Received: 9 October 2018 | Revised: 21 January 2019 | Accepted: 15 March 2019
DOI: 10.1002/env.2570

SPECIAL ISSUE PAPER

WILEY

Changepoint analysis of Klementinum temperature series

D. Jarušková[1] | J. Antoch[2]

[1]Department of Mathematics, Faculty of Civil Engineering, Czech Technical University in Prague, Prague, Czech Republic

[2]Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

Correspondence
J. Antoch, Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, CZ 186 75 Praha 8, Prague, Czech Republic.
Email: antoch@karlin.mff.cuni.cz

Funding information
European Regional Development Fund, Grant/Award Number: CZ.02.1.01/0.0/0.0/16_019/0000778 ; GACR, Grant/Award Number: P403/19/02773S

Abstract

This paper presents a statistical analysis of the stationarity of the Prague–Klementinum temperature series. In the first part, the stationarity of annual means is rejected, and several nonstationary models are suggested to assess the increase in temperature in the most recent decades. The analysis shows that positions of the changepoint estimates depend largely on our decision as to whether we apply a discontinuous or a continuous piecewise linear model. In the second part, we study the stationarity of the seasonal behavior of the series, particularly the stationarity of the mean annual profiles. The analysis of the seasonal cycle shows that the increase in temperature has not the same character throughout the calendar year. Rather, an increase in temperature in winter days is larger than an increase in summer days, and the temperature difference between summer and winter days decreases.

KEYWORDS

abrupt-shift, change in seasonal behavior, changepoint detection, gradual and multiple changes, Klementinum data, likelihood and Bayesian estimators, max-type and sum-type test statistics, MCMC, permutation principle

1 | INTRODUCTION

A question as to whether long-term observed temperature, precipitation, and other climatological series are stationary has been a central point of interest for both climatologists and statisticians for a long time. An answer to this question became more pertinent when, in many temperature series, an increasing trend appeared. Many scientists started to connect this increase to emissions of greenhouse gases, such as carbon dioxide, methane, and nitrous oxide (Camuffo & Jones, 2002). It is worth mentioning that Arrhenius (1896) has already suggested a link between carbon dioxide and temperature before the trend started to show.

With the goal of deciding whether the series nowadays behave similarly as those from when the measurements began, statisticians have developed many procedures for studying their stationarity. Roughly speaking, a time series is stationary if its statistical properties, for example, mean, variance, etc., remain constant through time.

Typically, the change in mean temperature is assessed by fitting a linear trend. However, as has been also recognized in Chapter 2 of the Intergovernmental Panel on Climate Change (Hartmann et al., 2013), which specifically states how climate change is quantified in observations, *Climatic time series often have trends for which a straight line is not a good approximation*. As such, there is a need for more complicated models for a description of climatological series that enable the estimation of both abrupt-shift and gradual changes, to detect changes not only in the mean but also in the variance, in occurrence of extreme events or changes in seasonal behavior, etc. All methods that aim at the detection and estimation of changes in stochastic models are referred to as "changepoint methods." These models have gained popularity lately and have been applied to the analysis of climatological and environmental data. Let us mention, among many others, papers by Drijfhout et al. (2015), El-Shaarawi (2013), Jandhyala, Fotopoulos, and El-Shaarawi (2006), Jarušková (1997), Reeves, Chen, Wang, Lund, and Reeves (2007), and Beaulieu and Killick (2018).

In our paper, we analyze the Prague–Klementinum temperature series, the longest series of Czech temperature measurements, using several changepoint methods. Even if we take into account that the temperature measurements are taken in the center of Prague and the increase in temperature may be partially explained by a heat island effect, the permanent increase in the most recent few decades is surprising. This has been also noticed by Brázdil et al. (2012). Due to space limitations, we focus here on the stationarity of the mean and seasonality.

This paper is organized as follows. In Section 2, the Prague–Klementinum data are described. In Section 3, we test the stationarity and look for models enabling to describe annual temperature means containing both gradual and abrupt-shift changes. We consider models with one and two changes. In Section 4, we test the stationarity of the annual mean cycle using three procedures for a reduction in dimension. All procedures—the Fourier series approximation, the method of principal components, and the method based on monthly means—conclude that an increase in temperature is not the same throughout the calendar year: The increase in the winter period starts earlier and is larger than in the summer period.

2 | DATA

In Prague–Klementinum, we can find the oldest meteorological station in Central Europe. Its measurements form the most important source of historical data about the climate evolution in Central Europe.

Astronomical tower of Klementinum (P. Příhoda, pen-and-ink drawing, 2008)

First, allow us a short reminiscence. The beginnings of regular measurements of temperature, air pressure, precipitation, solar activities, and other environmental characteristics in Prague date back to 1752. They are closely connected with the eminent astronomer and mathematician Joseph Stepling (cs.wikipedia.org/wiki/Joseph_Stepling), a Jesuit who, at that time, initiated the construction of a new astronomic tower (51 m high) in a Jesuit college of Klementinum (en.wikipedia.org/wiki/Clementinum) placed in the very center of Prague just across the Charles Bridge. Until 1775, the measurements were not regular, and sometimes, the precise measurements were replaced by a guess. Since January 1, 1775, the temperature measurements were taken regularly at 7, 14, and 21 hr (local time, GMT+1) and are complete. In the available data set are included the minimum, maximum, and average daily values. For our analysis, we use the average daily values. This means that, per year, we have 365 observations, respectively 366 in the leap years.

Despite the fact that the station is in the center of Prague, outside the grassland, and the box with instruments is located on the building wall, thus not satisfying the regulations concerning the placement of current meteorological stations, its long-term measurements form an invaluable and unique source of information about the variations of temperatures and precipitation in Central Europe. Several times the thermometers, as well as other measuring instruments, have been replaced, always after a long calibration. The original instruments are kept intact in the museum of Klementinum. The current operator of the station is the Czech Hydrometeorological Institute (www.chmi.cz). Data are provided by the National Climatic Data Center (www.ncdc.noaa.gov; for details, see Klein Tank et al., 2002). Partial data are also available from other sources.

The annual mean temperatures are plotted in Figure 1. Looking at the time series, it appears that the series may be nonstationary with an increasing trend starting around the beginning of the 20th century, which has intensified in recent decades.

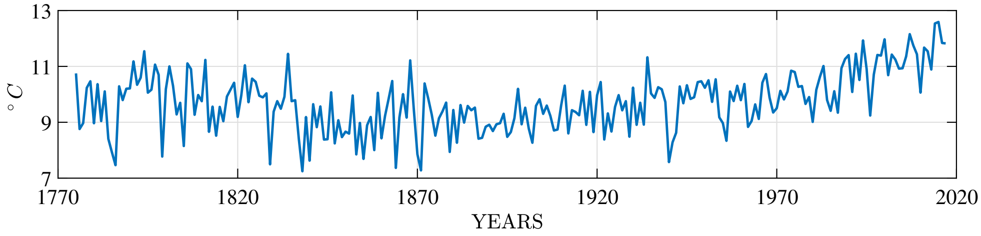

FIGURE 1 Klementinum annual mean temperatures

3 | STATIONARITY OF THE ANNUAL MEAN SERIES

In the scope of mathematical statistics, a decision as to whether the series is stationary is usually made by hypothesis testing. The null hypothesis claims that the process is stationary, whereas the alternative hypothesis claims that the process is nonstationary, and the stationarity is violated in a specific way. In our case, we will mainly be interested in stationarity in the mean of the observed series.

3.1 | One change

Suppose that the annual mean temperatures $\{Y_i,\ i = 1, \dots, n\}$ form a sequence of independent random variables with the corresponding expectations $\mathrm{E}(Y_i) = \mu_i, i = 1, \dots, n$, and the same variance $\mathrm{var}(Y_i) = \sigma^2, i = 1, \dots, n$. The null hypothesis claims that the mean values of all observations remain the same, that is,

$$H_0 : \ \mu_1 = \cdots = \mu_n.$$

When assuming at most one change, the simplest alternatives studied in the changepoint literature are an abrupt-shift and a gradual appearance of a linear trend. Note that the “at most one change” scheme can be considered to be a generalization of the two-sample problem.

The alternative hypothesis A_1, which corresponds to the abrupt-shift in the mean, may be set as follows:

$$\begin{aligned} A_1 : \exists k \in \{1, \dots, n-1\} \quad & \text{such that} \\ \mu_1 = \cdots = \mu_k \neq \mu_{k+1} = \cdots = \mu_n. & \end{aligned}$$

The alternative A_2, which corresponds to the appearance of a gradual linear trend arising at an unknown moment of time, the so-called ice-hockey-stick model, has the form

$$\begin{aligned} A_2 : \exists k \in \{1, \dots, n-1\} \quad & \text{such that} \\ \mu_1 = \cdots = \mu_k & \\ \mu_{k+i} = \mu_k + b \cdot \frac{i}{n}, \quad & \text{for} \quad i = 1, \dots, n-k. \end{aligned}$$

The most frequently used test statistics for testing the existence of a changepoint are of either maximum type or sum type. When developing them, we usually start with a two-sample test statistic that might be applied if the changepoint was known, fixed, and equal to k. In the case of the alternative A_1, the corresponding two-sample test statistic has the form

$$T_k = \frac{1}{\hat{\sigma}} \sqrt{\frac{k(n-k)}{n}} \left| \overline{Y}_k - \overline{Y}_k^{\star} \right|, \tag{1}$$

where $\overline{Y}_k$ is the average of the first k observations, $\overline{Y}_k^{\star}$ is the average of the last $n-k$ observations, and $\hat{\sigma}$ is a consistent estimate of the standard deviation σ. When we decide to use a maximum-type test statistic, we may use either the global maximum of the two-sample statistics T_k over all possible time points, the maximum over a trimmed portion of time points, or the maximum of weighted statistics $\{w_k T_k\}$ with weights $w_k = \sqrt{k(n-k)/n^2}$. Some authors use the term “penalized” maximum-type test statistic here. However, the terminology is not uniform, and we will use the term “weighted” throughout this paper. Provided the observations $\{Y_i\}$ are independent and normally distributed random variables, then the test statistic T_k may be obtained using the maximum likelihood principle. For a detailed derivation, see Section 3 in Antoch, Hušková, and Jarušková (2002).

Similarly, when we decide to use the sum-type test statistic, we may use either a sum of statistics $\{T_k\}$ or a sum of weighted statistics $\{w_k T_k\}$. A more detailed description of the test statistics and selected critical values may be found in Jarušková (1997).

In the case of the alternative A_2, the two-sample test statistic, supposing again that the changepoint is known, fixed, and equal to k, has the form

$$\widetilde{T}_k = \frac{1}{\widehat{\sigma}} \cdot \frac{\sum_{i=k+1}^{n}(Y_i - \overline{Y})(i-k)}{\sqrt{\frac{(n-k)(n-k+1)(2n-2k+1)}{6} - \frac{(n-k)^2(n-k+1)^2}{4n}}}. \tag{2}$$

Similarly as in the case of A_1, we may use maximum-type or sum-type test statistics for testing H_0 against A_2. For more details and selected critical values, see Jarušková (1997).

We would like to remark that the test statistics introduced above are aimed at the detection of the behavior described by the corresponding alternatives. However, they may reject the null hypothesis even in situations when the trend has another character: if, for instance, there are more abrupt-shifts or the trend is an increasing function of time, etc. However, all of the above-discussed test statistics lose their power for the detection of such nonstationarities.

In Kiefer (1959), MacNeill (1974), Jandhyala and MacNeill (1991), Csörgő and Horváth (1997), or Antoch et al. (2002), among others, the asymptotic distributions of the maximum-type and sum-type test statistics, as $n \to \infty$, are presented. These asymptotic distributions may be used for finding approximate critical values.

Besides that, there exists a different way for obtaining approximate critical values of the considered test statistics, which is based on the permutation principle. The basic idea, in our setting, can be applied as follows.

1. Test statistic is fixed.
2. The original data series $Y_1, \ldots, Y_n$ is permuted many times. For each permutation of the data, the corresponding test statistic is calculated and stored.
3. From the stored values, the empirical distribution function is constructed, and sample quantiles are calculated.

For the significance level α, the $(1 - \alpha)100\%$ empirical quantile can be used as an approximate $\alpha 100\%$ critical value, and the empirical distribution function may be used for estimating p values. Antoch and Hušková (2001) have shown that, for a large number of permutations, critical values obtained by the permutation principle are close to the true ones.

Analyzing the Klementinum annual mean temperatures ($n = 243$) by the hypothesis testing procedures described above, the p values obtained with the help of 100,000 random permutations for both alternatives A_1 and A_2 and all considered max-type and sum-type test statistics are smaller than 10^{-5} (supposing that the annual means are independent variables). It follows that the null hypothesis is clearly rejected by all considered test statistics.

Furthermore, we may estimate the changepoint $k^\star$ by the argument of the maximum of the nonweighted test statistics (1) (respectively, (2)). The fit of the models corresponding to alternative A_1 (respectively, A_2) may be judged by Figure 2.

In Bai and Perron (1998), Csörgő and Horváth (1997), or Stryhn (1996), an asymptotic distribution is given for a changepoint estimator in an abrupt-shift model together with the selected quantiles. Using this distribution for construction of the 95% confidence interval, we get the interval 1987 $\mp$ 2.

Supposing the ice-hockey-stick model is described by alternative A_2, the corresponding 95% confidence interval is much broader. Using the approach of Feder (1975), the asymptotic distribution of the estimate of changepoint $\tau^\star = k^\star/n$ satisfies

$$\sqrt{n}\left(\frac{\widehat{k}^\star}{n} - \tau^\star\right) \sim N\left(0, \frac{1}{\widehat{b}^2}\frac{(1+3\tau^\star)}{(1-\tau^\star)\tau^\star}\widehat{\sigma}^2\right).$$

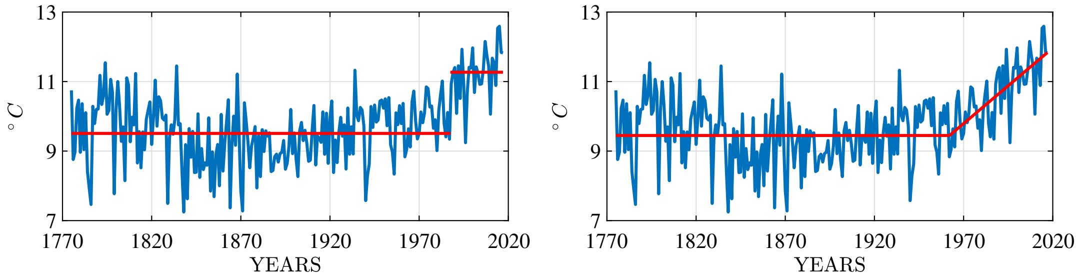

FIGURE 2 The "optimal" abrupt-shift model and the "optimal" ice-hockey-stick model

As $\hat{\tau}^\star = 0.774$, $\hat{b} = \hat{b}_n \cdot n = 0.0433 \cdot 243 = 10.519$, and $\hat{\sigma} = 0.849$, the limit distribution yields the approximate 95% confidence interval for the changepoint as 1962 ∓ 11.

The asymptotic distribution of the statistics (1) and (2) as well as of the changepoint estimates can be adapted for the case when the data are modeled by a linear process, for example, by an autoregressive–moving-average sequence (see Antoch, Hušková, & Prášková, 1997). The adaptation consists in multiplying the critical values of the statistics (1) and (2) derived for independent and identically distributed random variables by a constant that is derived from a correlation structure of the series. The correlation structure is estimated using the values of the sample autocorrelation function of the time series after removing an estimated trend. The effect of dependence of the statistics (1) and (2) was studied, for example, by Jarušková (1998). For the Klementinum annual temperature series, the values of the test statistics are so large that the null hypothesis is rejected even if we suppose that the series forms a linear process. When we suppose that the series is a long-memory process, distinguishing between a series with a trend and a long-memory model is more complicated (for a discussion, see Beaulieu & Killick, 2018).

3.2 | Multiple changes

We have also been interested in finding an optimal model with multiple changes, which minimizes a residual sum of squares. First, we consider a piecewise linear model with abrupt-shift changes and, second, a piecewise linear model with gradual changes. Figure 3 presents an optimal model with two abrupt-shift changes estimated in the years 1836 and 1988. The slope parameters of the second and third segments are 0.0091 and 0.0355, which correspond to the growth of almost one-and-a-half degrees during the period of 1836–1987, dramatically quadrupling that speed in the three most recent decades. These results confirm all the previous findings.

Figure 4 presents an optimal model with two gradual changes estimated in the years 1879 and 1980. The slope parameters of the second and third segments are 0.0217 and 0.0389, which correspond to the growth of more than two degrees during the hundred-year period 1879–1980, dramatically doubling that speed during the last four decades. This result confirms the findings for the ice-hockey-stick model discussed in Section 3.1.

An overview of the methods suitable for inference of multiple changepoints can be found in Jandhyala, Fotopoulos, MacNeill, and Liu (2013). Computational issues and efficient algorithms for the estimation of multiple changes are discussed, for example, by Antoch and Jarušková (2013), Kim, Yu, and Feuer (2008), or Maidstone, Hocking, Rigaill, and Fearnhead (2017). Moreover, we would like to point out that totally different approach for finding eventual multiple changepoints offer moving sum statistics and local linear estimators applied to the observed time series (for details, see, e.g., Antoch et al., 2002; Antoch, Gregoire, & Hušková, 2007).

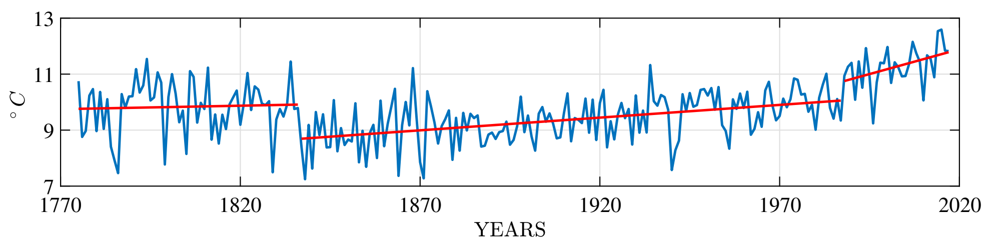

FIGURE 3 Optimal piecewise linear model with two abrupt-shift changes

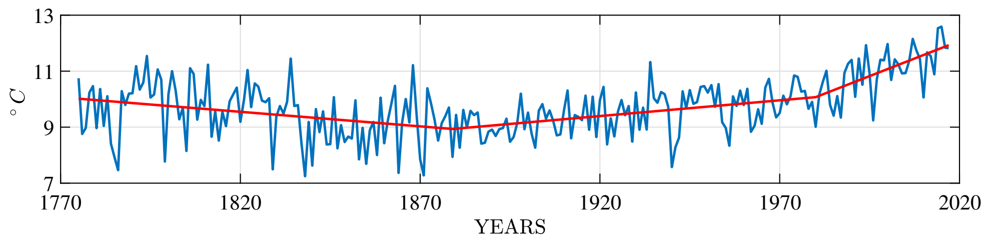

FIGURE 4 Optimal piecewise linear model with two gradual changes

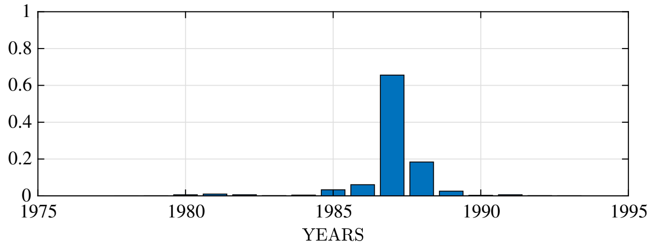

FIGURE 5 Posterior distribution of the changepoint $k^{\star}$ in the abrupt-shift model

3.3 | Bayesian approach

The estimates presented above are of the maximum likelihood type when the observations are supposed to be normally distributed. Less usual, but very interesting, is the Bayesian approach to estimating changes in time series studied for the Klementinum data by Antoch and Legát (2008). In the Bayesian approach, it is supposed that the changepoint is a parameter that has a prior distribution over all time points (often a uniform) and that all parameters describing the behavior of the mean before and after the change also have certain prior distributions. The goal of statistical inference is to find a posterior distribution for both the changepoint and the remaining parameters.

For instance, when assuming a sudden-shift model, the annual temperature means follow a normal-inverse-gamma prior for μ_1, μ_2, and σ^2, and the posterior distribution of the parameters can be obtained using Markov chain Monte Carlo.

We run the Markov chain Monte Carlo algorithm of Antoch and Legát 10^6 times and discarded the beginning 10^5 members. In Figure 5, we present a probability histogram of the posterior distribution of the changepoint parameter $k^{\star}$. It is evident that the distribution is concentrated around the year 1987, being in full agreement with the results of Section 3.1.

It is interesting to compare our results obtained using 243 observations from the years 1775–2017 with those of Antoch and Legát calculated using 218 observations from the years 1775–1992. The modus of the posterior distribution of the changepoint has shifted from the year 1942 to 1988. The reason for a shift in the posterior distribution of the changepoint is a gradually increasing trend in the last decades. As the model is a piecewise constant with one change only, the estimate of the change is shifted toward the end of the series where the gradual increase is extremely large. This is true both for a Bayesian and a frequentist approach.

In the Bayesian context, we considered, analogously as Antoch and Legát (2008), also more complicated models, for example, two-phase and three-phase linear models with both abrupt-shift and gradual changes. Trying to keep the size of the paper acceptable, we do not present here the results of the Bayesian approach with two or more changes. In any case, the results are parallel to the presented results for multiple changes based on the frequentist approach. Interested readers should look, for example, to Fearnhead (2006), where exact and efficient Bayesian inference for multiple changepoint problems is considered. The scope of this paper does not allow us to discuss an abundant Bayesian literature in the field of changepoints. In any case, we would like to draw attention to influencing papers by Perreault, Bernier, Bobée, and Parent (2000), Green (1995), Barry and Hartigan (1993), or Carlin, Gelfand, and Smith (1992).

Moreover, we would like to point out that in the domain of changepoint analysis, *Bayesian-like procedures*, combining both the Bayesian and the frequentist approach, have been suggested and shown to be an efficient tool for solving the problems under study. These estimators have a smaller variance than related argmax-type estimators and can be also viewed as one-step estimators, where the argmax-type estimators are used as the preliminary ones (for details, see, e.g., Antoch & Hušková, 2000).

4 | STATISTICAL METHODS FOR DETECTION OF NONSTATIONARITIES IN THE MEAN ANNUAL CYCLE

We have seen that the Klementinum annual mean temperature series is not stationary. Applying the same statistical tests as in Section 3, the null hypothesis on stationarity is rejected for all 12 monthly mean temperature series as well. Supposing that the monthly means follow either an abrupt-shift or an ice-hockey-stick model, we estimated a changepoint for all individual months. The results of the study hint that an increase in temperature starts earlier for the winter months than for the summer months. For an abrupt-shift model, Jarušková (2018) derived an asymptotic distribution of the changepoint estimates when the changes do not occur simultaneously. For the ice-hockey-stick model, the approach of Feder (1975) can be applied for obtaining an asymptotic distribution of the changepoint estimates. Unfortunately, neither of these models fit the behavior of the monthly mean sequences quite well, as their trends seem to be more complex. Therefore, we do not present here the results of those statistical analyses.

On the other hand, as the monthly mean series are not stationary, because they almost certainly exhibit an increasing trend in the most recent few decades, a natural question may arise as to whether the increase in temperature has the same character across all calendar days or whether, in some period(s) of the year, the trend is larger than in others. The question may be also formulated as follows: *Has the mean annual profile a change in the form of a shift only or is the change more complex?*

As the temperature is not measured continuously, our decision is based on the vectors of the deviations of the daily temperature from the corresponding calendar annual mean. We start the analysis by organizing daily data into an $n \times 365$ matrix, with n denoting the number of years in which the measurements are taken (the averages for the 29th of February are omitted). We suppose that the daily mean temperature D_{ij} in the jth calendar day of the ith year of measurements satisfies the model

$$D_{ij} = \nu_i + \delta_{ij} + \epsilon_{ij},$$

where $\{\nu_i,\ i = 1, \dots, n\}$ represent the annual mean temperatures and $\{\delta_{ij}, i = 1, \dots, n, j = 1, \dots, 365\}$ represent the deviations of the daily mean temperatures from the annual mean, supposing $\sum_{j=1}^{365} \delta_{ij} = 0, i = 1, \dots, n$. The random error vectors $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \dots, \epsilon_{i,365})^\top$ are supposed to be zero mean. For $i = 1, \dots, n$, we estimate the annual mean temperature by the corresponding average, that is, $\widehat{\nu}_i = \sum_{j=1}^{365} D_{ij}/365$, and subtract it from the daily averages, that is,

$$DD_{ij} = D_{ij} - \widehat{\nu}_i, \quad i = 1, \dots, n,\ j = 1, \dots, 365.$$

The variables $\{DD_{ij}\}$ are introduced for keeping a seasonality but getting rid of an overall trend.

Clearly, $\sum_{j=1}^{365} DD_{ij} = 0$ for $i = 1, \dots, n$, and the vector of the deviations of the observed daily temperatures from the corresponding annual mean $\boldsymbol{DD}_i = (DD_{i1}, \dots, DD_{i,365})^\top$ is an estimate of the vector $\boldsymbol{\delta}_i = (\delta_{i1}, \dots, \delta_{i,365})^\top, i = 1, \dots, n$, which we call a mean annual cycle. The goal of statistical inference is to decide whether the mean annual cycle remains the same in all n years, that is, whether $\boldsymbol{\delta}_1 = \cdots = \boldsymbol{\delta}_n$.

4.1 | Tests for an abrupt-shift in the mean of a random vector

The question on the stationarity of the mean annual cycle may be decided by changepoint analysis, more precisely, by a test for the existence of a changepoint in the means of random vectors.

Suppose that we observe a sequence of random vectors $\boldsymbol{X}_i = (X_{i1}, \dots, X_{ip})^\top, i = 1, \dots, n$, and that the observations obey the model

$$X_{ij} = \mu_{ij} + e_{ij}, \quad i = 1, \dots, n,\ j = 1, \dots, p,$$

where $\boldsymbol{\mu}_i = (\mu_{i1}, \dots, \mu_{ip})^\top, i = 1, \dots, n$, are vectors of some unknown constants, and $\boldsymbol{e}_i = (e_{i1}, \dots, e_{ip})^\top$ are zero-mean independent and identically distributed random variables with a covariance matrix $\boldsymbol{\Sigma}$.

We consider a hypothesis testing problem for testing the null hypothesis H_0 against the alternative A, that is,

$$\begin{aligned} H_0 &: \ \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_n \\ A &: \ \exists k \in \{1, \dots, n-1\} \quad \text{such that} \\ & \quad\ \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_k \neq \boldsymbol{\mu}_{k+1} = \cdots = \boldsymbol{\mu}_n. \end{aligned}$$

The test statistics for testing the problem introduced above may be derived in the same way as in the one-dimensional case. First, supposing that the changepoint is known, fixed, and equal to k, the two-sample statistic for testing the equality of all mean vectors $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_n$ has the form

$$\chi^2(k) = \frac{k(n-k)}{n} \left(\overline{\boldsymbol{X}}(k) - \overline{\boldsymbol{X}}^\star(k)\right)^\top \boldsymbol{\Sigma}^{-1} \left(\overline{\boldsymbol{X}}(k) - \overline{\boldsymbol{X}}^\star(k)\right),$$

where $\overline{\boldsymbol{X}}(k) = (\overline{X}_1(k), \dots, \overline{X}_p(k))^\top$ is the vector of averages over the first k years, whereas $\overline{\boldsymbol{X}}^\star(k) = (\overline{X}_1^\star(k), \dots, \overline{X}_p^\star(k))^\top$ is the vector of averages over the last $(n-k)$ years. If matrix $\boldsymbol{\Sigma}$ is unknown, it may be replaced by any consistent estimate.

Let us now address the situation in which the changepoint is unknown. In such a case, we calculate the statistics $\{\chi^2(k)\}$ for all possible values of k. Here, the test statistics most frequently used are again either a maximum or a sum of the statistics $\{\chi^2(k)\}$, or their weighted versions. In what follows, we will use the test statistics

$$\begin{aligned} T &= \max_{\lfloor \beta n \rfloor \le k \le \lfloor (1-\beta) n \rfloor} \chi^2(k), \\ TW &= \max_{1 \le k \le n} \frac{k(n-k)}{n^2} \chi^2(k), \\ MW &= \frac{1}{n} \sum_{k=1}^{n} \frac{k(n-k)}{n^2} \chi^2(k). \end{aligned} \tag{3}$$

The value of β is usually set to a small positive number; in our case, we use $\beta = 0.01$.

The statistics T, TW, and MW attain a large value when an abrupt-shift change exists in the analyzed mean vector. However, they may also attain a relatively large value when there are several changepoints or even when the coordinates of the mean vectors $\boldsymbol{\mu}_i$, $i = 1, \dots, n$, are monotonous functions of the time index i.

The presented test statistics were suggested and studied by MacNeill (1974) and Csörgő and Horváth (1997), among others, and later applied to temperature series by Horváth, Kokoszka, and Steinebach (1999). They derived the asymptotic distribution under the null hypothesis of the stationarity of the test statistics T, TW, and MW as $n \to \infty$. The way to calculate approximate asymptotic critical values may be found in Kiefer (1959), MacNeill (1974), and Jandhyala and MacNeill (1991). The approximate critical values may also be obtained by the permutation principle described in Section 3.

4.2 | Dimension reduction

Theoretically, it would be possible to apply any of the testing procedures described above to the vectors $\{(DD_{i,1}, \dots, DD_{i,364})^\top\}$. However, if the number of years n is substantially smaller than 364, a better approach can be adopted for discovering possible nonstationarities in the mean annual cycles. This approach applies the testing procedure to vectors of linear combinations of the daily deviations rather than to the daily deviations themselves. If a change in the mean of chosen linear combinations is detected, then the mean annual cycle is not stationary. Obviously, the inverse implication is not true. The test for a change in the mean of $l \ll n$ linear combinations cannot detect all possible changes in the mean annual cycle. Notice also that, when performing such a dimension reduction, we have subjectively to decide how many linear combinations to use. This choice fully depends on the character of the problem. Our results suggest that if the linear combinations are chosen to capture the main features of the series, then, for detecting nonstationarities in the temperature mean cycle, the number of applied linear combinations may be quite small, and the results are not extremely sensitive to that number.

4.3 | Fourier series approximation

Due to the Earth's rotation around the Sun, the mean vectors $\mathrm{E}\,\boldsymbol{DD}_i = (\mathrm{E}\,DD_{i1}, \dots, \mathrm{E}\,DD_{i,365})^\top$, $i = 1, \dots, n$, may be supposed to be periodic with the frequency $2\pi/365$ and to resemble a shifted cosine function. This is illustrated in Figure 6, presenting the vector $(\overline{DD}_{.,1}, \dots, \overline{DD}_{.,365})^\top$ with $\overline{DD}_{.,j} = \sum_{i=1}^{n} DD_{ij}/n$, $j = 1, \dots, 365$, together with the first term of the Fourier expansion

$$\hat{A} \cos \frac{2\pi j}{365} + \hat{B} \sin \frac{2\pi j}{365}, \quad j = 1, \dots, 365,$$

where $\hat{A} = \frac{2}{365} \sum_{j=1}^{365} \overline{DD}_{.,j} \cos \frac{2\pi j}{365}$ and $\hat{B} = \frac{2}{365} \sum_{j=1}^{365} \overline{DD}_{.,j} \sin \frac{2\pi j}{365}$.

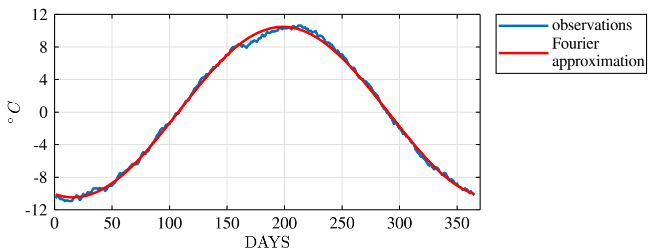

FIGURE 6 The periodic behavior of the average of the mean annual cycles

TABLE 1 Values of the test statistics T, TW, and MW for one, two, and three smallest Fourier frequencies with the corresponding p values (in brackets). In the last two rows, we may find the estimated differences between the maximal summer temperature and the minimal winter temperature in the same year for the years before and after the estimated changepoint. The corresponding p values were computed using 50,000 random permutations

Statistic	**1 freq.**	**2 freq.**	**3 freq.**
T	23.58	28.40	34.27
	(3.8×10^{-4})	(3.6×10^{-4})	(1.6×10^{-4})
TW	5.90	7.10	8.67
	(10^{-4})	(2×10^{-5})	(4×10^{-5})
MW	2.12	2.54	3.30
	(10^{-4})	(8×10^{-5})	(6×10^{-5})
max–min difference before $k^\star = 121$	21.6	21.6	21.3
max–min difference after $k^\star = 121$	20.3	20.3	20.3

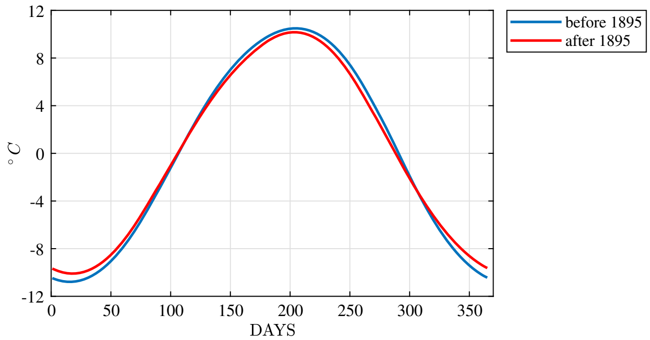

FIGURE 7 Estimated mean annual cycles before and after the year 1895

Since the agreement is very good, we may assume that the vectors $\{(\hat{A}_i, \hat{B}_i),\ i = 1, \dots, n\}$ with $\hat{A}_i$ and $\hat{B}_i$, computed analogously for the individual n annual cycles, will also capture a large part of the information on the behavior of the annual cycles so that they are suitable for detecting a change in their mean. Nevertheless, in spite of the fact that annual means are periodic with the period $2\pi/365$, they do not necessarily exactly follow the form of a shifted cosine function. Therefore, we also consider the vectors $\{(\hat{A}_i(1), \hat{B}_i(1), \dots, \hat{A}_i(l), \hat{B}_i(l)), i = 1, \dots, n\}$, corresponding to the l smallest Fourier frequencies $2\pi/365, \dots, 2l\pi/365$. Notice that the variance–covariance matrix of the vectors $\{(\hat{A}_i(1), \hat{B}_i(1), \dots, \hat{A}_i(l), \hat{B}_i(l))\}$ may easily be obtained from the variance–covariance matrix of the vectors $\{\boldsymbol{DD}_i = (DD_{i,1}, \dots, DD_{i,365})^\top\}$.

We have applied the procedure based on the Fourier series approximation for one, two, and three smallest Fourier frequencies. Table 1 shows the values of test statistics (3) together with the p values obtained by the permutation principle using 50,000 random permutations.

We see that the null hypothesis H_0 is rejected by all test statistics and that the p values do not substantially differ from either the different test statistics or the model with one, two, or three smallest Fourier frequencies. Moreover, all tests indicate the changepoint to occur in the year 1895, corresponding to $k^\star = 121$.

Figure 7 presents the estimated mean annual cycle before the year 1895 ($k^\star = 121$), and after it, using three smallest Fourier frequencies. We see that the mean annual cycle has mostly been changed in winter periods, where the increase in temperature is larger than in the summer periods. Figure 7 presents the estimated temperature difference before and after the year 1895. We again see that the estimated temperature difference after the year 1895 is smaller than the range before this year.

4.4 | Method of principal components

In multivariate analysis, vectors with many components are often replaced by shorter vectors whose elements are linear combinations of the original components with the largest variances. More precisely, let $\boldsymbol{u}_1, \dots, \boldsymbol{u}_l$ be the eigenvectors corresponding to the l largest eigenvalues $\lambda_1, \dots, \lambda_l$ of the estimated variance–covariance matrix $\boldsymbol{\Sigma}$. We replace the vectors $\{\boldsymbol{DD}_i, i = 1, \dots, n\}$ by vectors $\{(\boldsymbol{u}_1^\top \boldsymbol{DD}_i, \dots, \boldsymbol{u}_l^\top \boldsymbol{DD}_i)^\top\}$ and look for a change in the mean of these new vectors. As

TABLE 2 Values of the test statistics T, TW, and MW for three, four, and five eigenvectors with corresponding p values (in brackets). The corresponding p values were computed using 50,000 random permutations

Test statistic	3 eig. v.	4 eig. v.	5 eig. v.
T	19.50	28.94	29.07
	(0.0068)	(3.2×10^{-4})	(5.2×10^{-4})
TW	4.87	7.23	7.27
	(0.0013)	(2×10^{-5})	(4×10^{-5})
MW	1.76	2.55	2.59
	(0.0012)	(1.2×10^{-4})	(2.6×10^{-4})

TABLE 3 Values of test statistics with corresponding p values for testing the stationarity of monthly deviation means. The corresponding p values were computed using 50,000 random permutations

	Test statistic		
	T	TW	MW
Value of test statistic	43.55	10.88	4.53
p Value	(2.4×10^{-4})	(2.0×10^{-5})	(1.2×10^{-4})

the variance–covariance matrix $\boldsymbol{\Sigma}$ is practically always unknown, so are its eigenvalues and eigenvectors. Therefore, for computing the test statistics, we replace the vectors $\{(\boldsymbol{u}_1^\top \boldsymbol{DD}_i, \dots, \boldsymbol{u}_l^\top \boldsymbol{DD}_i)^\top\}$ by vectors $\{(\widehat{\boldsymbol{u}}_1^\top \boldsymbol{DD}_i, \dots, \widehat{\boldsymbol{u}}_l^\top \boldsymbol{DD}_i)^\top\}$, where $\widehat{\boldsymbol{u}}_1, \dots, \widehat{\boldsymbol{u}}_l$ are eigenvectors that correspond to the l largest eigenvalues of the estimated variance–covariance matrix of $\{\boldsymbol{DD}_i\}$. The described method was suggested by Horváth et al. (1999) and discussed by Aston and Kirch (2012).

We applied the procedure based on the principal component approximation described above to the daily mean temperature deviations. Studying the eigenvectors of the sample covariance matrix of $\{\boldsymbol{DD}_i\}$ more closely, we observed that the eigenvectors corresponding to the largest eigenvalues put the largest weights to the winter days where the change is more apparent.

Table 2 presents the values of the test statistics together with the corresponding p values computed using 50,000 random permutations for $l = 3, 4$, and 5 eigenvectors. Notice that the null hypothesis is rejected by all three types of statistics and that the p values do not differ substantially. Moreover, all maximum-type tests indicate the changepoint to occur in the year 1896 ($k^\star = 122$). We can see that this method is able to discover nonstationarities of the mean annual cycle even for relatively small values of l.

4.5 | Monthly means

For the mean annual cycle nonstationarity detection, we have analyzed the monthly deviations instead of the daily deviations from the mean annual temperature. The monthly deviations are linear combinations of the daily deviations, where the constants of these linear combinations are equal to 1 for the days belonging to the chosen month, whereas they are equal to 0 for all other days. The test statistics are based on the vector $\{\boldsymbol{MD}_i, i = 1, \dots, n\}$ with $\boldsymbol{MD}_i = (MD_{i1}, \dots, MD_{i,11})^\top$, where MD_{ij} is the monthly deviation of the jth month in the ith year.

We have applied this procedure for an abrupt-shift detection to monthly deviations. Table 3 presents the values of the test statistics together with the corresponding p values calculated using 50,000 random permutations. The results are again in good agreement, and the maximum of both maximum test statistics is attained in the year 1892 ($k^\star = 118$).

5 | BEHAVIOR OF THE SEASONAL CYCLE

All tests rejected the null hypothesis claiming that an increase in temperature is the same in all days during a calendar year. Moreover, the results suggested that an increase is larger in winter days than in summer days. Therefore, for all years, we computed a mean summer temperature by averaging temperatures of July, August, and September days and a

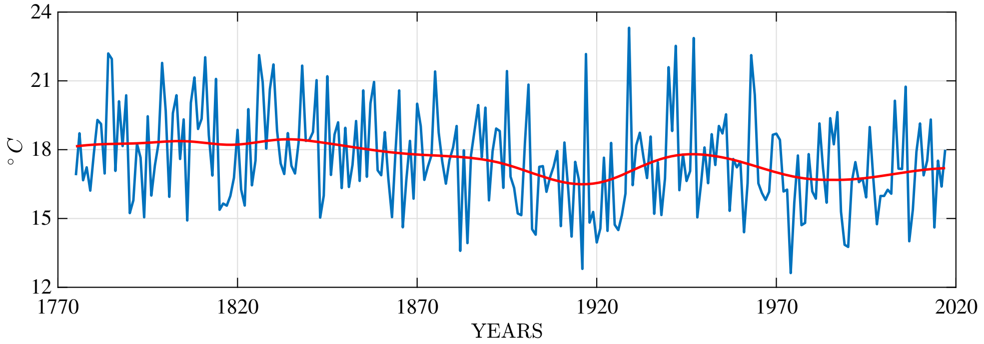

FIGURE 8 Difference between the mean summer and mean winter temperatures

mean winter temperature by averaging temperatures of January, February, and March days. Figure 8 presents a difference between the mean summer temperatures and mean winter temperatures for all years together with its kernel estimate obtained using a Gaussian filter. We see that the difference between the summer and winter temperatures is smaller in the 20th century than in the 19th century, which might be a reason why the stationarity of a seasonal cycle was rejected.

6 | CONCLUSIONS

We have studied the stationarity of the annual mean series and the stationarity of the seasonal cycle of the Klementinum temperature series. The basic tests for detecting a change clearly showed that the annual mean series is nonstationary. The study continued by modeling an expected behavior of the annual mean series by piecewise linear functions. Our goal was also to show that the changepoint estimates differ substantially whether the continuous or discontinuous piecewise linear model is assumed and according to how many changepoints the model contains. The choice of the model is given by our prior belief on how the expected mean annual series behaves, that is, whether it varies smoothly or may contain sudden jumps. If not only stochastic model is chosen according to our belief on the data-generating process but some additional knowledge on the parameters is available, then the application of the Bayesian methods applied to changepoint estimation can provide us with appealing results.

The analysis of the seasonal cycle showed that, in spite of the fact that the tests for an existence of a changepoint detected nonstationary behavior of all monthly mean temperature series, the change has not a form of a shift of the expected seasonal cycle. Rather, an increase in temperature in winter days is larger than an increase in summer days so that a difference between summer and winter days decreases.

In addition to the study of nonstationarities in the mean of the annual mean temperature series and in the mean of seasonal cycle, we may be interested in changes of variability, dependence structure, or some other characteristics of the Klementinum series. The more detailed analysis would be interesting to discuss as future research avenues.

ACKNOWLEDGEMENTS

The research of D. Jarušková has been performed in the Center of Advanced Applied Sciences and financially supported by the European Regional Development Fund No. CZ.02.1.01/0.0/0.0/16_019/0000778. The work of J. Antoch was partially supported by Grant P403/19/02773S. The authors are grateful to the associated editor and two unknown reviewers for their valuable comments that considerably improved the contents of this paper.

ORCID

J. Antoch https://orcid.org/0000-0002-5970-0306

REFERENCES

Antoch, J., Gregoire, G., & Hušková, M. (2007). Tests for continuity of regression functions. *Journal of Statistical Planning and Inference*, *137*, 753–777.

Antoch, J., & Hušková, M. (2000). Bayesian-type estimators of change points. *Journal of Statistical Planning and Inference*, *91*, 195–208.

Antoch, J., & Hušková, M. (2001). Permutation tests in change point analysis. *Statistics and Probability Letters*, *53*, 37–46.

Antoch, J., Hušková, M., & Jarušková, D. (2002). Off-line statistical process control. In C. Lauro, J. Antoch, V. E. Vinzi, & G. Saporta (Eds.), *Multivariate total quality control* (pp. 1–86). Heidelberg, Germany: Physica-Verlag.

Antoch, J., Hušková, M., & Prášková, Z. (1997). Effect of dependence on statistics for determination of change. *Journal of Statistical Planning and Inference*, *60*, 291–310.

Antoch, J., & Jarušková, D. (2013). Testing for multiple change points. *Computational Statistics*, *28*, 2161–2183.

Antoch, J., & Legát, D. (2008). Application of MCMC to change point detection. *Applications of Mathematics*, *53*, 281–296.

Arrhenius, S. (1896). On the influence of carbonic acid in the air upon the temperature of the ground. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *41*, 237–276.

Aston, J. A. D., & Kirch, C. (2012). Detecting and estimating changes in dependent functional data. *Journal of Multivariate Analysis*, *109*, 204–220.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, *66*, 47–78.

Barry, D., & Hartigan, J. (1993). A Bayesian analysis for change-point problems. *Journal of American Statistical Association*, *88*, 309–319.

Beaulieu, C., & Killick, R. (2018). Distinguishing trends and shifts from memory in climate data. *Journal of Climate*, *31*, 9519–9543.

Brázdil, R., Zahradníček, P., Pišoft, P., Štěpánek, P., Bělínová, M., & Dobrovolný, P. (2012). Temperature and precipitation fluctuations in the Czech Republic during the period of instrumental measurements. *Theoretical and Applied Climatology*, *110*, 17–34.

Camuffo, D., & Jones, P. (2002). *Improved understanding of past climatic variability from early daily European instrumental sources*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Carlin, B. P., Gelfand, A. E., & Smith, A. F. M. (1992). Hierarchical Bayesian analysis of changepoint problems. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *41*, 389–405.

Csörgő, M., & Horváth, L. (1997). *Limit theorems in change-point analysis*. New York, NY: John Wiley & Sons.

Drijfhout, S., Bathiany, S., Beaulieu, C., Brovkin, V., Claussen, M., Huntingford, C., … Swingedouw D. (2015). Catalogue of abrupt shifts in Intergovernmental Panel on Climate Change climate models. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, E5777–E5786.

El-Shaarawi, A. H. (2013). Changepoint methods, *The Encyclopedia of Environmetrics*. New York, NY: John Wiley & Sons.

Fearnhead, P. (2006). Exact and efficient Bayesian inference for multiple changepoint problems. *Statistics and Computing*, *16*, 203–213.

Feder, P. I. (1975). On asymptotic distribution theory in segmented regression problems – identified case. *The Annals of Statistics*, *3*, 49–83.

Green, P. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, *82*, 711–732.

Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., Alexander, L. V., Brönnimann, S., Abdul, Y., … Zhai, P. (2013). Observations: Atmosphere and surface. In Stocker, T. F. (Ed.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom: Cambridge University Press.

Horváth, L., Kokoszka, P., & Steinebach, J. (1999). Testing for changes in multivariate dependent observations with an application to temperature changes. *Journal of Multivariate Analysis*, *68*, 96–119.

Jandhyala, V. K., Fotopoulos, S. B., & El-Shaarawi, A. H. (2006). *Change-point methods and their environmental applications*, The Encyclopedia of Environmetrics. New York, NY: John Wiley & Sons.

Jandhyala, V. K., Fotopoulos, S. B., MacNeill, I., & Liu, P. (2013). Inference for single and multiple change-points in time series. *Journal of Time Series Analysis*, *34*, 423–446.

Jandhyala, V. K., & MacNeill, I. B. (1991). Test for parameter changes at unknown times in linear regression models. *Journal of Statistical Planning and Inference*, *27*, 291–316.

Jarušková, D. (1997). Some problems with application of change-point detection methods to environmental data. *Environmetrics*, *8*, 469–483.

Jarušková, D. (1998). Change-point detection methods for dependent data and application to hydrology. *Istatistik: Journal of the Turkish Statistical Association*, *1*(2), 9–21.

Jarušková, D. (2018). Estimating non-simultaneous changes in the mean of vectors. *Metrika*, *81*, 721–743.

Kiefer, J. (1959). K-sample analogues of the Kolmogorov–Smirnov and Cramér–von Mises tests. *The Annals of Mathematical Statistics*, *30*, 420–447.

Kim, H.-J., Yu, B., & Feuer, E. J. (2008). Inference in segmented line regression: A simulation study. *Journal of Statistical Computation and Simulation*, *78*, 1087–1103.

Klein Tank, A. M. G., Wijngaard, J. B., Können, G. P., Böhm, R., Demarée, G., Gocheva, A., … Petrovic, P. (2002). Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *International Journal of Climatology*, *22*, 1441–1453. Data and metadata are available at http://www.ecad.eu/. For Klementinum data, see files: TG_STAID000027.txt, TN_STAID000027.txt, TX_STAID000027.txt and RR_STAID000027.txt.

MacNeill, I. B. (1974). Tests for change of parameter at unknown times and distributions of some related functionals on Brownian motion. *The Annals of Statistics*, *2*, 950–962.

Maidstone, R., Hocking, T., Rigaill, G., & Fearnhead, P. (2017). On optimal multiple changepoint algorithms for large data. *Statistics and Computing*, *27*, 519–533.

Perreault, L., Bernier, J., Bobée, B., & Parent, E. (2000). Bayesian change-point analysis in hydrometeorological time series. Parts 1 and 2. *Journal of Hydrology*, *235*, 221–263.

Reeves, J., Chen, J., Wang, X., Lund, R., & Reeves, Q. Q. (2007). A review and comparison of changepoint detection techniques for climate data. *Journal of Applied Meteorology and Climatology*, *46*, 900–915.

Stryhn, H. (1996). The location of the maximum of asymmetric two-sided Brownian motion with triangular drift. *Statistics and Probability Letters*, *29*, 279–284.

How to cite this article: Jarušková D, Antoch J. Changepoint analysis of Klementinum temperature series. *Environmetrics*. 2020;31:e2570. https://doi.org/10.1002/env.2570