Time series modeling and central European temperature impact assessment of phenological records over the last 250 years

Christoph Schleip, 1 This Rutishauser, 2 Jürg Luterbacher, 3, 4 and Annette Menzel 1

Received 3 November 2007; revised 15 April 2008; accepted 7 August 2008; published 9 December 2008.

[1] 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 unprecedented rate of change of 0.08°C/a 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.


1. Introduction

[2] 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 organized 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, 2009]. 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, 2008]. 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, 2009]. 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].

[3] Phenological information from documentary sources, such as dates of grape harvests, sea-ice-free periods in harbors 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; Intergovernmental Panel on Climate Change, 2007] and indicate warm or cold periods in particular regions (see Brûzidil 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; Pijster, 1999; Bradley et al., 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 the work of Meier et al. [2007]).

[4] 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.


[5] Methodologically, trends in time series are often analyzed 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], and Menzel et al. [2006]). The slope of the linear regression equation indicates the average rate of change in phenology (days/a) or temperature (°C/a). 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, e-Handbook of Statistical Methods, 2006, available at http://www.itl.nist.gov/div898/handbook/index2.htm).

[6] 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 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., 2007] was another attempt to overcome the shortages of a priori decisions of window length for linear trend analyses.

[7] 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 analyzing 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.

3. Temperature Impact on Changing Phenology

[8] 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 3-monthly averaged temperature impacts. They used a simulated annealing optimization algorithm to receive a coherence factor and temperature weights [Schleip et al., 2008].

[9] In this paper we apply the linear trend analysis and Bayesian model comparison to an investigation of three unique, multidecadal, 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 homogenized 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 long-term phenological records but also uses rigorous, robust statistics to assess
temperature as the main environmental forcing factor for the first time.

4. Material

[10] 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: 1 January = 1. The altitudes of observation and measurement sites are between 200 and 800 m above sea level.

5. Phenological Observations

5.1. Swiss “Spring Plant” 1753–2006

[11] 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 2 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 modeling 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 timescales of ±10 days and of ±3.6 days at decadal timescales.

5.2. Swiss and Burgundy Grape Harvest Dates

[12] 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].

[13] 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, personal communication, 2007). We used the post-1753 data of both time series, which were overlapping with the instrumental temperature data from Switzerland.

5.3. Temperature Measurements

[14] Monthly and seasonal mean temperature measurements from Geneva (starting in 1753) and Basel (starting in 1755; Schüpp [1961] and 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).

[15] 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) [van Oldenborgh et al., 2005] for the overlapping period 1951–2000, except for October (76%), most likely because of an erroneous outlier in the Dijon record (not shown).

6. Methods

6.1. Time Series Models

[16] 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.

[17] 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.

6.2. Model-Averaged Rate of Change, Change-Point Probability Distribution, and Moving Linear Trend

[18] 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 behavior 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 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 1-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 timescales. 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 toward later spring onset / harvest dates) are expected to be synchronous to periods with negative temperature trends or cooling periods.

### 6.3. 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 $M_0$: temperature and blossom onset time series evolved independently and $M_1$: temperature and blossom onset time series exhibit coherence. The results are the probabilities for $M_0$ and $M_1$ 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 $\rho$. The correlation coefficient between a variable $y$ and a variable $x$ is a measure for the linear dependence between $x$ and $y$. $\rho^2$, the square of the correlation coefficient, describes the degree of explained variance. The range of values of $\rho^2$ is $0 \leq \rho^2 \leq 1$ with $\rho^2 = 1$ signaling 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 non-negligible, 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].

### 7. Results

#### 7.1. 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 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”).

#### 7.2. Time Series Models

Figures 2a–2c and 3a–3d present the functional descriptions of the constant, linear and change-point models for each selected phenological and temperature series, respectively. In Figures 2a–2c, 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 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 (Figures 3a–3d).

The linear model of the “spring plant” reveals a positive slope of 0.005 ± 0.004 days/a over the whole record
(a–c) Functional behavior 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), mean Swiss seasonal winter (December–February), spring (March–May), summer (June–August), and autumn (September–November) temperatures for 1753–2006.

(Figure 2a) indicating a delay of spring onset over the whole record. Both Swiss and Burgundy grape harvest dates (Figures 2b and 2c) show a consistent negative slope of $-0.04 \pm 0.007$ days/a and $-0.03 \pm 0.007$ days/a. For long-term temperature trends (1753–2006) the winter season exhibits the strongest positive slope of $0.01 \pm 0.001$°C/a. In comparison, spring, summer and autumn season show a less pronounced slope of $0.004 \pm 0.001$°C/a, $0.003 \pm 0.001$°C/a and $0.006 \pm 0.001$°C/a, respectively (Figures 3a–3d).

[27] 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 (Figures 2a–2c). 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.

[28] 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 (Figures 3a–3d). The function estimation for spring temperature declines from 1753 until 1850 and then from 1850 onward 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.

7.3. Model-Averaged Rates of Change

[29] All phenological time series exhibit increasingly negative rates of change (Figures 2d–2g, top 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 2f). 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 2e). Swiss and Burgundy grape harvest dates exhibit for 2003 exceptional early harvest dates (Figures 2b and 2c).

[30] All investigated temperature records show an increasingly positive rate of change from 1753 to 2006 (Figures 3e–3h). In 2006, winter, spring and autumn season exhibit a rate of temperature change of approximately 0.01°C/a. 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 3e). 1882 (Figure 3f) and 1900 (Figure 3h), respectively. Since the 1980s, the positive rate of change increased from almost zero to 0.08°C/a in 2006 (Figure 3g).

7.4. Moving Linear Trend Analysis

[31] 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 (Figures 4a–4c, top). 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 occurring 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 nonsignificant phenological trends (Figures 4a–4c, bottom). There is a distinct trend toward 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.

[32] The “spring plant” shows two distinct periods with trends toward earlier spring development of up to $-0.3$ days/a for the center year of the 30-year windows between 1976–1991 and 1939–1954 both with low error probabilities (Figure 4a). Periods of trends toward later spring development
Figure 2
Figure 3. As in Figure 2 but for winter (a, e) December–February, spring (b, f) March–May, summer (c, g) June–August, and autumn (d, h) September–November temperatures in the period 1753–2006. Functional model behavior is shown in Figures 3a–3d, and model-averaged trend and change-point probability is shown in Figures 3e–3h.
of +0.1 to +0.2 days/a 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 toward earlier dates of 0.4 days/a or more at the end of the 20th century, between 1830–1860, around 1810 and in the 1760s (Figure 4b). 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 4c). 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.

### Change-Point Analysis

The most likely change point for the “spring plant” time series is found in 1984 (probability = 2%) (Figure 2d, bottom line). In comparison, the most likely change point for Swiss grape harvest dates is 1986 (probability = 7%) (Figure 2e). 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 2f).

Winter and spring temperatures exhibit the highest probability for a change at the beginning of the 1850s (Figures 3e and 3f). Within the summer season, the highest probability for a change is found in 1978 with a maximum probability of 10% (Figure 3g). The autumn season reveals the highest probability of a change point in 1912 (Figure 3h).

### Coherence Factors, Temperature Weights, and Linear Correlation

Swiss “spring plant” and Swiss grape harvest dates reveal a high coherence with temperature time series (Figures 5a and 5b). The monthly resolution exhibits high coherence factors for both Swiss and Burgundy grape harvest dates with the highest temperature weight in June (Figure 5b). 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 5b). 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 1). Mean temperatures of March to July increasingly explain more variance of the grape harvest date records from 10 to around 25%. Afterward 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.

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 5a). Linear correlation indicates significant temperature impact of single months only from February to April.

### 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 behavior [e.g., Root et al., 2003; Parmesan and Yohe, 2003; Menzel et al., 2006]. The simplicity of the least

---

**Figure 4.** Moving linear trend analysis for (a) Swiss “spring plant”, (b) Swiss grape harvest dates, and (c) Burgundy grape harvest dates showing slope coefficients of the linear regression of phenology against time for 30-year periods. Bold lines show phonological trends, and thin lines show corresponding spring temperature trends. Note that the left axis represents the phonological trend, and the right axis represents 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.
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] reanalyzed 125,000 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 multicentury phenological records (Figure 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 toward 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 to 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].

8.2. Bayesian Model Comparison

[39] 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 behaviors 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; Schlep 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.

[40] 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/a 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 behavior 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.

The functional behavior 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 multidecadal timescales 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 onward (Figure 2a). The Swiss and Burgundy grape harvest dates show a continuous advance of the onset of harvests from 1753 onward (Figures 2b and 2c).

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 21st 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 (Figures 2f and 2g). 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 2e).

### 8.3. 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 (Figures 2d–2f and 3e–3h). 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 20th 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 20th 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 multimodal 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 modeling 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

---

**Table 1.** Pearson Correlation and Associated Error Probabilities Between Phenological Series and Preceding Monthly Mean Temperatures

<table>
<thead>
<tr>
<th>“Spring Plant”</th>
<th>cor</th>
<th>p-val</th>
<th>R² [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>pJan</td>
<td>−0.13</td>
<td>0.045</td>
<td>2</td>
</tr>
<tr>
<td>pJul</td>
<td>−0.07</td>
<td>0.301</td>
<td>0</td>
</tr>
<tr>
<td>pAug</td>
<td>−0.15</td>
<td>0.017</td>
<td>2</td>
</tr>
<tr>
<td>pSep</td>
<td>0</td>
<td>0.998</td>
<td>0</td>
</tr>
<tr>
<td>pOct</td>
<td>−0.01</td>
<td>0.923</td>
<td>0</td>
</tr>
<tr>
<td>pNov</td>
<td>−0.04</td>
<td>0.558</td>
<td>0</td>
</tr>
<tr>
<td>pDec</td>
<td>−0.07</td>
<td>0.271</td>
<td>0</td>
</tr>
<tr>
<td>Jan</td>
<td>−0.05</td>
<td>0.0448</td>
<td>0</td>
</tr>
<tr>
<td>Feb</td>
<td>−0.33</td>
<td>&lt;0.001</td>
<td>11</td>
</tr>
<tr>
<td>Mar</td>
<td>−0.46</td>
<td>&lt;0.001</td>
<td>21</td>
</tr>
<tr>
<td>Apr</td>
<td>−0.56</td>
<td>&lt;0.001</td>
<td>31</td>
</tr>
<tr>
<td>May</td>
<td>−0.18</td>
<td>&lt;0.004</td>
<td>3</td>
</tr>
</tbody>
</table>

**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      |

*p² 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). Cor, Pearson correlation; p-val, associated error probabilities.
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.

### 8.4. Temperature Impact

[45] 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, 2003; 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: that is, 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].

[46] 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). 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 1) [e.g., Sparks and Carey, 1995], this result is statistically much less evident and has never been discussed.

[47] 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 (Figures 5a and 5b). 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 artifact as seasonal temperature weights (Figure 5a) and linear correlation (Table 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. Afterward, 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 5b) 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.

[48] *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 not be 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.

### 9. Conclusions

[49] Unique long-term temperature and phenological data series for central Europe back to 1753 were analyzed with different approaches. A simplistic linear approach illustrates valuable information regarding the impact of temperature on multidecadal 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 behaviors 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-week 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.

Acknowledgments. CS was supported by a travel grant of COST725. TR was supported by the Swiss National Science Foundation (project “Past, Present and Future Climate Impact on Spring and Summer Vegetation: A Phenological Approach” contract number 205251-105691/1). TR and JL were supported by the Swiss NSF through its National Center of Competence in Research on Climate (NCCR-Clim) and the Oeschger Centre for Climate Change Research, University of Bern, Switzerland. We thank Abraham Miller-Rushing for his native English corrections.

References


J. Luterbacher, Oeschger Centre for Climate Change Research and NCCR Climate, University of Bern; Institute of Geography, Climatology and Meteorology, University of Bern, Hallerstrasse 12, CH-3012 Bern, Switzerland. (juerg@giub.unibe.ch)

A. Menzel and C. Schleip, Fachgebiet für Okoklimatologie, Technische Universität München, Am Hochanger 13, D-85354 Freising-Weihenstephan, Germany. (menzel@forst.tu-muenchen.de; christoph.schleip@web.de)

T. Rutishauser, Institute of Geography, University of Bern, Hallerstrasse 12, CH-3012 Bern, Switzerland. (rutis@giub.unibe.ch)