight of the eleven stations ended with an average rate around - 0.57 days/year. For the other stations, their trends exploded in the 1990s with trend values of - 7, - 4 or even - 12 days/year for *Aesculus hippocastanum* (Figure 2.6 b) and + 6.6 as well as – 5.6 days/year for *Fagus sylvatica* (Figure 2.6 d), with extremely high uncertainty margins which are not displayed in Figure 2.6.

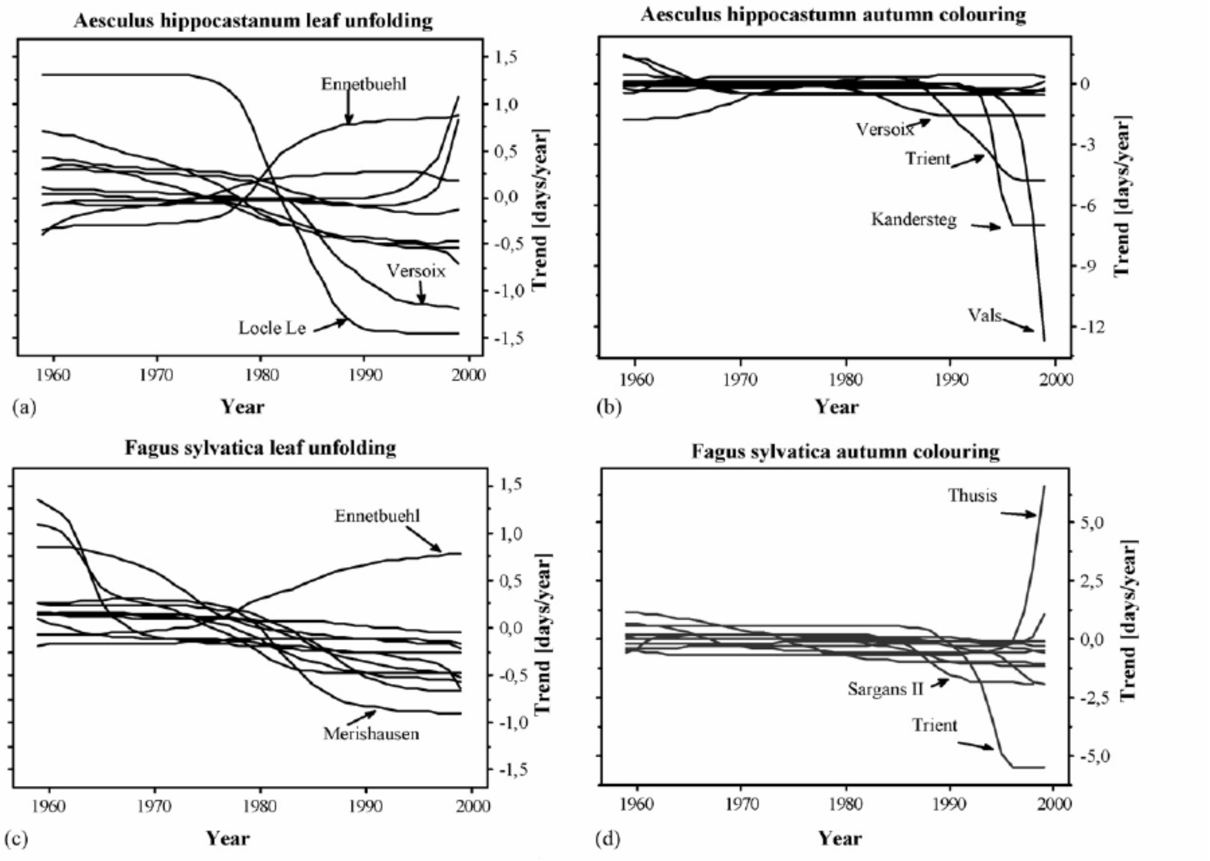


Figure 2.6: One change point model analysis at 11 stations in Switzerland for Aesculus hippocastanum (a) leaf unfolding, (b) autumn colouring and for Fagus sylvatica, (c) leaf unfolding and (d) autumn colouring. Note that only some extreme trend curves are labelled. Confidence intervals are not displayed.

For the spring phases, a few apparently deviating trend courses were identified. In Figure 2.7 and Figure 2.8, two of these stations, Versoix (6°17’ E, 46°28’ N, 445 m) and Ennetbuehl (9°22’ E, 47°23’ N, 900 m), with rather high maxima of change point probabilities are displayed, including the respective trends of the time series and their confidence intervals.

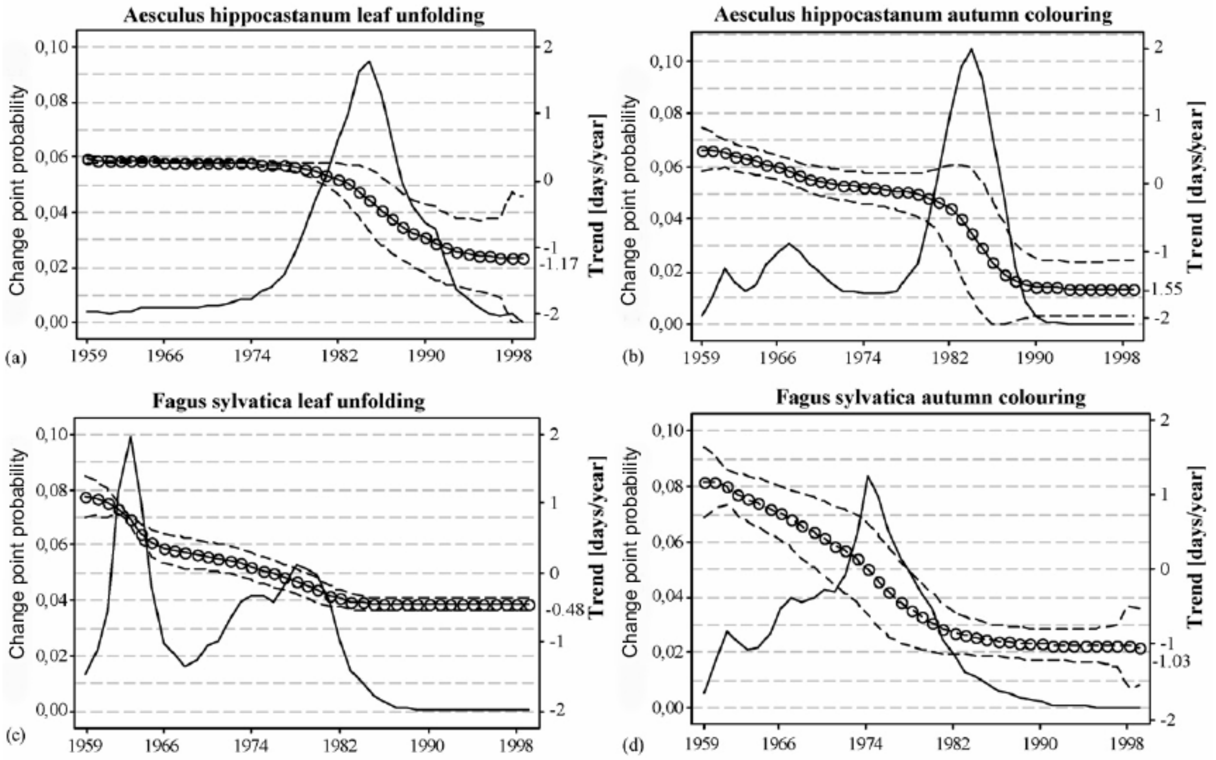


Figure 2.7: Results of the one change point model for Aesculus hippocastanum (a) leaf unfolding, (b) autumn colouring and for Fagus sylvatica, (c) leaf unfolding and (d) autumn colouring at Versoix, Switzerland. Trends are shown as lines with circles, confidence intervals as dashed lines and change point probability curves as continuous lines.

At Versoix (Figure 2.7 a, c), leaf unfolding of *Aesculus hippocastanum* and *Fagus sylvatica* exhibited delayed onset in 1959, which dropped to - 1.17 and - 0.48 days/year respectively in 1999. Autumn phases displayed a similar trend pattern at this location.

At Ennetbuehl (Figure 2.8 a, c), the spring events presented a completely inverse trend course to final positive rates of + 0.88 and + 0.77 days/year indicating also a delayed onset. At both stations, autumn phases were advancing as indicated by their negative trends. In general, the confidence intervals of the trends were relatively narrow. Moving to the end of the time series, the confidence range widened since the model estimate is supported by less data. This is not an artefact of Bayesian non-parametric function estimation but an example that Bayesian probability thereby is in fact “common sense reduced to calculation”.

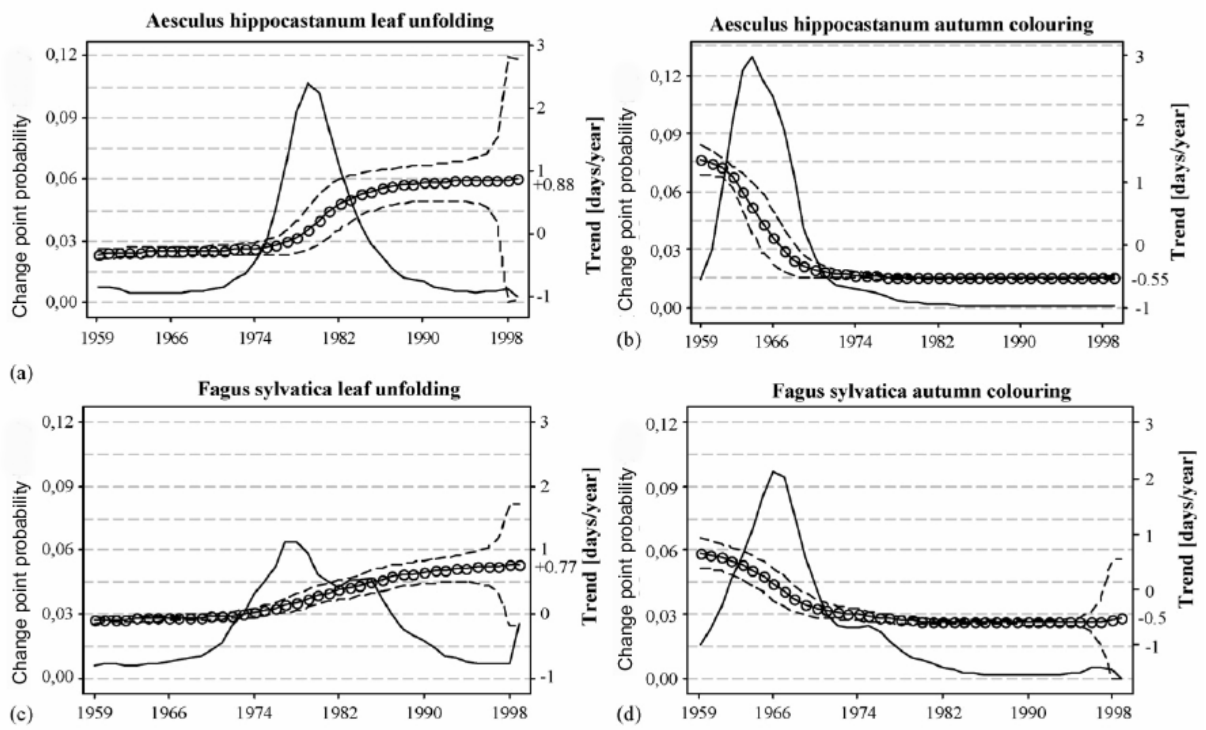


Figure 2.8: Results of the one change point model at Ennetbuehl, Switzerland (for description see Figure 2.6).

2.4 Discussion

The detection of climate change related responses in observational data remains a topic of increasing interest and importance. For various phenological networks across Europe a nearly coherent pattern of progressively earlier onset of spring and summer phases and, less homogeneously, a later end is found. Former studies were mostly based on linear trend analyses. Here, we compare the option of three models, the constant, linear and one change point models, using a new method for the analysis of phenological time series based on Bayesian concepts, recently presented by Dose and Menzel (2004).

The model comparison option gave new insight in the type of changes. Two model types were identified as less probable, thus less important; the linear model type which suggests a constant change, perhaps due to a gradual change in climate and the constant model, which can be linked to the idea of no changes in the underlying triggering factors or expressed time lags in their reaction.

The clear preference for the one change point model identified from this phenological data set signifies that there exist predominantly non-linear changes. Semenov et al. (2004) already emphasised that phenological data are non-linear functions of the temperature-regime characteristics of surface air. In this study, we prove that long-term phenological time series of key phases (Figure 2.3) for four seasons (very early spring, early spring, mid-summer / early autumn, and late autumn) could be described better by the one change point model than by traditional methods. In the model comparison, the one change point model outperforms the conventional linear and constant models by far. It is important to note that we identified relevant and mostly abrupt changes in all phenological seasons. We even reveal higher one change point model probabilities for autumn events than those in the spring. Thus, to answer question (1) in the introduction, we can conclude that the one change point model was superior in all cases. Seasons at the end of the year exhibited the highest probabilities for the one change point model.

This tendency was confirmed for two selected species, horse chestnut and common oat (Figure 2.4) with a nearly continuous increase of one change point model probabilities from stages at the beginning till those at the end of the year. These results may suggest that subsequent phases later in the growing season integrate more and longer positive temperature anomalies.

Generally, greater trends in spring were not necessarily related to a higher preference for one change point models in spring (answer to question 2 in the introduction).

Spring events at the eleven Swiss observation stations (Figure 2.5) presented a rather coherent picture of their change point probabilities; the majority of peaks was grouped somewhere in the mid 1980s where the first influences of climate warming were documented (IPCC, 2001). Since 1970 Swiss temperatures increased on average by 1.5 °C (global 0.5 °C) (BUWAL 2002). The results of Bader and Bantle (2004), who investigated the temperature and precipitation development in Switzerland (1864-2001) support our findings very well and confirm the inter-relationship between meteorology and plant phenology in Switzerland. In particular they detected highly increased temperatures in summer and winter periods in the second half of the 1980s. They highlighted that summer and winter seasons show clearly distinct temperature developments. Since 1910, winter seasons had rather constant temperatures until in 1987/1988 an increase to a pronounced warm winter period occurred. Summer periods, however, already had greater variations of increased temperatures before the 1980s.

The maxima of the change point probabilities indicated the period when the phenotypic reaction of a species to the change was most pronounced. In general, we conclude that plants received a clear temperature signal to change their phenological behaviour in spring and that their onset dates probably reflect the climatic impact more directly.

Our Swiss subset confirms the findings of our initial comprehensive model comparison that we do need non-linear functions to describe satisfactorily the respective time series in all seasons. Relevant changes seem to take place during later periods of the year as well. However, the dispersed pattern of peaks illustrated the multiplicity of possible impacts on autumn phases (Figure 2.5 b, d). All eleven Switzerland stations started with similar positive trends for the onset of spring and autumn, and reached moderate negative trends for spring and even slightly stronger ones for autumn (Figure 2.6).

In recent publications, where single or average time series have been studied by linear regression models (Sparks et al., 2000), no such detailed information was given. When comparing the regional trends of spring phases in Switzerland with European results of Menzel (2000), Menzel and Estrella (2001) and Ahas et al. (2002), we find only partly identical results. In our analysis, average trends in 1999 were - 0.14 and – 0.34 days / year for leaf unfolding of *Aesculus* and *Fagus*, whereas the trends for autumn colouring indicated advanced onset in recent years.

Menzel (2000) found, in her analysis of observational data from the International Phenological Gardens in Europe (1959- 1996), that spring events advanced by - 0.21 days / year, whereas autumn events were delayed by + 0.15 days / year. According to Menzel (2000), the mean annual growing season has lengthened by + 0.36 days / year or by 10.8 days since the early 1960s. Defila and Clot (2005) reported that the proportion of the significant trends in Switzerland is clearly higher in the alpine regions (42%) than in the lowland (33%) which is mirrored by the higher change point probabilities at higher elevation stations in our study. They also found that the phenophases occurred 20 and 15 days earlier in spring, 19 and 16 days earlier in summer, and 7 and 4 days earlier in autumn in the lowland and the alpine regions respectively.

Trends in Versoix and Ennetbuehl (Figure 2.7, Figure 2.8) have additionally been plotted with associated confidence intervals. For Versoix, the spring and autumn trend development resembles or even exaggerates the average trend at the other Switzerland stations. Horse chestnut (*Aesculus hippocastanum*) exhibited stronger changes in its phases. Beside abiotic influencing factors, biotic parameters, such as pests or diseases, should also be considered in a causal interpretation. For example, the prominent horse chestnut leaf miner *Cameraria ohridella* could have caused some odd trends in the horse chestnut time series. The multiple change point patterns for *Fagus sylvatica* phenology might suggest more confounding factors, apart from temperature triggers. Ennetbuehl is located at 900 m altitude, which is almost twice the elevation of Versoix (see Figure 2.1). Here, both spring phases were delayed by more than half a day per year whereas, in autumn, their colouring of leaves was advanced by about half a day. This means that in Ennetbuehl the response of trees to the climate change is a clear shortening of the vegetation period, perhaps due to a more sensitive reaction of mountain ecosystems. This clearly suggests that there are differences among different altitudes, which must be taken into account (third question in the introduction). The major findings of Defila and Clot (2005) concerning the trends of leaf unfolding and autumn colouring are confirmed by our study, which also emphasizes the strong regional and altitudinal differences occurring in Switzerland.

In general, the Bayesian approach provides a unique tool for phenological time series analysis as the reliability of the functional behaviour and of the trend is represented by confidence intervals, which are diagnostically useful as observed odd changes in our study emphasise. Plant phenological research provides fundamental knowledge on the development rhythm of plant species and of influencing factors. Plants have different sensitivities and responses to climate changes which may lead to changes in population dynamics. Differences in their phenological response may affect the competition among species (Kramer et al., 2000) and promote those with a better adaptive response. Changes in species distribution and abundance are expected to result from climate change, which may have positive and negative effects. New crop varieties can become productive in specific regions and unknown diseases or weeds may appear in certain areas. Thus, the phenological responses should be examined at all levels of the plant-environment system (Beaubien, 1996), as a continued trend towards earlier development is very likely to happen.

The frequent observations of a non-linear change underline the importance of Bayesian analyses. Often the strongest changes in the length of the growing season, which is a central clue for the variation of carbon exchange of terrestrial ecosystems, take place at the beginning of the vegetation period. The earlier spring phenophases are the most consistent indicators, however late spring species profit more from a prolongation in spring time than early spring species because they exploit longer days (Schaber, 2002). Bayesian nonparametric function estimation is the method of choice for the identification of changes in phenological records and will allow a better large scale attribution of these changes to the temperature evolution in the second half of the 20th century.

2.5 References

Ahas R., (1999) Long-term phyto-, ornitho- and ichthyophenological time-series analyses in Estonia. *Int. J. Biometeorol.* **42**, 119–123.

Ahas R., Aasa A., Menzel A., Fedotova V.G., Scheifinger H., (2002) Changes in European spring phenology. *Int. J. Climatol.* **22** (14), 1727–1738.

Bader S., Bantle H., (2004) Das Schweizer Klima im Trend. Temperatur- und Niederschlagsentwicklung 1864-2001. Veröffentlichung der *MeteoSchweiz* **68**, 45.

Beaubien E.G., (1996) Plantwatch, a model to initiate phenology in school classes. *Phenol. Season* **1**, 33–35.

Bradley N.L., Leopold A.C., Ross J., Huffaker W., (1999) Phenological changes reflect climate change inWisconsin. In: Proceedings of the National Academy of Sciences USA *Ecology* **96**, 9701–9704.

BUWAL (Bundesamt für Umwelt,Wald und Landschaft) (2002) *Das Klima in Menschenhand—neue Fakten und Perspektiven.* Bern.

Defila C., Clot B., (2001) Phytophenological trends in Switzerland. *Int. J. Biometeorol.* **45** (4), 203–207.

Defila C., Clot B., (2005) Phytophenological trends in the Swiss Alps, 1951–2002. *Meteorologische Zeitschrift.* **14** (2), 191–196.

Dose V., Menzel A., (2004) Bayesian analysis of climate change impacts in phenology. *Global Change Biol.* **10**, 259–272.

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

IPCC, (2001) Climate Change 2001—The Scientific Basis. Contribution of Working Group I to the Third assessment Report of the Intergovernmental Panel on Climate Change. Houghton, J.T., Ding, Y., Griggs, D.J., Noguer,M., van der Linden, P.J., Dai, X., Maskell, K., Johnson. (CA Eds.), Cambridge University Press, UK

Jones G.V., Davis R.E., (2000) Climate influences on grapevine phenology, grape composition, and wine production and quality for Bordeaux, France. *Am. J. Ecol. Viticult.* **51**, 249–261.

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

Lauscher F., (1978) Neue Analysen ältester und neuerer phänologischer Reihen. *Archiv für Meteorologie Geophysik 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.

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., Edwards, P.J. (Eds.), Fingerprints of Climate Change—Adapted Behaviour and Shifting Species Ranges. Kluwer Academic/Plenum Publishers, NewYork and London, pp. 123–137.

Menzel A., Estrella N., Fabian P., (2001) Spatial and temporal variability of the phenological seasons in Germany from 1951 to 1996. *Global Change Biol.* **7** (6), 657–666.

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

Menzel A., Estrella N., Heitland W., Susnik A., Dose V., (2008) Species specific effects on the lengthening of the growing season in two European countries. *Int. J. Biometeorol.*, **52**, 209–218.

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

Peñuelas J., Filella I., Comas P., (2002) Changed plant and animal life cycles from 1952 to 2000 in the Mediterranean region. *Global Change Biol.* **8** (6), 531–544.

Root T.L., Price J.T., Hall K.R., Schneider S.H., Rosenzweig C., Pounds J.A., (2003) Fingerprints of global warming on wild animals and plants. *Nature 421*, 57–60.

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

Schaber J., (2002) Phenology in Germany in the 20th Century: methods, analyses and models. Ph.D. Thesis, University of Potsdam, Potsdam, Germany. http://pub.ub.uni-potsdam.de/2002/0022/schaber.pdf.

Scheifinger H., Menzel A., Koch E., et al., (2002) Atmospheric mechanisms governing the spatial and temporal variability of phenological observations in central Europe. *Int. J. Climatol.* **22,** 1739–1755.

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

Schnelle F., (1955) Pflanzen-Phänologie (Plant Phenology)—Akademische Verlagsgesellschaft Geest and Portig K.-G.

Schwartz M.D., Reiter, B.E., (2000) Changes in North American Spring. *Int. J. Climatol.* **20**, 929–932.

Semenov S.M., Koukhta, B.A., Gelver, E.S., (2004) Nonlinearity of climate-driven changes in phenological dates in woody plants. *Doklady Biol. Sci.* **396**, 221–223.

Sivia D.S., (2005) Data Analysis—A Bayesian Tutorial. Clarendon, Oxford.

Sparks T.H., Carey, P.D., (1995) The responses of species to climate over two centuries: an analysis of the Marsham phenological record. *J. Ecol.* **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. *Int. J. Biometeorol.* **44** (2), 82–87.

Sparks T.H., Menzel, A., (2002) Observed changes in seasons: an overview. *Int. J. Climatol.* **22**, 1715–1725.

Walther G.R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.C.J., Fromentin, J.M., Hoegh-Guldberg, O., Bairlein, F., (2002) Ecological responses to recent climate change. *Nature* **416**, 389–395.

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

Christoph **Schleip** [a], Annette **Menzel** [a] and Volker **Dose** [b]

[a] *Fachgebiet für Ökoklimatologie, Technische Universität München, Am Hochanger 13, D-85354 Freising - Weihenstephan, Germany*

[b] *Max-Planck-Institut für Plasmaphysik, Boltzmannstraße 2, D-85748 Garching bei München, Germany*

Agricultural and Forest Meteorology (2008), doi:10.1016/j.agrformet.2007.11.008.

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 a nonlinear way. 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. In these temperature and Norway spruce bud burst time series we detected 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 besides commonly used sums of daily mean temperatures, forcing temperatures in phenology models should also include solutions where weighted effective temperatures in a sensitive time span are considered.

Keywords: Bayesian analysis; Simulated annealing; Norway spruce; Bud burst; Time series; Nonlinearity; Temperature response; Prompt response.

3.1 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 (*P. abies* (L.) Karst.), 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; 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 explanations of the rates of forcing. For an improved understanding of ecological responses to climate change we seek methods that are equally applicable to nonlinear changes in time series and to 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 temperature 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 and 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 3-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 an overlay of a constant, linear and a change point rate of change weighted by their respective model probabilities.

3.2 Material and Methods

3.2.1 Climatic data

Daily temperature data from 18 meteorological stations (met stations) in Germany (Figure 3.1) 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 to 567m 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 number in the range 0–53. Monday was considered as the first day of the week.

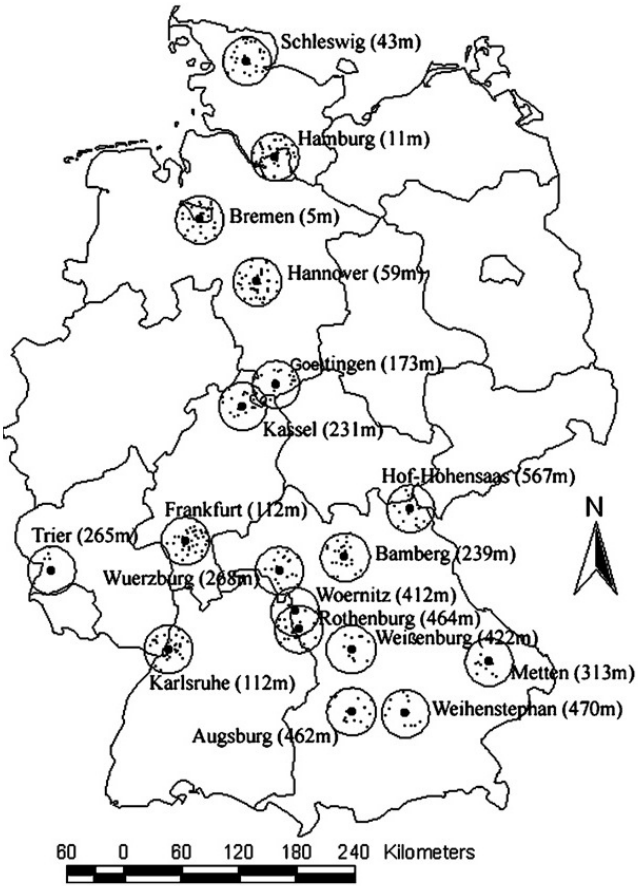


Figure 3.1: Distribution and altitude of the climate stations in Germany (big dots) and corresponding phenological stations (small dots). The radius of the circles around each climate station is 25 km.

3.2.2 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 1600 stations observe defined plant development stages (DWD, 1991). The phenological data were obtained by averaging time series from at least eight phenological stations within a 25 km radius of the met station (Figure 3.1), that 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 the influence of local microclimate. In this study, we focused on the phenophases of bud burst of Norway spruce (*P. 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 (Figure 3.2). The DWD observer manual (DWD, 1991) defines a bud as opened when the fresh green needles are clearly visible and separated.

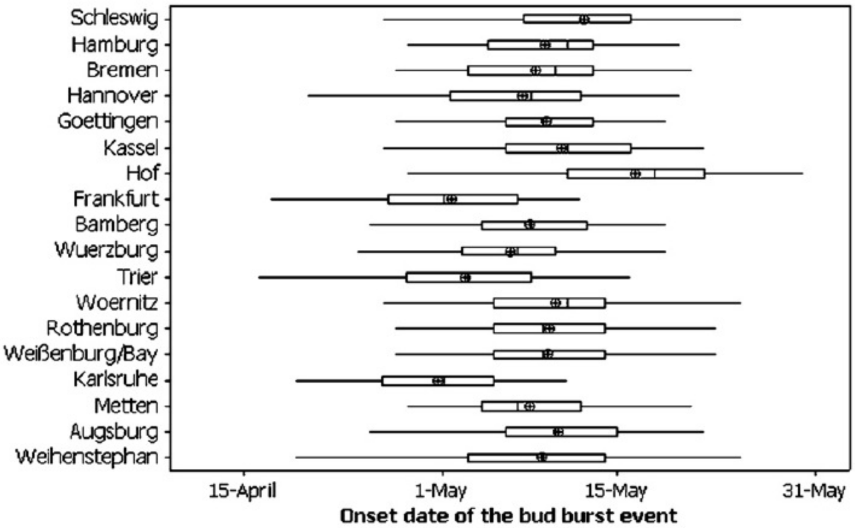


Figure 3.2: Horizontal boxplots of the onset date of bud burst at all 18 climate stations. The 25th percentile is found at the left end and the 75th percentile is found at the right end of the box. The range is marked as black horizontal line, the median as black vertical line in the boxes. The mean is marked as circle with cross.

3.2.3 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 used. Here we briefly summarize the main features of the Bayesian approach. An important feature of Bayesian probability theory is that it allows the 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: a constant, a linear and a change point model. The constant model assumes a functional behaviour constant in time with an associated rate of change of zero. 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 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 Figure 3.3 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 discussion of Schleip et al., 2008) 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 Occam`s razor because of overestimation.

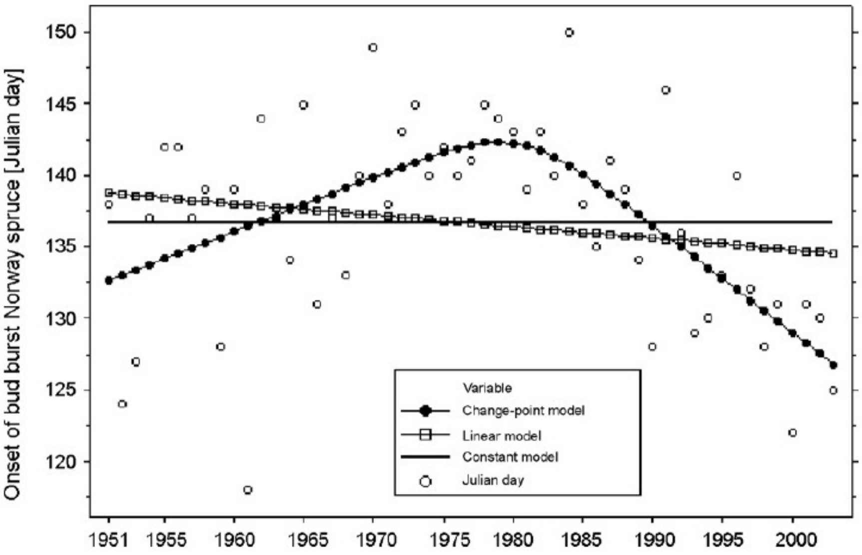


Figure 3.3: Bayesian change point, linear and constant model estimation of the onset of bud burst Norway spruce (Picea abies L.) in Hof. In this example the change point model exhibits a probability of 100%.

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 marginalization rule overlays all possible triangular functions and then weigh 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 Figure 3.4 a and Figure 3.4 b. 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 for 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 marginalization 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.

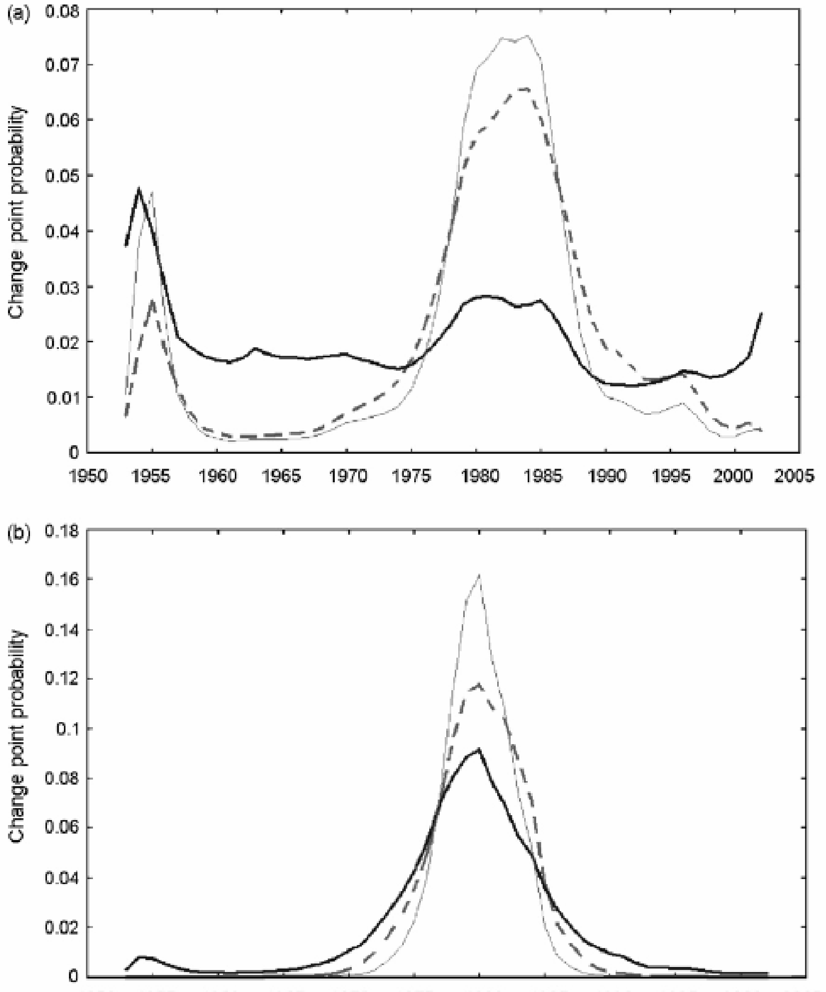


Figure 3.4: Distributions of temperature, bud burst and joint (temperature and bud burst) change point probability of Norway spruce bud burst (Picea abies L.) in Schleswig (a) and in Hof (b). In the upper panel the coherence factor has a value of 1.2 and in the lower panel a value of 3.3. Note that the 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.

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 that 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

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

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 $C_{\text{new}} < C_{\text{old}}$ 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. Figure 3.5 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 (Figure 3.5), 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.

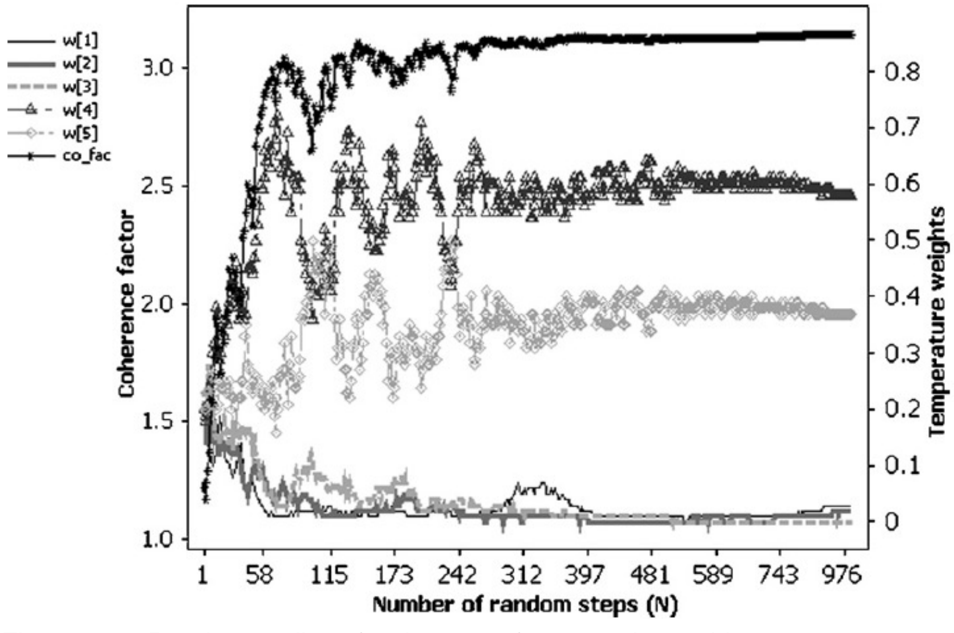


Figure 3.5: Random walks of coherence factor and monthly mean temperature weights using the simulated annealing approach for Norway spruce (Picea abies L.) in Hof, Germany. w[1]–w[5] are weights of January–May mean temperatures, respectively, co_fac = coherence factor. Note that the x-axis shows the 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.

3.3 Results

3.3.1 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 (Figure 3.6 a). 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 over all stations of the monthly temperature time series increased from 39% in January to a maximum probability of 61% in May (Figure 3.6 b). The linear and constant models exhibited a considerably large probability in the months January–March. The mean probability of the constant model of all stations had its maximum for March temperatures (47%).

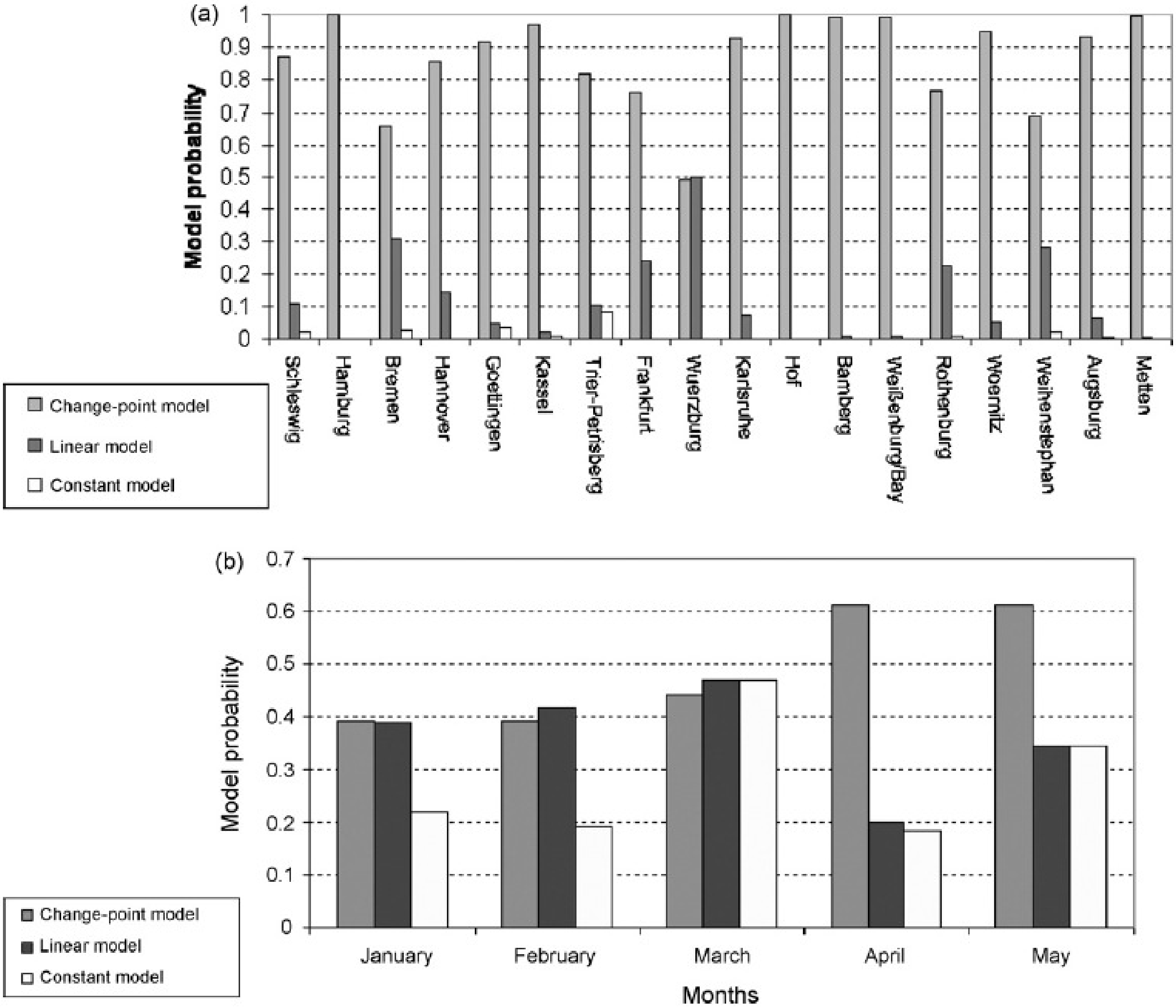


Figure 3.6: Bayesian model probabilities of the change point, linear and constant model of (a) Norway spruce bud burst at 18 phenological stations in Germany and of (b) mean temperatures from January to May at 18 corresponding climate station.

3.3.2 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 Figure 3.7 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 (Figure 3.7 d).

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% (Figure 3.7 a). 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 (Figure 3.7 a).

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 Figure 3.7 b and c. The joint (temperatures and phenological) change point probability distribution exhibited a clear single peak at the beginning of the 1980s (Figure 3.7 d).

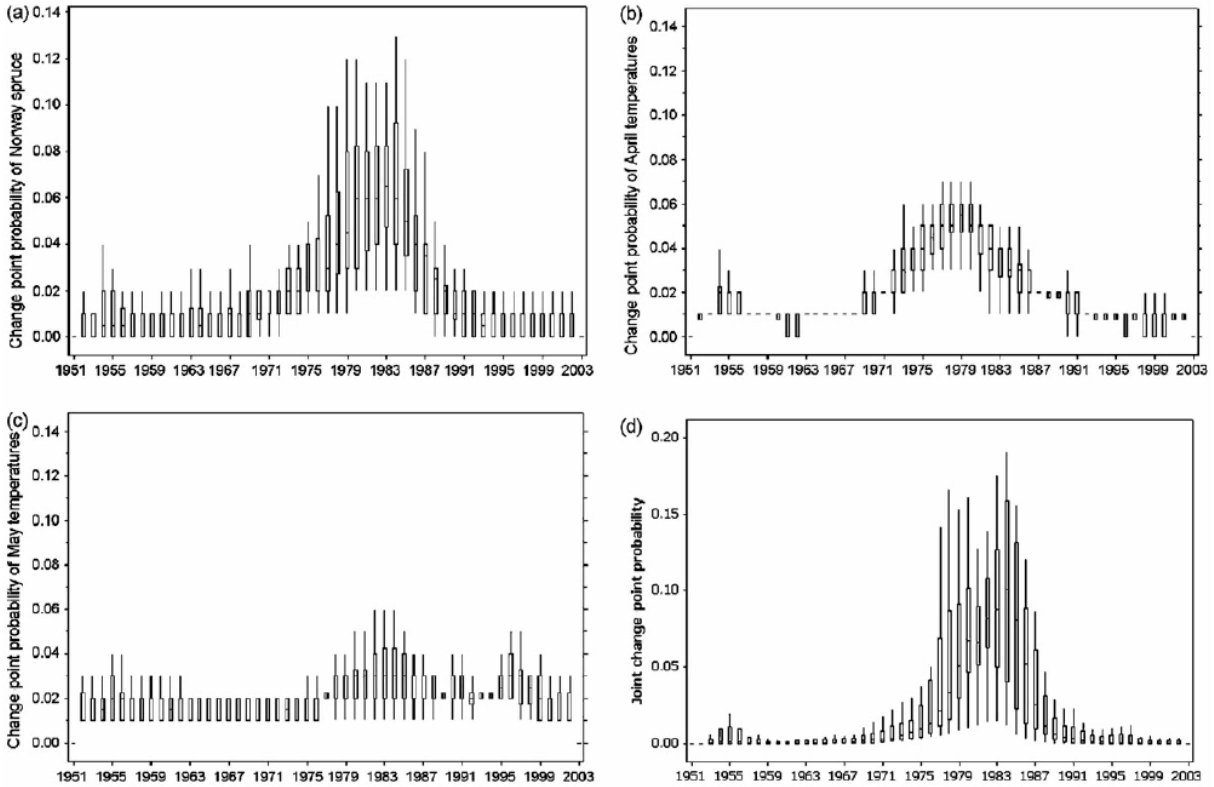


Figure 3.7: Box plots of change point probability distributions of (a) Norway spruce bud burst at 18 phenological stations and of (b) April mean temperature time series and of (c) May mean temperature time series and of (d) joint (temperature and phenological) change point probability at the corresponding 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.

3.3.3 Coherence factors

To answer the second question of whether temperature and Norway spruce time series evolved independently or coherently we calculated coherence factors for each station in Germany. At 9 of 18 stations, the coherence factors had a value of two or higher (Figure 3.8 a). 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, 2.40 (not shown in Figures).

3.3.4 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 (Figure 3.8 a). 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 (Figure 3.8 a). Similar results can be seen in the weekly resolution. Within April (week numbers 14–17) and May (week numbers 18 and 20) the temperature weights had their highest values (Figure 3.8 b). 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.

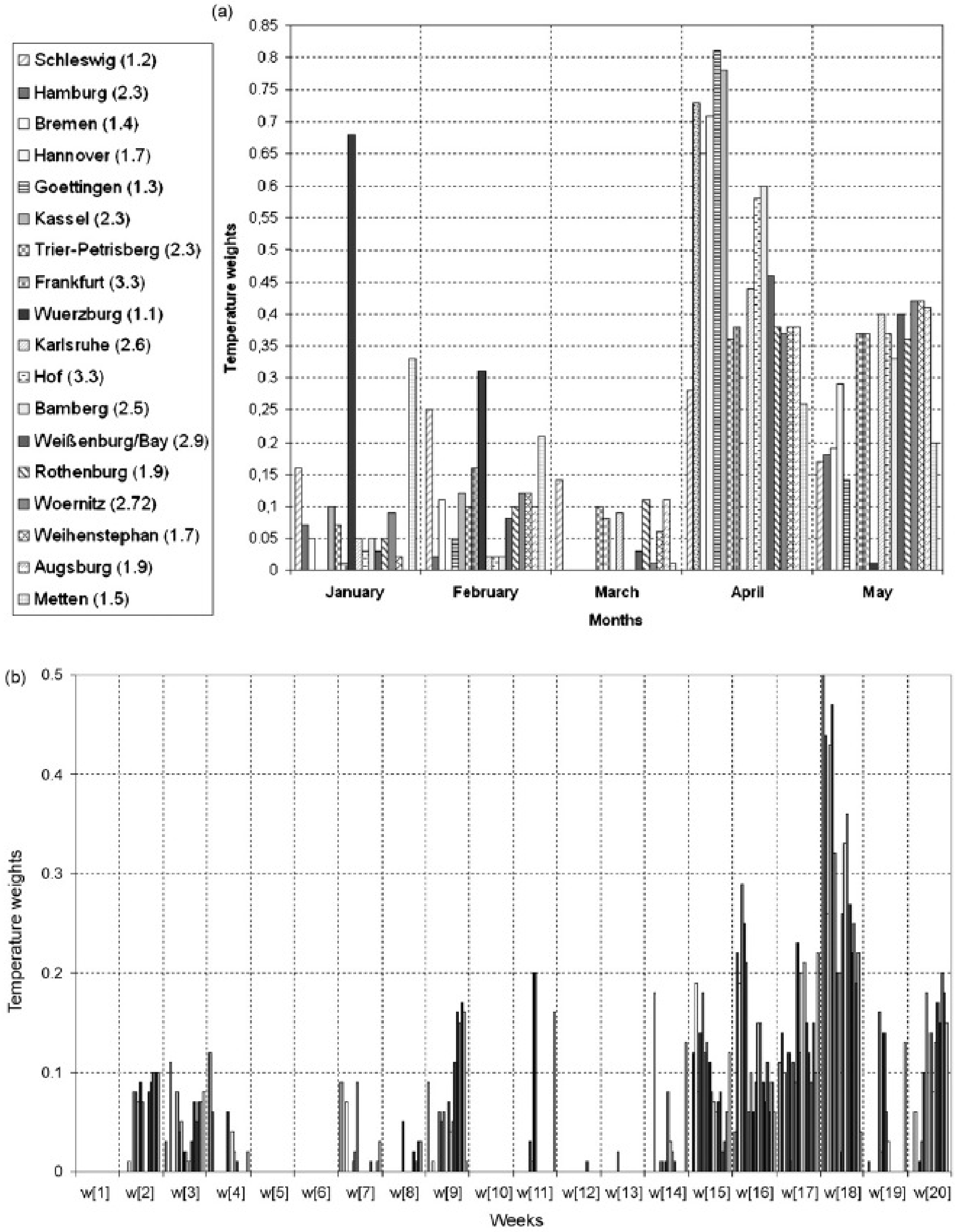


Figure 3.8: Coherence factors and (a) monthly and (b) weekly temperature weights of bud burst Norway spruce in Germany. In (a) the coherence factors are in brackets following the names of the climate stations. The bars represent the temperature weights for (a) the months January–May and for (b) the weeks since the beginning of the year. Temperature weights were obtained by the simulated annealing optimization.

3.3.5 *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 (Figure 3.9 a), 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} (Figure 3.9 a). At the beginning of the time series in 1951 the rates of change ranged from 0.35 to -0.02 days year^{-1} which means 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, rates of change 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 all investigated 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}.

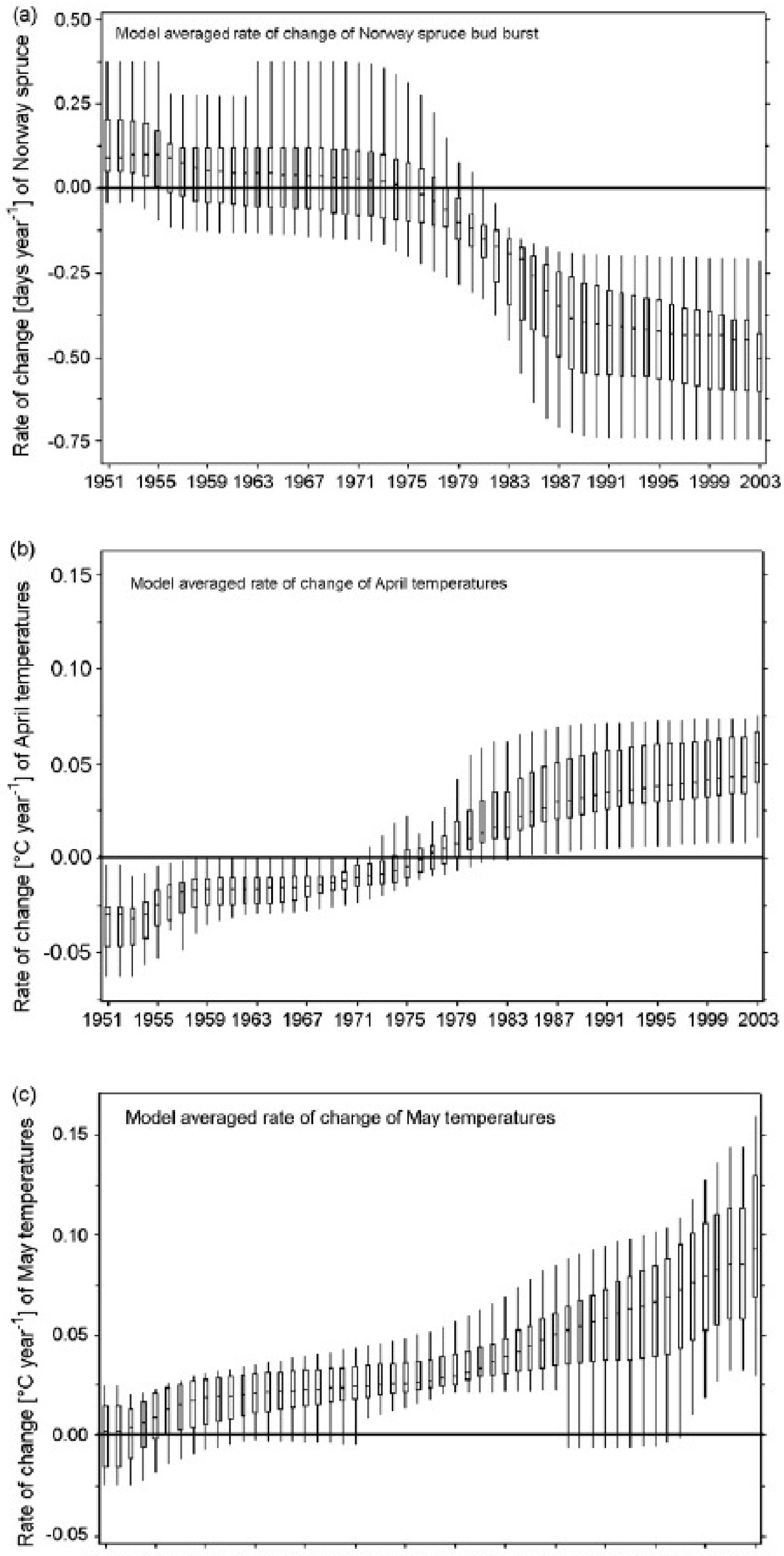


Figure 3.9: Box plots of Bayesian model averaged rates of change of (a) Norway spruce bud burst at 18 phenological stations in days year $^{-1}$ and of (b) April mean temperature time series and of (c) May mean temperature time series in °C year^{-1} at the corresponding 18 climate stations. Model averaged rates of change were calculated for the period 1951–2003.

3.4 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 into 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. Using the approach of Dose and Menzel (2004) we showed 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 showed 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 change 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 Figure 3.4 b). 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 drawn 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 respond 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 rates of change of 18 climate stations ranged from 0.03 to 0.17 °C year^{-1} whereas in 1951 May temperature rates of change ranged from -0.03 to 0.03 °C year^{-1}. In contrast, the rate of change of bud burst of Norway spruce in 2003 was estimated from -0.25 to -0.75 days year^{-1}, but 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–May temperatures. Menzel (2003) calculated the subset regression between phenological anomalies of bud burst of Norway spruce and the 3-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 a fixed temperature sum or a 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 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 might 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 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.

3.5 Conclusion

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 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 the temperature respond pattern for each species. 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 (Garrett, 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.

3.6 References

Anscombe F.J., (1973). Graphs in statistical analysis. *Am. 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., Smith R.I., (1983) Thermal time, chill days and prediction of budburst in *Picea sitchensis*. *J. Appl. Ecol.* **20**, 951–963.

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 Res.* **19**, 257– 264.

Chuine I., Cour P., Rousseau D.D., (1998) Fitting models predicting dates of flowering of temperature-zone trees using simulated annealing. *Plant Cell Environ.* **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 Biol.* **12**, 1451–1459.

Dose V., Menzel, A., (2004) Bayesian analysis of climate change impacts in phenology. *Global Change Biol.* **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) Ockham's Razor. In: Grandy, W.T., Schick, L.H. (Eds.), *Maximum Entropy and Bayesian Methods*. Kluwer, Dordrecht, pp. 357–364.

Gregory P.C., (2005) Bayesian logical data analysis for the physical sciences. In: *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., (1990) 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: Dixon, R.K., Meldahl, R.S., Ruark, G.H., Warren, W.G. (Eds.), Process Modeling of Forest Growth Responses to Environmental Stress. Timber Press, Portland, pp. 159–165.

Häkkinen R., Linkosalo T., Hari P., (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 Physiol*. **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., Lechowicz M.J., (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 Environ***.** **17**, 367–377.

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

Linkosalo T., (2000. Mutual regularity of spring phenology of some boreal tree species: predicting with other species and phenological models. *Can. J. Forest Res*. **30**, 667–673.

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

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 Biol.* **12** (10), 1969–1976.

Murray M.B., Cannel M.G.R., Smith R.I., (1989) Date of bud burst of fifteen tree species in Britain following climatic warming. *J. Appl. Ecol.* **26**, 693–700.

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. *Int. J. Climatol.* **22**, 1739–1755.

Schnelle F., (1955) Pflanzenphänologie. Probleme der Bioklimatologie. 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. *Agric. Forest Meteorol.* **141**, 179–191.

Schleip C., Rutishauser T., Luterbacher J., A. Menzel., (2008) Time series modelling and central European temperature impact assessment of phenological records in the last 250 years, *J. Geophys. Res.,* **113**, G04026.

Sparks T.H., Carey P.D., (1995) The responses of species to climate over two centuries: an analysis of the Marsham phenological record. *J. Ecol.* **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. *Int. J. Biometeorol.* **44** (2), 82–87.

4 Time series modelling and central European temperature impact assessment of phenological records in the last 250 years

Christoph **Schleip** [a], This **Rutishauser** [b], Jürg **Luterbacher** [b] and Annette **Menzel** [a]

[a] *Fachgebiet für Ökoklimatologie, Technische Universität München, Am Hochanger 13, D-85354 Freising - Weihenstephan, Germany*

[b] *Oeschger Centre for Climate Change Research (OCCR) and NCCR Climate, University of Bern, Institute of Geography, Climatology and Meteorology, CH-3012 Bern, Switzerland*

Journal of Geophysical Research- Biogeosciences (2008), 113, G04026, doi:10.1029/2007JG000646

Abstract

Long-term spring and autumn phenological observations from Switzerland and Burgundy (eastern France) as well as long-term Swiss monthly and seasonal temperature measurements offer a unique possibility to evaluate plant phenological variability and temperature impacts over the last 250 years. We compare Pearson correlation coefficients and linear moving window trends of two different lengths with a Bayesian correlation and model comparison approach. The latter is applied to calculate model probabilities, change point probabilities, functional descriptions and rates of change of three selected models with increasing complexity and temperature weights of single months.

Both approaches the moving window trends as well as the Bayesian analysis, detect major changes in long-term phenological and temperature time series at the end of the 20th century. Especially for summer temperatures since the 1980s, Bayesian model averaged trends reveal a warming rate that increased from an almost zero rate of change to an umprecenednted rate of change of 0.08°C/year in 2006. After 1900, temperature series of all seasons show positive model-averaged trends. In response to this temperature increase, the onset of phenology advanced significantly. We assess the linear dependence of phenological variability by a linear Pearson correlation approach. In addition we apply the Bayesian correlation to account for nonlinearities within the time series. Grape harvest dates show the highest Bayesian correlations with June temperatures of the current year. Spring phenological phases are influenced by May temperatures of the current year and summer temperatures of the preceding growing season. For future work we suggest testing increasingly complex time series models such as multiple change point models.

Keywords: Bayes; phenology; historical climatology; Bayesian model comparison; Bayesian correlation; linear trend; change point probability; temperature impact; long-term record; grape harvest dates.

4.1 Introduction

Phenology has traditionally consisted of the study of the rhythm of biological phenomena mainly related to climate (e.g. Schnelle, 1955; Schwartz, 2003). In many European countries, National Meteorological Services have been organise phenological recordings since the second half of the 20th century (Menzel, 2003a). Some networks already existed at the beginning of the 20th century (Schnelle, 1955; Menzel, 2003a). In earlier centuries, phenological knowledge improved the understanding of the variability of life cycle events for agricultural purposes (e.g. Pfister, 1999; Burri and Rutishauser, 2007). The longest written phenological record is probably the record of the beginning of flowering of cherry at the royal court of Kyoto, Japan, which dates back to AD 705 (Sekiguti, 1969; Menzel, 2002; Aono and Kazui, 2007). One of the oldest and longest European sets of phenological observations is the Marsham family record in Norfolk, UK (1736-1947) (Sparks and Carey, 1995). The Economical Society of the State of Bern (Switzerland) established the first longer running phenological network comparable with present monitoring networks in 1759 (Pfister, 1975; Burri and Rutishauser, 2007). The first European-wide phenological networks were initiated and installed by the Societas Meteorologica Palatina at Mannheim (1781-1792) and by Hoffmann and Ihne (1881-1941) (Schnelle, 1955).

Phenological information from documentary sources, such as dates of grape harvests, sea-ice-free periods in harbours and diaries describing the occurrence of frost or heat waves, have been included in multiproxy climate reconstructions (e.g. Luterbacher et al., 2004, 2007; Xoplaki et al., 2005; Guiot et al., 2005; IPCC WG1, 2007) and indicate warm or cold periods in particular regions (see Brázdil et al., 2005 for a review). Recently, records of grape harvest dates from Burgundy (Chuine et al., 2004; Menzel, 2005; Guiot et al., 2005; Le Roy Ladurie et al., 2006) and from Switzerland (Pfister, 1992; Meier et al., 2007) were used to reconstruct late spring-summer temperatures for the last couple of centuries. The timing of agricultural work, such as grape harvest, is tightly related to temperatures over the preceding months (Menzel, 2005; Chuine et al., 2004; Pfister, 1999; Bradley, 1999). For this reason harvest dates are used as a climate indicator for summer and growing season temperatures. Grape harvest dates provide very long, precisely dated and uninterrupted series of regional temperature anomalies without chronological uncertainties and thus, provide a good method for understanding interannual temperature variability. Grape harvest dates are precisely documented but underlie changing viticultural traditions, war times that attracted mercenary soldiers from farmer communities, diseases and other environmental influences than temperature (see discussion and references in Meier et al., 2007).

For spring season, Rutishauser and Studer (2007) compared reconstructed temperature measurements and a unique compilation of cherry tree flowering dates for the Swiss Plateau region 1721–2003. The flowering record was used to assess the impact of spring temperature on phenology. Subsequently, Rutishauser et al. (2007) reconstructed a so-called statistical “Spring plant” from several spring phenological phases back to 1702 in order use the large number of historical phenological records in archives.

Trends and changing trends in phenological time series

Methodologically, trends in time-series are often analysed using simple linear regression where phenological dates or temperatures are plotted against time (for phenology, see e.g. Menzel and Fabian, 1999; Walther et al., 2002; Parmesan and Yohe, 2003; Root et al., 2003;

Menzel et al., 2006). The slope of the linear regression equation indicates the average rate of change in phenology (days/year) or temperature (°C/year). This method can be easily applied to a large number of sites to compare differences between species and sites. The main disadvantages of this least squares method are their restriction to time series exhibiting more or less linear performance, possibly poor extrapolation properties, and sensitivity to outliers/ extremes and boundary values. For inherently nonlinear processes it becomes difficult to find a linear model that fits the data well as the range of the data increases. Finally, while the method of least squares often provides optimal estimates of unknown parameters, it is very sensitive to the presence of outliers in the data used to fit a statistical model. One or two outliers can sometimes seriously skew the results of a least squares analysis (von Storch and Zwiers, 2001, NIST/SEMATECH, 2006).

Tomé and Miranda (2004) identified change points in several climatological time series by the least squares method and determined the most appropriate continuous set of straight lines. Their method fitted a given time series with the conditions of a minimum 15 year interval between breakpoints and of changing sign of slope between two consecutive trends. Since their results strongly depend on the selected length of each linear segment, Tomé and Miranda (2004) tested the best continuous set of straight lines also with conditions of a minimum year period of 10, 20 or 30 years. Moving linear trend window analysis (e.g. Menzel et al., 2004, Rutishauser et al., 2007a) was another attempt to overcome the shortages of a priori decisions of window length for linear trend analyses.

In consequence, Dose and Menzel (2004) introduced a Bayesian approach of model comparisons to evaluate the fit of a constant, a linear regression, and a change point model on time series. The Bayesian analysis has the great advantage of analysing varying changes, model probabilities and change point probabilities of time series. Along with rates of change, rigorously calculated uncertainties of model averaged rates of change and linear trends can be described.

Temperature impact on changing phenology

Environmental impact on plant phenology has been studied by correlation and regression analysis (e.g. Menzel, 2003b, Menzel et al., 2006, Rutishauser et al., 2008). Studies of Sparks and Carey (1995), Sparks et al. (2000), Menzel (2003b), Dose and Menzel (2006) and Rutishauser and Studer (2007) also pointed to the fact that the timing of spring phenophases such as the date of first flowering, bud break, unfolding of first leaf or first bird migration is clearly correlated with climate variables and responds mainly to temperature (for review, see also Rosenzweig et al. 2007). Dose and Menzel (2006) used a conceptually new Bayesian correlation approach that was methodically improved by Schleip et al. (2008). In this method, the coherence of long-term temperature and phenological time series is estimated to determine and weight single monthly and seasonal three-monthly averaged temperature impacts. They used a simulated annealing optimization algorithm to receive a coherence factor and temperature weights (Schleip et al., 2008).

In this paper we apply the linear trend analysis and Bayesian model comparison to an investigation of three unique, multi-decadal, phenological time series from Switzerland and France from 1753 to the present (Rutishauser et al., 2007; Chuine et al., 2004; Meier et al., 2007). Additionally we compare the phenological records with independent Swiss instrumental temperature measurements starting in 1753 in order to assess the monthly temperature impact on phenological variability of the past three centuries by Pearson and by Bayesian correlation. We compare the ability of different methods of analysis to describe variation in temperature measurements from Switzerland. Our unique analysis of three long phenological records starting in the middle of the 18th century allows us to study the temporal evolution of the phenological records and to assess the key environmental factor affecting phenology, namely temperature. Whereas grape harvest dates have already figured prominently in historical phenology and provide a proxy record for climate reconstruction, the statistical "Spring plant" is the first homogenised compilation of a long-term phenological record from different plant species that spans several centuries. We aim to show detailed analyses of past ecological information that provide an important source for understanding long-term ecological change (Cheke, 2007). Our analysis not only applies unique longterm phenological records but also uses rigorous, robust statistics to assess temperature as the main environmental forcing factor for the first time.

4.2 Material and Methods

4.2.1 Material

We selected long-term phenological observations and temperature measurements from Switzerland and Burgundy (eastern France) for the time period 1753–2003/2006. All phenological dates were transformed into Julian days January 1st=1. The altitudes of observation and measurement sites are between 200 and 800 m a.s.l..

Phenological observations

Swiss “Spring plant” 1753-2006

We use a sub period (1753-2006) of a reconstructed statistical “Spring plant” that describes Swiss plant phenological spring variability for 1702-2006 (Rutishauser et al. 2007). The ”Spring plant“ is defined as the weighted mean of apple and cherry tree flowering and first leaves of beech. Each selected phase represents a spring event within two weeks at the end of April and beginning of May. Historical observations were extracted from the Euroclimhist database (Pfister and Dietrich-Felber, 2006) and from Vassella (1997). For 1951-2006 the phenological data were extracted from the Swiss Phenological Database (Defila and Clot, 2005). To construct the “Spring plant”, a mixed linear modelling approach (Schaber and Badeck, 2002) was applied to estimate a representative averaged index value out of up to 23 Swiss phenological observation sites. This method accounts for systematic differences when averaging several single phenological series into a regionally representative mean chronology. Subsequently, linear regression models were calibrated as transfer functions to estimate the ”Spring plant” from single phenological series. The availability of the historical phenological records varies from year to year (see *Rutishauser et al.,* 2007 for details). Finally, Rutishauser et al. (2007) provide an annual estimate of the onset of spring including an uncertainty range at interannual time scales of +/- 10 days and of +/- 3.6 days at decadal time scales.

Swiss and Burgundy grape harvest dates

The Swiss grape harvest date records were compiled from 15 single village series with a total number of 1435 records for the 1480–2006 period (Pfister and Dietrich-Felber, 2006; Meier et al., 2007). Annual median values were selected as representative values for the Swiss plateau region following the methodology of Chuine et al. (2004). There are missing observations in 1876, 1927, and between 1879 and 1884. The longest period is the consequence of severe, wide spread grape diseases (Mullins, 1992), such as the phylloxera which also heavily affected Swiss vineyards (Meier et al., 2007).

The Burgundy series 1370–2003 was taken from Chuine et al. (2004) (downloaded from http://www.ncdc.noaa.gov/paleo/pubs/chuine2004/chuine2004.html). The Burgundy phenological data set is not regularly updated (I. Chuine, pers. comm., 2007). We used the post-1753 data of both time series, which were overlapping with the instrumental temperature data from Switzerland.

Temperature measurements

Monthly and seasonal mean temperature measurements from Geneva (starting in 1753) and Basel (starting in 1755; Schüepp, 1961; Begert et al. 2005, updated) were averaged into a Swiss mean series. After 1864 the data can be considered as homogeneous (Begert et al., 2005). Three monthly means represent the traditional climatological seasons winter (December/January/February), spring (March/April/May), summer (June/July/August) and autumn (September/October/November).

We used our derived Swiss temperature measurements also for an analysis of the impact on the phenology of Burgundy harvest. Comparisons of monthly and seasonal means reveal that Swiss temperatures explain approximately 90% of the temperature variability at Burgundy station at Dijon (climate explorer; climexp.knmi.nl; Oldenborgh et al., 2005) for the overlapping period 1951–2000, except for October (76%), most likely due to an erroneous outlier in the Dijon record (not shown).

4.2.2 Methods

Time-series models

Analogous to the procedure of Dose and Menzel (2004) we use the Bayesian approach to describe long-term phenological and temperature time series with three implemented models. We refer to Dose and Menzel (2004, 2006) for computational and mathematical formulae details. In addition, Schleip et al. (2006, 2008) demonstrated the flexible application of the Bayesian procedure on different climate change detection issues.

The simplest model is a constant model associated with no rate of change and represents just the mean value of the data. The second model used in this study is a linear regression with a constant rate of change over time. The third model, the change point model, involves the selection of two linear segments matching at a particular time. The change point model provides a time varying rate of change. We calculate the model probabilities of the three models with the Bayesian approach of Dose and Menzel (2004). However, our inferences are derived from the results of the individual models weighted by their respective model probability.

Model averaged rate of change, change point probability distribution and moving linear trend

We aim to find the most probable functional description and rate of change given by three models. This is obtained from a marginalization over the constant, linear and change point model. Marginalization is a very powerful device in data analysis because it enables us to deal with nuisance parameters; that is, quantities which necessarily enter the analysis but are of no intrinsic interest (Dose and Menzel, 2004). While the more complicated model—i.e., the change point model—certainly provides a better fit, it does not necessarily have a higher model probability. Bayesian probability theory selects a model by considering the trade-off between lower misfit and higher complexity, also known as Ockham`s razor (Garrett, 1991). This means that the chosen model should be as complex as required to explain the data and as simple as necessary to avoid fitting the model to noise. The model average for the rate of change is calculated by averaging the rates of change of the three models weighted by their respective probabilities.

The change point model allows for nonlinearities in the description of functional behaviour and rate of change. The change point model is made up of triangular functions consisting of two linear segments defined by the endpoints of the series and a change point in between. The variables of these triangular model functions are the unknown functional values at the endpoints and at the change point as well as the temporal position of the change point. Bayesian probability theory estimates the probabilities of all possible change point positions by marginalization over the functional values at the endpoints and the change point of the series. Most often there is no change point with overwhelming probability, but the range of substantial change point probabilities extends over several years. Change point probability distributions exhibit the change point probabilities as a function of time for a temperature or a phenological time series.

Finally our Bayesian results are compared to moving linear trend analysis as used by Menzel et al. (2004), Matti et al. (2008), Rutishauser et al. (2007) and Rebetez and Reinhard (2008). Slopes of linear regression are calculated for each 30-year and 60-year period around a center year that is shifted with a one-year time step. For the trend estimation within each window, we also estimate the error probability (p-value) of the linear trend at the 95% significance level. This indicates the statistical certainty of artificially assumed trends for the selected window length. We arbitrarily choose two window lengths of 30 and 60 years in order to distinguish between shorter and longer time scales. The selected window lengths are applied to both phenological and temperature series in order to detect periods of common trends and matching trend signs. Periods with positive phenological trends (trend towards later spring onset / harvest dates) are expected to be synchronous to periods with negative temperature trends or cooling periods.

Temperature impact model

For the analysis of the coherence of long-term temperature and phenological time series, we use a Bayesian correlation approach, proposed by Dose and Menzel (2006) and recently used by Schleip et al. (2008). The calculation of the coherence factor relates to the change point distributions of the temperature and phenological time series. We use the long-term time series to test whether the temporal evolution of the phenological observations can be attributed to temperature changes. This is a simple case of Bayesian model comparison. We compare two alternative models M_a: temperature and blossom onset time series evolved independently and M_b: temperature and blossom onset time series exhibit coherence. The results are the probabilities for M_a and M_b or alternatively the odds ratio of the two probabilities.

Following the procedure of Dose and Menzel (2006) we calculate the ratio of probabilities p (coherent)/p (independent), also called coherence factor henceforth. A coherence factor above one signifies that the synchronous evolution of the two time series is more probable than the independent one. We determine temperature weight coefficients with an implemented simulated annealing algorithm by maximizing the coherence between temperature and phenology time series (Dose and Menzel, 2006; Schleip et al., 2008). The higher the estimated temperature weights for a certain month, the more overlap can be expected in the change point distributions of temperatures and the phenological event. For the Swiss ”Spring plant“ we considered monthly mean temperature of June of the previous year until May of the actual year. This is the last month of the observed phenological spring onset. For grape harvest dates we chose the mean temperatures of the previous November until October of the current year of harvest. As initial months we selected June and November temperatures because they mark the end of the same phenological process in the previous year.

Finally we compared our results to the results of a traditional statistical analysis. The traditional statistical analysis quantifies the relation between temperature and phenological data by a correlation coefficient ρ. The correlation coefficient between a variable y and a variable x is a measure for the linear dependence between x and y. ρ^2, the square of the correlation coefficient, describes the degree of explained variance. The range of values of ρ^2 is $0 \leq \rho^2 \leq 1$ with $\rho^2 = 1$ signalling perfect linear dependence. Values of $\rho^2 < 1$ arise for different reasons. If the data, neither x nor y are affected by noise then $\rho^2 < 1$ indicates a more complicated relationship. If the noise on the data is nonnegligible, then $\rho^2 < 1$ is even if the data generating mechanism a linear relation. In general, therefore, $\rho^2 < 1$ includes both, the noise and deviation from a linear interdependence (Dose and Menzel, 2006).

4.3 Results

4.3.1 Model selection results

The by far highest model probabilities (from 51% to 100%) are generally found for the change point model, when describing phenological and temperature data (Figure 4.1). It is thus the most suitable model to describe the long-term evolution of phenology and temperature. The linear model is an appropriate alternative to describe mean Swiss winter and autumn temperature (50% model probability for time series of these variables). The constant model is the least preferred one to describe temperature and phenological time series (maximum of 6% , “Spring plant”).

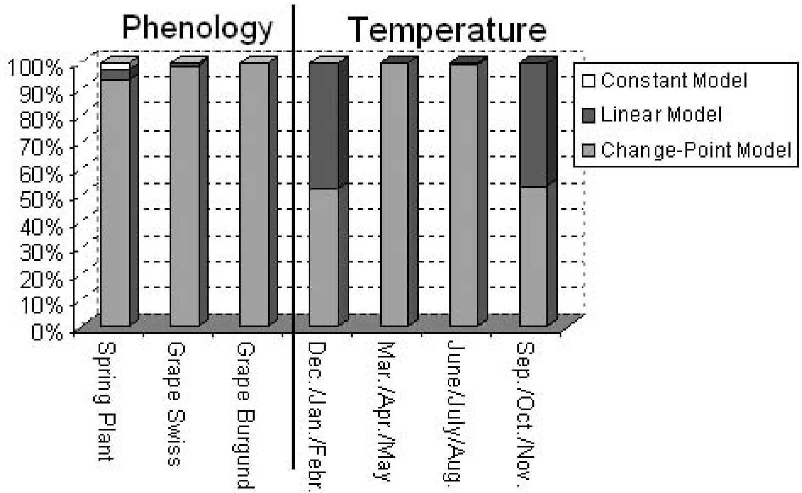


Figure 4.1: Bayesian model comparison of the constant, linear and one-change point model. From left to right: Swiss "Spring plant“ (1753-2006), Swiss grape harvest dates (1753-2006), Burgundy grape harvest dates (1753-2003), mean Swiss seasonal winter (December–February), spring (March-May), summer (June-August) and autumn (September-November) temperatures for 1753-2006.

4.3.2 Time series models

Figure 4.2 a-c and Figure 4.3 a-d present the functional descriptions of the constant, linear and change point models for each selected phenological and temperature series, respectively. In Figure 4.2. a-c, we plotted the functional description of each model for the ”Spring plant“ and the grape harvest records. The constant model represents the mean of the time series. The functional description is shown as a straight line which intercepts the y-axis at the mean value of the time series. For the period 1753-2006 the mean onset of the ”Spring plant“ is on day 118 (Julian day), Swiss grape harvest dates on day 284, and Burgundy mean grape harvest dates (1753-2003) on day 270. The mean values for the winter, spring, summer and autumn temperatures 1753-2006 are 0.8°C, 9.0°C, 17.7°C and 9.3°C, respectively (Figure 4.3 a-d).

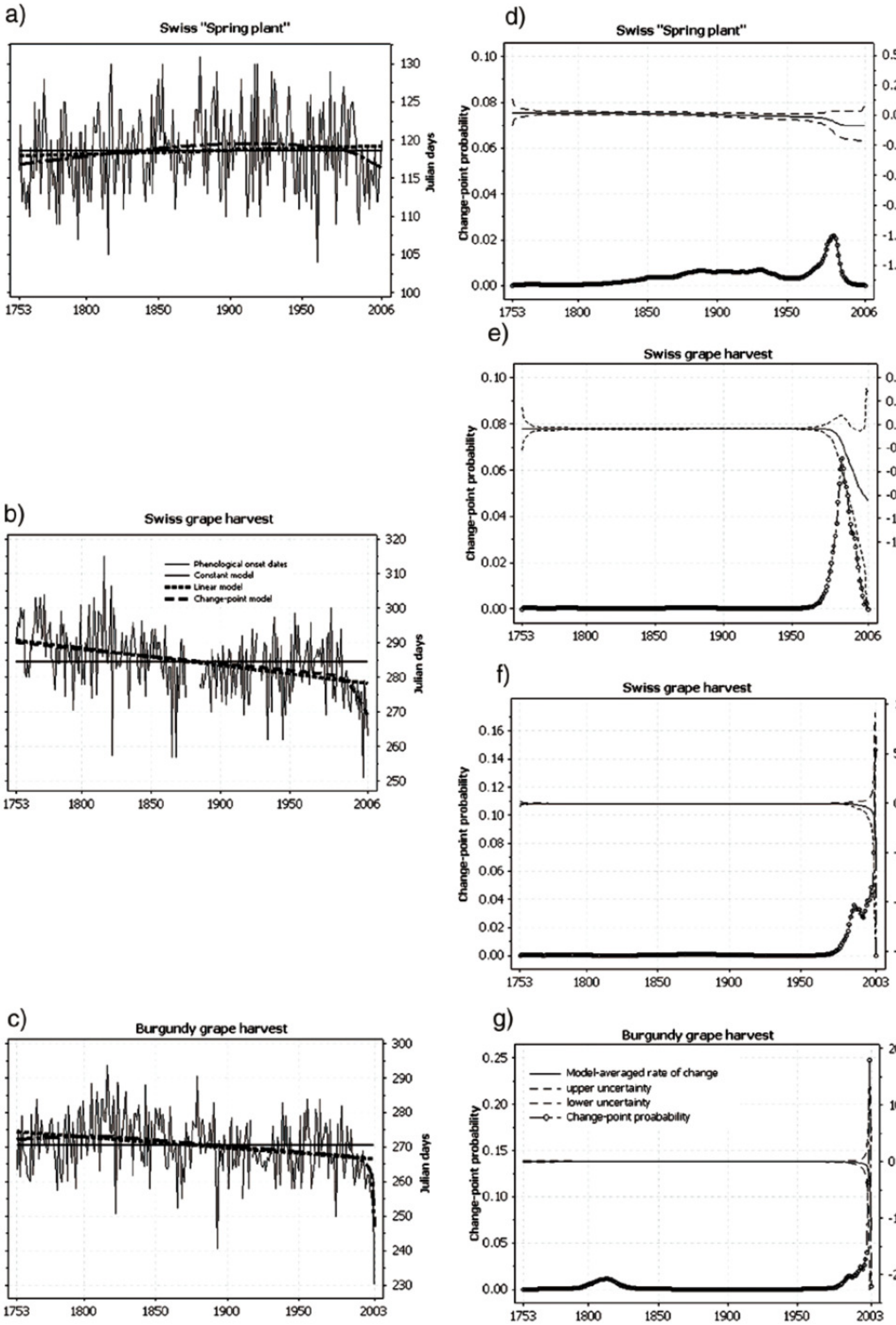


Figure 4.2: a-c) Functional behaviour of the constant, linear and change point model to describe the Swiss "Spring plant" (1753-2006), Swiss grape harvest dates (1753-2006) and Burgundy grape harvest dates (1753-2003). Legend is shown as inset in Figure 4.2 b. The thin black line indicates mean onset day.

d-g) Bayesian model averaged rates of change in days/year (top thin black line) and confidence intervals (dashed lines, right y-axis). Probabilities of change point model (bottom circles and line, left y-axis). Legend is shown as inset in Figure 4.2 g. Figure 4.2 f shows the Swiss grape harvest dates comprising the same data length as the Burgundy grape harvest dates (1753-2003).

The linear model of the “Spring plant” reveals a positive slope of 0.005 ± 0.004 days/year over the whole record (Figure 4.2 a) indicating a delay of spring onset over the whole record. Both Swiss and Burgundy grape harvest dates (Figure 4.2 b, c) show a consistent negative slope of -0.04 ± 0.007 days/year and -0.03 ± 0.007 days/year. For long-term temperature trends (1753-2006) the winter season exhibits the strongest positive slope of 0.01 ± 0.001°C/year. In comparison, spring, summer and autumn season show a less pronounced slope of 0.004 ± 0.001 °C/year, 0.003 ± 0.001 °C/year and 0.006 ± 0.001 °C/year, respectively (Figure 4.3 a-d).

The functional description of the change point model shows a sharp decline at the end of the 20th century for all three phenological time series (Figure 4.2 a-c). The decline is stronger in the grape harvest date records (more than 10 days for 1980-2003/6) than in the ”Spring plant“ (about 5 days 1980-2006). In the case of the ”Spring plant“ we find rising functional values over centuries before a bend down at the end of the 20th century. For Swiss and Burgundy grape harvest dates, the change point function estimation is very close to the linear model function estimation for almost two and a half centuries except at the end of the 20th century.

The function estimation for winter and autumn temperatures describes increasing values with a slower temperature increase until 1850 and since then, a steep temperature increase until 2006 of approximately 2°C (Figure 4.3 a-d). The function estimation for spring temperature declines from 1753 until 1850 and then from 1850, onwards a warming of about 1.5°C is detected. The change point function estimation for summer exhibits a very sharp increase of about 2.5°C from the 1980s until 2006.

4.3.3 Model averaged rates of change

All phenological time series exhibit increasingly negative rates of change (Figure 4.2 d-g; upper lines, right scale). For comparison to the Burgundy grape harvest dates, we additionally calculated for the Swiss grape harvest dates model averaged trends for the shorter time period 1753-2003 (Figure 4.2 f). Note that for the shorter time series model averaged rates of change at the end of the time series are associated with high uncertainty intervals that are as high as or higher than the absolute point estimate itself. For the Swiss grape harvest time series these high uncertainties are reduced substantially. In the longer time series the estimation of rates of change and associated change point probabilities is supported by further years of data (Figure 4.2 e). Swiss and Burgundy grape havest dates exhibit for 2003 exceptional early harvest dates (Figure 4.2 c).

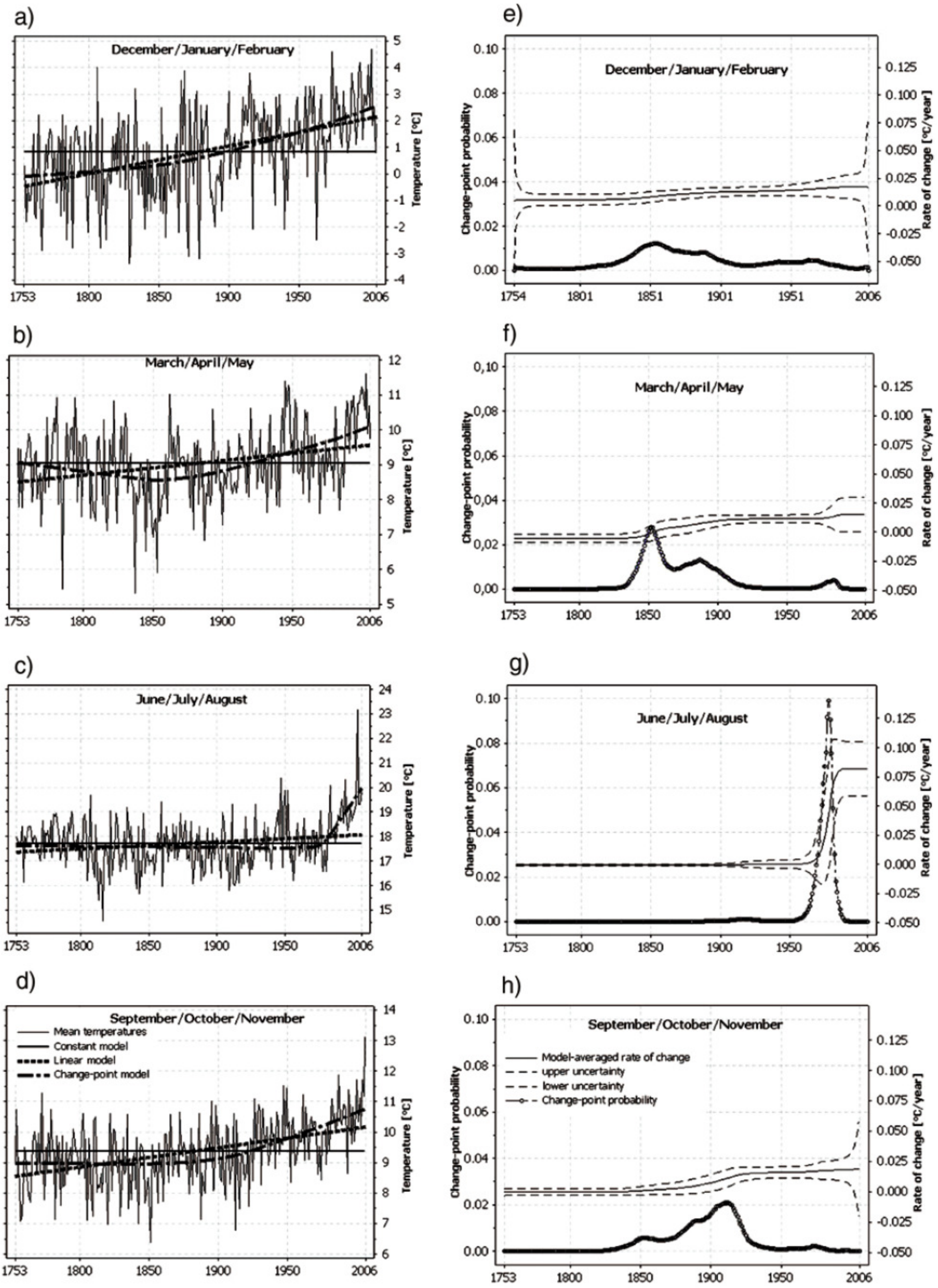


Figure 4.3: As Figure 4.2 but for winter (December–February, Figure 4.3 a, e), spring (March-May,Figure 4.3 b, f), summer (June-August, Figure 4.2 c, g) and autumn (September-November, Figure 4.3 c, h) temperatures in the period 1753-2006. Functional model behaviour Figure 4.3 a-d, model averaged trend and change point probability Figure 4.3 e-h.

All investigated temperature records show an increasingly positive rate of change from 1753-2006 (Figure 4.3 e-h). In 2006, winter, spring and autumn season exhibit a rate of temperature change of approximately 0.01° C/year. The most abrupt change that was significantly different form zero occurred in the summer season. However, continuing warming trends significantly different from zero can also be found in winter, spring and autumn after 1772 (Figure 4.3 e), 1882 (Figure 4.3 f) and 1900 (Figure 4.3 h), respectively. Since the 1980s, the positive rate of change increased from almost zero to 0.08°C/year in 2006 (Figure 4.3 g).

4.3.4 Moving linear trend analysis

Moving linear trend analysis for the three phenological time series show alternating periods of positive and negative trends throughout the period 1753-2003/6 for 30-year time windows (Figure 4.4, upper panels). The phenological trends (bold lines) follow the temperature trends of the spring season (March/April/May, thin lines) of the current growing season. The decisive temperature period for the date of grape harvest lies between the flowering and véraison development stage occuring in late spring. Summer temperature however influences the sugar content of the grape and not the vintage date (Mullins, 1992; Meier et al., 2007). Significance tests (F-test) show that there is a high proportion of low or non significant phenological trends (Figure 4.4 a-c, lower panels). There is a distinct trend towards earlier harvest and spring dates in all series at the end of the 20th century but only highly significant with a window length of 30 years.

The ”Spring plant“ shows two distinct periods with trends towards earlier spring development of up to -0.3 days/year for the center year of the 30-year windows between 1976-1991 and 1939-1954 both with low error probabilities (Figure 4.4 a). Periods of trends towards later spring development of +0.1 to +0.2 days/year can be seen for the center years between 1955-1975, 1916-1929 and around 1870 with only the latest showing high significance. Applying a 60-year window, significant negative slopes are found at the end of the 20th century (not shown). However, error probabilities are generally higher and trend signs may differ except at the end of the 20th century. Applying a 30-year window the Swiss grape harvest record reveals trends towards earlier dates of -0.4 days/year or more at the end of the 20th century, between 1830-1860, around 1810 and in the 1760s (Figure 4.4 b). These periods all show very low error probabilities with a 30-year window. The 1800s show a very long period of significant delay and advance when using a 60-year window (not shown).

The Burgundy series largely corresponds to the Swiss grape harvest record (Figure 4.4 c). The exception is a more pronounced delay around 1880 and a less pronounced peak in 1960. The longest period of low error probability is seen before 1800.

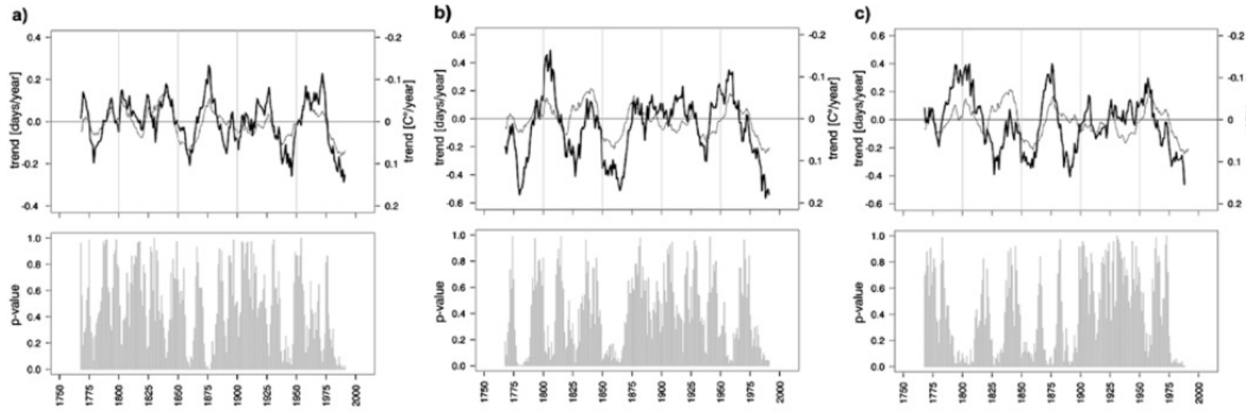


Figure 4.4: Moving linear trend analysis for Swiss "Spring plant" (a), Swiss grape harvest dates (b) and Burgundy grape harvest dates (c) showing slope coefficients of the linear regression of phenology against time for 30-year periods. Bold lines show phenological, thin lines corresponding spring temperature trends. Note that the left axis represents the phenological trend and the right axis the temperature trend. The values are plotted at the middle year of the respective windows. The lower panels are the error probability estimates (p-values) from the regression of the phenological records.

4.3.5 Change point analysis

The most likely change point for the "Spring plant" time series is found in 1984 (probability = 2%) (Figure 4.2 d lower line). In comparison, the most likely change point for Swiss grape harvest dates is 1986 (probability = 7%) (Figure 4.2 e). Burgundy grape harvest dates show an extremely high probability (24%) of having a change point in the year 2002 and a smaller probability (2%) in 1812. If we cut the Swiss grape harvest dates to the same length as the Burgundy grape harvest dates ending in 2003, Swiss grape harvest dates exhibit an extremely high probability (14%) of having a change point in the year 2002 and a smaller probability in 1987 (Figure 4.2 f).

Winter and spring temperatures exhibit the highest probability for a change at the beginning of the 1850s (Figure 4.3 e, f). Within the summer season, the highest probability for a change is found in 1978 with a maximum probability of 10% (Figure 4.3 g). The autumn season reveals the highest probability of a change point in 1912 (Figure 4.3 h).

4.3.6 Coherence factors, temperature weights and linear correlation

Swiss "Spring plant" and Swiss grape harvest dates reveal a high coherence with temperature time series (Figure 4.5 a, b). The monthly resolution exhibits high coherence factors for both Swiss and Burgundy grape harvest dates with the highest temperature weight in June (Figure 4.5 b). For Burgundy grape harvest dates the highest temperature weights are seen in the months of June and September. For Swiss grape harvest dates, high temperature weights are revealed for February to June of the current year (Figure 4.5 b). Remarkable is that for Swiss grape harvest dates (1753-2003) June temperatures exhibit the highest and only weight. Linear correlations after Pearson indicate the same result (Table 4.1). Mean temperatures of March to July increasingly explain more variance of the grape harvest date records from 10 to around 25%. Afterwards temperature explains only about 5% of the variance until the date of the grape harvest. Unlike the coherence factors, linear correlation does not indicate an impact of mean September temperatures. Seasonal averages of spring and summer temperature explain one third of the variance whereas autumn temperatures are statistically not relevant for the date of grape harvest.

Table 4.1: Pearson correlation (cor) and associated error probabilities (p-val) between phenological series and preceding monthly mean temperatures. R^2 indicates the percentage of variance in the phenological records explained by temperature for the periods 1753–2006 (Swiss Spring Plant and grape harvest dates) and 1753–2003 (Burgundy grape harvest dates).

"Spring plant"	**cor**	**p-val**	**R^2 [%]**
pJun	-0.13	0.045	2
pJul	-0.07	0.301	0
pAug	-0.15	0.017	2
pSep	0	0.098	0
pOct	-0.01	0.923	0
pNov	0.04	0.558	0
pDec	-0.07	0.271	0
Jan	-0.05	0.0448	0
Feb	-0.33	<0.001	11
Mar	-0.46	<0.001	21
Apr	-0.56	<0.001	31
May	-0.18	<0.004	3
Swiss Grape Harvest Dates			
pNov	-0.17	0.009	3
pDec	-0.26	<0.001	7
Jan	-0.10	0.120	1
Feb	-0.13	0.046	2
Mar	-0.28	<0.001	8
Apr	-0.31	<0.001	10
May	-0.47	<0.001	22
Jun	-0.48	<0.001	23
Jul	-0.45	<0.001	20
Aug	-0.24	<0.001	6
Sep	-0.25	<0.001	6
Oct	-0.23	<0.001	5
Burgundy Grape Harvest Dates			
pNov	-0.23	<0.001	5
pDec	-0.28	<0.001	8
Jan	-0.11	0.095	1
Feb	-0.13	0.035	2
Mar	-0.29	<0.001	8
Apr	-0.36	<0.001	13
May	-0.46	<0.001	21
Jun	-0.54	<0.001	29
Jul	-0.46	<0.001	21
Aug	-0.28	<0.001	8
Sep	-0.25	<0.001	6
Oct	-0.19	0.002	4

For the ”Spring plant“, the highest temperature weights are found within the spring of the current year and of the previous summer. July of the previous year and May of the harvest year exhibit the highest weights (Figure 4.5 a). Linear correlation indicates significant temperature impact of single months only from February to April.

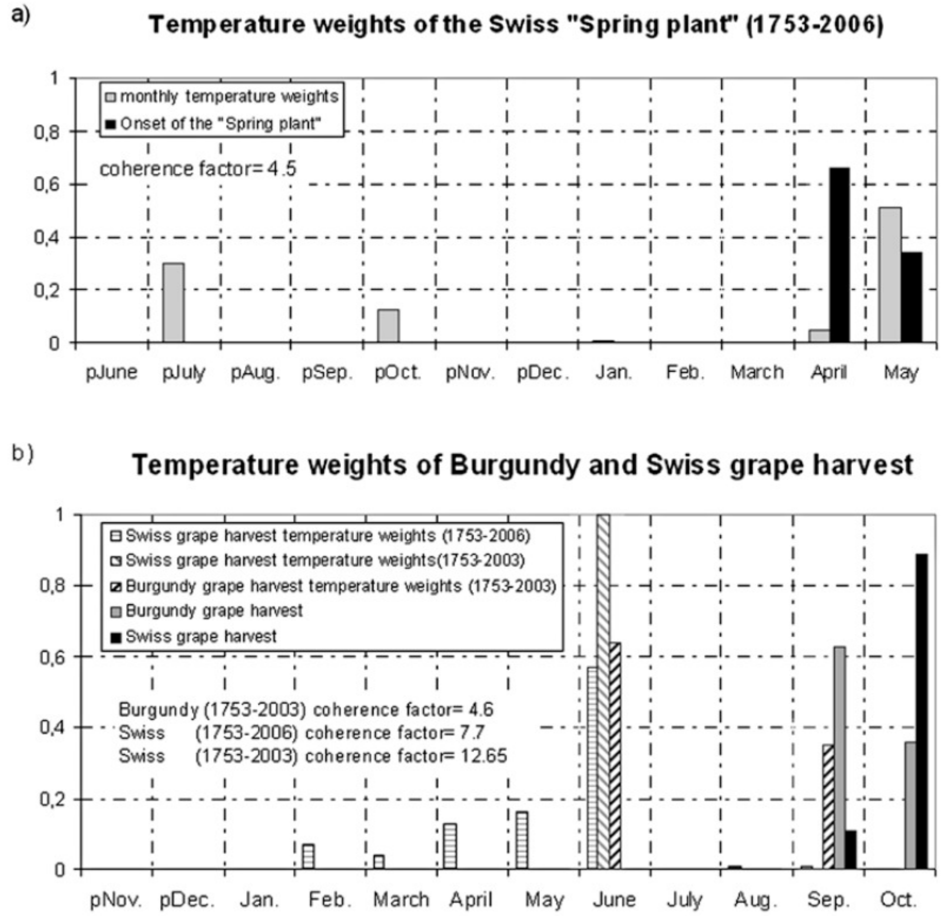


Figure 4.5: a) Temperature weights estimated by the simulated annealing process for the Swiss ”Spring plant“ (1753-2006) and corresponding coherence factors and weights for monthly temperatures from the previous June (pJune) until the current year's May. b) Temperature weights estimated by the simulated annealing process for the Swiss and Burgundy grape harvest (1753-2006 and 1753-2003) and corresponding coherence factors and weights for monthly mean temperature from November of the previous year (pNov) until October of the present year.

4.4 Discussion

Linear regression approach

Simple approaches such as the description of linear trends derived from regression models have proven to be a valuable tool for initial descriptions of phenological time series behaviour (e.g. Root et al., 2003; Parmesan and Yohe, 2003; Menzel et al., 2006). The simplicity of the least squares linear regression model makes it possible to compare linear trends from around the world and gathered from previously published literature (Root et al., 2003; Parmesan and Yohe, 2003). For example Menzel et al. (2006) reanalysed 125000 European phenological time series for the period 1971–2000. We conducted a moving trend window approach as applied by Rutishauser et al. (2007). Applying the moving window technique on multi-century phenological records (Figure 4.4), the trends of the last and 20th century are put in the long-term perspective back to 1753. The three phenological records all show clear trends towards earlier spring onset and grape harvest dates at the end of the 20th century respectively. However, grape harvest trends at the beginning of the 19th century indicate even stronger advancing trends than in the recent decades. Here significant negative trends are seen from 1940–1950 and with end years after 1990. Positive trends, however, are statistically significant only for a small number of periods. But different results are calculated if we choose different window lengths as we demonstrated with an additional 60 year window. The rate of change strongly depends on the underlying time period and no distinct rate of change for single years can be given (Dose and Menzel, 2004).

Bayesian model comparison

Bayesian analysis offers the possibility to overcome the shortcomings of linear regression models. To assess the potential value of the estimated records we applied a Bayesian model averaging approach to detect changes in temperature and phenology by estimating different model probabilities, functional behaviours and model-averaged rates of change. The Bayesian model comparison provided an excellent opportunity to judge and compare different model estimates. The results of our Bayesian time series analysis are more informative than results based on single model approaches. The description of the data in terms of only one model is often unsatisfactory (Dose and Menzel, 2004, Schleip et al., 2006, 2008). Our Bayesian model comparison showed that if we had concentrated on just a single model such as the commonly used linear model our final inferences may be incorrect.

For example on one hand our results reveal that for the winter temperatures the linear model has the strongest positive slope of 0.01°C/year compared to the other seasons. On the other hand, if we look at the summer season, the linear model fails to reflect the real nature of the time series especially at the end of the 20th century. The functional behaviour of the change point model suggests a considerable increase in summer temperatures since 1978. Compared to the other seasons, model averaged rates of change of summer temperatures show the most pronounced warming. Since the 1980s the model-averaged rate of change increased from almost zero to a rate of change of 0.08°C/year in 2006 (Figure 4.3 g).

The functional behaviour of the change point model and the model-averaged rates of change of the corresponding phenological phases show that these phases have advanced considerably, particularly at the end of the 20th century. This abrupt nonlinear advance of the onset of spring and harvest dates does not appear in the linear model when it is used on multi-decadal time scales longer than 30 years. The change point function allows for more detail. The change point function of the “Spring plant“ shows that the onset of spring tended to occur progressively later from 1753 to approximately 1940 and progressively earlier from the 1980s onwards (Figure 4.2 a). The Swiss and Burgundy grape harvest dates show a continuous advance of the onset of harvests from 1753 onwards (Figure 4.2 b, c).

In the case of the Swiss and Burgundy grape harvest dates extreme onset dates at the end of the time series cause high trend uncertainties. The beginning of the 21th century and especially the summer of 2003 was extremely warm (Luterbacher et al., 2004; Schär et al., 2004). The extreme early grape harvest date after the widespread European heat wave summer in 2003 had a noticeable impact on the Swiss and Burgundy record. Despite the robustness of the Bayesian statistic method, these findings suggest that extreme outliers at the end of the time series can also lead to large uncertainties. Outliers at the end of the time series do not have more corresponding observational data and therefore exhibit a rapid widening of the confidence range (Figure 4.2 f, g). For Swiss and Burgundy grape harvest dates this rapid widening is far too large to make any reliable conclusion at the end of the time series. Note that the available updated Swiss grape harvest dates until 2006 exhibit a better model averaged calculation of the advancing (Figure 4.2 e).

Bayesian change point probability distributions

All investigated 250-year phenological time series reveal the highest change point probability at the end of the 20th century (Figure 4.2 d-f, Figure 4.3 e-h). These findings suggest that three spring and autumn phenological records from close spatial origin all show concurrent and unique changes at the end of the 20^{th} century. Accumulation of pronounced change points can also be found in many phenological time series (Schleip et al., 2006, 2008). The change point analysis of seasonal temperatures exhibits a more dispersed pattern of high change point probabilities indicating a greater variability of temperature data compared to phenological data. The summer season exhibits a change point probability distribution, which is relatively narrow with a clear single peak at the end of the 20^{th} century. A clear narrow single peak indicates a very strong abrupt temperature change within a defined time period. Change point probability curves for winter, spring and autumn temperature reveal much broader change point distributions. A broad multi-modal change point distribution indicates that several major nonlinear changes occurred consecutively in a certain time span. Our results do not reveal change point distributions with several separated high change point peaks within one time series. Only Burgundy grape harvest and spring temperatures show a second minor peak.

Analysis of the change point probabilities provides the advantage of visualizing and quantifying major changes in our long-term time series. Thus it filters out low frequency variations in the long-term time series. Our one-change point model is capable of identifying the major relevant changes in long-time series. In the future, a multiple change point model would be capable of modelling a more detailed structure in a time series and therefore would mirror several minor changes within the last 250 years. But each added change point adds two more variables, which may make the model unnecessarily complex. Whether the complex multiple change point model will really provide a better description of temperature and phenological time series should be tested in future work.

Temperature impact

Our phenological and temperature time series exhibit nonlinear changes, so we applied the Bayesian correlation approach of Dose and Menzel (2006). Many publications of recent years have pointed to the coherence of phenological spring phases and temperature using classical statistical methods such as correlation analysis, linear and multiple regression methods (Sparks and Carey, 1995; Sparks et al., 2000; Menzel, 2003b; Luterbacher et al., 2007; Rutishauser et al., 2008). Additional experiments have shown the link between temperature and phenology to be causal in many plant species—i.e., warmer temperatures generally lead to earlier spring phenology (e.g. Saxe et al., 2001). Plant phenophases may also respond to many other meteorological and environmental factors such as light, photoperiod, temperature, precipitation, humidity, wind, soil conditions etc. (Schnelle, 1955; Menzel, 2002).

The analyses by Dose and Menzel (2006) assumed that temperatures of the previous year of the phenological event can be neglected with regard to the phenological onset in the year of interest. Our results show that summer temperatures very likely influence not only summer phases of the current year but also spring phases of the following year as seen in the onset of the Swiss ”Spring plant“ (Figure 4.4). For ”Spring plant“ phenology, temperatures during the spring season of the year of budburst and temperatures during the summer season of the previous year appeared particularly important, especially temperatures in the previous July and the following May. With simple linear approaches (Table 4.1, e.g. Sparks and Carey, 1995), this result is statistically much less evident and has never been discussed.

The summer phases of Swiss and Burgundy grape harvest are mainly influenced by the season's spring and early summer and less by temperatures in the autumn (Figure 4.5 a, b). Only in the case of Burgundy grape harvest did the monthly resolution show the influence of September temperatures, which occurred simultaneously with the grape harvest event. However, September still seems a statistical artefact as seasonal temperature weights (Figure 4.5 a) and the majority of findings in the literature (Pfister, 1981, Pfister, 1988, Mullins, 1992) suggest that the harvest date of the grape vine is hardly influenced by the three months preceding the harvest. According to Alleweldt (1967), the development of the grapevine is shut down in August, meaning that the temperature of this month and thereafter does not carry weight for the harvest date of the grapevine. During maturation, solar irradiance mainly acts to control the amount of sugar accumulation in the grapes (Jones, 2003), and temperature signal is moreover stored in the accumulated sugar content in the grape (Pfister, 1981). Grape harvest dates are most strongly influenced by temperatures between flowering and ripeness of the grape development (Pfister, 1981, Meier et al., 2007). This finding is also supported by simple linear correlations (Table 4.1). Swiss grape harvest reveals the increasing influence of temperatures from February to June, as shown by temperature weights (Figure 4.5 b) and linear correlation (Table 4.1). Burgundy grape harvest dates do not show such a continuous increase from February to June but exhibit the highest temperature weight in June. Swiss grape harvest (1753-2003) identifies June as the highest temperature weight, too, indicating the importance of June temperatures for grape harvest phenology. Afterwards, July shows lower correlations and a dramatic decrease in August and September. In general we assume that the process of maturation is also promoted by temperature sums which are accumulated in the preceding months of June. However, temperature weights for Burgundy (Figure 4.5 b) indicates a September temperature impact on the harvest date. We hypothesize that different viticultural traditions in France and Switzerland such as vintage ban (Meier et al., 2007, and references therein) might contribute to Bayesian statistical findings.

Menzel's (2005) estimates of the correlation between dates of grape harvest and monthly mean temperatures differed from our own with no significant correlations between temperatures of the winter months January to March as well as of temperatures of the summer months June and August and the grape harvest dates. Menzel (2005) mentioned that the low correlations of the summer months may have been due to the `biologically artificial' separation of the growing season into single calendar months. In our analysis the earliest harvest in Switzerland started in August. The entire month August is not always part of the growing season and mean temperatures might be not relevant in all years. In further investigation it would be interesting to examine the influence of biweekly rather than monthly temperatures on plant phenology. In addition, we suggest the application of a phenology model developed for the Pinot Noir grape variety by Chuine et al., (2004) also to Swiss grape harvest date observations.

4.5 Conclusion

Unique long-term temperature and phenological data series for central Europe back to 1753 were analysed with different approaches. A simplistic linear approach illustrates valuable information regarding the impact of temperature on multi-decadal phenological records despite the well-known limitations—e.g., *a priori* selection of window lengths. We used a Bayesian model comparison to provide for the first time a detailed description of preferred models, change point probabilities, functional behaviours and estimates of the rate of change of the Swiss temperature and phenological time series as well as of the Burgundy grape harvest dates. Results show that the model-averaged rates of change of the phenological phases show a considerable advance of the onset of spring and harvest dates. Additionally, the summer temperature time series shows an abrupt temperature increase at the end of the 20th century. For all phenological time series the change point model is the preferred model to describe the time series. The linear model provides an adequate alternative for describing the temperature time series for winter and autumn. In the context of the last 250 years the end of the 20th century represents a period with unique major increases in temperatures of all seasons and earlier grape harvest phenology as derived from model-averaged trends.

For the first time we also investigated the relationship of phenological records with temperatures of the previous year by Bayesian-methods. Coherence factors and temperature weights indicate that spring phenological variability is not only influenced by forcing temperatures of the current year but also by temperatures of the preceding June and October. For grape harvest dates, we could not detect temperature impacts of the termination of the previous growing season. However, June temperatures of the year of harvest appear significantly related to harvest dates.

Future work should address the assessment of increasingly complex time series models such as multiple change point models in addition to the simplistic linear approaches. Following the Bayesian coherence approach of Dose and Menzel (2006) analyses of the impact of temperature on phenology should include temperature forcing periods other than calendar months—e.g., shifting 4-weeks period—or should include precipitation and drought—e.g., PDSI (Drought severity indices). It would also be intriguing to investigate the possible role of temperatures in the previous year to influence future plant phenology.

4.6 References

Alleweldt G., (1967) Der Einfuss des Klimas auf Ertrag und Mostqualität der Reben, *Rebe und Wein.* **20**, 312–317.

Aono Y., Kazui K., (2007) Phenological data series of cherry tree flowering in Kyoto, Japan, and its application to reconstruction of springtime temperatures since the 9th century. *Int. J. of Clim.* **28**, 7, 905-914.

Begert M., Schlegel T., Kirchhofer W., (2005) Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000. *Int. J. of Clim.* **25**, 65-80.

Bradley N.L., Leopold A.C., Ross J., Huffaker W., (1999) Phenological changes reflect climate change in Wisconsin. Proceedings of the National Academy of Sciences USA, *Ecology* **96**, 9701–9704.

Brazdil R., Pfister C., Wanner H., von Storch H. , Luterbacher J., (2005) Historical climatology in Europe, The state of the art. *Climatic Change* **70** , 363-430.

Burri, M., and T. Rutishauser (2009), Johann Jakob Sprünglis Beobachtungen als Quelle der Historischen Klimaforschung, in press, edited by M. Stuber et al., pp.1759–2009, Oekon. und Gemeinnützige Ges. des Kantons Bern Bern.

Cheke R.A., (2007) Thinking Long Term. *Science* **318**, 577-578.

Chuine I., Yiou P., Viovy N., Séguin B., Daux V., Le Roy Ladurie E., (2004) Grape ripening as a past climate indicator. *Nature* **432**, 289-290.

Chuine I., Cour P., (1999) Climatic determinants of budburst seasonality in four temperate zone tree species. *New Phytologist* **143**, 339-349.

Defila C., Clot B., (2005) Phytophenological trends in the Swiss Alps, 1951—2002. *Meteorologische Zeitschrift* **14** , 191-196.

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

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

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

Guiot J. A., Nicault C., Rathgeber J., Edouard F., Giubal G., Pichard C. T., (2005) Last millennium summer-temperature variations in Western Europe based on proxy data. *Holocene* **15**, 489-500.

IPCC 2007 WG 1(2007). Solomon, S.; Qin, D.; Manning, M., Marquis, M., Averyt, K., Tignor, M., Le Roy Miller, H. and Z. Chen, (ed.): Climate Change 2007. The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change IPCC Cambridge University Press.

Jones G., (2003) Winegrape Phenology. In: Schwartz, M. D. (ed). *Phenology: An Integrative Environmental Science*, Kluwer, New York and Amsterdam, 523–539.

Le Roy Ladurie E.; Daux V., Luterbacher J. (2006) Le climat de Bourgogne et d'ailleurs XIVe--XXe siècle. *Histoire Economie Societé* **3** , 421-436.

Luterbacher J., Dietrich D., Xoplaki E., Grosjean M., Wanner H., (2004) European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* **303**, 1499-1503.

Luterbacher J., Liniger M.A., Menzel A., Estrella N., Della-Marta P.M., Pfister C., Rutishauser T., Xoplaki E., (2007) The exceptional European warmth of Autumn 2006 and Winter 2007: Historical context, the underlying dynamics and its phenological impacts. *Geophysical Research Letters* **34**, L12704.

Matti C., Pauling A., Küttel, M., Wanner H., (2008) Winter precipitation trends for two selected European regions over the last 500 years and their possible dynamical background. *Theoretical and Applied Climatology*, online first, doi:10.1007/s00704-007-0361-x.

Meier N., T. Rutishauser, C. Pfister, H. Wanner, Luterbacher J., (2007) Grape harvest dates as a proxy for Swiss April to August temperature reconstructions back to AD 1480. *Geophysical Research Letters,* **34**, L20705.

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

Menzel A., (2002) Phenology: Its Importance to the Global Change Community. *Climatic Change* **54**, 379-385.

Menzel A., (2003a) Phenological Data, Networks and Research: Europe. In: Schwartz, M.D. (ed.), Phenology – an intergative environmental science. Kluwer Academic Publishers, Dordrecht, Boston, London: 45-56.

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

Menzel A., Estrella, N. and P. Fabian (2004), Pflanzen in einer wärmeren Welt. *Münchner Rück* 144-155.

Menzel A., (2005) A 500 year pheno-climatological view on the 2003 heatwave in Europe assessed by grape harvest dates. *Meteorologische Zeitschrift* **14**, 75-77.

Menzel A., Sparks T.H., Nicole E., Koch E. Aasa A., Ahas R., Alm-Kübler K., Bissolli P., Braslavska O., Briede A., Chmielewski F., Crepinsek Z., Curnel Y., Dahl A., Defila C., Donnelly A., Filella Y., Jatczak K., Mage F., Mestre A., Nordli Ø., Peñuelas J., Pirinen P., Remisova V., Scheifinger H., Striz M., Susnik A., Van Vliet A., Wielgolaski F., Zach S. and A. Zust (2006) European phenological response to climate change matches the warming pattern. *Global Change Biol.* **12**, 1969–1976.

Mullins M., (1992) Biology of the grape vine, Cambridge Univ. Press, Cambridge, U.K.

Oldenborgh G.J. van, M.A Balmaseda, L. Ferranti, T.N. Stockdale, Anderson L.T., (2005) Evaluation of atmospheric fields from the ECMWF seasonal forecasts over a 15 year period. *Journal of Climate* **18**, 2970-2989

NIST/SEMATECH, (2006), e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/index2.htm

Parmesan C., Yohe I., (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**, 37-42.

Pfister C., (1975) Agrarkonjunktur und Witterungsverlauf im westlichen Schweizer Mittelland zur Zeit der Ökonomischen Patrioten 1755-1797 University of Bern. *Geographica Bernesia* G 2.

Pfister C., (1981) Die Fluktuationen der Weinmosterträge im schweizerischen Weinland vom 16. Jh. bis ins frühe 19. Jh. Klimatische Ursachen und sozioökonomische Bedeutung. *Schweizerische Zeitschrift für Geschichte* **31**, 445-491.

Pfister C., (1992) Monthly temperature and precipitations in Central Europe 1525-1979: quantifying documentary evidence on weather and its effects. London: Routledge.

Pfister C., (1999) Wetternachhersage, Verlag Paul Haupt, Bern

Pfister C., Dietrich-Felber U. (Eds.) (2006) Euro-Climhist A data base on past weather and climate in Europe and its human dimension University of Bern, version 2006-02-01, www.euroclimhist.com.

Rebetez M., Reinhard M., (2008) Monthly air temperature trends in Switzerland 1901–2000 and 1975–2004. *Theoretical and Applied Climatology* **91**, 27-34.

Rosenzweig C., G. Casassa D.J., Karoly A., Imeson, C. Liu, A. Menzel, S. Rawlins, T.L. Root, B. Seguin, P. Tryjanowski, Hanson C.E., (2007) Assessment of observed changes and responses in natural and managed systems. In Climate Change (2007): Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. M.L. Parry, O.F. Canziani, J.P. Palutikof, and P.J. van der Linden, Eds. Cambridge University Press, pp. 79-131.

Root T., Price J., Hall K., Schneider S., Rosenzweig C., Pounds A., (2003) Fingerprints of global warming on wild animals and plants. *Nature* **421**, 57-60.

Rutishauser T., Studer S., (2007) Klimawandel und der Einfluss auf die Frühlingsphänologie. *Schweizerische Zeitschrift für das Forstwesen* **158**, 105-111.

Rutishauser T., Luterbacher J., Jeanneret F., Pfister C. Wanner H., (2007) A phenology-based reconstruction of interannual changes in past spring seasons. *Journal of Geophysical Research – Biogeosciences* **112**, G04016.

Rutishauser T., Luterbacher J., Defila C., Frank D., Wanner H., (2008) Swiss Spring Plant Phenology 2007: Extremes, a multi-century perspective and changes in temperature sensitivity. *Geophysical Research Letters* **35**, L05703.

Saxe H., Cannell M., Johnsen Ø., Ryan M., Vourlitis G., (2001) Tree and forest functioning in response to global warming. *New Phytologist* **149**, 369-400.

Schaber J.A., Badeck F., (2002) Evaluation of methods for the combination of phenological time series and outlier detection. *Tree Physiology* **22** , 973-98

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.

Schleip C., Menzel A., Dose V., (2008) Norway spruce (*Picea abies*): Bayesian analysis of the relationship between temperature and bud burst. Agricultural and Forest Meteorology, **148**, 631-643.

Schnelle F., (1955) Pflanzenphänologie. Leipzig, Akademische Verlagsgesellschaft, Geest and Portig KG.

Schwartz M. D., (2003) Phenology. An integrated environmental science. Dordrecht, The Netherlands: Kluwer Academic Publishers. *Biometeorology* **42**, 139–145.

Schär C., P.L. Vidale, D. Lüthi, C. Frei, C. Häberli, M.A. Liniger and C. Appenzeller, (2004) The role of increasing temperature variability in European summer heatwaves. *Nature* **427**, 332-336.

Schüepp M., (1961) Lufttemperatur. Langjährige Temperaturreihen. *Annalen der Schweizerischen Meteorologischen Zentralanstalt* **2**, pp. 48.

Sekiguchi T., (1969) Historical Dates of Japanese Cherry Festivals since the 8th Century and her Climatic Changes. *Tokyo Geography Papers* **13**, 175-190.

Sparks T. H., Carey P.D., (1995) The responses of species to climate over two centuries: An analysis of the Marsham phenological record, 1736–1947. *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. *Int. J. of Clim.* **44**, 82–87.

Tomé A., Miranda P., (2004) Piecewise linear fitting and trend changing points of climate parameters. *Geophysical Research Letters* **31** , L02207.

Vassella A., (1997) Phänologische Beobachtungen des Bernischen Forstdienstes von 1869 bis 1882. Witterungseinflüsse und Vergleich mit heutigen Beobachtungen. Bern: Bundesamt Umwelt Wald Landschaft. *Umweltmaterialien* **73**. pp. 9–75.

von Storch H., Zwiers F. W.,(2001) Statistical Analysis in Climate Research, Cambridge University Press, Cambridge.

Walther G.R., Post E., Convey P., Menzel A., Parmesan C., Beebee T.C.J., Fromentin J.M., Hoegh-Guldberg O., Bairlein F., (2002) Ecological responses to recent climate change. *Nature* **416,** 389–395.

Xoplaki E., Luterbacher J., Paeth H., Dietrich D., Steiner N., Grosjean M., Wanner H., (2005) European spring and autumn temperature variability and change of extremes over the last half millennium. *Geophysical Research Letters* **32**, L15713.

5 Bayesian analysis of changes in Radiosonde Atmospheric Temperature

Christoph **Schleip** [a], Annette **Menzel** [a] , Volker **Dose** [b],

[a] *Fachgebiet für Ökoklimatologie, Technische Universität München, Am Hochanger 13, D-85354 Freising-Weihenstephan, Germany*

[b] *Centre of Interdisciplinary Plasma Science, Max-Planck-Institut für, Plasmaphysik, Boltzmannstraße 2, D-85748 Garching bei München, Germany*

Journal of Climatology, doi: 10.1002/joc.1790, article online in advance of print

Abstract

This paper describes long-term changes of global atmospheric temperature, using a strict Bayesian approach which considers three different models to describe the time-series: the constant model, the linear model and a change point model. The change point model allows the description of nonlinear annual rates of change with associated confidence intervals. We calculate the probabilities of each of the three models and average finally over these models to obtain the expected functional behaviour and rate of change in temperature with annual resolution. We apply this procedure to a new homogenized Radiosonde Atmospheric Temperature Product for Assessing Climate (RATPAC-A) data set. Annual mean temperature for 13 pressure levels from the surface to 30 hPa are examined. Residual sums of squares reveal that Bayesian model averaged function descriptions and rates of changes are especially useful and informative for the surface, troposphere and tropopause and less appropriate for the stratosphere. From the surface up to the tropopause (200-100 hPa) results reveal that the change point model provides the best data fit. Despite the occurrence of two volcanic eruptions El Chicón (1982) and Mt. Pinatubo (1991) the stratosphere (70-30 hPa) shows a preference for the linear model (60%). The near surface changes exhibit comparatively high change point probability around 1985 and 1995 while those at the tropopause level are highest between 1995 and 2000. For the surface and troposphere the model averaged functional behaviour increases quite constantly, whereas the model averaged functional behaviour for the tropopause decreases until the end of the 1990s and increases from 2000 onwards. The limitations of the currently used radiosonde data render interpretation of the observed changes difficult. Additionally undetected change points may result from our limited model space. In future it should be tested whether a multiple change point model provides a better data description for the stratosphere.

Keywords: Radiosonde Atmospheric Temperature data; tropospheric warming; stratospheric cooling; Bayesian analysis; Bayesian model averaging.

Introduction

The analysis of long-term changes in the climatic system is a fundamental task in studies on climate change detection. While trends can characterize current behaviour, they must be used with caution. Despite their common use as a diagnostic of climate change, linear trends may not be the best paradigm given the nature of the time series (Seidel and Lanzante, 2004). Describing atmospheric temperature trends that are statistically modelled using the simplest functional form, the straight line are in most cases questionable (Wigley et al., 2006). Although estimated linear trends analyses are regularly used, they are often inappropriate for the study of natural time series. Most time series of naturally occurring events, such as phenological records, do change their slopes and /or intercepts over time (Schleip et al., 2006, 2008; Dose and Menzel, 2004). However, not only phenological and surface temperature data sets reveal that linear trends may not be the best description of the time series. Trenberth and Hurrell (1994) emphasised a rapid rise of the temperature in the troposphere and concomitant changes in global climate variables in the mid 1970s. Pawson et al. (1998) highlighted in their study the nonlinear nature of recent stratospheric temperature changes. So it is widely recognized that atmospheric temperatures do not follow monotonic linear changes. Temperature changes have been modelled by a piecewise linear behaviour and even step-like changes (e.g., Karl et al., 2000; Lindzen and Giannitsis, 2002; Christiansen, 2003; Lanzante et al., 2003b; Tomé and Miranda, 2004; Rodionov, 2004; Seidel and Lanzante, 2004). Tomé and Miranda (2004) proposed for the description of the North Atlantic Oscillation, temperature and precipitation time series a least-squares approach. Their least-square approach computes the best combination of continuous line segments that minimize the residual sum of squares. A pair of conditions was used to define the method uniquely: (a) the interval between change points must equal or exceed a given value, and (b) a changing sign between two consecutive trends is required.

Furthermore beside the linear trend model Seidel and Lanzante (2004) explored three other statistical models. Each statistical model involves change points (abrupt changes): (1) a flat steps model, in which all changes occur abruptly; (2) a piecewise linear model; (3) a sloped steps model, incorporating both abrupt changes and slopes during the period between change points.

The three change point models differ in complexity. Let N be the number of change points then (2N+1) variables will enter the flat steps model, (2N+2) the piecewise linear model and (3N +2) the sloped steps model. In a nested model scenario the more complicated model never offers a worse fit. Therefore the goodness of fit is a monotonic function of model complexity. For this reason Seidel and Lanzante (2004) employed the Schwarz Bayesian information criterion (Schwarz, 1978; Priestley, 1981). The Schwarz Bayesian information criterion balances the best fit against complexity to rank their models. It turned out that in many cases models more complicated than the linear trend were superior. The results of this paper are obtained with a fully Bayesian treatment starting with three models: the constant model which describes a temperature stable in time, the linear model with constant rate of change and the change point model with a time varying rate of change. The associated degrees of freedom are one, two and four respectively. This approach has proven to be flexible in the past (Dose and Menzel, 2004; Schleip et al., 2006).

The types of observational data available for an investigation into atmospheric temperature changes are diverse. They differ in type of measurement, length of time period, and space time sampling. Trends and variations in global atmospheric temperatures are an integral part of the changes occurring in the Earth's climate system. Changes in upper air temperatures are an important indicator of climate change (Santer et al., 1996; Tett et al., 1996 Thorne et al., 2005). Data sets for analyzing long-term (a decade and more) changes in atmospheric temperatures consist of radiosonde, satellite, lidar, and rocketsonde measurements; meteorological analyses based on radiosonde and/or satellite data; and products based on assimilating observations using a general circulation model. Each of these contain varying degrees of uncertainties that influence the interpretation and significance of trends (Ramaswamy et al., 2001). These uncertainties arise from different complications of each data set. For example one complication with satellite data is the discontinuities in the time series owing to the measurements being made by different satellite monitoring the atmospheric changes since 1979 (Ramaswamy et al., 2001).

For our Bayesian analysis we use a new data set containing large-scale regional mean upper air temperatures based on adjusted global radiosonde data. The Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC) that we use in this paper are datasets created through a collaborative effort involving National Oceanic and Atmospheric Administration (NOAA) Climatic Data Center scientists from the Air Resources Laboratory (Silver Spring, Maryland), the Geophysical Fluid Dynamics Laboratory (Princeton, New Jersey), and the National Climatic Data Center (NCDC) (Asheville, North Carolina). The temporal homogeneity of many radiosonde time series is suspect due to historical changes in instruments and measurement practices. Such changes may introduce artificial inhomogeneities to the time series, making them unsuitable for the study of long-term climate variations, such as through trend analysis. In this study we use historical Lanzante, Klein, Seidel (LKS) homogeneity-adjusted radiosonde temperatures that are less influenced by any such inhomogeneities. Considerable effort has been devoted to assessing and improving the quality of the radiosonde temperature records. A particular aim of previous studies has been to remove artificial changes arising from instrumental and procedural developments during the radiosonde record (Free et al. 2005; Thorne et al., 2005, Karl et al. 2006). The aim of this study is to improve the applied methods in the field of change detection by the use of a Bayesian approach. We introduce an alternative approach to assess more of the potential of current adjusted radiosonde data.

We provide results of temperature data for 13 atmospheric pressure levels from the surface up to the lower stratosphere. Such data at discrete vertical levels provide unique information for assessing changes in the structure of the atmosphere (Christy et al., 2006). We calculate model probabilities and derive our inferences from the results of the individual respective model weighted by the model probability. We examine and compare model probabilities, change points and rates of change of global annual mean temperature anomalies in the period 1979-2004.

5.1 Data and Methods

5.1.1 Data

A new data set of large-scale regional mean atmospheric air temperatures anomalies based on adjusted global radiosonde data is analysed. The NOAA Climatic Data Center provides the RATPAC data set that is freely available at http://www.ncdc.noaa.gov/oa/cab/ratpac/index.php.

The data starting in 1958 extends to the present and are updated on a monthly basis. The RATPAC datasets represent two different combinations of on one hand the historical

Lanzante, Klein, Seidel (LKS) homogeneity-adjusted data and on the other hand of recent data from the Integrated Global Radiosonde Archive (IGRA).

From 1958 to 1995, the data is derived from spatial averages of temperature data of LKS-adjusted 87-stations (Lanzante et al., 2003a, 2003b). In this period temperature anomalies are defined as departures from the entire period of record. Homogenization is referred to as the task of removing artificial change points from time series. The adjusting procedure by LKS was done by a visual examination of temperature time series at multiple levels, including night-day temperature differences. Indicators such as metadata, statistical change points, the Southern Oscillation Index and the dates of major volcanic eruptions were used to identify inhomogeneities.

After 1995, the RATPAC data are derived from the Integrated Global Radiosonde Archive (IGRA) station data. The IGRA anomalies are computed with respect to the mean for 1970–1999. Free et al. (2004) developed the ‘first difference’ method to extend the LKS data and to combine the station time series of IGRA station data. This method incorporates monthly averages and calculates differences in temperature from one year to the next, resulting in a time series of year-to-year differences for each month at each station. These first difference time series were then combined to form large-scale means, and mean temperature time series were constructed from the first difference series. The “first difference” method is independent of satellite or reanalysis data and is useable in the absence of appropriate near neighbour stations. The first difference method provides a substantial reduction of inhomogeneities in large-scale mean time series without making adjustments to the individual time series. Uncertainty introduced by the first difference method as estimated by Free et al. (2005) with Monte Carlo experiments is 0.02-0.04 K/decade in the troposphere and up to 0.15 K/decade in the stratosphere for the global annual mean trend for 1979-2004. Since the “first difference” method is used in only approximately one third of the time period 1979-2004 we conducted a comparison of RATPAC-A and RATPAC-B data sets to test the sensitivity of our results to observational biases related to the “first difference” method. RATPAC-B data sets contains data for individual stations as well as large-scale arithmetic averages corresponding to areas used for RATPAC-A. The station data consist of adjusted data produced by LKS for the period 1958-1997 and unadjusted data from IGRA after 1997. The regional-mean time series in RATPAC-B are based on arithmetic averaging of these station data, rather than the first difference method used to create RATPAC-A. The comparison of RATPAC-A and RATPAC-B results of Bayesian model probabilities for all investigated pressure levels shows that our approach is robust to the biases (*Figure 5.1* a, b). Estimates of uncertainty arising from the “first difference method” indicate that trends from RATPAC-A for layer means are more reliable than those for individual levels, and those for annual mean time series are more reliable than seasonal or monthly values (see also Free et al., 2004).

In our analysis we used the RATPAC-A data set that is recommended for analysis of interannual long-term changes over large areas. It contains homogenized adjusted global mean temperature anomalies. Typically, radiosonde-based data sets are developed for specific atmospheric pressure surfaces. Temperature anomaly time series are provided on 13 atmospheric pressure levels: the surface, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50 and 30 hPa. We applied the full Bayesian analyse to the period 1979-2004 which resembles at the same time the satellite era. This time period closely resembles the quantity of interest to most users of the data set. Most studies focus their results on this period and allow a comparison of our results.

5.1.2 Methods of analysis

In analogy to the procedure adapted to describe phenological time series (Dose and Menzel 2004, Schleip et al., 2006) we use the Bayesian approach to model RATPAC-A data sets. We refer to this publication for computational and mathematical formulae details. In this chapter and in the appendix we summarize the main features of the Bayesian approach. Dose and Menzel (2004, 2006), Schleip et al. (2006, 2008), Raftery et al. (2005) and Sloughter et al. (2006) demonstrated already the flexible attribution of the Bayesian procedure on different climate issues.

The first model the simplest model is a constant model associated with no rate of change and represents just the mean value of the data. The second model is the commonly used linear regression model with a constant rate of change over time. The third model, the change point model involves the selection of two linear segments matching at a particular time. The change point model provides a time varying rate of change. In our work the change point model allows for nonlinearities of the description of functional behaviours and rates of change. The change point model is made up of triangular functions consisting of two linear segments defined by the endpoints of the series and a change point E in between. The variables of these triangular model functions are the unknown functional values at the endpoints and at the change point and the temporal position of the change point. Bayesian probability theory estimates the probabilities of all possible change point positions by marginalization over the functional values at the endpoints and the change point of the series. The Bayesian marginalization rule is employed to remove the change point variable by marginalization (see appendix). This extremely important rule tells how to remove an unwanted ‘nuisance’ parameter from a Bayesian calculation (Dose and Menzel 2004). The change point variable is unwanted because we do not consider it for example in a root mean square error (RMSE) method the best solution (RMSE minimum) or the two or three best triangular functions, but all of them. The advantage is that the marginalization rule superposes all possible triangular functions and 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 a particular E. In cases of more gradual change, appreciable change point probabilities are also observed for E+/-1, E+/-2….

An example is shown in panel a) of *Figure 5.1*. Hence, there is no single model function that explains the data outstandingly. In such a case Bayesian probability theory requires that the proper function is constructed by a weighted average of all possible triangular model functions with their respective probabilities as weights. The resulting functional description and rate of change are shown in panels b) and c) of *Figure 5.1*. Note that these inferences are presented with their standard deviation. Panels d) and e) show the predictions from the constant and the linear model.

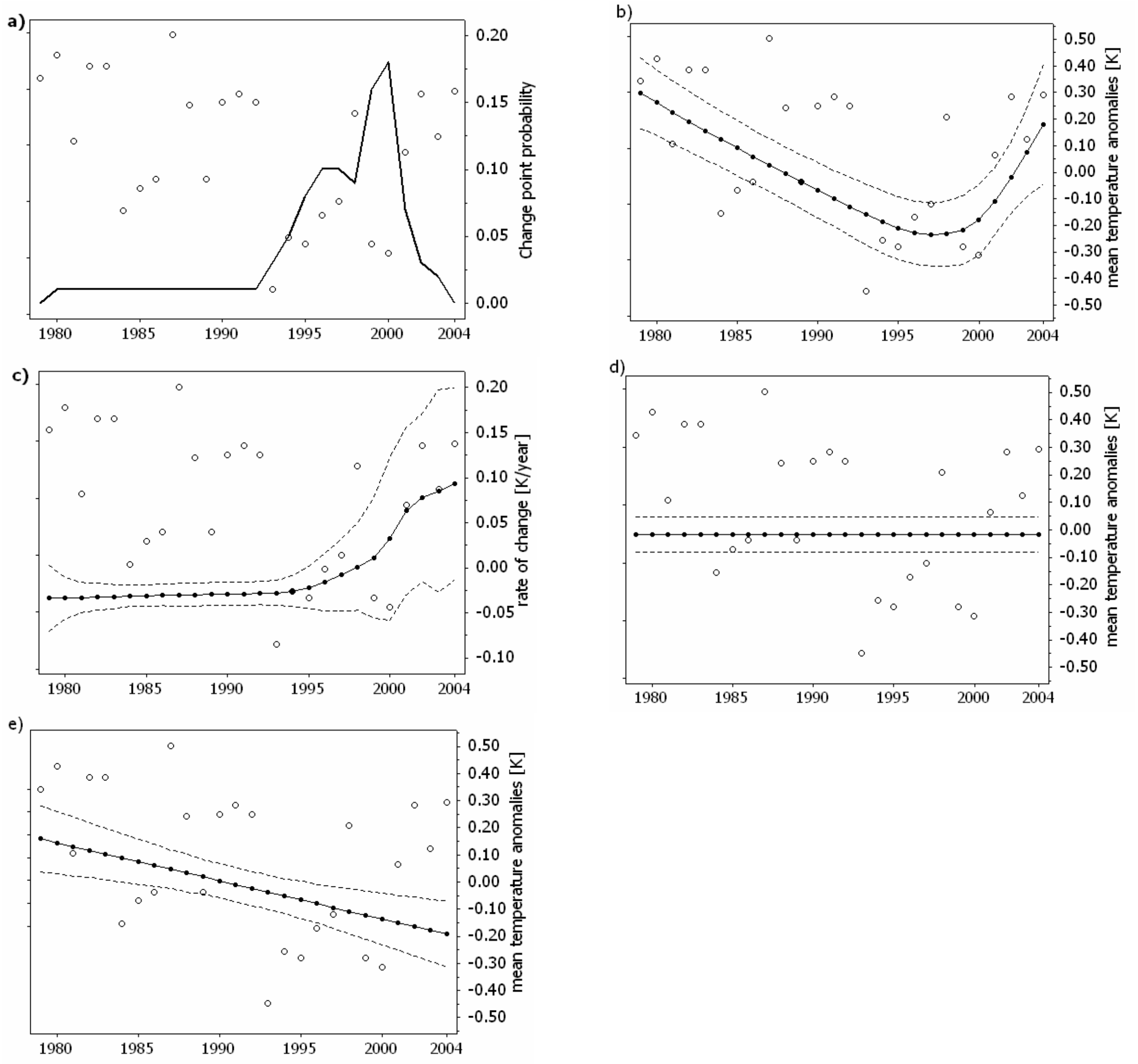


Figure 5.1: Bayesian model fits of the global annual mean temperature anomalies of RATPAC-A at the 150 hPa atmospheric pressure level over 1979-2004. (a) The change point probability distribution, (b) model-averaged functional behaviour (c) model-averaged rate of change (d) the constant model and (e) the linear model. Confidence intervals (standard deviations) are shown for each model as dashed lines. Open circles represent the data of the annual mean temperature anomalies in Kelvin [K]. On the left side the scale for the temperature anomalies ranges from -0.75 to 0.50 Kelvin.

A final word about the associated uncertainties of the change point model is in order. The predictive function for the change point model is itself a straight line between successive years of observations. The function extrapolates, therefore, linearly and the rate correspondingly at a constant level. Extrapolation into the region that is no longer supported by observational data exhibits a rapid widening of the confidence range in the end of each time series.

Our aim is not to achieve a result of a particular single model. We consider the most probable functional description and rate of change given by the space of three models. This is obtained from a final marginalization over the models. While the more complicated model certainly provides a better fit it has not necessarily the higher model probability. Bayesian probability theory selects a model by considering the tradeoff between lower misfit and higher complexity. This built-in property is also known as Ockham`s razor (Garret, 1991). Following the Ockham`s razor, a model should be chosen as complex as required for an explanation of the data and as simple as necessary to avoid the fitting of noise.

While the average over models is easily understood, the model average for the rate of change deserves a further comment. The constant model offers zero rate of change, the linear model a constant offset and only the change point model provides a time varying rate of change. The average over models does therefore depend on the probability of the constant model since it reduces the amplitudes of the linear and change point model contributions.

In the present paper we first analysed the global model preferences for the change point, linear and constant model. In a next step we analysed the global change point probability distributions to identify possible change points. Then for each pressure level the model-averaged global rates of change with associated uncertainties are presented.

5.2 Results

5.2.1 Model Preferences for the functional behaviour of Global Temperature Anomaly Data

Figure 5.2 shows the model preferences for the mean temperature anomaly of RATPAC-A and RATPAC-B data at each atmospheric pressure level on a global view and the residual sum of squares for RATPAC-A data sets. For both RATPAC-A and RATPAC-B data sets the model probabilities of each model are very analogous. From the surface up to the tropopause pressure level at 100 hPa the change point model fits the temperature time series best with change point model probabilities of mostly above 50 %. For both data sets at an air pressure level of 150 hPa the highest change point model probability of 90% for RATPAC-A and 78% for RATPAC-B is detected. At stratospheric pressure levels (70-30 hPa) the linear model provides with model probabilities of about 60% the best data fit. Different from the other pressure levels we see within the upper troposphere at 250 hPa the highest constant model probability of 20% for RATPAC-A and 28% for RATPAC-B data sets. Figure 5.2 c presents the residual sum of squares for the RATPAC-A data sets. The residual sum of squares increases continuously for each model from the surface up to the stratosphere. Within the stratosphere (70-30 hPa) the residual sums of squares for the constant model indicate that the data is not well represented by the constant likelihood function. Compared to the surface and troposphere also the change point and linear model reveal higher residual sums of squares.

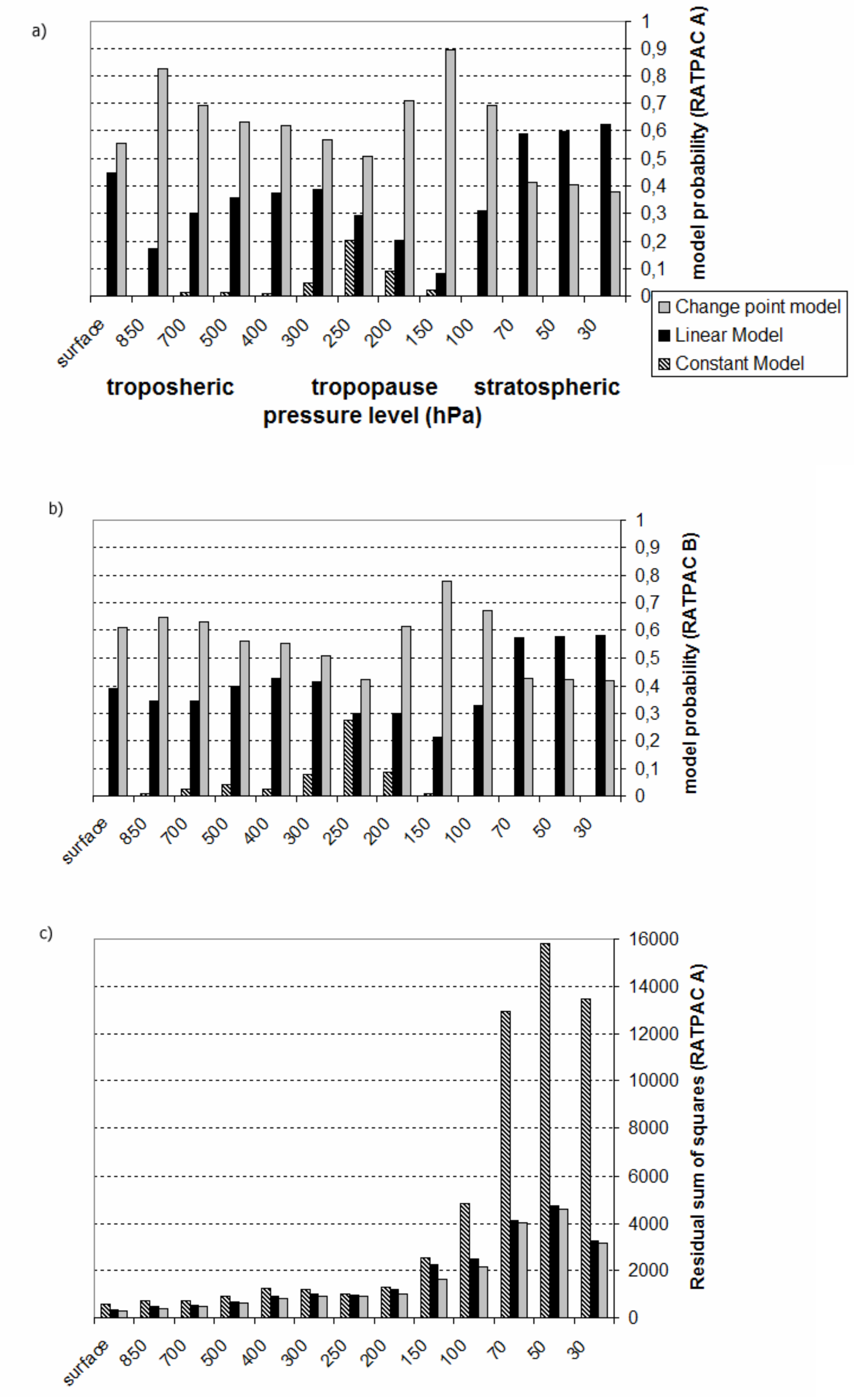


Figure 5.2: Bayesian model comparison of the change point, linear and constant model to describe the global annual mean temperatures time series over 1979-2004 at different pressure levels. In a) model probabilities of the RATPAC-A data set, b) model probabilities of the RATPAC-B data set and in c) the residual sum of squares of the RATPAC-A data set are presented.

5.2.2 Change point probability distributions

The results of the global change point probability distributions for the period 1979-2004 are depicted in Figure 5.3. From the surface pressure level to the upper troposphere high change point probabilities of about 8% are detected in 1982 (Figure 5.3 a and b). The highest change point probability of 10% is seen in 1993 at the 850 hPa pressure level (Figure 5.3 a). Since the beginning of the 1990s, change point probabilities in the tropopause pressure levels increase (Figure 5.3 c). In the year 2000, the highest change point probability peak of 18% at 150 hPa is reached which indicate an abrupt change. Generally stratospheric change point probability distributions exhibit almost no peaks above 5% (Figure 5.3 d).

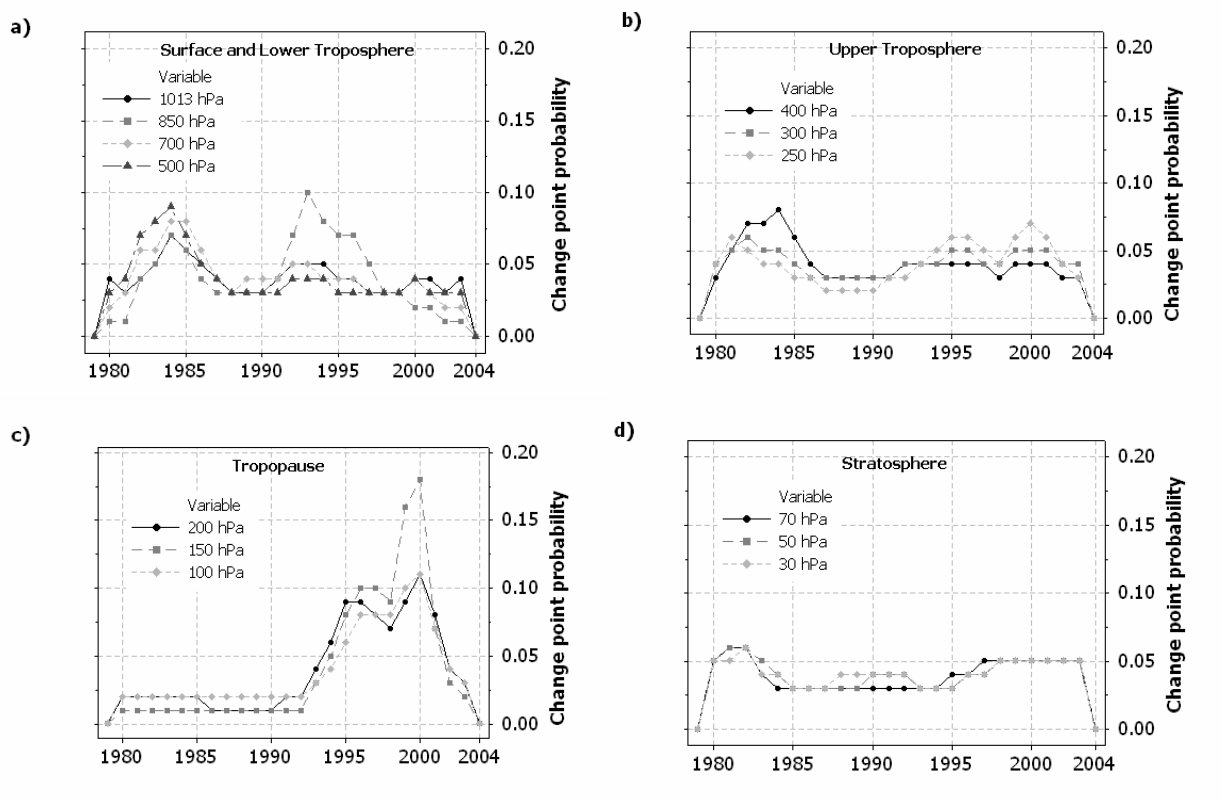


Figure 5.3: Change point probability distribution of global annual mean temperature anomalies over 1979-2004 presented for each pressure level. In panel a) Surface and Lower Troposphere b) Upper troposphere c) Tropopause d) Stratosphere change point probability distributions are shown. Note, that the symbols for 50 hPa and for 70 hPa exhibit a large overlap.

5.2.3 Global temperature rates of change and model averaged functional behaviours

Model averaged rates of change are obtained with annual resolution. Global rates of temperature change with associated confidence intervals are offered for all available atmospheric pressure levels (Figure 5.4. and Figure 5.5). For a better understanding of annual change rates we display additionally to the model averaged rate of change the model averaged function estimation of the mean temperature anomalies (Figure 5.4. and Figure 5.5 upper lines within each panel).

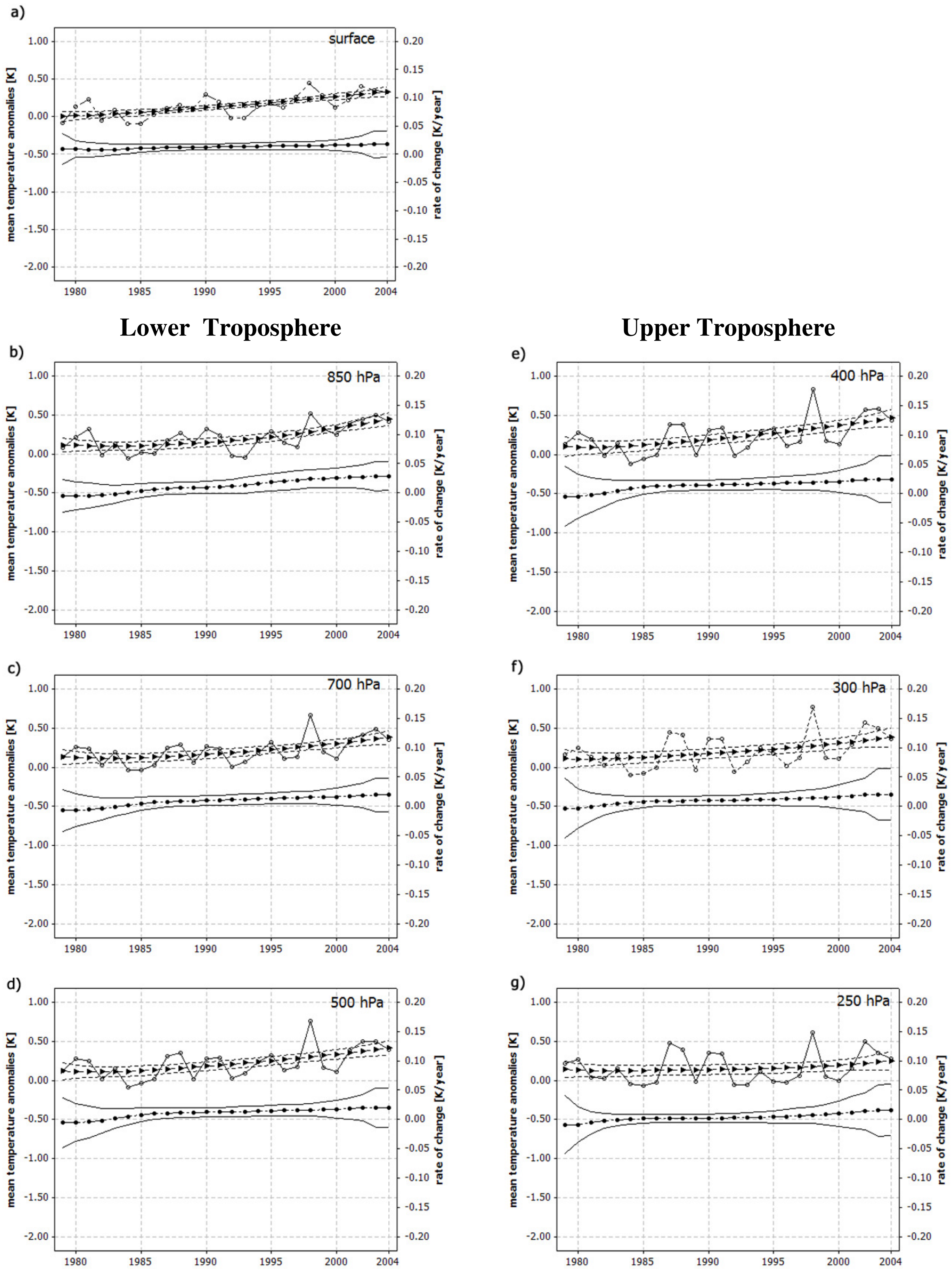
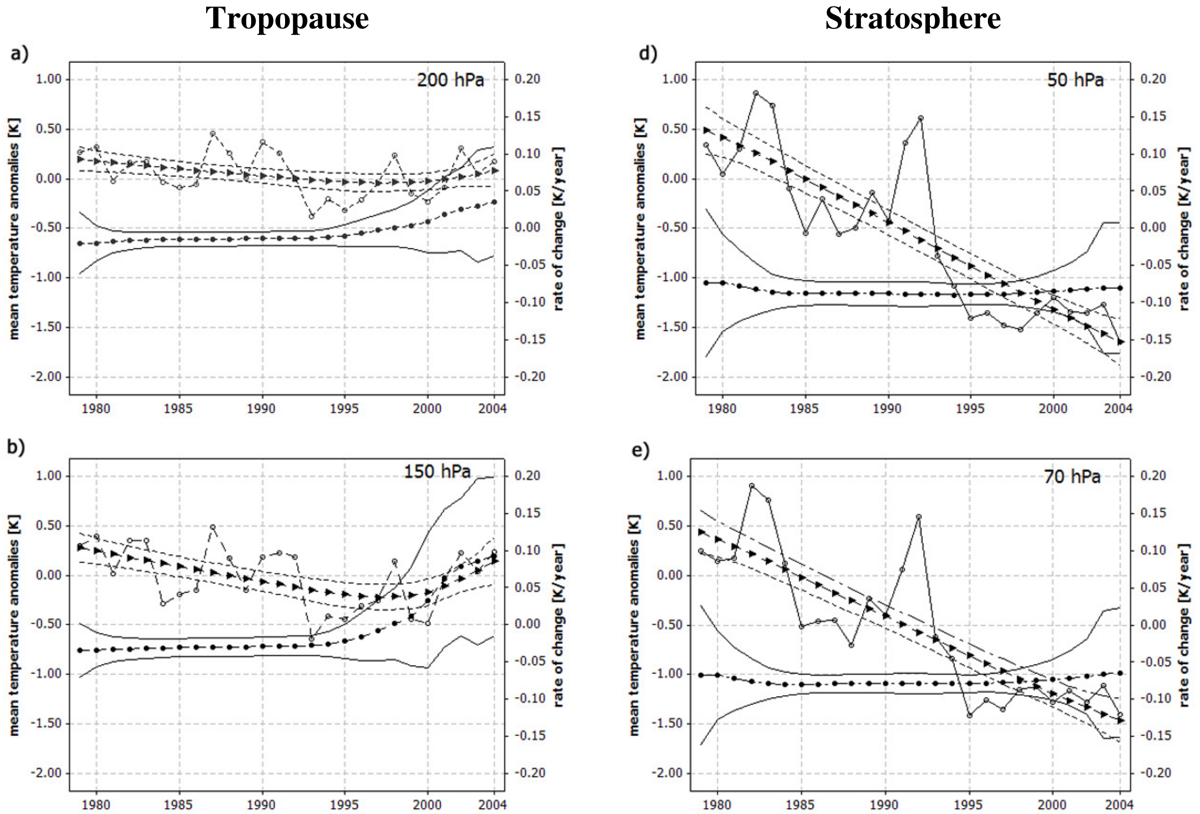


Figure 5.4: Bayesian model averaged rates of change (K/year) (line with full circles) and the model averaged functional behaviour (line with triangles) of global annual mean temperature anomalies (line with open circles) in annual resolution with associated confidence intervals (dashed line) over 1979-2004 for a) Surface, b-d) Lower Troposphere, e-g) Upper Troposphere pressure levels.

From the surface pressure level up to the upper troposphere the model averaged rates of change are very uniformly. On average, over all surface and troposphere pressure levels (surface to 250 hPa), the model averaged rates of change do gradually rise from -0.03 K/year in 1979 to 0.02 K/year in 2004. The associated confidence intervals of each pressure level at the end of the time series do not allow to give significant answers. The confidence intervals of the model averaged function estimation are narrower and therefore underline the results about the change directions. For the troposphere the model averaged functional behaviour increases quite constantly, whereas the model averaged functional behaviour for the tropopause (200-100 hPa) decreases until the end of the 1990s and increases from 2000 onwards. The rates of change in the tropopause (200- 100 hPa) are on average -0.03 K/year in 1979 and 0.04 K/year in 2004. Results for the 150 hPa pressure level have already been introduced within the method section and are shown in detail with a one year resolution in Figure 5.5 b. At 150 hPa we find a constant change of -0.03 K/year from 1979 to 1995 and since 1995 the rate of change has risen to 0.1 K/year (Figure 5.5 b). Especially within the tropopause the confidence intervals are relatively large in the end of the time series (Figure 5.5 b). In contrast, the stratosphere exhibits over the entire investigated period nearly constant negative rates of change of on average -0.07 K/year in 1979 and 2004, indicating cooling. But also here the confidence interval of each pressure level are quite large and do not remain in any case negative. In the stratosphere the model averaged functional behaviour decreases constantly. A visual inspection of the mean temperature data variations reveal for all stratosphere layers pronounced peaks at the two major volcanic eruptions- 1982 El Chicón and 1991 Pinatubo.



Tropopause **Stratosphere**

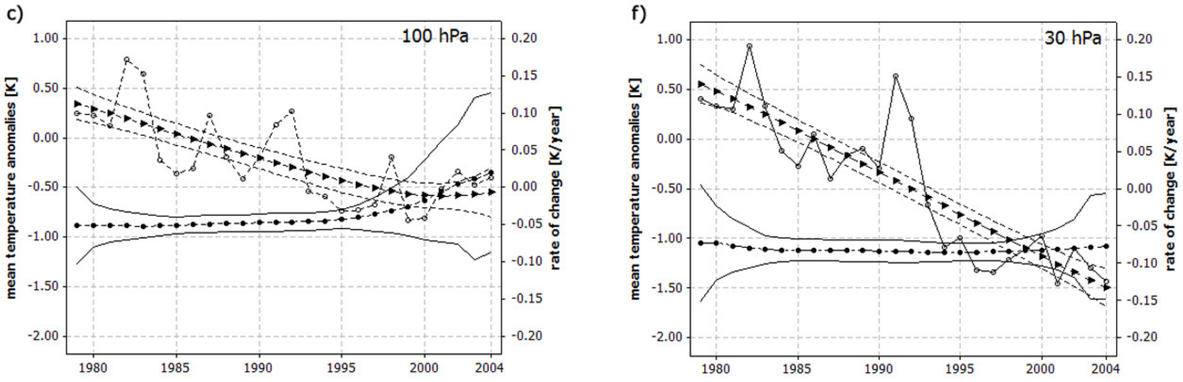


Figure 5.5: Bayesian model averaged rates of change (K/year) (line with full circles) and the model averaged functional behaviour (line with triangles) of global annual mean temperature anomalies (line with open circles) in annual resolution with associated confidence intervals (dashed line) over 1979-2004 for a-c) Tropopause, d-f) Lower Stratosphere pressure levels.

5.3 Discussion

Although radiosonde temperature records have great potential value as a climate change indicator, their use has been limited by the presence of many inhomogeneities due to changes in instruments and practices (Gaffen, 1994). Within the community that constructs and actively analyses satellite- and readiosonde-based temperature records there is agreement that the uncertainties about long-term change are substantial (Trenberth et al., 2007). Inhomogeneities are difficult to remove because of the frequent lack of neighbouring homogeneous reference series, incorrect or incomplete station history information, and the need to maintain vertical consistency in adjusted time series (Free et al. 2005). The Intergovernmental Panel on Climate Change (Trenberth et al., 2007) stated that it is difficult to make quantitatively defensible judgements as to which, if any, of the multiple, independently derived estimates is closer to the true climate evolution. The limitations of the currently used radiosonde data render comparison and interpretation of the observed changes extremely difficult. Thus, there is a strong need of improving on one hand the quality of the radiosonde temperature record but on the other hand also of improving recently applied methods of change detection in radiosonde time series. Standard statistical practice ignores model uncertainty. Data analysts typically select a model from some class of models and then proceed as if the selected model had generated the data. This approach ignores the uncertainty in model selection, leading to over-confident inferences and decisions that are more risky than one thinks they are (Hoeting et al., 1999).

To assess the potential value of the actual adjusted radiosonde temperature records we suggest and apply a Bayesian model averaging approach for the change detection by estimating different model probabilities, functional behaviours and model averaged rates of change. The Bayesian model comparison provides an excellent possibility to judge and incorporate different model estimates. If all three model alternatives exhibit similar probabilities as our results reveal for the 250 hPa level, Bayesian probability theory calculates the probability of a given model, i.e. in our case the change point model compared to one or more alternatives, such as the constant and linear models. Unless one of the models sticks out with a very high probability as e.g. the change point model in the 150 hPa level, the description of the data in terms of only one model is again unsatisfactory. Instead, the rigorous application of Bayesian probability theory describes that the proper functional behaviour and the proper trend are obtained by superposition of a constant, a linear function and the change point model function again weighted with their respective model probabilities. The estimated model probabilities provide an indication of the general change patterns (no change, linear change or nonlinear change) of atmosphere temperature. The residual sum of squares reveals that the representation of Bayesian model averaged functional behaviour and rates of changes are especially useful for the surface, troposphere and tropopause. On one hand that means that our chosen models are adequately to represent the changes. Bayesian results of these pressure levels are more informative and improved than results based on single model approaches. On the other hand within stratosphere levels the residual sums of squares indicate that our three chosen models are suboptimal to describe adequately the changes in the stratosphere. But the model averaging is still more reliable than results based just on a single model as the linear regression model.

The availability of various temperature observations and the progressively acquired capability to perform improved models enables to find realistic and causal explanations of the observed changes. Trenberth et al. (2007) stated that an anthropogenic influence on free atmosphere temperatures has been detected in analyses of satellite data since 1979, although this finding has been found to be sensitive to which analysis of satellite data is used. Anthropogenic influence on free atmosphere temperatures has been robustly detected in a number of different studies analysing various versions of the Hadley Centre Radiosonde Temperature data set (Parker et al., 1997) by means of variety of different diagnostics and fingerprints estimated with the HadCM2 and HadCM3 models (Jones and Moberg, 2003).

Free et al. (2005) found with the help of the linear model a noticeable levelling off of the stratospheric cooling for the 100-50 hPa layer in the global mean. For 1979-2004 they present a stratospheric cooling of approximately –0.5 to -0.9 K/decade. The model averaged rates of change for the global mean temperature anomalies reveal similar values as the linear approach of Free et al. (2005) for stratospheric cooling from 1979 to 2004. Our Bayesian results reveal on average of all stratosphere pressure levels a constant negative derivative of -0.07 K/year but associated with high confidence intervals (Figure 5.5 d-f).

Estimates from adjusted radiosondes, satellites (MSU channel 4) and reanalyses are in qualitative agreement, suggesting a lower-stratospheric cooling of between 0.3°C and 0.6°C per decade since 1979. Model calculations indicated that the observed ozone losses had the potential to yield substantial cooling of the global lower stratosphere (Ramaswamy et al., 2001).

Longer radiosonde records (back to 1958) also indicate cooling but the rate of cooling has been significantly greater since 1979 than between 1958 and 1978. A reason for the high linear contribution to the model averaged rates of change can be induced by increasing greenhouse gas concentrations and stratospheric ozone depletion. Climate model simulations of Ramaswamy et al. (2006) indicate that the observed continued cooling is largely attributable to the combined effect of changes in anthropogenic factors such as ozone depletion and increase in well-mixed greenhouse gases and natural factors such as solar irradiance variation and volcanic aerosols. Anthropogenic contributions to aerosols (primarily sulphate, organic carbon, black carbon, nitrate and dust) together produce a cooling effect (Trenberth et al., 2007). Ramaswamy et al. (2006) described in their study that the anthropogenic cooling influence becomes an important global feature by the late 1980s and a dominant one in the 1990s.

It is likely that radiosonde records overestimate stratospheric cooling, owing to residual effects of changes in instruments and processing (such as for radiation corrections; Lanzante et al., 2003b; Randel et al., 2006) and possibly increased sampling of cold conditions owing to stronger balloons (Parker and Cox, 1995). In radiosonde instruments the largest discontinuities appear to be related to solar heating of the temperature sensor. These discontinuities have greatest impact at stratosphere levels, where direct sunlight can cause radiosonde-measured temperatures to rise several °C above ambient temperatures (Christy et al., 2006). Generally the newly developed RATPAC approaches (Lanzante et al., 2003a, 2003b; Free et al., 2005) reduce most of the inhomogeneities in used radiosonde data.

However many studies detected a stratospheric warming in ~1982-1984 and ~1991-1993 after the El Chicón (1982) and Pinatubo (1991) volcanic eruptions (e.g. Ramaswamy et al., 2006, Pawson et al., 1998). This warming comes from the radiative effects of the stratospheric sulphate aerosols formed after eruptions (Seidel and Lanzante, 2004, Santer et al., 2003). Because of the stratospheric warming episodes following major volcanic eruptions, the trends are far from being linear. A linear trend over the long-term is often not a very good approximation of what has occurred (Seidel and Lanzante, 2004; Thorne et al., 2005); alternative interpretations are to factor in the abrupt climate regime shift (Trenberth, 1990) and episodic stratospheric warming and tropospheric cooling for the two years following major volcanic eruptions. Numerous studies (Ramaswamy et al., 2006, Thompson et al., 2005, Santer et al., 2003) reveal that substantial cooling of the global lower stratosphere occurred in two pronounced steplike transitions over 1979-2003. For stratospheric time series Seidel and Lanzante (2004) identified the best description with a nonlinear type of model. They found that the steplike model fitted the radiosonde and microwave sounding unit data, but only when they did not exclude the influence of major volcanic eruptions. Our Bayesian results reveal a preference for the linear model of about 60% and for the change point model of about 40%. The change point distributions for the 70 to 30 hPa pressure stay below 5% and do not mirror high change point probabilities at the years during the two major volcanic eruptions. As quite high residual sums of squares indicate that non of our three chosen models is suitable to describe adequately the changes in the stratosphere the Bayesian results have to be interpreted with cautious. The Bayesian model averaged procedure filter out interannual variations and thus the change point probability, model averaged functional behaviour and rate of change do not represent the influence of the two major volcanic events. In our analysis we use a one change point model. Thus undetected change points may result from our limited model space. A multiple change point model is of course capable of modelling more detailed structure in a time series and therefore would mirror more adequately the observed step like changes of other studies. But each added change point adds two more variables to the likelihood, that may be punished by the Ockham`s razor. Whether the multiple change point model really provides a better data description has to be tested in future work.

The tropopause represents the boundary between the troposphere and the stratosphere and is characterised by large changes in the thermal, dynamical, and chemical structure of the atmosphere (Santer et al., 2003). The height of the tropopause is affected by the heat balance of both the troposphere and the stratosphere. For example, when the stratosphere warms owing to absorption of radiation by volcanic aerosol, the tropopause is lowered. Conversely, a warming of the troposphere raises the tropopause, as does a cooling of the stratosphere. Santer et al. (2003) found in their analysis that human induced changes in well-mixed greenhouse gases and tropospheric and stratospheric ozone account for 80% of the simulated rise in tropopause height over 1979-1999. The Santer et al. (2003) results provide support for warming of the troposphere and cooling of the lower stratosphere over the last four decades of the 20th century, and indicate that both of these changes in atmospheric temperature have contributed to an overall increase in tropopause height. The radiosonde-based analyses of Randel et al. (2000), Seidel et al. (2001) and Highwood et al. (2000) also show increases in tropical tropopause height.

Bayesian results reveal for the pressure levels of 200 hPa, 150 hPa and 100 hPa change point model probabability of 70%, 90% and 70% respectively. These high change point model probabilities indicate strong nonlinear changes since the beginning of the 1990s. Within the tropopause levels the change point distributions are very synchronous and rise since the beginning of the 1990s and exhibit maximum peaks in 2000 (Figure 5.3 c). On the global view in the tropopause pressure levels (100-200 hPa) we identify a tendency to a warming with maximum rates of 0.1 K/year in 2004. The model averaged functional behaviour for the tropopause decreases until the end of the 1990s and increases from 2000 onwards.

The dominant change can be very likely associated with the El Niño-Southern Oscillation phenomenon (ENSO). The major ENSO warming effect in 1998 would match with the detected high change point probability distributions in the tropopause and the indication of warming of both the model averaged function and rates of changes since the mid of the 1990s. The cooling following the eruption of Mt. Pinatubo in 1991 is evident and also reflected by our model averaged functional behaviour but the cooling effect of El Chicón (1982) is masked by an ENSO warming that occurred at the same time (Wigley et al., 2006). High nonlinear changes at a certain tropopause level can be also possibly induced by changes in vertical thermal structure of the atmosphere, e.g. through a rise in the tropopause height. For example if the pressure level is above the tropopause in the beginning of the record and local temperature rates of changes are small, a systematic increase in tropopause height will lead to a change point once the tropopause passes the pressure level under consideration. The change point in the vertical thermal structure would be translated into a change point in time.

We detected an accumulation of change point probability peaks at surface and troposphere pressure levels in 1984 and in the beginning of the 1990s right after the volcanic eruptions. The two volcanic eruptions might affect change point probability in the troposphere but very likely these changes are additionally caused by other major influences as e.g. ENSO. Results suggest that a nonlinear model describes the troposheric data better than the alternative linear and constant model. For the troposphere the change point model probability are mostly clearly above 60%. Thus the linear model contribution is not as dominant as in the stratosphere. Seidel and Lanzante (2004) concluded that the linear regression model provides the best data description for the microwave sounding unit data and radiosonde data at a pressure level of 850-300 hPa in the satellite data period (1979-2001). But Seidel and Lanzante (2004) already indicated in their publication a flat steps model as a reasonable alternative. Santer et al. (2003) concluded that the troposphere warming is predominantly due to increases in well mixed greenhouse gases. Global time series of satellite-born microwave sounders show a global cooling of the stratosphere of -0.32°C to -0.47°C per decade and a global warming of the troposphere of 0.04°C to 0.20°C per decade for the period 1979 to 2004 (Trenberth et al., 2007). The large spread in tropospheric trends stems from differences in the inter-statellite calibration and merging technique, and differences in the corrections for oribital drift, diurnal cycle change. In the paper of Free et al. (2005) the RATPAC-A annual mean temperature series exhibit a warming of 0.13 K/decade in the global mean troposphere for 1979-2004. In agreement with Free et al. (2005) our global rates of change exhibit a uniform trend of on average 0.02 K/year in 2004 from the surface to the upper troposphere but again associated with high confidence intervals.

In our analysis we have found that confidence intervals are often larger than single-model based confidence intervals; the latter tend to be too narrow. Model averaging results are robust to model choice. Model averaging also allows user to incorporate several competing models in the estimation process; thus model averaging may offer a committee of scientists a better estimation method than the traditional approach of trying to get the committee to agree on one best model (Hoeting et al., 1999). Another potential concern is that model averaging tends to produce higher estimates of variance than do estimates that ignore model uncertainty (Hoeting et al., 1999). Why we use model averaging when they are less likely to get significant results? The simple answer is that model averaging is more correct, because it takes account of a source of uncertainty that analyses based on model selection ignore. The implification is that standard analyses probably tend to find significant results too often. Also if results are significant under model averaging, then conclusions are more robust than those that depend upon the particular model that has been selected.

5.4 Conclusion

The predominant observations of non-linear temperature changes in nearly all pressure levels underline the importance of alternatives to linear models. The Bayesian approach used here affords not only description of nonlinearity but offers in addition the estimation of model probabilities functional behaviours and model averaged rates of change in annual resolution. This bundle of results provides for each pressure level further details into the changes of radiosonde data. The model probabilities and functional behaviours provide information about the general change patterns (no change, linear change or nonlinear change) of atmosphere temperature. The Bayesian model comparison provides an excellent possibility to judge and incorporate different model estimates. Bayesian model averaged functional behaviours and rates of changes are especially useful for the surface, troposphere and tropopause. For these pressure levels our three chosen models are adequately to represent the changes. Bayesian results of these pressure levels are more informative and improved than results based on single model approaches. Bayesian model averaging provides better average predictive performance than any single model that could be selected.

However in the stratosphere none of our three chosen models is complex enough to describe adequately the changes. Thus in future work it should be tested with the Ockham's razor whether e.g. a multiple change point model provides a better data description. There are many open research questions related to Bayesian model averaging. These include the investigation of the performance of Bayesian model averaging when the true model is not in the model class. Further more it would be interesting to develop especially for the stratosphere a Bayesian model averaging methodology for model classes not considered here.

5.5 *Appendix*

In response to a request during the review procedure we outline in this appendix the Bayesian calculation without going into details of the algebra. Those interested in these details are referred to Dose and Menzel (2004). Bayesian calculations rest on the application of two theorems. The first is Bayes' theorem which drives from the product rule for manipulating conditional probabilities

$$p\left(\vec{\theta}\middle|\vec{d},\vec{I}\right)=p\left(\vec{\theta}\middle|\vec{I}\right)*p\left(\vec{d}\middle|\vec{\theta},\vec{I}\right)\Big/p\left(\vec{d}\middle|\vec{I}\right) \qquad \text{(A1)}$$

$p\left(\vec{\theta}\middle|\vec{d},\vec{I}\right)$ is called the posterior probability or probability density for a set of parameters $\vec{\theta}$ taking into account the information of new data $\vec{d}$. $p\left(\vec{\theta}\middle|\vec{I}\right)$ is called the prior probability on $\vec{\theta}$ before taking into account the data. $p\left(\vec{d}\middle|\vec{\theta},\vec{I}\right)$ is the sampling distribution of the data given $\vec{\theta}$. When considered as a function of $\vec{\theta}$ for fixed $\vec{d}$ it is called the likelihood of $\vec{\theta}$. $p\left(\vec{d}\middle|\vec{I}\right)$ is the normalisation integral for the right hand side of (A1)

$$p\left(\vec{d}\middle|\vec{I}\right)=\int p\left(\vec{\theta}\middle|\vec{I}\right)*p\left(\vec{d}\middle|\vec{\theta},\vec{I}\right)d\vec{\theta} \qquad \text{(A2)}$$

The operation in (A2) is called marginalization of $\vec{\theta}$. $p\left(\vec{d}\middle|\vec{I}\right)$ is then the global probability of the data given the information I which incorporates the model applied for formulating the likelihood, irrespective of the numerical values of the parameters $\vec{\theta}$. It is the key quantity in Bayesian model comparison.

We shall now adapt these abstract concepts to the problems of this paper and consider the change point model. The other two models are obtained from the same analysis by reductions of the dimension of $\vec{\theta}$.

The likelihood of the change point model is given by

$$p\left(\vec{d}\middle|\vec{f},\sigma,t_E,M,I\right) \qquad \text{(A3)}$$

where $\vec{d}$ are the data, $\vec{f}$ the three support functional values at the end of the time series and at the change point t_E and σ the uncertainty of the data assumed unknown but common to all data. t_E is of course a discrete variable. Our first goal is to derive the probability of the change point time t_E given the change point model (this information has been split off from I in A3) and the data $p\left(t_E\middle|\vec{d},M,I\right)$. This distribution is obtained from $p\left(t_E,\vec{f},\sigma\middle|\vec{d},M,I\right)$ by mariginalisation

$$p\left(t_E,\vec{f},\middle|\vec{d},M,I\right)=\int d\sigma d\vec{f}\,p\left(t_E,\vec{f},\sigma\middle|\vec{d},M,I\right) \qquad \text{(A4)}$$

We decompose the probability under the integral by application of the product rule

$$p\left(t_E,\vec{f},\middle|\vec{d},M,I\right)=\int d\sigma d\vec{f}\,p\left(t_E,\vec{f},\sigma\middle|M,I\right)*p\left(\vec{d}\middle|t_E,\vec{f},\sigma,M,I\right) \qquad \text{(A5)}$$

The first factor under the integral (A5) is the joint prior distribution for $t_E,\vec{f}$ and σ. The second factor is the likelihood (A3). This concludes the calculation of the change point probability distribution. It can also be used to obtain the evidence of the data. From Bayes theorem

$$p\left(t_E \middle| \vec{d}, M, I\right) = p\left(t_E \middle| M, I\right) * p\left(\vec{d} \middle| t_E, M, I\right) \Big/ p\left(\vec{d} \middle| M, I\right) \qquad \text{(A6)}$$

We obtain by summation over t_E on both sides

$$p\left(\vec{d} \middle| M, I\right) = \sum_{t_E} p\left(t_E \middle| M, I\right) * p\left(\vec{d} \middle| t_E, M, I\right) \qquad \text{(A7)}$$

Where the likelihood $p\left(\vec{d} \middle| t_E, M, I\right)$ is obtained from (A3) by marginalization over $\vec{f}$ and σ. The marginal likelihood (A7) forms also the terms for model selection

$$p\left(M \middle| \vec{d}, I\right) = p\left(M \middle| I\right) * p\left(\vec{d} \middle| M, I\right) \Big/ p\left(\vec{d} \middle| I\right) \qquad \text{(A8)}$$

where

$$p\left(\vec{d} \middle| I\right) = \sum_{M} p\left(M \middle|, I\right) * p\left(\vec{d} \middle| M, I\right) \qquad \text{(A9)}$$

Our final task is the estimation of a function $\varphi\left(t \middle| \vec{f}, t_E\right)$depending on parameters $\vec{f}$ and t_E as a function of t. We generalize to

$$\left\langle \varphi^n(t) \right\rangle = \sum_{t_E} \int d\vec{f} \left\{\varphi\left(t \middle| \vec{f}, t_E\right)\right\}^n p\left(\vec{f}, t_E \middle| \vec{d}, M, I\right) \qquad \text{(A10)}$$

The estimations $\left\langle \varphi^n(t) \right\rangle$ are obtained from a marginalization over $\vec{f}$ and t_E using the posterior probability of these parameters. (A10) is employed to derive mean $\left\langle \varphi\ (t) \right\rangle = \mu_1$ and second moment $\mu_2 = \left\langle \varphi^2(t) \right\rangle =$of $\varphi^2(t)\rangle$ from which we obtain the standard deviation of φ in the usual way

$$\left\langle \Delta\varphi^2(t) \right\rangle = \left\{\mu_2(\varphi) - \mu_1^{\,2}(\varphi)\right\}^{1/2} \qquad \text{(A11)}$$

This completes the formal description of our calculation. For details on the various probability distributions and the discussion of logical independencies from one or several parameters the reader is again referred to Dose and Menzel (2004).

5.6 References

Christiansen, B., (2003) Evidence for nonlinear climate change: Two stratospheric regimes and a regime shift. Journal of Climate **16**, 3681–3690.

Christy J.R., Seidel D.J., Sherwood S.C., (2006) What kinds of atmospheric temperature variations can the current observing systems detect and what are their strengths and limitations, both spatially and temporally? *in Temperature Trends in the Lower Atmosphere:Steps for Understanding and Reconciling Differences. Karl TR, Hassol SJ, Miller CD et al. editors. A Report by the Climate Change Science Program and the Subcommittee on Global Change Research, Washington, DC.*

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.

Free M., Angell J.K., Durre I., Lanzante J., Peterson T.C., Seidel D.J., (2004) Using first differences to reduce inhomogeneity in radiosonde temperature datasets. *Journal of Climate* **21**, 4171– 4179.

Free M., Seidel D.J., Angell J.K., Lanzante J., Durre I., Peterson T.C., (2005) Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC): A new dataset of large-area anomaly time series. *Journal of Geophysical Research* **110**, D22101.

Gaffen D.J., (1994) Temporal inhomogeneities in radiosonde temperature records. *Journal of Geophysical Research* **99**, 3667–3676.

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.*

Highwood E.J., Hoskins B.J., Berrisford P., (2000) Properties of the Arctic tropopause. *Q. J. Roy. Meteor. Soc.* **126**, 1515-1532.

Hoeting J.A., Madigan D.M., Raftery A.E., Volinsky C.T., (1999) Bayesian model averaging: A tutorial (with discussion). *Statistical Science* **14**, 382–401. [A corrected version is available online at www.stat.washington.edu/www/research/online/ hoeting1999.pdf.]

Jones P.D., Moberg A., (2003) Hemispheric and large-scale surface air temperature variations: An extensive revision and update to 2001. *Journal of Climate* **16**, 206–223.

Karl T.R., Knight R.W., Baker B., (2000) The record breaking global temperatures of 1997 and 1998: Evidence for an increase in the rate of global warming? *Geophysical Research Letters* **27**(5), 719– 722.

Karl T.R., Miller C.D., Murray W.L., editors., (2006) in Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences. T. R. Karl, S. J. Hassol, C. D. Miller, and W. L. Murray, editors. A Report by the Climate Change Science Program and the Subcommittee on Global Change Research, Washington, DC.

Lanzante, J.R., Klein S.A., Seidel D.J., (2003a) Temporal homogenization of monthly radiosonde temperature data. Part I: Methodology. *Journal of Climate* **16**, 224– 240.

Lanzante J.R., Klein S.A., Seidel D.J., (2003b) Temporal homogenization of monthly radiosonde temperature data. Part II: Trends, sensitivities, and MSU comparison. *Journal of Climate* **16**, 241– 262.

Lindzen R.S., Giannitsis C., (2002) Reconciling observations of global temperature change. *Journal of Climate* **29**(12).

Parker D.E., Cox D.I., (1995) Towards a consistent global climatological rawindsonde database. *International Journal of Climatology* **15,** 473–496.

Parker D.E., Gordon M., Cullum D.P.N., Sexton D.M.H., Folland C.K., Rayner N., (1997) A new global gridded radiosonde temperature data base and recent temperature trends. *Geophysical Research Letters* **24**, 1499-1502.

Pawson S, Labitzke K, Leder S. 1998. Stepwise Changes in Stratospheric Temperature. *Geophysical Research Letters* **25**, 12, 2157-2160.

Priestley M., (1981) Spectral Analysis and Times Series. *Univariate Series* **1**, Academic, San Diego, Calif.

Raftery A.E., Gneiting T., Balabdaoui F., Polakowski M., (2005) Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review* **133**, 1155– 1174.

Ramaswamy V., Chanin M.L., Angell J., Barnett J., Gaffen D., Gelman M., Keckhut P., Koshelkov Y., Labitzke K., Lin J.J.R., O'Neill A., Nash J., Randel W., Rood R., Shine K., Shiotani M., Swinbank R., (2001) Stratospheric temperature trends: observations and model simulations. Reviews of Geophysics ***39**, 71-122. Karl TR, Hassol SJ, Miller CD et al. editors. A Report by the Climate Change Science Program and the Subcommittee on Global Change Research, Washington, DC.*

Ramaswamy V., Hurrell J.W., Meehl G.A., Phillips A., Santer B.D., Schwarzkopf M.D., Seidel D.J., Sherwood S.C., Thorne P.W., (2006) Why do temperatures vary vertically (from the surface to the stratosphere) and what do we understand about why they might vary and change over time? *in Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences. Karl TR, Hassol SJ, Miller CD et al. editors. A Report by the Climate Change Science Program and the Subcommittee on Global Change Research, Washington, DC.*

Randel W.J., Wu F., V.Ã. mel H., Nedoluha G.E., Forster P., (2006) Decreases in stratospheric water vapor after 2001: Links to changes in the tropical tropopause and the Brewer-Dobson circulation. *Journal of Geophysical Research* **111**(d12), D12312.

Rodionov S.N., (2004) A sequential algorithm for testing climate regime shifts. *Geophysical Research Lett*ers **31**, L09204.

Santer B.D., Taylor K.E., Wigley T.M.L., Johns T.C., Jones P.D., Karoly D.J. , Mitchell J.F.B., Oort A.H., Penner J.E., Ramaswamy V., Schwarzkopf M.D., Stouffer R.J., Tett S., (1996) A search for human influences on the thermal structure of the atmosphere. *Nature* **382**, 39–46.

Santer B.D., Sausen R., Wigley T.M.L., Boyle J.S., AchutaRao K., Doutriaux C., Hansen J.E., Meehl G.A., Roeckner E., Ruedy R.G., Schmidt G., Taylor K.E., (2003) Behavior of tropopause height and atmospheric temperature in models, reanalyses, and observations: Decadal changes. *Journal of Geophysical Research* ***108***(D1), 4002, doi:10.1029/2002JD002258.

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.

Schleip C., Menzel A., Dose V., (2008) Norway spruce (*Picea abies*): Bayesian analysis of the relationship between temperature and bud burst. *Agricultural and Forest Meteorology* **148**, 4, 631-643.

Schwarz G., (1978) Estimating the dimension of a model. *Annals of Statistic* **6**, 461– 464.

Seidel J.D., Ross R.J., Angell J.K., Reid G.C., (2001) Climatological characteristics of the tropical tropopause as revealed by radiosondes. *Journal of Geophysical Research* **106**, 7857-7878.

Seidel D.J., Lanzante J.R., (2004) An assessment of three alternatives to linear trends for characterizing global atmospheric temperature changes. *Geophysical Research Lett*ers **109**, D14108.

Sloughter J.M., Rafery A., Gneiting T., (2006) Probabilistic Quantitative Precipitation Forecasting Using Bayesian Model Averaging. Technical Report no. 496, Department of Statistics, University of Washington.

Tett S.F.B., Mitchell J.F.B., Parker D.E., Allen M.R., (1996) Human influence on the atmospheric vertical temperature structure: Detection and observations. *Science* **274**, 1170–1173.

Thorne P.W., Parker D.E., Tett S.F.B., Jones P.D., McCarthy M., Coleman H., Brohan P., (2005) Revisiting radiosonde upper air temperatures from 1958 to 2002. *Journal of Geophysical Research*. **110**, D18105.

Thompson, DWJ and Solomon S., (2005) Recent stratospheric climate trends: Global structure and tropospheric linkages. *Journal of Climate*, **18**, 4785–4795

Tomé A., Miranda P.M.A., (2004) Piecewise linear fitting and trend changing points of climate parameters. *Geophysical Research Letters* **31**, L02207.

Trenberth K.E., (1990) Recent observed interdecadal climate changes in the Northern Hemisphere. *Bull. Am. Meteorol. Soc*. **71**, 988–993.

Trenberth K., Hurrell J., (1994) Decadal atmosphere–ocean variations in the Pacific. *Climate Dynamics* **9**, 303–319.

Trenberth K.E., Jones P.D., Ambenje P., Bojariu R., Easterling D., Klein Tank A., Parker D., Rahimzadeh F., Renwick J.A., Rusticucci M., Soden B., Zhai P., (2007) Observations: Surface and Atmospheric Climate Change. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Wigley T.M.L., Ramaswamy V., Christy J.R., Lanzante J.R., Santer B.D., (2006) Executive Summary *in Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences. Karl TR, Hassol SJ, Miller CD et al. editors. A Report by the Climate Change Science Program and the Subcommittee on Global Change Research, Washington, DC.*

6 General and Summarizing Discussion

This chapter summarizes the previous chapters in which I have examined influences of recent climate change on natural systems. I particularly detect recent temperature increases and analyse the relationship to the bio-indicator plant phenology. In the newest fourth IPCC report (AR4), Working Group II found broad observational evidence that recent regional climate changes, particularly temperature increases, have already affected many physical and biological systems (Rosenzweig et al., 2007). Among the affected systems phenology is one where the changes are quite obvious, strongly temperature related and where almost no non-climate driving factors exist. These phenological responses are manifest in a variety of ways, originating at ecophysiological scales and expressed through individual organisms, populations, communities, and, notably, at landscape and regional scales. The results of this dissertation highlight the complexity of broad scale phenological responses to climatic changes. Also Inouye (2008), Kudo et al. (2008), Miller-Rushing and Primack (2008), Post et al. (2008) illustrate the importance of phenology, especially in the context of global warming and increasing climate variability, in a special feature in *Ecology*.

Below I recapitulate the major findings of the single chapters and give concluding answers structured under the three leading research questions introduced in Chapter 1.4.

6.1 First leading Research question

What are the advantages and disadvantages of the Bayesian approach compared to conventional statistical methods when analysing climate change impacts on natural systems?

Bayesian methods and classical methods both have advantages and disadvantages. In general the dissertation illustrates that the Bayesian analysis is a practical and theoretically appropriate tool for drawing inferences about climate change and for making decisions based on those inferences. During the review processes of all articles introduced here in the previous chapters I made the experience that many scientific readers tend to believe results that support their preconceptions and disbelieve results that surprise them. Bayesian methods seem to encourage this undisciplined mode of thinking. And, probably worse, Bayesian techniques motivate even the best-intentioned researchers to get caught in the principles of prior beliefs (Gelman, 2008).

The fundamental objections to Bayesian methods imply that Bayesian methods are presented as an automatic inference engine, and this raises suspicion in anyone with applied experience, who realizes that different methods work well in different settings (see, for example, Little, 2006). "Bayesians" promote the idea that a multiplicity of parameters can be handled via hierarchical models, but it seems implausible that this could really work automatically. Compared to classical inference, which focuses on how to extract the information available in data, Bayesian methods seem to quickly move to elaborate computation. Much of the work in modern non-Bayesian statistics is focused on developing methods that give reasonable answers using minimal assumptions (Gelman, 2008). Bayesian probability measures the degree of belief that you have in a random event. By this definition, probability is highly subjective. It follows that all priors are subjective priors. A prior distribution of a parameter is the probability distribution that represents your uncertainty about the parameter before the current data are examined. Not everyone agrees with this notion of subjectivity when it comes to specifying prior distributions. There has long been a desire to obtain results that are objectively valid. See Berger (2006) and Goldstein (2006) for discussions about objective Bayesian versus subjective Bayesian analysis.

The Bayesian approach requires specifying a prior distribution for all unknown parameters. Multiplying the prior distribution and the likelihood function together leads to the posterior distribution of the parameter. One uses the posterior distribution to carry out all inferences. Thus one can not carry out any Bayesian inference or perform any modelling without using a prior distribution. When there is concrete prior knowledge about the parameters, it can be done, and should be done! But, in many cases, particularly in climate change research, prior knowledge is either vague, or non-existent, and that makes it very difficult to specify a unique prior distribution. Within the Bayesian paradigm, this can be achieved by using prior distributions that are "objective" (that means that they have a minimal impact on the posterior distribution). Such distributions are called objective or noninformative priors. However, in all studies presented here noninformative priors are used. This is the case in the analysis of the Bayesian model comparison or the estimation of the change point probability distributions. A noninformative prior is e.g. a flat prior, which is a prior distribution that assigns equal likelihood on all possible values of the parameter. More precisely we use a prior that is called Jeffreys' prior (Jeffreys, 1961). It satisfies the local uniformity property: a prior that does not change much over the region in which the likelihood is significant and does not assume large values outside that range. Jeffreys' prior is locally uniform and hence noninformative. Many statisticians favour noninformative priors because they appear to be more objective. However, it is unrealistic to expect that noninformative priors represent total ignorance about the parameter of interest. In some cases, noninformative priors can lead to improper posteriors (nonintegrable posterior density). One can not make inferences with improper posterior distributions. While Jeffreys' prior provides a general recipe for obtaining noninformative priors, it has some shortcomings: the prior is improper for many models and it can lead to an improper posterior in some cases; and the prior can be uncomfortable to use in high dimensions. For future work in climate detection in phenology I propose to verify for the Bayesian model comparison informative priors. An informative prior is a prior that is not dominated by the likelihood and that has an impact on the posterior distribution. If a prior distribution dominates the likelihood, it is clearly an informative prior. But these types of distributions must be specified with care in actual practice. On the other hand, the proper use of prior distributions illustrates the power of the Bayesian method: information gathered from the previous study, past experience or expert opinion can be combined with current information in a natural way.

Results of the present PhD provide specific information about change point probabilities in the field of phenology that could serve as informative priors in future studies. General results reveal that at least for phenological phases and for most of the temperature time series a clear nonlinear change is detected by the Bayesian model comparison. The nonlinear change is mirrored by high change point probabilities occurring predominantly in the last two decades (1980–2000) (see **Chapter 2** for spring phases, **Chapter 3**).

Thus in further Bayesian analysis in phenology studies an informative prior could include the knowledge of a clear preference for the change point model. But in single cases for distinct species it occurs that the change point model probability is almost equal to the linear model probability maybe due to a different type of plant reaction. Thus a predefined prior would be only appropriate when analysing several phenological phases, groups or indices. The prior information could even include the probable time span of maximal change point probabilities.

Often researchers argue that they could calculate the change point distributions or model averaged trends with conventional statistical methods as well. This opinion is only partly true. It would be incorrect to assume that the Bayesian framework is simply a probabilistic re-interpretation of classical methods, just because the maxima a posteriori and the penalised least-squares estimates are equivalent. The difference is that the Bayesian approach removes for example the change point variable to estimate the probability distribution or remove the model variable to estimate the model averaged trend by marginalization. The distinguishing element of Bayesian methods is marginalization, where we attempt to integrate out all 'nuisance' variables (Dose and Menzel, 2004). The change point variable is an example for such a 'nuisance' variable 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 marginalization rule overlays all possible triangular functions and then weighs them with their respective change point probability. By using the Bayesian marginalization rule the supporting functional values can be eliminated resulting in a probability, $p(E)$, for a particular change point choice.

Bayesian methods typically involve high-dimensional integrals. If the statistical problem involves 4 parameters (e.g. change point model), then the inference involves 4-dimensional integration. Bayesian methods allow fitting complex models without resorting to large sample approximation. Exact inferences (e.g. confidence intervals) that do not rely on large sample approximations are available by the Bayesian approach. Elimination of nuisance parameters is conceptually straightforward and is also practical due to advances in Bayesian computing. This convenience is a result of Bayesian analysis being a logically simple and practical approach.

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 by only one model is often insufficient (Dose and Menzel, 2004). The Bayesian model comparison analysis allows estimating a reliable model averaged rate of change. The proper functional behaviour and the proper trend are obtained by superposition of a constant, a linear function and the change point model function weighted with their respective model probabilities. 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. Thus Bayesian model averaged results are more informative than results based on single model approaches. On the other hand we have found that confidence intervals are often larger than single-model based confidence intervals; the latter tend to be too narrow (see **Chapter 2, 4 or 5**). Also Hoeting et al. (1999) stated that model averaging tends to produce higher estimates of variance than estimates that ignore model uncertainty. The high uncertainty intervals have often been criticized by reviewers because we had difficulties in presenting significant climate change patterns. But model averaging is more correct, because it takes account of a source of uncertainty that analyses based on model selection ignore. Standard analyses probably tend to find significant results too often. Also if results are significant under model averaging, then conclusions are more robust than those that depend upon the particular model that has been selected. Model averaging results are robust to model choice. Model averaging also allows users to incorporate several competing models in the estimation process; thus to a committee of scientists model averaging may offer a better estimation method than the traditional approach of trying to get the committee to agree on one best model (Hoeting et al., 1999).

When individual time series are compared it often happens that start or end year differ, as well as intermediate years are often missing. Linear trend analysis is very susceptible to the values observed in start and end years (see Figure 8 of Dose and Menzel, 2004). Changing start and end year by one year can already drastically change the resulting trend especially in highly variable phenological time series of a few decades. Thus results of time series with varying start and end dates are critical to be compared. Reporting observed trends in days per year rather than absolute change (Menzel et al., 2001) does not solve the problem of incomparability. Sparks (2000) and Chmielewski (2001) tried to solve this problem with conventional statistical methods by averaging several phenological time series over the same period. Also e.g. Menzel and Estrella, (2001), Sparks and Menzel (2002) and Scheifinger et al. (2002) addressed this problem and illustrated the variation of resulting changes with the altered periods of interest. Chmielewski (2001) showed that when using yearly averages the resulting trends can be biased by unequal distribution of the stations in time. Reviews of phenological trend studies suggest that only about 40% of the reported trends have proved to be statistically significant (Dose and Menzel, 2004). 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). Very few studies also report the standard error of the slope (e.g. Sagarin, 2001). Moving linear trend window analysis (e.g. Menzel et al., 2004; Rutishauser et al., 2007) was another attempt to overcome the shortages of a priori decisions of window length for linear trend analyses.

A conventional method that overcomes such limitations is analysing combined time series or a composed phenological index (e.g. Häkkinen et al., 1995; Rutishauser et al., 2007). Rutishauser et al. (2007) constructed a so called ‘‘spring plant’’ which is defined as the weighted mean of apple and cherry tree flowering and first leaves of beech. This method accounts for systematic differences when averaging several single phenological series into a regionally representative mean chronology. Subsequently, linear regression models were calibrated as transfer functions to estimate the ‘‘spring plant’’ from single phenological series. The availability of the historical phenological records varies from year to year (see Rutishauser et al., 2007 for details). Finally, Rutishauser et al. (2007) provided an annual estimate of the onset of spring including an uncertainty range at interannual timescales of ±10 days and of ±3.6 days at decadal timescales. Combining several fragmentary phenological time series for a certain region not only reduces the influence of individual time series and the noise in the time series but also gives the opportunity to apply proper outlier detection methods and normalize the combined time series to a common observation period. But if one is interested in regional differences of trends combined time series are inappropriate because important informations are merged together. Here the robust estimation of annual Bayesian rates of change, functional behaviours and change point probability distributions possess a great advantage compared to conventional statistical solutions. Particularly the annual resolution of trend and change point probabilities allows comparing single time series in detail. But due to the high confidence intervals of model averaged rates of change the separation of regional differences still remains difficult. In **Chapter 4** (see Figure 4.2 e) and f)) I show that outliers especially at the end of the time series are able to seriously influence Bayesian change point distributions and rates of change. Results of such time series are nearly not interpretable due to extraordinary high uncertainties. Therefore in **Chapter 4** we illustrate the difference of the result by removing such an extreme outlier from the time series.

Another critique is that reviewers see the application of our three selected models (constant, linear and change point model) limited especially when we analysed long-term data (over 100 years) or atmospheric radiosonde temperature data (**Chapter 4** and **5**). They argue that it is very unlikely that in these time series just one major change point occurred. Thus undetected change points may result from our limited model space. A multiple change point model is of course capable of modelling more detailed structure in a time series and therefore would mirror more adequately the observed changes of other studies. Each added change point adds two more variables to the likelihood, that may be punished by the Ockham`s razor (Garret, 1991). Following the Ockham`s razor, a model should be chosen as complex as required for an explanation of the data and as simple as necessary to avoid the fitting of noise. Whether the multiple change point model really provides a better data description has to be tested in future work. This includes the investigation of the performance of Bayesian model averaging when the true model is not in the model class. We conclude that the model averaging is still more reliable than results based just on a single model as e.g. the linear regression model. Moreover the shape of the change point probability distribution indicates whether we have more than one pronounced abrupt change in our time series. A clear narrow single peak indicates a very strong abrupt temperature change within a defined time period. A broad multi-modal change point distribution indicates that several major nonlinear changes occurred consecutively in a certain time span (see **Chapter 4** for winter, spring and autumn temperatures). Further more we are able to test the residual sum of squares for each model. In **Chapter 5** within stratosphere levels the residual sums of squares indicate that the three chosen models are suboptimal to describe adequately the temperature changes.

The major advantages of this Bayesian statistical approach is a provision of several useful outputs such as model probabilities, model comparisons, change point probabilities, model averaged rates of change including uncertainty intervals and functional behaviours.

For more in-depth treatments of the pros and cons of Bayesian analysis, see Robert (2001, Chapter 11). Robert (2001) presents a concluding chapter, a justification of the Bayesian approach that summarizes various arguments.

6.2 Second leading Research question

Which potentials of the Bayesian approach (such as model probabilities, functional behaviours, model averaged rates of change, confidence intervals and time spans of elevated change point probability) contribute to an accurate assessment of climate change impacts on natural systems?

In each study the Bayesian feature gives us important instruments for the accurate assessment of the climate change. The model probabilities and functional behaviours primarily provide information about the general change patterns (no change, linear change or nonlinear change) of phenological and temperature records. The Bayesian model comparison serves as an option to test and incorporate different model estimates.

Results of the Bayesian model comparison reveal a clear preference of the change point model in all chapters. **Chapter 2** demonstrates the use of the Bayesian approach to detect recent changes in phenological time series throughout the year. In this study the change point model delivers the best description of 2600 observational time series in Europe from all seasons (very early spring, early spring, mid summer and early autumn, late autumn) and for 90 phenological phases. The dominance of the change point model is most pronounced for phases of summer to late autumn.

In **Chapter 3** time series of the phenological phase bud burst in Norway spruce and mean monthly/weekly temperatures of corresponding climate stations in Germany are best described by the change point model. In this study not only phenological events exhibit nonlinear changes but also the mean temperature times series particularly of April and May.

Chapter 4 includes long-term spring and autumn phenological observations covering 250 years from Switzerland and Burgundy (eastern France) as well as long-term Swiss monthly and seasonal temperature measurements. Here the by far highest model probabilities (from 51% to 100%) are generally found for the change point model for both phenological and temperature time series. It is thus the most suitable model to describe the long-term evolution of phenology and temperature. The conventional linear model is left as an appropriate alternative to describe mean Swiss winter and autumn temperature (50% model probability for time series of these variables).

The reseach subject of **Chapter 5** considers atmospheric temperature change analyses. In **Chapter 5** the Bayesian procedure is firstly applied to a new homogenized Radiosonde Atmospheric Temperature Product for Assessing Climate (RATPAC-A) data set. Annual mean temperature anomalies (1979–2004) for 13 pressure levels from the surface to 30 hPa are examined. From the surface up to the tropopause (200–100 hPa) our results reveal that the change point model provided the best data fit. The predominant observations of nonlinear temperature changes in nearly all pressure levels from the surface up to the stratosphere underline the importance of alternatives to linear models. The near surface changes exhibit comparatively high change point probabilities around 1985 and 1995 while those at the tropopause level are highest between 1995 and 2000.

Supporting results concerning the supremacy of the change point model can be found in Menzel et al. (2006a) where phenological shifts of agricultural crops, fruit trees and wild plants are analysed. In phenological time series of 93 phases in Germany (1951–2004) the change point model reaches a mean probability of 0.79; for 34 time series 0.95 or higher, which is a satisfactory description of the data (Menzel et al., 2006a). This important result implies that the traditional method of change detection in phenological time series, the maximum likelihood linear regression approach would not be appropriate in the majority of cases (phenological time series). There even exist significant differences in mean change point probabilities between phenological groups ($p < 0.001$).

Similar results are detected in Menzel et al. (2007) where data from Germany (1951–2003) and Slovenia (1961–2004) are used to study if changes in the start, end and length of the growing season differ among four deciduous broad-leaved tree species. In Germany, Menzel et al. (2007) find a strong preference for the change point model to describe phenological spring (between 61% for birch and 98% for oak). In Slovenia oak displayed the highest preference for the change point model among all species. Therefore this examined time series have relevant nonlinear changes too. For horse chestnut and birch however, the linear model is almost as well as or even better than the one change point model.

In one example Menzel et al. (2008) display a 1900–2000 time series of observed flowering dates of redcurrant at Geisenheim, near Frankfurt am Main. In this study the change point model turns out to be the best in explaining the functional behaviour of the century-long record.

Uniform preferences for the change point model for phenology and temperature data can also be found by Schleip et al. (in review, see Appendix for Abstract). This study involves a unique Bayesian temperature sensitivity analysis of 43 different recorded phenological phases and 30 species covering more than 50 years within the time span 1951-2006. Results for the late spring, summer and early autumn temperature months exhibit a clear preference for the change point model (>50%) indicating a nonlinear change over the last 5 decades. April and August show exceptionally high nonlinearities with change point model probabilities of 78% and 81%. Phenophases occurring in summer and early autumn reveal the highest change point model probabilities up to 100% whereas early spring and late autumn phenophases possesses lower change point probabilities, but still over 50%.

In conclusion the model comparison feature provides the ability to evaluate and include outputs of competing statistical models. This helps to identify the prevailed change type (in our case nonlinear, linear or no change) of the time series.

Beside the estimation of general change patterns which are principally done by the model probabilities Bayesian features like the rates of change and change point probability distributions are able to illustrate temporal and spatial changes. One of the most powerful features for phenological time series is the possibility of providing annual estimates for change point probabilities, functional behaviours and rates of changes.

Thus the Bayesian approach allows an intensified comparison of regional changes in phenology. An example for comparisons of spatial and temporal changes of temperature and phenological time series is illustrated in **Chapter 3**. In this study comparable time records of a single species are analysed at 18 meteorological stations. 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. A further example for a comparison of temporal and spatial changes is shown in **Chapter 5** for atmospheric temperature changes in different pressure levels.

The reliability of the functional behaviour and of the rates of change is represented by confidence intervals, which are diagnostically useful. However in **Chapter 2** (Figure 2.7 or 2.8), **Chapter 4** (Figure 4.2) and **Chapter 5** (Figure 5.5) rather high uncertainty intervals appear predominantly at the end of the time series. Moving to the end of the time series, the confidence range widens since the model estimate is supported by less data. Therefore we have difficulties in presenting significant climate change patterns particularly for recent years. But if results exhibit small confidence intervals as it appears in each study for point estimates that are not influenced by marginal effects the results of model averaged rates of change are more robust than estimates of single models (see also Section 6.1).

Beside the confidence intervals the residual sum of squares of each model (constant, linear or change point model) shows if the respective rates of change are useful for an accurate assessment of the change. In **Chapter 5** the residual sums of squares indicate that in the stratosphere none of our three chosen models is complex enough to describe the changes adequately. Despite the occurrence of two volcanic eruptions El Chicón (1982) and Mt. Pinatubo (1991) the stratosphere (70-30 hPa) shows a preference for the linear model (60%). But still the model averaging proofs to be more reliable than results based on just a single model like the linear regression model.

The change point probability distributions are an appropriate tool to identify nonlinear changes. Change point probability distributions mostly exhibit a marked peak in the last two decades (1980–2000) with a maximum around 1986. This is quite commonly observed in phenological records (see **Chapter 2** for spring phases, **Chapter 3**). Our results support the findings of exceptional changes in phenological and temperature records in the late 1980s

(Chmielewski and Rötzer, 2002; Scheifinger et al., 2002; Dose and Menzel, 2004; Studer et al., 2005). Chmielewski and Rötzer (2002) primarily examined the rapid change of the mean air temperature since 1989 and thus the shift in the average beginning of the growing season in Europe. They used a correlation analysis and concluded that changes in air temperature correspond well with changes in the air circulation over Europe. Also Scheifinger et al. (2002) found that the trend, the standard deviation and the discontinuity of the phenological time series at the end of the 1980s can, to a great extent, be explained by the North Atlantic oscillation. Studer et al. (2005) developed a robust multispecies estimate for spring phenology for the period 1965–2002 by applying empirical orthogonal function analysis on a combination of 15 spring phases. This multispecies estimate proves to be an alternative approach to assess the pattern of spring appearance during the last 40 years. It reveals an earlier onset of spring in recent years, mainly since 1988. Also Dose & Menzel (2004) detect for three single species a maximum change point probability near 1985.

In most cases spring and summer phases of wild plants, fruit trees as well as spring events in agricultural plants revealed discontinuous changes at the beginning of the 1980s (**Chapter 2**, Menzel et al. 2006a; Schleip et al., in review, see Appendix for Abstract). This finding is in accordance with an observed accelerated temperature increase in spring and summer records in the mid 1980s (**Chapter 3 and 4**; Dose and Menzel, 2006). By contrast autumn events such as leaf colouring of wild plants as well as all other events in agriculture (spring sowing, yellow ripeness, harvest, autumn sowing and emergence, see Menzel et al., 2006a) experienced such discontinuous changes mainly inhomogeneously within a respective decade (see also **Chapter 2**, Figure 2.1).

Chapter 3 and 4 as well as Schleip et al. (in review) suggest even a synchronous development of nonlinear changes in temperature and phenological time series. It can be concluded that change point probability distributions derived from two data sets (joint change point distribution) are more informative (e.g. better localized in time than that obtained from a single series of data) (see **Chapter 3**). Further more the implemented simulated annealing method allows determining temperature weight coefficients that identify the month or even the week with nonlinear temperature changes that support phenological change points (see **Chapter 3 and 4**)

It has been emphasized that a key to a better understanding of changes in phenology and its causes is the availability of long-time series of phenological observations (Peñuelas and Filella, 2001). Schönwiese and Rapp (1997) point out that for climatological studies as far as the causal aspects are concerned it is reasonable to look at time series of about 100 years. For seasonal temperature data they showed that trends can become random when time series of less than 30 years are analysed. Menzel et al. (2001) and Gornik (1995) indicated that average trends stabilize for time periods over 20 years. Therefore in all presented studies time series of 30 years and more are applied. The general question was whether new aspects to the phenology of Central Europe evolve when we broaden our temporal window from a few decades to a whole century and more. Our study in **Chapter 4** confirms the possibility to derive meaningful results from long-term temperature and phenological data about the prevailed change type. In the last 250 years the end of the 20th century represents a period with unique major increases in temperatures of all seasons as derived from model averaged trends and change point probability distributions. The highest probability for a change-point in three long-term grape harvest phenology records from France and Switzerland are also found at the end of the 20th century. Rutishauser et al. (2009) put these detected changes in a historical perspective. Applying a Pettitt's test and a Bayesian model comparison approach Rutishauser et al. (2009) analyzed two historical plant phenological records from northern Switzerland and the UK from 1753 to 1958. Rutishauser et al. (2009) explicitly used the same Bayesian model comparison set up as in **Chapter 4** as the change-point distributions were also able to identify the existence of possible minor changes in long term time series. For a Swiss Plant Phenological Spring Index record a dominating change point was found in 1984 with 2%, but also potential change points between 1850 and 1950 with weak peaks around 1890 and 1935. Also the change point analysis of **Chapter 4** indicates broader, even multi-modal change point distributions especially for seasonal temperature records. But the lack of change point probability distributions that possess several separated high narrow peaks supports the hypothesis that recent changes in phenology even in long-term time series are exceptional.

The frequent detection of nonlinear changes underlines the importance of the Bayesian model comparison incorporating the change point model with all its additional features. In all chapters and publications the combination of each Bayesian feature allows for an accurate description of the prevailed change-type of the studied species or climate factor.

6.3 Third leading Research question

What kind of biological insights into the triggering climate change factor temperature and its influence on phenology can be gained by the Bayesian concept?

We use the Bayesian concept primarily to assess insights into to the triggering climate change factor temperature and the influence on phenological responses.
In question two I have mentioned the identification of nonlinear changes in plant phenology time series starting mostly uniformly in the 1980s. As we will see in the following paragraphs there are discernible differences in the temperature response among species and seasons. Particularly in **Chapter 2 and 4** we conduct studies of multiple species demonstrating differences among species in the extent of the phenological change. Results of **Chapter 2 and 4** reveal an earlier start of spring in the last five decades and more heterogeneous changes in autumn. In a subanalysis of **Chapter 2** we investigate two species and two phases (*Aesculus hippocastanum* leaf unfolding/ autumn colouring and *Fagus sylvatica* leaf unfolding/ autumn colouring) in detail, with complete records (1959–1999) at eleven stations in Switzerland. In this subanalysis we detect a nonlinear trend evolution over time. Model averaged rates of change in 1999 are - 0.14 and - 0.34 days year^{-1} for leaf unfolding of *Aesculus* and *Fagus*, whereas the trends for autumn colouring indicate even advanced onsets in recent years.
In **Chapter 4** unique long-term temperature and phenological data series for central Europe back to 1753 are analysed with a Bayesian and conventional linear approach. Bayesian results show that the model averaged rates of change of the phenological phases exhibit a considerable advance of the onset of spring and harvest dates. Additionally, the summer temperature time series show an abrupt temperature increase at the end of the 20th century. Coherence factors and temperature weights estimated by a simulated annealing approach indicate that spring phenological variability is not only influenced by forcing temperatures of the current year but also by temperatures of the preceding June and October.

Menzel et al. (2006b) mentioned that the difference in degree of change between spring and fall phenophases could reflect either less climatic change during summer or fall, or less sensitivity of the phenophases to the changes that are occurring. The former explanation partly seems to explain the results from Estonia, where fall temperatures have been changing less than those in spring (Ahas and Aasa, 2006). In contrary to this first explanation the results of **Chapter 2** reveal increasing change point model probabilities and maximum rates of change for summer and fall months indicating a strong climate change. Also in **Chapter 4** the most abrupt change that is significantly different from zero occurs in the summer season (see Figure 4.3c). In line with the second explanation are observations by Estrella and Menzel (2006) that leaf colour changes in four deciduous trees in Germany are only weakly correlated with summer or fall temperatures.

Menzel et al. (2006a) concluded that: delayed leaf colouring in autumn is clearly different from other events in autumn, such as fruit ripening or harvest and sowing of winter crops; and there is apparent evidence that recent warming had clearly advanced a significant part of the agricultural calendar. Menzel et al. (2006a) observed a clear signal of advancing spring, summer and autumn phenology across all plant types studied. Furthermore the time series analyses by Bayesian nonparametric function estimation reveal a clearly different functional behaviour: (1) between wild plants/fruit trees and annual crops and (2), within crop production, among farmers' activities, such as seed drilling, tilling, harvesting (false phases), and (3) climate triggered true phases (partly emergence but especially shooting, flowering and ripening). Growth stages of wild plants and fruit trees like perennial crops have strongly advanced since the beginning of the 1980s and began, together with true agricultural phases, 1 to 2 week earlier in the 2000s. False agricultural phases seem to respond to a lesser degree to climate, but reflect plant-production and crop-management practices.

Menzel (2000) presents in an analysis of 751 springtime phenological series recorded in Europe's International Phenological Gardens that 22% significantly advance and 5% are significantly delayed. Of 100 species examined from 1970–1999 in Washington, D.C., flowering advances in 89 (significantly in 76, the greatest advance being 46 days) and retreat in 11, the greatest delay being 10 days (Abu-Asab et al., 2001). Among 385 species observed in central England, 16% show a significant advance in flowering time and 3% show a significant delay (Fitter and Fitter, 2002). Of 60 regressions of phenophase timing on year in central Europe, 11 show significant advances and 2 show significant delays (Rötzer et al., 2000). In the comprehensive European study by Menzel et al. (2006b), 31% of leafing and flowering stages show significant advances and 3% show significant delays during the period 1971-2000.

Menzel et al. (2007) detects for leaf unfolding of four species an onset of on average 8 days earlier, leaf colouring 10 days later and the growing season 18 days longer in Slovenia than in Germany. However, these differences in phenological means across countries are not stable in time. For birch, beech and oak, Menzel et al. (2007) found a stronger advance of leaf unfolding in Germany than in Slovenia in recent decades. In spring, oak changes most, followed by horse chestnut, beech and birch. Changes in leaf colouring are most similar for all countries and species, except for horse chestnut. Consequently, the lengthening of the growing season is most marked for oak, followed by birch, beech and finally horse chestnut, and this ranking is the same in both countries. Furthermore this study clearly detects an influence of the damage caused by the horse chestnut leaf-miner on leaf colouring dates in Slovenia and Germany. Results of Menzel et al. (2007) demonstrate that first reported occurrence dates of the pest are generally associated with years of maximum change point probability of the Bayesian analysis. In this study the Bayesian analysis also allows interpreting major differences between leaf unfolding of birch/horse chestnut and beech/oak in Germany. Depending on the mean onset dates, time series of these observations reflect temperature changes of different periods in spring. Birch and horse chestnut leaf unfolding are mainly triggered by March temperatures that gradually increase during recent decades, leading to a smoother advancing of leaf unfolding of those species. Beech and oak leaf unfolding occur on average at the beginning of May but are predominantly triggered by April temperatures. April monthly mean temperatures exhibit a cooling from 1951 till the end of the 1970s and a strong warming from then onwards. The change point model probability helps to underline these biological insights. The change point model is highly preferred in the case of beech and oak (96% and 98% model probability). The highest change point probabilities can always be identified in 1978/79; however higher values are found for beech and oak (maximum of 8.1% and 10.1%, respectively) than for birch and horse chestnut (3.8% and 3.5%).

The timing of phenological phases is an important factor in analysing changes in net production of trees in response to interannual variation and long-term changes in climate (Kramer et al.; 1996, 2000; Chuine and Beaubien, 2001). Kramer et al. (1996) showed that a difference of one week in bud burst corresponds to a difference of about 5% in the radiation absorbed during the growing season. Tree species react different to climate change not only with respect to changed growth patterns but also frost damage (Hänninen, 1991; Murray et al., 1994; Kramer et al., 1996) resulting in altered competitive relationships between tree species (Kramer 1995). Models for the phenological development of trees are developed to assess the impact of climate change on forest ecosystems (Cannell and Smith, 1983; Hänninen, 1994, Chuine, 2000). In **Chapter** 3 we suggest the use of the Bayesian method for phenology models to incorporate specific forcing temperature patterns for each phase. The influence of forcing temperatures may be defined by daily temperature sums but also by weighted effective temperatures in a sensitive time span. **Chapter 3** illustrates that Norway spruce bud burst and temperature time series both reveal nonlinear changes at the beginning of the 1980s. In this study with a simulated annealing approach we correlate the change point distributions of monthly and even weekly temperature means with change point distributions of the phenological event. Change point distributions of Norway spruce bud burst exhibit the highest Bayesian correlations with temperatures at the end of February and in April and May. Since the beginning of the 1980s, April and May temperature rates of change of all 18 investigated stations increase to positive values (warming) and Norway spruce bud burst time series reveal an enhanced advancing of the phenological phase.

6.4 Summary and Conclusion

Methodical advantages and disadvantages of the Bayesian analysis

1. The distinguishing element of Bayesian methods is the marginalization, where we attempt to integrate out all ‘nuisance’ variables. The Bayesian approach removes for example the change point variable to estimate the probability distribution or the model variable to estimate the model averaged trend by marginalization.
2. Bayesian analysis considers the inability to prefer one model against another that enforces the collection of new data. Model averaging allows users to incorporate several competing models in the estimation process. Compared to the commonly used linear regression approach, we are able to provide model averaged rates of change at annual resolution. This helps to describe discontinuities and to quantify the direction and speed of the changes. Thus Bayesian model averaged results are more informative than results based on single model approaches.
3. Model averaged confidence intervals are often larger than single-model based confidence intervals; the latter tend to be too narrow. If results are significant under model averaging, then conclusions are more robust than those that depend upon the particular model that has been selected.
4. Due to the high confidence intervals of model averaged rates of change the separation of regional differences remains difficult. In **Chapter 4** (see Figure 4.2 e) and f)) outliers especially at the end of the time series seriously influence Bayesian change point distributions and rates of change.
5. The application of our three selected models (constant, linear and change point model) is limited especially when we analyse long-term data (over 100 years) or atmospheric radiosonde temperature data (**Chapter 4 and 5**). But the shape of the one-change point probability distribution indicates whether we have more than one pronounced abrupt change in our time series. A broad multi-modal change point distribution indicates that several major nonlinear changes occurred consecutively in a certain time span (see **Chapter 6** for winter, spring and autumn temperatures).

In future Bayesian studies of phenology an informative prior could be tested instead of a noninformative prior. An informative prior could include the knowledge of a clear preference for the change point model or even the probable time span of maxima change point probabilities. A predefined prior would only be appropriate when analysing several phenological phases or phenological indices.

Another future task could be the examination whether a multiple change point model provides a better data description for long-term time series. This includes the investigation of the performance of Bayesian model averaging when the true model is not in the model class.

Results derived from the Bayesian features:

1. **Bayesian Model comparison**

Results of all chapters reveal a clear preference for the change point model by the Bayesian model comparison. The by far highest model probabilities (from 51% to 100%) are generally found for the change point model when describing phenological and temperature data. The dominance of the change point model was most pronounced for phases in summer to late autumn (**Chapter 2**). In **Chapter 3** the Bayesian model comparison reveals that at 17 phenological stations the time series of Norway spruce are described best with a change point model (87% average model probability). The average change point model probability over all stations of the monthly temperature time series increases from 39% in January to a maximum probability of 61% in May.

In **Chapter 4** the linear model is an appropriate alternative to describe mean Swiss winter and autumn temperature (50% model probability for time series of these variables). In **Chapter 5** the predominant observations of nonlinear temperature changes in nearly all pressure levels from the surface up to the stratosphere underline the importance of alternatives to linear models. The residual sums of squares indicate that for stratosphere levels our three chosen models are suboptimal to describe adequately complex changes in the stratosphere. In the stratosphere none of our three chosen models is complex enough to describe adequately the changes.

2. Functional behaviours

In most cases spring and summer phases of wild plants as well as spring events in agricultural plants reveal discontinuous changes at the beginning of the 1980s (**Chapter 2**). This finding is in accordance with an observed accelerated temperature increase in spring and summer records in the mid 1980s (**Chapter 3 and 4**). By contrast autumn events such as leaf colouring of wild plants as well as all other events in agriculture (harvest) experience such discontinuous changes mainly inhomogeneously within a certain decade (**Chapter 2**, **Chapter 4**).

For a better understanding of annual rates of change we display in **Chapter 5** in additon to the model averaged rate of change the model averaged function estimation of the mean temperature anomalies.

3. Change point probability distributions

The change point probability distributions are another appropriate tool to identify accurate nonlinear changes. Change point probability distributions mostly exhibit a marked peak in the last two decades (1980–2000) with a maximum around 1986. In **Chapter 2** spring events at the eleven Swiss observation stations present a rather coherent picture of their change point probabilities; the majority of peaks is grouped somewhere in the mid 1980s. However, in autumn the dispersed pattern of peaks illustrates the multiplicity of possible impacts on autumn phases. In **Chapter 3** change point distributions of Norway spruce bud burst exhibit the highest Bayesian correlations with temperatures of the end of February, and of April and May. Since the beginning of the 1980s, April and May temperature rates of change of all 18 investigated stations increase to positive values (warming) and Norway spruce bud burst time series reveal an enhanced advancing of the phenological phase.

Chapter 3 and 4 suggest a synchronous development of nonlinear changes in temperature and phenological time series. It can be concluded that change point probability distributions derived from two data sets (joint change point distribution) are more informative (e.g. better localized in time than those obtained from a single series of data) (see **Chapter 3**). The implemented simulated annealing method allows determining temperature weight coefficients that identify the month or even the week with nonlinear temperature changes that support phenological change points (see **Chapter 3 and 4**). In **Chapter 5** the near surface temperature changes exhibit comparatively high change point probabilities around 1985 and 1995 while those at the tropopause level are highest between 1995 and 2000. In Chapter 5 we detect an accumulation of change point probability peaks at surface and troposphere pressure levels in 1984 and in the beginning of the 1990s right after the volcanic eruptions. The change point probability distributions for the 70 to 30 hPa pressure levels stay below 5% and do not show potential high change point probabilities in the years during the two major volcanic eruptions. Undetected change points may result from our limited model space.

4. Model averaged rates of change

Analysing the change patterns of phenology over the last century in Europe reveals that the strong advancement of spring phases that has been observed in former studies is not a singular event in the course of the 20th century. Distinct periods of varying trend behaviour for spring and particularly for autumn phases can be distinguished.

Bayesian results reveal that the model averaged rates of change of the phenological phases show a considerable advance of the onset of spring (**Chapter 2, 3 and 4**) and harvest dates (**Chapter 4**). In **Chapter 3** 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 rates of change of 18 climate stations range from 0.03 to 0.17 °C year^{-1} whereas in 1951 May temperature rates of change range from -0.03 to 0.03 °C year^{-1}. In contrast, the rates of change of bud burst of Norway spruce in 2003 is estimated from -0.25 to -0.75 days year^{-1}, but show a delay in 1951. Over most of the investigated period, there is essentially a zero rate of change; but from the 1980s onwards the rate of change is negative for Norway spruce bud burst.

Additionally, the summer temperature time series show an abrupt temperature increase at the end of the 20th century (**Chapter 4**). For long term time series (> 250 years) we have difficulties in presenting significant climate change patterns particularly for recent years. On the other hand if results exhibit small confidence intervals as it appears in each study for point estimates, that are not influenced by marginal effects, the results of model averaged rates of change are more robust than estimates of single models (see also Section 6.1).

Gained biological insights of the triggering climate factor temperature and its impacts on phenology

Investigating the causes of the observed phenology patterns, the dominating influence of temperature on spring and its diverse effects on autumn phenology are confirmed in this PhD (particularly in **Chapter 3 and 4**). The advantage of the Bayesian probability method with all its features presented above allows an accurate analysis of the relationship between phenology and temperature observations. In nearly all cases investigated here, the results clearly suggest a coherent development of temperature and phenological time series. Therefore 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 (**Chapter 3 and 4**). Weight coefficients are obtained from an optimization of the coherence factor by simulated annealing.

In **Chapter 3** the temperature weights reveal that at nearly all investigated stations the change point distributions of April and May temperatures correlate with the change point distributions of bud burst of Norway spruce. The weekly resolution shows similar results. We suggest for phenology models to incorporate specific forcing temperature 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.
In **Chapter 4** for ”Spring plant“ phenology, temperatures during the spring phase as well as temperatures during the summer season of the previous year appear particularly important for the onset of the phase. With linear approaches this result is statistically much less evident and has never been discussed. The summer phases of Swiss and Burgundy grape harvest are mainly influenced by the season’s spring and early summer temperatures and less by temperatures in autumn. Swiss grape harvest reveals the increasing influence of temperatures from February to June, as shown by temperature weights and linear correlation.

In conclusion the separations of temporal and spatial changes in phenology and temperature time series as well as the ability to judge and incorporate outputs of competing statistical models are an important helpful contribution in the field of climate change detection. Beside this PhD work the reader is also referred to the soon published Book chapter: Bayesian methods in phenology (Schleip et al., forthcoming). Finally I conclude that the Bayesian approach is a very practical and appropriate tool to detect climate change impacts in natural systems. Its strengths enable us to reanalyse nearly each study that was conducted with conventional method. Especially the analysis of atmospheric changes from the surface up to the top of the stratosphere shows the transferability of our method to other research fields beyond phenology. Probabilistic depictions, particularly at the regional level, are new to climate change science. As we have highlighted, the awareness of this issue is growing in the field of climate research, however, it is yet to be thoroughly addressed.

6.5 References

Abu-Asab M.S., Peterson P.M., Shelter S.G., Orli S.S., (2001) Earlier plant flowering in spring as a response to global warming in the Washington, DC, area. *Biodivers. Conserv.* **10**, 597-612.

Ahas R., Aasa A., (2006) The effect of climate change on the phenology of selected Estonian plant, bird and fish populations. *Int. J. Biometeorol.* **51**, 17-26.

Berger J.O., (2006) “The Case for Objective Bayesian Analysis,” *Bayesian Analysis* **3**, 385–402, http://ba.stat.cmu.edu/journal/2006/vol01/issue03/berger.pdf.

Bernardo J.M., Smith A.F.M., (1994) Bayesian Theory, New York: John Wiley and Sons.

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

Chmielewski F.M., Rötzer T., (2001) Response to tree phenology to climate change across Europe. *Agricultural and Forest Meteorology* **108**, 101-112.

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

Chuine I., Beaubien E.G., (2001) Phenology is a major determinant of tree species range. *Ecology Letters* **4**, 500-510.

Chuine I., (2000) A unified model for bud burst of trees. *J. Theor. Biol.* 207, 3, 337-347.

Defila C., Clot B., (2001) Phytophenological trends in the Switzerland, 1951–2002. *Int. J. Biometeorol.* **45**, 203–207.

Dose V., Menzel A., (2004) Bayesian analysis of climate change impacts in phenology. *Global Change Biol.* **10**, 259–272.

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

Estrella N., Menzel A., (2006) Responses of leaf colouring in four deciduous tree species to climate and weather in Germany. *Clim. Res.* **32**, 253-267.

Fitter A.H., Fitter R.S.R., (2002) Rapid changes in flowering time in British plants. *Science* **296,** 1689-1691.

Garrett A.J.M., (1991) Ockham’s Razor. In: Grandy, W.T., Schick, L.H. (Eds.), Maximum Entropy and Bayesian Methods.Kluwer, Dordrecht, pp. 357–364.

Gelman A., (2008) Why I don't like Bayesian statistics. Statistical Modeling, Causal Inference, and Social Science blog, 1 April 2008.

Goldstein M., (2006) "Subjective Bayesian Analysis: Principles and Practice". *Bayesian Analysis* **3**, 403–420.

Gornik W., (1995) Further investigations into the problems of finding a mean value in phenological data series. *Arboreta Phaenologica* **40**, 26-27.

Häninnen H., (1991) Does climate warming increase the risk of frost damage in northern trees? *Plant, Cell and Enviroment* **14**, 449-454.

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

Hoeting J.A., Madigan D.M., Raftery A.E., Volinsky C.T., (1999) Bayesian model averaging: A tutorial (with discussion). *Statistical Science* **14**, 382–401. [A corrected version is available online at www.stat.washington.edu/www/research/online/ hoeting1999.pdf.]

Inouye D.W., McGuiRE A.D., (1991) Effects of snowpack on timing and abundance of flowering in Delphinium nelsonii (Ranunculaceae): implications for climate change. *Am. J. Bot.* **78**, 9971001.

Jeffreys H., (1961) Theory of Probability. third Edition, Oxford: Oxford University Press.

Kramer K., (1995) Modelling comparison to evaluate the importance of phenology for the effects of climate change in groth of temperature-zone deciduous trees. *Clim. Res.* **4**, 119-130.

Kramer K., Friend A., Leinonen I., (1996) Modelling comparison to evaluate the importance of phenology and spring frost damage for the effects of climate change on growth of mixed temperate-zone deciduous forests. *Clim. Res.* **7**, 31-41.

Kramer K., Leinonen I., Loustau D., (2000) The importance of phenology for the evaluation of impact of climate change on growth of boreal, temperate and Mediterranean forests ecosystems: an overview. *Int. J. Biometeorol.* **44**, 2, 67-75-

Kudo G., Nishikawa Y., Kasagi T., Kosuge S., (2004) Does seed production of spring ephemerals decrease when spring comes early? *Ecol. Res.* **19**, 255-259.

Little R.J., (2006) Calibrated Bayes: a Bayes/frequentist roadmap. *American Statistician* **60**, 213–223.

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

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

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., (2004) Pflanzen in einer wärmeren Welt, *Münchner Rück* **144**, 144– 155.

Menzel A., von Vopelius, J., Estrella N., Schleip C., Dose V., (2006a) Farmers’ annual activities are not tracking the speed of climate change. *Clim. Res.* **32**, 3, 201-207.

Menzel A., Sparks T., Estrella N., Koch E., Aasa A., Ahas R., Alm-Kubler K., Bissolli P., et al., (2006b) European phenological response to climate change matches the warming pattern. *Global Change Biol.* **12** (10), 1969–1976.

Menzel A., Estrella N., Heitland H., Susnik A., Schleip C., Dose V., (2007) Bayesian analysis of the species-specific lengthening of the growing season in two European countries and the influence of an insect pest. *Int J Biometeorol.* **52**, 3*,* 209-218.

Menzel A., Estrella N., Schleip C., (2008) Impacts of Climate Variability, Trends and NAO on 20th Century European Plant Phenology. In: Climate variability and extremes during the past 100 years. Series: *Advances in Global Change Research* , Vol. 33 Brönnimann, S.; Luterbacher, J.; Ewen, T.; Diaz, H.F.; Stolarski, R.S.; Neu, U. (Eds.), XVI, 364 p. 146 illus., 50 in color., Hardcover ISBN: 978-1-4020-6765-5. Springer 221-233.

Miller-Rushing A., J., Primack. R., (2008) Global warming and flowering times in Thoreau's Concord; A community perspective. *Ecology* **89**, 332-341.

Murray M.B, Smith I.D., Leith D., Fowler D., Lee H.S.J., Friend D., Jarvis P.G., (1994) Effects of elevated CO_2, nutrition and climatic warming on bud phenology in sitka spruce (*Picea sitchensis*) and their impact on the risk of frost damage. *Tree Physiology* **14**, 691-706.

Peñuelas J., Filella I., (2001) Phenology- Responses to a warming world. *Science* **294**, 793-794.

Post E.S., Pedersen C., Wilmers C.C., Forchhammer M.C., (2008) Phenological sequences reveal aggregate life history response to climatic warming. *Ecology* **89**, 363–370.

Robert C.P., (2001) The Bayesian Choice, Second Edition, New York: Springer-Verlag

Rosenzweig C., Casassa G., Karoly D.J., Imeson A., Liu C., Menzel A., Rawlins S., Root T.L., Seguin B., Tryjanowski P., (2007) Assessment of observed changes and responses in natural and managed systems. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental

Panel on Climate Change, Parry M.L., Canziani O.F., Palutikof J.P., van der Linden P.J. and Hanson C.E., Eds., Cambridge University Press, Cambridge, UK, 79-131.

Rutishauser T., Schleip C., Sparks T., Nordli Ø., Menzel A., Wanner H., Jeanneret F., Luterbacher J., (2009) The Temperature Sensitivity of Swiss and British Plant Phenology 1753–1958. *Clim. Res.* Special Issue, COST-725, accepted.

Rutishauser T., Luterbacher J., Jeanneret F., Pfister C., Wanner H., (2007) A phenology-based reconstruction of interannual changes in past spring seasons. *J. Geophys. Res.* **112**, G04016.

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

Scheifinger H., Menzel A., Koch E., Peter C., Ahas R., (2002) Atmospheric mechanisms governing the spatial and temporal variability of phenological observations in central Europe. *Int. J. Clim.* **22**, 1739–1755.

Schleip C., Menzel A., Dose V., (forthcoming) Bayesian methods in phenology in Hudson IL & Keatley M.R., (Eds) Phenological Research: Methods for environmental and climate change analysis. Springer. Berlin, Heidelberg, New York.

Schleip C., Rais A., Menzel A., Bayesian analysis of temperature memory effects of plant phenology in Germany. *Agricultural and Forest Meteorology,* in review.

Schönwiese C.D., Rapp J., (1990) Climate Trend Atlas of Europe: Based on obervations 1991-1990. Kluwer Academic Publishers, Dortrecht. 228p.

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 UK. *International Journal of Biometeorol.* **44**, 88-87.

Sparks T.H., Menzel A., (2002) Observed changes in seasons: an overview. *J. Clim.* **22**, 1715–1725.

Studer S., Appenzeller C., Defila C., (2005) Interannual variability and decadal trends in Alpine spring phenology. A multivariate analysis approach. *Clim. Chan.* **73**, 395-583 414.

7 Peer-reviewed Scientific Papers and books

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**, 2-4, 179-191, doi: 10.1016/j.agrformet.2006.09.013.

Menzel A., von Vopelius J., Estrella N., **Schleip C**., Dose V., (2006) Farmers' annual activities are not tracking the speed of climate change.*Climate Research,* **32**, 3, 201-207, doi: 10.3354/cr032201.

Menzel A., Estrella N., Heitland H., Susnik A., **Schleip C.,** Dose V., (2007) Bayesian analysis of the species-specific lengthening of the growing season in two European countries and the influence of an insect pest. *Int J Biometeorol.* **52**, 3*,* 209-218. *doi: 10.1007/s00484-007-0113-8.*

Schleip C., Menzel A., Dose V., (2008) Bud burst in Norway spruce (*Picea abies*): Bayesian analysis of the relationship between temperature and bud burst. *Agricultural and Forest Meteorology,***148**, 4, 631-643. doi:10.1016/j.agrformet.2007.11.008.

Schleip C., Rutishauser T., Luterbacher J., Menzel A., (2008) Time series modelling and temperature impact assessment of phenological records in the last 250 years, J. Geophys. Res., **113**, G04026, doi: 10.1029/2007JG000646.

Rutishauser T., **Schleip C.,** Sparks T., Nordli Ø., Menzel A., Wanner H., Jeanneret F., Luterbacher J., (2008) The Temperature Sensitivity of Swiss and British Plant Phenology 1753–1958, *Climate Research* Special Issue, COST-725, accepted.

Schleip C., Rais A., Menzel A.: Bayesian analysis of temperature memory effects of plant phenology in Germany, *Agricultural and Forest Meteorology,* in review.

Schleip C., Menzel A., Dose V., (2008). Bayesian analysis of changes in Radiosonde Atmospheric Temperature, *J. Climatology*, doi: 10.1002/joc.1790, article online in advance of print.

Menzel A., Estrella N., **Schleip C**., (2008). Impacts of Climate Variability, Trends and NAO on 20th Century European Plant Phenology. In: Climate variability and extremes during the past 100 years. Series: Advances in Global Change Research , Vol. 33 Brönnimann, S.; Luterbacher, J.; Ewen, T.; Diaz, H.F.; Stolarski, R.S.; Neu, U. (Eds.), XVI, 364 p. 146 illus., 50 in color., Hardcover ISBN: 978-1-4020-6765-5. Springer, 221-233.

Schleip C., Menzel A., Dose V., (forthcoming) Bayesian methods in phenology in Hudson IL & Keatley M.R., (Eds) Phenological Research: Methods for environmental and climate change analysis. Springer. Berlin, Heidelberg, New York.

8 Candidate's individual contribution

First Publication Chapter 2:

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**, 2-4, 179-191, doi:10.1016/j.agrformet.2006.09.013.

The candidate developed independently the research question and story of this paper. He conducted the Bayesian analysis, interpreted and discussed the results. He composed the graphs and the first draft. As a corresponding author he handled the review and revision process. Overall, the candidate contributed about **90%** to the completion of the entire article.

Second Publication Chapter 3:

Schleip C., Menzel A., Dose V., (2008). Bud burst in Norway spruce (*Picea abies*): Bayesian analysis of the relationship between temperature and bud burst. *Agricultural and Forest Meteorology,***148**, 4, 631-643. doi:10.1016/j.agrformet.2007.11.008.

Similar as in Chapter 2: Overall, the candidate contributed about **90%** to the completion of the entire article.

Third Publication Chapter 4:

Schleip, C., Rutishauser, T., Luterbacher, J., Menzel, A., (2008). Time series modelling and temperature impact assessment of phenological records in the last 250 years, J. Geophys. Res., **113**, G04026, doi:10.1029/2007JG000646.

Similar as in Chapter 2 with the only differene that the candidate developed the research question and story of this paper in cooperation with Dr. This Rutishauser in an external visit (STSM of the COST-Project 705). Overall, the candidate contributed about **80%** to the completion of the entire article.

Fourth Publication Chapter 5:

Schleip C., Menzel A., Dose V., (2008). Bayesian analysis of changes in Radiosonde Atmospheric Temperature, *J. Climatology*, conditionally accepted.doi: 10.1002/joc.1790, article online in advance of print.

Similar as in Chapter 2: Overall, the candidate contributed about **90%** to the completion of the entire article.

Acknowledgements

I would like to acknowledge many people for helping me during my doctoral work. I would especially like to thank my advisor, Prof. Dr. Annette Menzel, for her generous time and commitment. Throughout my doctoral work she encouraged me to develop independent thinking and research skills. She continually stimulated my analytical thinking and greatly assisted me with scientific writing.

I am extremely grateful for the assistance, generosity, and methodical advice I received from Prof. Dr. Volker Dose at the Max-Planck-Institut für Plasmaphysik, Garching München. I feel blessed to have worked with both. Without their belief in me I could not have developed confidence in my abilities as a Bayesian analyst and for that I am truly grateful. I have learned a great deal from them and I will never forget the valuable lessons they taught me.

I am also very grateful to Prof. Dr. Peter Fabian for supporting me to get the scholarship. I appreciate the financial support from the Universität Bayern e. V. that funded the research. I extend many thanks to my colleagues especially to Dr. Nicole Estrella who helped me with her special skills to generate and format data sets, Dr. Christian Heerdt, Dr. Michael Leuchner, Dipl.-Ing. Andreas Rais, Brigitte F. Fleischner, Nikolaus Hofmann and Anton Knötig for their various forms of support during my graduate study.

I would like to acknowledge the group at the Climatology and Meteorology Research Group, University of Bern who were very cooperative and friendly during my visit in Bern. I especially like to thank Dr. This Rutishauser and PD Dr. Jürg Luterbacher for the interesting exchange and discussions.

This dissertation could not have been written without the support and friendship found at Freising and elsewhere. The love of family and friends provided my inspiration and was my driving force. It has been a long journey and completing this work is definitely a high point in my academic career. I'm especially grateful to my wife and my son Benjamin for helping me to keep my life in proper perspective and balance.

Appendix

Candidate's Papers and Book Chapters

Abstracts of Peer-reviewed Scientific Papers or Book Chapters, published

Menzel A., von Vopelius J., Estrella N., ***Schleip*** *C., Dose V., (2006) Farmers' annual activities are not tracking the speed of climate change. Climate Research, 32, 3, 201-207, doi: 10.3354/cr032201.*

Abstract

Global climate change impacts are already tracked in many physical and biological systems and they reveal a consistent picture of changes, e.g. an earlier onset of spring events in mid and higher latitudes and a lengthening of the plant growing season. However, available results are mainly based on the study of wild plants, whereas only a few studies have hinted at an earlier spring onset for agricultural plants. So far, no comprehensive study has compared phenological shifts between agricultural crops, fruit trees and wild plants. We analysed phenological time series of 93 phases in Germany (1951–2004) employing Bayesian nonparametric function estimation, and found that events related to the production of annual crops clearly differ from spring and summer events in wild plants and fruit trees. While non-farmer driven agricultural events and spring and summer growth stages of wild plants and fruit trees advanced (i.e. occurred earlier) by 4.4 to 7.1 d decade^{-1}, farming indicators, such as sowing and subsequent emergence of spring and winter crops, as well as harvesting, advanced by only 2.1 d decade^{-1}. The estimated functional behaviour and emergence of discontinuous changes are clearly different between the 2 groups. We conclude that phenological responses to temperature changes are only reflected in data of wild plants, fruit trees and those spring growth stages of winter crops and later growth stages of spring crops which are exclusively triggered by climate, while other changes due to agricultural production are subject to management practice alterations.

Menzel A., Estrella N., Heitland H., Susnik A., ***Schleip*** *C., Dose V., (2007) Bayesian analysis of the species-specific lengthening of the growing season in two European countries and the influence of an insect pest. Int J Biometeorol. 52, 3, 209-218. doi: 10.1007/s00484-007-0113-8.*

Abstract

A recent lengthening of the growing season in mid and higher latitudes of the northern hemisphere is reported as a clear indicator for climate change impacts. Using data from Germany (1951–2003) and Slovenia (1961–2004), we study whether changes in the start, end, and length of the growing season differ among four deciduous broad-leaved tree species and countries, how the changes are related to temperature changes, and what might be the confounding effects of an insect attack. The functional behaviour of the phenological and climatological time series and their trends are not analysed by linear regression, but by a new Bayesian approach taking into account different models for the functional description (one change point, linear, constant models). We find advanced leaf unfolding in both countries with the same species order (oak > horse chestnut, beech, and birch). However, this advance is nonlinear over time and more apparent in Germany with clear change points in the late 1970s, followed by marked advances (on average 3.67 days decade^{-1} in the 2000s). In Slovenia, we find a more gradual advance of onset dates (on average 0.8 days decade^{-1} in the 2000s). Leaf colouring of birch, beech, and oak has been slightly delayed in the last 3 decades, especially in Germany, however with no clear functional behaviour. Abrupt changes in leaf colouring dates of horse chestnut with recent advancing onset dates can be linked across countries to damage by a newly emerging pest, the horse chestnut leaf-miner (*Cameraria ohridella*). The lengthening of the growing season, more distinct in Germany than in Slovenia (on average 4.2 and 1.0 days decade^{-1} in the 2000s, respectively), exhibits the same species order in both countries (oak > birch > beech). Damage by horse chestnut leaf-miner leads to reduced lengthening (Germany) and drastic shortening (Slovenia) of the horse chestnut growing season (-12 days decade^{-1} in the 2000s). Advanced spring leaf unfolding and lengthening of the growing season of oak, beech and birch are highly significantly related to increasing March temperatures in both countries. Only beech and oak leaf unfolding in Germany, which is generally observed later in the year than that of the other two species, is more closely correlated with April temperatures, which comparably exhibit marked change points at the end of the 1970s.

Menzel A., Estrella N., ***Schleip*** *C., (2008) Impacts of Climate Variability, Trends and NAO on 20th Century European Plant Phenology. In: Climate variability and extremes during the past 100 years. Series: Advances in Global Change Research , Vol. 33 Brönnimann, S.; Luterbacher, J.; Ewen, T.; Diaz, H.F.; Stolarski, R.S.; Neu, U. (Eds.), XVI, 364 p. 146 illus., 50 in color., Hardcover ISBN: 978-1-4020-6765-5. Springer, 221-233.*

Abstract

We provide here a brief overview of the impacts of climate variability and recent climate change on the European plant phenology across the 20th century. Facing recent climate changes, phenology has two major functions. Firstly, it reveals measurable impacts of climate change on nature, which at the same time clearly demonstrate global climate change in people's backyards. Secondly, long-term phenological data allow the reconstruction of temperature and its variability in the last centuries. The most prominent temperature driven changes in plant phenology are an earlier start of spring in the last three to five decades of, on average, 2.5 days/decade, mainly observed in midlatitudes and higher latitudes of the northern hemisphere. More heterogeneous changes in autumn are not as pronounced as in spring and cannot be linked to climate factors. A marked spatial and temporal variability of spring and summer onset dates and their changes can be mainly attributed to regional and local temperature. In this context, we discuss the temperature responses of the growing season and other phenological phases and their relation to the North Atlantic Oscillation. These results

illustrate main feedbacks in biogeochemical cycles and land–surface interactions of the climate system.

Abstract of Peer-reviewed Scientific Papers or Book Chapters, accepted or in press

*Rutishauser T., **Schleip** C., Sparks T., Nordli Ø., Menzel A., Wanner H., Jeanneret F., Luterbacher J. (2008) The Temperature Sensitivity of Swiss and British Plant Phenology 1753–1958, Climate Research Special Issue, COST-725, accepted.*

Abstract

Recent changes in springtime plant phenological 22 records are likely unprecedented and have been attributed to anthropogenically induced temperature change. In Europe, a major synchronous break in phenological time series in the 1980s was found in numerous studies. However, very few studies put these breaks in a historical perspective. We present evidence from two historical plant phenological records from northern Switzerland and the UK from 1753 to 1958. Monthly mean temperature measurements are available for the same regions. We assess whether synchronised changes in temperature and plant phenological records and recent temperature impacts are unprecedented at the end of the 20th century. We compare the temporal evolution of Plant Phenological Spring Indices (PPSI) and temperature series, search for common shifts and change-points by applying Pettitt's test and a Bayesian model comparison approach, and discuss changing temperature sensitivity for both localities. Results show that the Swiss records contain half the phenological variability (standard deviation 5 days) compared to the UK observations but higher temperature variability in winter and spring. There is a lack of synchronous shifts and one-point-changes in phenological and temperature series prior to 1958 in contrast to the widespread changes in Europe since the 1980s. However, there are indications of phenological shifts between 1840 and 1870 (Pettitt test) and changes in 1930 (highest change-point probability). Finally, we found a greater and more variable temperature sensitivity of the UK PPSI with respect to seasonal temperatures (2-15 days 42 change/°C) than in Switzerland (-2.5 to -5 days change/°C).

***Schleip** C., Menzel A., Dose V., (forthcoming) Bayesian methods in phenology in Hudson IL & Keatley M.R., (Eds) Phenological Research: Methods for environmental and climate change analysis. Springer. Berlin, Heidelberg, New York.*

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. The first example of this chapter will focus on the model comparison option of the Bayesian approach that was used to compare three different types of models (constant, linear, and one change point) for the analysis of three species in Germany. In addition to the functional behaviour, rates of change in terms of days per year were also calculated. The second example of this chapter illustrates the application of the Bayesian method to several phases throughout the year in two different countries. Thus we particularly investigate phenological changes of different phases and seasons.

Abstracts of Peer-reviewed Scientific Papers, revised, in revision or in review

***Schleip** C., Rais A., Menzel A.: Bayesian analysis of temperature memory effects of plant phenology in Germany, Agricultural and Forest Meteorology, in review.*

Abstract
A Bayesian approach enabled us to determine the temperature sensitivity of 43 phenological phases throughout the year in Germany within the period 1951-2006. First we conducted a Bayesian model comparison of monthly temperatures and phenological phases throughout the year. We analysed the data as constant (mean onset date), linear (constant trend over time) and as change point model (time varying change). The change point model involves the selection of two linear segments which match at a particular time. The matching point is estimated by an examination of all possible breaks weighted by their respective change point probability.

Results reveal that the late spring, summer and early autumn temperature months exhibit a clear preference for the change point model (>50%) indicating nonlinear change over the last 5 decades. April and August show exceptionally high nonlinearities with change point model probabilities of 78% and 81%. Phenophases occurring in summer and early autumn reveal the highest change point model probabilities up to 100% whereas early spring and late autumn phenophases possess lower change point probabilities but still over 50%. Secondly we used a Bayesian coherence analysis to investigate the relationship between phenological onset dates and an effective temperature generated as a weighted average of monthly means. Temperature weight coefficients were obtained from an optimization of a coherence factor by simulated annealing. For all phenophases the results suggest a coherent treatment of temperature and phenological change point distributions instead of an independent treatment, equivalent to a strong dependence of temperature and phenology. All spring and summer phenophases exhibit a nonlinear change predominantly in the 1980s. The most evident coherence of temperature and phenological nonlinearities is seen between fruit ripening/ picking stages and April temperatures. The autumn temperature signal remains vague. We can classify two main temperature response patterns of the studied phenological phases. On the one hand spring phenophases are particularly sensitive to temperatures in April, exhibiting a prompt response. On the other hand summer phenophases are less influenced by temperature during or right before the month of the onset. They exhibit a delayed response to nonlinear temperature changes mainly of April. We assume that especially abrupt temperature changes during the temperature sensitive stage of species cause a pronounced change in plant phenology regardless the time of onset.

Erklärung

Ich erkläre an Eides statt, dass ich die der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Promotionsprüfung vorgelegte Arbeit mit dem Titel:

„Climate change detection in natural systems by Bayesian methods"

am Fachgebeit für Ökoklimatolgie unter der Anleitung und Betreuung durch Prof. Dr. Annette Menzel ohne sonstige Hilfe erstellt und bei der Abfassung nur die gemäß § 6 Abs. 5 angegebenen Hilfsmittel benutzt habe.

Ich habe die Dissertation in keinem anderen Prüfungsverfahren als Prüfungsleistung vorgelegt.

Ich habe den angestrebten Doktorgrad noch nicht erworben und bin nicht in einem früheren Promotionsverfahren für den angestrebten Doktorgrad endgültig gescheitert.

Die Promotionsordnung der Technischen Universität München ist mir bekannt.

München, den ………….. ………………………
Unterschrift