Detecting relevant changes in time series models

Holger Dette
Ruhr-Universität Bochum, Germany

and Dominik Wied
Technische Universität Dortmund, Germany

[Received July 2014. Revised February 2015]

Summary. Most of the literature on change point analysis by means of hypothesis testing considers hypotheses of the form \( H_0 : \theta_1 = \theta_2 \) versus \( H_1 : \theta_1 \neq \theta_2 \), where \( \theta_1 \) and \( \theta_2 \) denote parameters of the process before and after a change point. The paper takes a different perspective and investigates the null hypotheses of no relevant changes, i.e. \( H_0 : \|\theta_1 - \theta_2\| \leq \Delta \), where \( \|\cdot\| \) is an appropriate norm. This formulation of the testing problem is motivated by the fact that in many applications a modification of the statistical analysis might not be necessary, if the difference between the parameters before and after the change point is small. A general approach to problems of this type is developed which is based on the cumulative sum principle. For the asymptotic analysis weak convergence of the sequential empirical process must be established under the alternative of non-stationarity, and it is shown that the resulting test statistic is asymptotically normally distributed. The results can also be used to establish similarity of the parameters, i.e. \( H_1 : \|\theta_1 - \theta_2\| \leq \Delta \), at a controlled type 1 error and to estimate the magnitude \( \|\theta_1 - \theta_2\| \) of the change with a corresponding confidence interval. Several applications of the methodology are given including tests for relevant changes in the mean, variance, parameter in a linear regression model and distribution function among others. The finite sample properties of the new tests are investigated by means of a simulation study and illustrated by analysing a data example from portfolio management.

Keywords: Change point analysis; Cumulative sum; Precise hypotheses; Relevant changes; Strong mixing; Weak convergence under the alternative

1. Introduction

The analysis of structural breaks in a sequence \((Z_t)_{t=1}^n\) of random variables has a long history. Early work on this problem can be found in Page (1954, 1955) who investigated quality control problems. Since these seminal papers numerous researchers have worked on the problem of detecting structural breaks or change points in various statistical models (see Chow (1960), Brown et al. (1975) and Krämer et al. (1988), among others). Usually methodology is firstly developed for independent observations and—in a second step—extended to more complex dependent processes. Prominent examples of change point analysis are the detection of instabilities in mean and variance (see Horváth et al. (1999) and Aue, Hörmann, Horváth and Reimherr (2009) among others). These results have been extended to more complex regression models (see Andrews (1993) and Bai and Perron (1998)) and to change point inference on the second-order characteristics of a time series (see Berkes et al. (2009), Wied et al. (2012) and Preuss et al. (2014)). Quite an extensive list of references can be found in the recent work of Aue and Horváth...
(2013) who described how popular procedures investigated under the assumption of independent observations can be modified to analyse structural breaks in data exhibiting serial dependence.

A large portion of the literature attacks the problem of structural breaks by means of hypothesis testing instead of directly focusing on for example estimating the potential break points (see the introduction in Jandhyala et al. (2013)). Usually the hypothesis of no structural break is formulated as

$$H_0 : \theta_1 = \theta_2 = \ldots = \theta_n$$

(1.1)

where $\theta(t)$ denotes a (not necessarily finite dimensional) parameter of the distribution of the random variable $Z_t$ ($t = 1, \ldots, n$), such as the mean and variance. The alternative is then formulated (in the simplest case of one structural break) as

$$H_1 : \theta_1 = \theta_2 = \ldots = \theta_k \neq \theta_{k+1} = \theta_{k+2} = \ldots = \theta_n = \theta_2,$$

(1.2)

where $k \in \{1, \ldots, n\}$ denotes the (unknown) location of the change point. If the null hypothesis of structural breaks has been rejected, the location of the change must be estimated (see Csörgő and Horváth (1997) or Bai and Perron (1998) among others) and the statistical analysis must be modified to address the different stochastic properties before and after the change point.

The present work is motivated by the observation that such a modification of the statistical analysis might not be necessary if the difference between the parameters before and after the change point is quite small. For example, in risk management situations, one is interested in fitting a suitable model for forecasting value at risk from ‘uncontaminated data’, which means from data after the last change point (see for example Wied (2013)). But, in practice, small changes in the parameter are perhaps not very interesting because they do not yield large changes in the value at risk. The forecasting quality might only improve slightly, but this benefit could be negatively overcompensated by transaction costs. However, as an illustration with real interest rates at the end of this paper indicates, a relevant difference can potentially be linked to significant real world events. One could also think of an application to inflation rates in the sense that only ‘large’ changes call for interventions of, for example, the European Central Bank. With this point of view it might be more reasonable to replace hypothesis (1.2) by the null hypothesis of no relevant structural break, i.e.

$$H_0 : \|\theta_1 - \theta_2\| \leq \Delta \quad \text{versus} \quad H_1 : \|\theta_1 - \theta_2\| > \Delta,$$

(1.3)

where $\theta_1$ and $\theta_2$ are the parameters before and after the change point, $\| \cdot \|$ denotes a (semi)norm on the parameter space and $\Delta$ is a prespecified constant representing the ‘maximal’ change accepted by statisticians without modifying the statistical analysis. Note that this formulation of the change point problem avoids the consistency problem as mentioned in Berkson (1938), i.e. any consistent test will detect any arbitrary small change in the parameters if the sample size is sufficiently large. Moreover, the ‘classical’ formulation of the change point problem in formula (1.1) does not allow us to control the type II error if the null hypothesis of no structural break cannot be rejected, and as a consequence the statistical uncertainty in the subsequent data analysis (under the assumption of stationarity) cannot be quantified. In contrast, a decision of ‘no small structural’ break at a controlled type I error can be easily achieved by interchanging the null hypothesis and alternative in expression (1.3), i.e.

$$H_0 : \|\theta_1 - \theta_2\| > \Delta \quad \text{versus} \quad H_1 : \|\theta_1 - \theta_2\| \leq \Delta.$$

(1.4)

The new approach requires the specification of the quantity $\Delta > 0$, which depends on the specific application. ‘Classical’ hypotheses tests simply use $\Delta = 0$, but we argue that from a practical point of view it might be very reasonable to think about this choice more carefully and to define the size
of change in which one is really interested. The relevance of testing hypotheses of the form (1.3), which are also called *precise hypotheses* in the literature (see Berger and Delampady (1987)), has nowadays been widely recognized in various fields of statistical inference including medical, pharmaceutical, chemistry or environmental statistics (see Chow and Liu (1992) and McBride (1999)). However—to our best knowledge—the problem of testing for relevant structural breaks has not been discussed in the literature so far.

In this paper we present a general approach to address this problem, which is based on the cumulative sum (CUSUM) principle. The basic ideas are illustrated in Section 2 for the problem of detecting a relevant change in the mean of a multivariate sequence of independent observations. The general methodology is introduced in Section 3 and is applicable to several other situations including changes in the variance, the parameter in regression models and changes in the distribution function (the non-parametric change point problem). Additionally, if it is difficult to specify the threshold $\Delta$, it can be used to estimate the magnitude $\|\theta_1 - \theta_2\|$ with a corresponding confidence interval.

It turns out that—in contrast with the classical change point problem—testing relevant hypotheses of the type (1.3) requires results on the weak convergence of the sequential empirical process under non-stationarity (more precisely under the alternative $H_1$), which—to our best knowledge—have not been developed so far. The reference which is most similar in spirit to investigations of this type is Zhou (2013), who considered the asymptotic properties of tests for the classical hypothesis of a change in the mean, i.e. $H_0 : \mu_1 = \mu_2$, under piecewise local stationarity. The present paper takes a different and more general perspective using weak convergence of the sequential empirical process in the case $\theta_1 \neq \theta_2$. These asymptotic properties depend sensitively on the dependence structure of the basic time series $(Z_t) \in \mathbb{Z}$ and are developed in section A.1 of the on-line supplement for the concept of strong mixing triangular arrays. Although the analysis of the sequential process under non-stationarities of the type (1.2) is very complicated, the resulting test statistics for the hypothesis of *no relevant structural break* have a very simple asymptotic distribution, namely a normal distribution. Consequently, statistical analysis can be performed by estimating a variance and using quantiles of the standard normal distribution.

In Section 4 we illustrate the methodology and develop tests for the hypothesis (1.3) of a relevant change in the mean, parameters in a linear regression model and distribution function. In particular, we consider the situation of testing for a change in the mean with possibly simultaneously changing variance, which occurs frequently in applications. Note that none of the classical change point tests can address this problem. In fact it was pointed out by Zhou (2013) that the classical CUSUM approach and similar methods are not pivotal in this case leading to severe biased testing results. Section 5 presents some finite sample evidence of the new test revealing appealing size and power properties. We also give an illustration in a data example from portfolio management. In the on-line appendix we provide some theoretical results, which demonstrate that the assumptions that are made in Section 3 are satisfied for strong mixing processes, give some of the more technical proofs and discuss the problem of detecting relevant changes in the variance and correlation.

The programs that were used to analyse the data can be obtained from

http://wileyonlinelibrary.com/journal/rss-datasets

2. Relevant changes in the mean—motivation

This section serves as a motivation for the general approach to detect relevant changes in time series which will be discussed in Section 3. For illustration we consider independent $d$-dimen-
sional random variables \(Z_1, \ldots, Z_n\) with common positive definite variance \(\text{var}(Z_i) = \Sigma\), such that for some unknown \(t \in (0, 1)\)

\[
\mu_1 = \mathbb{E}[Z_1] = \ldots = \mathbb{E}[Z_{[nt]}]; \mathbb{E}[Z_{[nt]+1}] = \ldots = \mathbb{E}[Z_n] = \mu_2.
\]

The case of a variance simultaneously changing with the mean will be discussed in Section 4. We are interested in the problem of testing for a relevant change in the mean, i.e.

\[
H_0 : \| \mu_1 - \mu_2 \| \leq \Delta \quad \text{versus} \quad H_1 : \| \mu_1 - \mu_2 \| > \Delta,
\]

(2.1)

where \(\| \cdot \|\) denotes the Euclidean norm on \(\mathbb{R}^d\). For this we consider the CUSUM statistic \(\{ \hat{\mathbb{U}}_n(s) \}_{s \in [0, 1]}\) defined by

\[
\hat{\mathbb{U}}_n(s) = \frac{1}{n} \sum_{j=1}^{[ns]} Z_j - \frac{s}{n} \sum_{j=1}^{n} Z_j = \frac{1}{n} \sum_{j=1}^{[ns]} Z_j - \frac{s}{n} \sum_{j=1}^{n} Z_j.
\]

A straightforward computation gives \(\mathbb{E}[\hat{\mathbb{U}}_n(s)] = (s - t)(\mu_1 - \mu_2)\{1 + o(1)\}\). A similar calculation yields

\[
\mathbb{E}[\| \hat{\mathbb{U}}_n(s) \|^2] = \left\{ \frac{\sigma^2}{n} s(1-s) + \| \mu_1 - \mu_2 \|^2 (s - t)^2 \right\} \{1 + o(1)\}.
\]

Consequently, we obtain

\[
\mathbb{E} \left[ \int_0^1 \| \hat{\mathbb{U}}_n(s) \|^2 \, ds \right] = \left\{ \int_0^1 \frac{\sigma^2}{n} s(1-s) + \| \mu_1 - \mu_2 \|^2 (s - t)^2 \, ds \right\} \{1 + o(1)\} = \left[ \frac{\sigma^2}{6n} + \| \mu_1 - \mu_2 \|^2 \frac{(t(1-t))^2}{3} \right] \{1 + o(1)\},
\]

(2.2)

and therefore it is reasonable to consider the statistic

\[
\frac{3}{(t(1-t))^2} \int_0^1 \| \hat{\mathbb{U}}_n(s) \|^2 \, ds
\]

as an estimator of the distance \(\| \mu_1 - \mu_2 \|^2\) (a bias correction addressing the term \(\sigma^2/(6n)\) will be discussed later). The following result specifies the asymptotic properties of this statistic. Throughout this paper the symbol \(\Rightarrow^D\) means weak convergence in the appropriate space under consideration.

**Theorem 1.** For any \(t \in (0, 1)\) we have as \(n \to \infty\)

\[
L_n = \sqrt{n} \left[ \frac{3}{(t(1-t))^2} \int_0^1 \| \hat{\mathbb{U}}_n(s) \|^2 \, ds - \| \mu_1 - \mu_2 \|^2 \right] \Rightarrow^D \mathcal{N}(0, \tau^2),
\]

(2.3)

where the asymptotic variance is given by

\[
\tau^2 = (\mu_1 - \mu_2)^T \Sigma (\mu_1 - \mu_2) \frac{4(1 + 2t(1-t))}{5t^2(1-t)^2}.
\]

(2.4)

**Proof:** We start calculating the covariance \(\text{cov}\{\hat{\mathbb{U}}_n(s_1), \hat{\mathbb{U}}_n(s_2)\}\) by using the decomposition

\[
\hat{\mathbb{U}}_n(s) = (1-s) \hat{\mathbb{U}}_n^{(1)}(s) - s \hat{\mathbb{U}}_n^{(2)}(s),
\]

where
\[
\mathcal{U}_n^{(1)}(s) = \frac{1}{n} \sum_{j=1}^{[ns]} Z_j,
\]
\[
\mathcal{U}_n^{(2)}(s) = \frac{1}{n} \sum_{j=\lfloor ns \rfloor + 1}^{n} Z_j.
\]

For this we first assume that \(s_1 \leq s_2\) and note that \(\text{cov}\{\mathcal{U}_n^{(1)}(s_1), \mathcal{U}_n^{(2)}(s_2)\} = 0\) in this case. Moreover, the remaining covariances are obtained as follows:
\[
\text{cov}\{\mathcal{U}_n^{(1)}(s_1), \mathcal{U}_n^{(1)}(s_2)\} = \frac{1}{n^2} \sum_{j,k=1}^{[ns_2]} \text{cov}(Z_j, Z_k) = \frac{s_1}{n} \Sigma\{1 + o(1)\},
\]
\[
\text{cov}\{\mathcal{U}_n^{(1)}(s_2), \mathcal{U}_n^{(2)}(s_1)\} = \frac{1}{n^2} \sum_{j=1}^{[ns_2]} \sum_{k=\lfloor ns_1 \rfloor + 1}^{n} \text{cov}(Z_j, Z_k) = \frac{s_2 - s_1}{n} \Sigma\{1 + o(1)\},
\]
\[
\text{cov}\{\mathcal{U}_n^{(2)}(s_1), \mathcal{U}_n^{(2)}(s_2)\} = \frac{1}{n^2} \sum_{j=\lfloor ns_1 \rfloor + 1}^{n} \sum_{s=\lfloor ns_2 \rfloor + 1}^{n} \text{cov}(Z_j, Z_k) = \frac{1 - s_2}{n} \Sigma\{1 + o(1)\},
\]
which gives
\[
\text{cov}\{\hat{\mathcal{U}}_n(s_1), \hat{\mathcal{U}}_n(s_2)\} = \frac{s_1(1 - s_2)}{n} \Sigma\{1 + o(1)\}
\]
if \(s_1 \leq s_2\). A similar calculation for the case \(s_1 \geq s_2\) finally yields
\[
\lim_{n \to \infty} n \text{cov}\{\hat{\mathcal{U}}_n(s_1), \hat{\mathcal{U}}_n(s_2)\} = (s_1 \land s_2 - s_1 s_2) \Sigma.
\]

It can be shown (note that for illustration the random variables \(Z_1, \ldots, Z_n\) are assumed to be independent and a corresponding statement under the assumption of a strong mixing process is given in the on-line appendix A.1) that an appropriately standardized version of the process \(\hat{\mathcal{U}}_n\) converges weakly, i.e.
\[
\{\sqrt{n}\{\hat{\mathcal{U}}_n(s) - \mu(s,t)\}\}_{s \in [0,1]} \overset{\mathcal{D}}{\to} \Sigma^{1/2}\{B(s)\}_{s \in [0,1]},
\]
where \(\mu(s,t) = (s \land t - st)(\mu_1 - \mu_2)\) and \(B\) denotes a vector of independent Brownian bridges on the interval \([0,1]\). This gives for the random variable \(L_n\) in equation (2.3)
\[
L_n = \frac{3}{\{t(1-t)\}^2} \left[ \int_{0}^{1} \|\hat{\mathcal{U}}_n(s)\|^2 ds - \|\mu_1 - \mu_2\|^2 \frac{\{t(1-t)\}^2}{3} \right]
\]
\[
= \frac{3}{\{t(1-t)\}^2} \left[ \int_{0}^{1} \left\{ \|\hat{\mathcal{U}}_n(s)\|^2 - \|\mu(s,t)\|^2 \right\} ds \right]
\]
\[
= \frac{3}{\{t(1-t)\}^2} \left[ \int_{0}^{1} \|\hat{\mathcal{U}}_n(s) - \mu(s,t)\|^2 ds + 2 \int_{0}^{1} \mu^T(s,t)\{\hat{\mathcal{U}}_n(s) - \mu(s,t)\} ds \right]
\]
\[
\overset{\mathcal{D}}{\to} \frac{6}{\{t(1-t)\}^2} \int_{0}^{1} \mu^T(s,t)\Sigma^{1/2} B(s) ds.
\]
It is well known that the distribution on the right-hand side is a centred normal distribution with variance
\[
\frac{36}{\{t(1-t)\}^4} \int_{0}^{1} \int_{0}^{1} \mu^T(s_1, t)\Sigma\mu(s_2, t)(s_1 \land s_2 - s_1 s_2) ds_1 ds_2,
\]
and it follows by a straightforward calculation that this expression is given in equation (2.4).

The test statistic for the hypothesis (2.1) is finally defined as

$$
\hat{M}_n^2 = \frac{3}{i(1-i)^2} \int_0^1 \|\hat{U}_n(s)\|^2 ds - \frac{\sigma^2}{6n}
$$

where $\hat{i}$ and $\hat{\sigma}^2$ are consistent estimators of $i$ and $\sigma^2$ respectively. This definition corrects for the additional bias in equation (2.2) which is asymptotically negligible. The null hypothesis of no relevant change point is finally rejected, whenever

$$
\hat{M}_n^2 \geq \Delta^2 + u_{1-\alpha} \frac{\hat{\tau}}{\sqrt{n}},
$$

where $u_{1-\alpha}$ is the $(1-\alpha)$-quantile of the standard normal distribution and $\hat{\tau}$ is an appropriate estimator of $\tau$. An estimator of the change point can be obtained by the argmax principle, i.e. $\hat{i} = \arg \max_{s \in [0,1]} \|\hat{U}_n(s)\|$ (see Carlstein (1988)). For the estimation of the residual variance we denote by

$$
\hat{\mu}_1 = \frac{1}{[n\hat{i}]} \sum_{i=1}^{[n\hat{i}]} Z_i,
$$

$$
\hat{\mu}_2 = \frac{1}{[(1-\hat{i})n]} \sum_{i=[n\hat{i}]+1}^{n} Z_i
$$

the estimates of the mean ‘before’ and ‘after’ the change point and define a variance estimator by

$$
\hat{\Sigma}_1 = \frac{1}{n} \left\{ \sum_{i=1}^{[n\hat{i}]} (Z_i - \hat{\mu}_1)(Z_i - \hat{\mu}_1)^T + \sum_{i=[n\hat{i}]+1}^{n} (Z_i - \hat{\mu}_2)(Z_i - \hat{\mu}_2)^T \right\}.
$$

This yields

$$
\hat{\tau} = \sqrt{\frac{2\hat{\nu}}{5\hat{i}(1-\hat{i})}} \sqrt{\{1 + 2\hat{i}(1-\hat{i})\}^2}
$$

as an estimation of $\tau$, where $\hat{\nu}^2 = (\hat{\mu}_1 - \hat{\mu}_2)^T \hat{\Sigma}_1 (\hat{\mu}_1 - \hat{\mu}_2)$. It will be shown in Section 3 that the test defined by condition (2.5) is consistent and has asymptotic level $\alpha$.

Remark 1. Our motivation for considering a statistic based on the integral $\int_0^1 \|\hat{U}_n(s)\|^2 ds$ is twofold. On the one hand, we want to consider a CUSUM-type statistic, which has some optimality properties for testing the ‘classical’ change point hypothesis $H_0 : \mu_1 = \mu_2$ (see for example Lorden (1971) and the subsequent literature for early references and Moustakides (2004) for a more recent reference). On the other hand, we are interested in a test statistic with a simple limit distribution, such as a normal distribution. This allows us to use classical results on uniformly most powerful unbiased tests for interval hypotheses in exponential families (see Lehmann (1986), chapter 7) to derive powerful tests for the hypothesis of a relevant change point.

As pointed out by a referee, an alternative test could be based on the statistic

$$
\hat{\Theta}_n = \max_{0 < \epsilon < 1} \left\| \frac{1}{[ns]} \sum_{j=1}^{[ns]} Z_j - \frac{1}{n-[ns]} \sum_{j=[ns]+1}^{n} Z_j \right\|,
$$

which estimates the jump size directly. The null hypothesis is rejected for large values of $\hat{\Theta}_n$. However, the asymptotic distribution of the statistic $\hat{\Theta}_n$ (appropriately standardized) is not
known in the case \( \mu_1 \neq \mu_2 \) and, as a consequence, the classical uniformly most powerful unbiased test theory for interval hypotheses is not applicable here.

3. Testing for relevant changes—a general approach

3.1. General formulation of the problem

Let \( Z_1, \ldots, Z_n \) denote \( d \)-dimensional random variables such that

\[
\begin{align*}
Z_1, \ldots, Z_{[n]} & \sim F_1, \\
Z_{[n]+1}, \ldots, Z_n & \sim F_2,
\end{align*}
\]  

(3.1)

where \( F_1 \) and \( F_2 \) denote continuous distribution functions before and after the change point. Let \( S \) be a Hilbert space with (semi)norm \( \| \cdot \| \), define \( L^\infty(\mathbb{R}^d|S) \) as the set of all bounded functions \( g: \mathbb{R}^d \rightarrow S \) and consider \( \mathcal{F} \subset L^\infty(\mathbb{R}^d|S) \). We denote by

\[
\theta: \mathcal{F} \rightarrow S; F \rightarrow \theta(F)
\]  

(3.2)

a given function defining the parameter of interest. Typical examples include the mean \( \theta(F) = \int z \, dF \) or the distribution function (here \( \theta \) is the identity map). We are interested in testing the hypothesis of no relevant change in the functional \( \theta(F) \), i.e.

\[
\begin{align*}
H_0 &: \| \theta(F_1) - \theta(F_2) \| \leq \Delta, \\
H_1 &: \| \theta(F_1) - \theta(F_2) \| > \Delta,
\end{align*}
\]  

(3.3)

where \( \Delta > 0 \) is a prespecified constant. If \( S \subset \mathbb{R}^k \) with \( k \leq d \), then \( \| \cdot \| \) denotes always the Euclidean norm, if not specified otherwise.

Our general approach will be based on an estimator of the distance \( \| \theta(F_1) - \theta(F_2) \|^2 \) by a CUSUM-type statistic. For this we assume for a moment linearity of the functional \( \theta \) in expression (3.2), i.e.

\[
\theta(\alpha F_1 + \beta F_2) = \alpha \theta(F_1) + \beta \theta(F_2)
\]  

(3.4)

for all \( \alpha, \beta \in \mathbb{R}, F_1, F_2 \in \mathcal{F} \). We introduce the statistic

\[
\hat{F}_n(s, z) = \frac{1}{n} \sum_{j=1}^{[ns]} I \{ Z_j \leq z \},
\]  

(3.5)

where \( s \in [0, 1] \), \( z \in \mathbb{R}^d \) and the inequality is understood componentwise. For fixed \( s \in (0, 1] \) the function \( (n/[ns])\hat{F}(s, \cdot) \) is a distribution function and a straightforward calculation yields

\[
\lim_{n \rightarrow \infty} \mathbb{E}[\hat{F}_n(s, z)] = E_{F_1, F_2,t}(s, z) := (s \wedge t) F_1(z) + (s - t) \mathbb{V} F_2(z).
\]  

(3.6)

We also introduce the function

\[
Z_{F_1, F_2,t}(s, z) := E_{F_1, F_2,t}(s, z) - s E_{F_1, F_2,t}(1, z) = (s \wedge t - st) \{ F_1(z) - F_2(z) \}
\]  

(3.7)

and note that \( Z_{F_1, F_2,t} \) vanishes on \([0, 1] \times \mathbb{R}^d \) if and only if \( F_1 = F_2 \). If \( \Phi_{\text{lin}}: L^\infty([0, 1] \times \mathbb{R}^d|\mathbb{R}) \rightarrow L^\infty([0, 1]|S) \) denotes the (linear) operator defined by

\[
\Phi_{\text{lin}}(E_{F_1, F_2,t})(s) := \theta \{ E_{F_1, F_2,t}(s, \cdot) - s E_{F_1, F_2,t}(1, \cdot) \} = \theta(Z_{F_1, F_2,t})(s),
\]  

we obtain from expression (3.4) and (3.7) for the function \( \mathbb{U} := \Phi_{\text{lin}}(E_{F_1, F_2,t}) \) the representation

\[
\mathbb{U}(s) := \Phi_{\text{lin}}(E_{F_1, F_2,t})(s) = (s \wedge t - st) \{ \theta(F_1) - \theta(F_2) \}.
\]  

(3.8)
Consequently, the norm of this function is given by
\[
\mathbb{T}^2(s) = \|\cup(s)\|^2 = \|\theta \{ Z_{F_1, F_2}(s, \cdot) \}\|^2 = (s \wedge t - st)^2 \|\theta(F_1) - \theta(F_2)\|^2,
\] (3.9)
which can be used as the basis for estimating the distance between the parameters \(\theta(F_1)\) and \(\theta(F_2)\). Before we explain the construction of this estimate in more detail, we ‘remove’ assumption (3.4) and consider more general non-linear functionals.

In this case the situation is slightly more complicated and we assume throughout this paper that there is a mapping
\[
\Phi : L^\infty([0, 1] \times \mathbb{R}^d ; \mathbb{R}) \rightarrow L^\infty([0, 1] \| S),
\] (3.10)
such that the difference between \(\theta(F_1)\) and \(\theta(F_2)\) can be expressed as a functional of the function \(E_{F_1, F_2, t}\) in equation (3.6), i.e.
\[
\cup(s) := \Phi(E_{F_1, F_2, t})(s) = (s \wedge t - st) \{ \theta(F_1) - \theta(F_2) \}.
\] (3.11)

For linear functionals such a representation is obvious as shown in the preceding paragraph. Other examples where assumption (3.11) is satisfied include linear regression models or the detection of relevant changes in the correlation and will be discussed in Section 4 and in section C.2 of the on-line supplement.

For the construction of an estimate of \(\|\theta(F_1) - \theta(F_2)\|^2\) we note that it follows by similar arguments to those given in Section 2 that the function \(\mathbb{T}(s) = \|\cup(s)\|\) satisfies
\[
\int_0^1 \mathbb{T}^2(s) \, ds = \int_0^1 (s \wedge t - st)^2 \|\theta(F_1) - \theta(F_2)\|^2 \, ds = \frac{t(1-t)^2}{3} \|\theta(F_1) - \theta(F_2)\|^2.
\] (3.12)

Observing equations (3.9) and (3.12) we see that the distance
\[
M^2 = M^2(F_1, F_2) = \|\theta(F_1) - \theta(F_2)\|^2 = \frac{3}{t(1-t)^2} \int_0^1 \|\Phi(E_{F_1, F_2, t}(s))\|^2 \, ds
\] (3.13)
between the parameters \(\theta(F_1)\) and \(\theta(F_2)\) can be expressed as a functional of \(E_{F_1, F_2, t}(\cdot, \cdot)\), which can easily be estimated by a sequential empirical process \(\hat{\mathbb{T}}_n\) defined in equation (3.5). The null hypothesis (3.3) is then rejected for large values of this estimator. In the following discussion we shall derive the asymptotic properties of this estimator, which can be used for the calculation of critical values for a test of the null hypothesis (3.3) of no relevant change.

### 3.2. Estimating the distance \(M(F_1, F_2) = \|\theta(F_1) - \theta(F_2)\|\)

To estimate the distance \(M^2(F_1, F_2) = \|\theta(F_1) - \theta(F_2)\|^2\) we recall the definition of the sequential empirical process (3.5) and its asymptotic expectation \(E_{F_1, F_2, t}\) defined in expression (3.6). Observing assumption (3.11) we consider the processes
\[
\hat{\cup}_n(s) = \Phi(\hat{\mathbb{T}}_n(s, \cdot)),
\] (3.14)
\[
\hat{\mathbb{T}}_n^2(s) = \|\hat{\cup}_n(s)\|^2 = \|\Phi(\hat{\mathbb{T}}_n(s, \cdot))\|^2.
\] (3.15)

Note that \(\hat{\cup}_n\) and \(\hat{\mathbb{T}}_n\) are \(S\)- and \(\mathbb{R}\)-valued processes. If \(S \subset \mathbb{R}^k\) we make the assumption
\[
\{\sqrt{n} \{ \hat{\cup}_n(s) - \cup(s) \}\}_{s \in [0, 1]} \overset{D}{\Rightarrow} \{D_{F_1, F_2, t}(s)\}_{s \in [0, 1]}
\] (3.16)
where \( \Rightarrow^D \) means weak convergence in \( L^\infty([0,1])\mathbb{R}^k \) and \( \{ D_{F_1,F_2,t}(s) \}_{s\in[0,1]} \) is a centred, \( k \)-dimensional Gaussian process with covariance kernel

\[
d_{F_1,F_2,t}(s_1,s_2) = \mathbb{E}[D_{F_1,F_2,t}(s_1)D_{F_1,F_2,t}(s_2)] \in \mathbb{R}^{k \times k}.
\]

Remark 2. The weak convergence results of the type (3.16) have been investigated for numerous types of stationary stochastic processes (see Horváth et al. (1999), Aue, Hörmann, Horváth and Reimherr (2009) or Dehling et al. (2013)). However, the detection of relevant change points by testing hypotheses of the form (3.3) requires weak convergence results in the non-stationary situation (3.1), for which—to our best knowledge—no results are available. In particular, as will be demonstrated in section A.1 of the on-line supplement, the distribution of the limiting processes \( D_{F_1,F_2,t} \) depends on the distribution functions \( F_1 \) and \( F_2 \) and the change point \( t \) in a complicated way. Only in the case \( F_1 = F_2 \) does it simplify to the standard situation, which is usually considered in change point analysis. Intuitively many results for stationary processes that are mentioned in the references cited at the beginning of this remark should also be available in the non-standard situation (3.1), but the limiting distribution is more complicated and this must be worked out for each case under consideration. In section A.1 of the on-line supplement we illustrate the arguments for this generalization in the case of a strong mixing process (satisfying condition (3.1)).

In the same section similar results will be established for the sequential process \( \hat{F}_{n,t} \), i.e.

\[
\{ \sqrt{n}\{ \hat{F}_{n,t}(s,z) - E_{F_1,F_2,t}(s,z) \} \}_{s\in[0,1],z\in\mathbb{R}^d} \Rightarrow^D \{ G_{F_1,F_2,t}(s,z) \}_{s\in[0,1],z\in\mathbb{R}^d}, \tag{3.17}
\]

where \( G_{F_1,F_2,t} \) denotes a centred \( (d+1) \)-dimensional Gaussian process on \([0,1] \times \mathbb{R}^d\) with covariance kernel

\[
g_{F_1,F_2,t}(s_1,z_1,s_2,z_2) = \mathbb{E}[G_{F_1,F_2,t}(s_1,z_1)G_{F_1,F_2,t}(s_2,z_2)] = k_t(s_1,s_2,z_1,z_2).
\]

Consequently, if the functional \( \Phi \) in equation (3.10) is (for example) Hadamard differentiable, weak convergence of the process \( \{ \sqrt{n}\{ \hat{U}_{n,t}(s) - \bar{U}(s) \} \}_{s\in[0,1]} \) is a consequence of the representation (3.14) and (3.17). Some details are given in remark 3 below. However, many functionals of interest in change point analysis (such as the mean or variance) do not satisfy this property, and for this reason we also state expression (3.16) as a basic assumption, which must be checked in concrete applications. An example where expression (3.17) can be used directly consists in the problem of detecting a relevant change in the distribution function and will be given in Section 4.

Theorem 2. If \( S \subset \mathbb{R}^k \) and the assumptions (3.16) and (3.4) are satisfied, then

\[
\sqrt{n}\left\{ \int_0^1 \hat{T}_n^2(s) \, ds - \int_0^1 \bar{T}^2(s) \, ds \right\} \Rightarrow^D \mathcal{N}(0,\sigma_{F_1,F_2,t}^2),
\]

where \( \bar{T}^2(s) \) and \( \int_0^1 \bar{T}^2(s) \, ds \) are given in equations (3.9) and (3.12) respectively. Here the asymptotic variance is given by

\[
\sigma_{F_1,F_2,t}^2 = 4(\theta(F_1) - \theta(F_2))^	op \Gamma(t,F_1,F_2) (\theta(F_1) - \theta(F_2)), \tag{3.18}
\]

where the matrix \( \Gamma \in \mathbb{R}^{k \times k} \) is defined by

\[
\Gamma(t,F_1,F_2) = \int_0^1 \int_0^1 (s_1 \wedge t - s_1)(s_2 \wedge t - s_2) \, dF_{1,F_2,t}(s_1,s_2) \, ds_1 \, ds_2. \tag{3.19}
\]

Proof. Let \( \langle \cdot, \cdot \rangle \) denote the inner product on \( \mathbb{R}^k \). Observing the representation
it follows from assumption (3.16) that
\[
\sqrt{n} \left\{ \int_0^1 \hat{\mathbf{T}}_n^2(s) \mathrm{d}s - \int_0^1 \mathbf{T}^2(s) \mathrm{d}s \right\} \overset{\mathcal{D}}{\rightarrow} 2 \left\{ \mathbb{E}(\mathbf{U}(s), \mathbb{D}_{F_1,F_2,t}(s)) \right\}_{s \in [0,1]}.
\]
Now the continuous mapping theorem yields
\[
\sqrt{n} \left\{ \int_0^1 \hat{\mathbf{T}}_n^2(s) \mathrm{d}s - \int_0^1 \mathbf{T}^2(s) \mathrm{d}s \right\} \overset{\mathcal{D}}{\rightarrow} 2 \int_0^1 \langle \mathbb{E}(\mathbf{U}(s), \mathbb{D}_{F_1,F_2,t}(s)) \rangle \mathrm{d}s,
\]
and standard arguments show that the random variable on the right-hand side is normally distributed with mean 0 and variance
\[
\sigma_{F_1,F_2,t}^2 = 4 \int_0^1 \int_0^1 \mathbb{E} \left[ \langle \mathbf{U}(s_1), \mathbb{D}_{F_1,F_2,t}(s_1) \rangle \langle \mathbf{U}(s_2), \mathbb{D}_{F_1,F_2,t}(s_2) \rangle \right] \mathrm{d}s_1 \mathrm{d}s_2
= 4(\theta(F_1) - \theta(F_2))^T \Gamma(t, F_1, F_2) (\theta(F_1) - \theta(F_2)).
\]

Remark 3. A similar statement can be derived under assumption (3.17) if the function \( \Phi \) in expression (3.11) is Hadamard differentiable at the point \( E_{F_1,F_2,t} \) (tangentially to an appropriate subset, if necessary). In this case it follows from expression (3.17) and the same arguments as given in the proof of theorem 2 that
\[
\sqrt{n} \left\{ \int_0^1 \hat{\mathbf{T}}_n^2(s) \mathrm{d}s - \int_0^1 \mathbf{T}^2(s) \mathrm{d}s \right\} \overset{\mathcal{D}}{\rightarrow} 2 \int_0^1 \langle \Phi' \{ G_{F_1,F_2,t}(s, \cdot) \}, \Phi \{ E_{F_1,F_2,t}(s, \cdot) \} \rangle \mathrm{d}s
\]
where \( \Phi' \) denotes the Hadamard derivative of \( \Phi \) and \( \langle \cdot, \cdot \rangle \) is the inner product on the (not necessarily finite dimensional) Hilbert space \( \mathcal{S} \). The details are omitted for brevity.

3.3. Testing for relevant changes
It follows from theorem 2 and expression (3.12) that
\[
\sqrt{n} \left[ \frac{3}{\{t(1-t)\}^2} \int_0^1 \hat{\mathbf{T}}_n^2(s) \mathrm{d}s - \| \theta(F_1) - \theta(F_2) \|^2 \right] \overset{\mathcal{D}}{\rightarrow} \mathcal{N}(0, \tau_{F_1,F_2,t}^2),
\]
where the asymptotic variance is given by
\[
\tau_{F_1,F_2,t}^2 = \frac{9 \sigma_{F_1,F_2,t}^2}{\{t(1-t)\}^4} = \frac{36(\theta(F_1) - \theta(F_2))^T \Gamma(t, F_1, F_2) (\theta(F_1) - \theta(F_2))}{\{t(1-t)\}^4}
\]
and \( \sigma_{F_1,F_2,t}^2 \) is defined in equation (3.18). In the following discussion let \( \hat{t} \) denote a consistent estimator of the change point \( t \), such that \( |\hat{t} - t| = o_p(1/\sqrt{n}) \), whenever \( \theta(F_1) \neq \theta(F_2) \) and \( \hat{t} \to \mathcal{D} T_{\text{max}} \) whenever \( \theta(F_1) = \theta(F_2) \), where \( T_{\text{max}} \) denotes a \([0,1]\)-valued random variable. Typically, the estimator \( \hat{t} = \text{arg max}_{s \in [0,1]} \| \hat{\mathbf{U}}_n(s) \| \) satisfies these assumptions with \( T_{\text{max}} = \text{arg max}_{s \in [0,1]} \| G(s) \| \) for some Gaussian process \( G \) (for a recent review on the relevant literature see Jandhyala et al. (2013)). Consequently, if \( \hat{\sigma}^2 \) is an estimator of \( \sigma_{F_1,F_2,t}^2 \), we obtain by \( \hat{\tau}^2 = 9\hat{\sigma}^2/\{\hat{t}(1-\hat{t})\}^4 \) an estimate of the asymptotic variance in equation (3.21). This yields for the statistic
\[
\hat{M}_n^2 = \frac{3}{\{\hat{t}(1-\hat{t})\}^2} \int_0^1 \hat{\mathbf{T}}_n^2(s) \mathrm{d}s
\]
the weak convergence
Detecting Changes in Time Series Models

\[
\sqrt{n} \left\{ \hat{M}_n^2 - \| \theta(F_1) - \theta(F_2) \|^2 \right\} \Rightarrow \mathcal{N}(0, 1),
\]  
(3.22)
whenever \( \theta(F_1) \neq \theta(F_2) \). However, if \( \theta(F_1) = \theta(F_2) \) we have

\[
\sqrt{n} \int_0^1 \hat{T}_n(s) \, ds \xrightarrow{p} 0.
\]  
(3.23)

**Theorem 3.** If assumption (3.22) is satisfied, then the test, which rejects the null hypothesis (3.3) of no relevant change, whenever

\[
\hat{M}_n^2 \geq \Delta^2 + u_{1-\alpha}\frac{\hat{\tau}}{\sqrt{n}},
\]  
(3.24)
is a consistent asymptotic level \( \alpha \) test.

**Proof.** Define \( \delta = \| \theta(F_1) - \theta(F_2) \| \) and assume that the null hypothesis \( \delta \leq \Delta \) holds. If \( \delta > 0 \) it follows from expression (3.22) that the probability of rejection by the decision rule (3.24) is given by

\[
\beta_n(\delta) = \mathbb{P}_\delta \left( \hat{M}_n^2 \geq \Delta^2 + u_{1-\alpha}\frac{\hat{\tau}}{\sqrt{n}} \right) = \mathbb{P}_\delta \left\{ \frac{\sqrt{n}(\hat{M}_n^2 - \delta^2)}{\delta} \geq \frac{\sqrt{n}(\Delta^2 - \delta^2)}{\delta} + u_{1-\alpha} \right\}
\]
\[
\leq \mathbb{P}_\delta \left\{ \frac{\sqrt{n}(\hat{M}_n^2 - \delta^2)}{\delta} \geq u_{1-\alpha} \right\} \xrightarrow{n \to \infty} \alpha.
\]  
(3.25)

Similarly, if \( \delta = 0 \) (which implies that \( \cup(s) \equiv 0 \)), we obtain from result (3.23)

\[
\beta(0) = \mathbb{P} \left\{ \sqrt{n} \int_0^1 \hat{T}_n(s) \, ds \geq \frac{\hat{T}^2(1-\hat{T})^2}{3} \left( \sqrt{n}\Delta^2 + u_{1-\alpha}\hat{\tau} \right) \right\} \xrightarrow{n \to \infty} 0.
\]  
(3.26)

Consequently, the test, which rejects the null hypothesis whenever condition (3.24) is satisfied, is an asymptotic level \( \alpha \) test. In contrast, under the alternative \( \delta > \Delta \), a similar argument shows that \( \beta_n(\delta) \xrightarrow{n \to \infty} 1 \), which proves consistency. \( \square \)

The choice of the estimators \( \hat{\tau}^2 \) and \( \hat{T} \) depends on specific examples under consideration and will be discussed in more detail in Section 5, where we illustrate the methodology by several examples.

**Remark 4.**

(a) It is worthwhile to mention that for the problem of testing the ‘classical’ hypothesis \( H_0 : \theta_1 = \theta_2 \) the test (3.24) that is proposed in this paper is usually less powerful than the classical CUSUM test independently of the size of \( \Delta^2 \). The reason for this consists in the fact that it follows from assumption (3.16) and the continuous mapping theorem that under the null hypothesis \( H_0 : \theta_1 = \theta_2 \) the statistic \( \sqrt{n} \int_0^1 \hat{T}_n(s) \, ds \) converges weakly to a non-degenerate random variable. Consequently, we observe from expression (3.26) that for reasonable sample sizes the level of the test is practically zero, where the classical CUSUM test has approximately level \( \alpha \). As a consequence, the power of the classical test for the hypothesis \( H_0 : \theta_1 = \theta_2 \) is usually larger than the power of the test (3.24). In contrast, the new test (3.24) has a substantially smaller type I error. Therefore (without any adjustment of the nominal level) both tests are not comparable and the tests for a relevant change should not be used for the classical hypothesis \( H_0 : \theta_1 = \theta_2 \).
(b) If the true change point is very close to 0 or 1 and the sample size is fixed it is intuitively clear that the new test (3.24) has a similar behaviour to that where there is no change point in the process. Observing expression (3.23) we therefore expect that the new test is conservative in this case. These observations have been confirmed in a simulation study, which is not displayed for brevity. Moreover these findings are in line with results in classical change point analysis. For example, Andrews (1993) recommended restricting the interval $[0, 1]$ to $[\varepsilon, 1 - \varepsilon]$ for a small constant $\varepsilon > 0$ to gain power of the CUSUM test for the hypothesis $H_0: \theta_1 = \theta_2$, and a similar strategy could be applied in the problem of testing for relevant changes in the process.

3.4. Further discussion

In this section we briefly mention two further applications of the new approach. First we note that testing hypotheses of the form (1.1) and (1.2) does not allow us to control the type II error if the null hypothesis of no (relevant) change point cannot be rejected and subsequent data analysis is performed under the assumption of no change point. If the statistician is interested in controlling the error of erroneously deciding for a non-relevant change point, we propose testing hypotheses of the form (1.4) for similarity of the parameters. The corresponding test is easily obtained from the previous discussion and rejects the null hypothesis $H_0 : \|\theta_1 - \theta_2\| > \Delta$ in favour of $H_1 : \|\theta_1 - \theta_2\| \leq \Delta$, whenever

$$\hat{M}_n^2 \leq \Delta^2 + u_\alpha \frac{\hat{\tau}}{\sqrt{n}}.$$  \hfill (3.27)

Secondly, $\hat{M}_n^2$ provides an estimate of the magnitude of the change and it is of particular importance to quantify the uncertainty of the estimate. This can easily be achieved by using the result on weak convergence that is specified in expression (3.22). For example a two-sided confidence interval for the squared distance $\|\theta_1 - \theta_2\|^2$ between the parameters $\theta_1$ and $\theta_2$ is given by

$$\left[\hat{M}_n^2 - u_{1-\alpha/2} \frac{\hat{\tau}}{\sqrt{n}}, \hat{M}_n^2 + u_{1-\alpha/2} \frac{\hat{\tau}}{\sqrt{n}}\right].$$ \hfill (3.28)

The coverage probabilities of this interval are investigated by means of a simulation study in Section 5.

4. Applications: detecting relevant change points

In this section we discuss several examples to illustrate the theory that was developed in Section 3. In particular, we concentrate on the detection of relevant changes in the mean, coefficients in a linear regression and a relevant change in the distribution itself. Further examples discussing changes in the variance and correlation are presented in section C of the on-line supplement. To be precise we assume that the assumptions of theorem A.1 and A.2 in section A.1 of the on-line supplement are satisfied. Similar results can be derived for alternative dependence concepts.

4.1. Relevant changes in the mean

The most prominent example of change point analysis in model (3.1) consists in the investigation of structural breaks in the mean $\mu = \theta_{\text{mean}}(F) = \int \rho \text{d} F$ (dz). Although the classical change point problem $H_0 : \mu_1 = \mu_2$ versus $H_1 : \mu_1 \neq \mu_2$ has been investigated by numerous researchers (see Csörgő and Horváth (1997) for a survey of methods for the independent case and Aue and Horváth (2013) for an extension to dependent data), we did not find any references on testing the hypotheses (2.1) of relevant change points in the mean. In contrast with the discussion of
Section 2 and with most of the literature, we do not assume that the stochastic features of the process besides the mean coincide before and after the break point. In particular, the variances or more generally the dependence structures before and after the change point can be different, although the means \( \mu_1 \) and \( \mu_2 \) are ‘close’, i.e. \( \| \mu_1 - \mu_2 \| \leq \Delta \). Theorem A.2 in the on-line supplement establishes condition (3.16), where the covariance kernel of the limiting process is defined in expressions (A.14) and (A.15) with \( \theta_{\text{lin}}[I\{W_k(l) \leq \cdot\}] = \theta_{\text{mean}}[I\{W_k(l) \leq \cdot\}] = W_k(l) \). Consequently, the corresponding asymptotic variance in expression (3.20) is given by

\[
\tau_{F_1,F_2,t}^2 = 4 \frac{4}{5\{\hat{t}(1-\hat{t})\}} \{t(5-10t+6t^2)\} V_{1,\text{mean}}^2 + (1-3\hat{t}+8\hat{t}^2-6\hat{t}^3) V_{2,\text{mean}}^2 \quad (\mu_1 - \mu_2)
\] (4.1)

where \( V_{1,\text{mean}} \) and \( V_{2,\text{mean}} \) are defined in theorem A.2 in the on-line supplement with \( \theta_{\text{lin}}[I\{W_k(l) \leq \cdot\}] = W_k(l) \) \((l = 1,2)\). Then, for a \( d \)-dimensional sample \( Z_1, \ldots, Z_n \) with \( Z_i = (Z_{i1}, \ldots, Z_{id}) \), \( i = 1, \ldots, n \), the test statistic is obtained as

\[
\hat{M}_n^2 = \frac{3}{\{\hat{t}(1-\hat{t})\}^2} \frac{1}{n} \sum_{i=1}^{n} T_n^2(i),
\] (4.2)

where \( T_n(i) = \sum_{k=1}^{d} T_k^2(i), T_k(i) = (1/n) \sum_{j=1}^{k} Z_{j,k} - (i/n^2) \sum_{j=1}^{n} Z_{j,k} \) and

\[
\hat{t} = (1/n) \arg \max_{1 \leq i \leq n} |T_n(i)|.
\]

The null hypothesis of no relevant change in the mean with potentially different variances before and after the change point is rejected whenever condition (3.24) holds. The estimator \( \hat{\tau}_{\tilde{F}_1,\tilde{F}_2,t}^2 \) of the asymptotic variance is obtained from formula (4.1) by replacing the unknown quantities \( t, \mu_1 \) and \( \mu_2 \) by their empirical counterparts \( \hat{t}, \hat{\mu}_1 \) and \( \hat{\mu}_2 \) (see formulae (2.6) and (2.7) in Section 2). For the estimation of the long-run variances \( V_{1,\text{mean}} \) and \( V_{2,\text{mean}} \) in equation (4.1) we must account for potential serial dependence, and we propose a kernel-based estimator as described in Andrews (1991) in the two different subsamples. More precisely we choose the Bartlett kernel and a data-adaptive bandwidth \( \gamma_n = 1.1477\{4\hat{\rho}^2 [n\hat{t}] / (1-\hat{\rho}^2)^2 \}^{1/3} \). Here, \( \hat{\rho} \) is the mean of the estimated auto-regressive (AR(1)) parameters for the \( k \) univariate series \{\( Z_{i,k} | i = 1, \ldots, n \)\} \((k = 1, \ldots, d)\) for the sample before the estimated break point (note that this choice of \( \hat{\rho} \) is optimal for an AR(1) process in a univariate context). The estimator of \( V_{1,\text{mean}} \) is then defined by

\[
\hat{V}_{1,\text{mean}} = \frac{1}{[n\hat{t}]} \sum_{i=1}^{[n\hat{t}]} (Z_i - \hat{\mu}_1)(Z_i - \hat{\mu}_1)^T + \frac{2}{[n\hat{t}]} \sum_{j=1}^{[n\hat{t}]-1} k(\gamma_n) \sum_{i=1}^{[n\hat{t}]-j} (Z_i - \hat{\mu}_1)(Z_{i+j} - \hat{\mu}_1)^T
\]

with \( k(x) = (1-|x|) I\{|x| \leq 1\} \) and an analogous expression is used for the estimation of the quantity \( V_{2,\text{mean}} \) in expression (4.1). The choice of the bandwidth has no big influence in the case of serial independence, but reduces size distortions if there is high serial dependence.

### 4.2 Relevant changes in the parameters of a regression

Early results on change point inference in linear regression models can be found in Kim and Siegmund (1989), Hansen (1992), Andrews (1993), Kim and Cai (1993) and Andrews et al. (1996). More recent work on this problem has been done by Chen et al. (2013) and Nosek and Skuznitnka (2014), among others. In this section we introduce the problem of testing for relevant changes in the parameters of a regression model. To be precise, we consider the common linear regression model

\[
Y_{n,i} = g^T(X_i)\beta(i) + \epsilon_i \quad i = 1, \ldots, n
\]
where \( \beta_1 = \ldots = \beta_{(n/2)} \neq \beta_{(n/2)+1} = \ldots = \beta_n = \beta_2 \) and \( (X_i)_{i=1,\ldots,n} \) are independent strictly stationary processes. In the notation of Section 3 and section A.1 in the on-line supplement we have \( Z_{n,1}, \ldots, Z_{n,|n/2|} = (X_1, Y_1), \ldots, (X_{|n/2|}, Y_{|n/2|}) \sim F_1 \), and \( Z_{n,|n/2|+1}, \ldots, Z_{n,n} = (X_{|n/2|+1}, Y_{|n/2|+1}), \ldots, (X_n, Y_n) \sim F_2 \), where \( F_1 \) and \( F_2 \) are the joint distribution functions before and after the change point. The marginal distribution \( F_X \) of the predictor \( X \) satisfies \( F_X = F_1(\cdot, \infty) = F_2(\cdot, \infty) \) by these assumptions.

To construct tests for the null hypothesis of no relevant change

\[
H_0 : \| \beta_1 - \beta_2 \| \leq \Delta \quad \text{versus} \quad H_1 : \| \beta_1 - \beta_2 \| > \Delta
\]  

we assume that the \( k \times k \) matrix

\[
B := \int_{\mathbb{R}^{d+1}} g(x) g^T(x) F(dx, dy) = \int_{\mathbb{R}^d} g(x) g^T(x) F_X(dx)
\]  

is non-singular and note that the parameter \( \beta_i \) can be represented as

\[
\beta_i = \theta(F_i) = \left\{ \int_{\mathbb{R}^{d+1}} g(x) g^T(x) F_i(dx, dy) \right\}^{-1} \left\{ \int_{\mathbb{R}^{d+1}} y g(x) F_i(dx, dy) \right\} \quad i = 1, 2.
\]  

For an illustration we consider the case \( k = 1 \) and \( g(x) = x \), i.e. \( Y_i = \beta X_i + \varepsilon_i \) \((i = 1, \ldots, n)\). The test statistic is defined by equation (4.2) where

\[
T_n(i) = \frac{1}{B_n} \left( \frac{1}{n} \sum_{j=1}^i X_j Y_j - \frac{i}{n^2} \sum_{j=1}^n X_j Y_j \right),
\]

\[
\hat{B}_n = \frac{1}{n} \sum_{i=1}^n X_i^2
\]

and \( \hat{i} = (1/n) \arg \max_{1 \leq i \leq n} |T_n(i)| \). The null hypothesis (4.3) of no relevant change in the parameter \( \beta \) is rejected whenever condition (3.24) is satisfied, where

\[
\tau^2_{F_1, F_2, \hat{i}} = \frac{4(\beta_1 - \beta_2)^2}{5B^2 \hat{B}_n^2 (1 - \hat{i})^2} (V_1 \{1 + 2(1 - \hat{i})\} + V_0 \{5(1 - \hat{i}) \{1 - \hat{i}\} \beta_1 + \hat{i} \beta_2 \}^2
\]

\[
+ \{\hat{i}^3 \beta_1^2 + (1 - \hat{i})^3 \beta_2^2 \})).
\]  

See section B.1 in the on-line supplement for a derivation of these formulae and the definition of \( V_1 \) and \( V_2 \). We replace the unknown quantities \( \hat{i}, B, \beta_1, \beta_2, V_0 \) and \( V_1 \) by \( \hat{i}, \hat{B}_n \), the ordinary least squares estimates \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \) from the two subsamples before and after the estimated change point \([n\hat{i}]\) and the estimators

\[
\hat{V}_0 = \frac{1}{n} \sum_{i=1}^n \left( X_i^2 - \frac{1}{n} \sum_{j=1}^n X_j^2 \right)^2,
\]

\[
\hat{V}_1 = \frac{1}{n} \sum_{i=1}^{[n\hat{i}]} \left( X_i \hat{\varepsilon}_i^{(1)} - \frac{1}{[n\hat{i}]} \sum_{j=1}^{[n\hat{i}]} X_j \hat{\varepsilon}_j^{(1)} \right)^2 + \frac{1}{n} \sum_{i=[n\hat{i}]+1}^n \left( X_i \hat{\varepsilon}_i^{(2)} - \frac{1}{n-[n\hat{i}]} \sum_{j=[n\hat{i}]+1}^n X_j \hat{\varepsilon}_j^{(2)} \right)^2,
\]

where \( \hat{\varepsilon}_i^{(1)} \) and \( \hat{\varepsilon}_i^{(2)} \) are the least squares residuals from the sample before and after the estimated change point. In the case of serial dependence the estimators \( \hat{V}_0 \) and \( \hat{V}_1 \) must be modified appropriately and the details are omitted for brevity. We finally mention that the results of this section can be generalized to error processes \((\varepsilon_i)_{i=1}^n\) with different strictly stationary phases before and after the change point.
4.3. Relevant changes in the distribution

To investigate the problem of a relevant change with respect to the distribution in a univariate sequence of the form (3.1) we consider the distance
\[
\|F_1 - F_2\| = \left[ \int_{\mathbb{R}} \{F_1(z) - F_2(z)\}^2 \, dz \right]^{1/2}
\]
on the set of all distribution functions with existing first moment. In this case the null hypothesis of no relevant change in the distribution function is formulated as
\[
H_0 : \|F_1 - F_2\| \leq \Delta,
\]
\[
H_1 : \|F_1 - F_2\| > \Delta.
\]
(4.8)

For a given sample \(Z_1, \ldots, Z_n\) of independent random variables the test statistic for the null hypothesis (4.7) of no relevant change in the distribution function is defined by equation (4.2), where
\[
T_n(i) = \sum_{k=1}^{n-1} (Z_{(k+1)} - Z_{(k)}) \left( \frac{1}{n} \sum_{j=1}^{i} 1\{Z_j \leq Z_{(k)}\} - \frac{i}{n^2} \sum_{j=1}^{n} 1\{Z_j \leq Z_{(k)}\} \right)^2.
\]
(4.9)

Here \(Z_{(1)}, \ldots, Z_{(n)}\) denotes the order statistic of \(Z_1, \ldots, Z_n\). The null hypothesis of no relevant change in the distribution function is rejected whenever condition (3.24) holds. For the definition of an estimator of the asymptotic variance we note that we assumed independent observations such that we have
\[
\tau_{F_1, F_2, t}^2 = \frac{4}{5t^2 (1-t)^2} \left[ it(5 - 10t + 6t^2) \int_{\mathbb{R}^2} \Delta(z_1, z_2) \{F_1(z_1) \wedge F_1(z_2) - F_1(z_1) F_1(z_2)\} \, dz_1 \, dz_2 
\right. 
+ (1 - 3t + 8t^2 - 6t^3) \left. \int_{\mathbb{R}^2} \Delta(z_1, z_2) \{F_2(z_1) \wedge F_2(z_2) - F_2(z_1) F_2(z_2)\} \, dz_1 \, dz_2 \right],
\]
where we use the notation \(\Delta(z_1, z_2) = \{F_1(z_1) - F_2(z_1)\} \{F_1(z_2) - F_2(z_2)\}\). The estimator \(\hat{\tau}_{F_1, F_2, t}^2\) is now obtained by plugging in \(t = (1/n) \max_{1 \leq i \leq n} |T_n(i)|\) and replacing the unknown distribution functions by \(F_1\) and \(F_2\) by the empirical distribution functions that were calculated from the subsample before and after the estimated change point.

5. Finite sample properties

In this section, we illustrate the application of the new testing procedure and provide some finite sample evidence. For brevity we investigate three cases: the detection of relevant changes in
the mean, the parameter of a linear regression model and a relevant change in the distribution function. Similar results can be obtained for the other testing problems that are considered in the on-line supplement but are not displayed here for brevity. In all examples under consideration, we performed 5000 replications of test (3.24) at significance level \(\alpha = 0.05\). We also note that it follows from the proof of theorem 3 that the power of test (3.24) is approximately given by
\[
\beta_n(\delta) \approx 1 - \Phi\left\{ \sqrt{n(\Delta^2 - \delta^2)} \over \hat{\tau}_{F_1, F_2, t} + u_{1-\alpha} \right\}. 
\]
(5.1)

Similarly, we obtain a formula for the \(p\)-value of the test, i.e.
\[
1 - \Phi\left\{ \sqrt{n} \over \hat{\tau}_{F_1, F_2, t} \right\} (\hat{M}_n^2 - \Delta^2),
\]
(5.2)
where \( \Phi \) is the distribution function of the standard normal distribution. These formulae will be helpful to understand some properties of test (3.24).

5.1. Relevant changes in the mean

First, we look at the test for changes in the mean as discussed at the beginning of Section 4 focusing on a one-dimensional sample \( Z_1, \ldots, Z_n \). In Fig. 1 we display the rejection probabilities of test (3.24) for sample sizes \( n = 200, 500, 1000 \) and independent normally distributed random variables with mean \( \mu_1 = 0 \) in the first half and mean \( \mu_2 = 1 \) in the second half of the sample, i.e. \( t = 0.5 \). The variance is constant and equal to 1. Fig. 1(a) presents the empirical rejection probabilities of test (3.24) for fixed \( \delta = 1 \), where the parameter \( \Delta \), which defines the size of a relevant change in hypothesis (2.1), varies in the interval \([0.2, 1.2]\). We observe that the power of the test decreases in \( \Delta \) as predicted by formula (5.1). For \( \Delta = 1 \), the power is approximately 0.05, which shows that the test keeps its nominal level.

Fig. 1(b) displays the power curve of test (3.24) for the same sample sizes and \( \Delta = 1 \), where the ‘true’ difference \( \delta = \mu_1 - \mu_2 \) varies in the interval \([-2, 2]\). As expected, the power curve is U shaped with a minimum at \( \delta = 0 \) (note that the power of the test converges to 0 in this case—see formula (3.26) in the proof of theorem 3). Again the nominal level is well approximated at the boundary of the null hypothesis, i.e. \( \delta = \pm 1 \). We also observe that the type I error is much smaller inside the interval \( \{\delta \in \mathbb{R}||\delta|<\Delta\} \).

Figures as displayed in Fig. 1(a) are useful to obtain the minimal size of the parameter \( \Delta \) in equation (2.1) such that the null hypothesis of no relevant change of size \( \Delta \) is rejected at controlled type I error, whereas Fig. 1(b) directly displays the power function of test (3.24). Both types essentially provide the same information and for brevity we focus in the following discussion only on the power function. Moreover, because of the obvious symmetry, we just present the values for \( \delta \geq 0 \).

In Fig. 2 we analyse the effect of changes in the variances on the testing procedure, where the sample size is fixed as \( n = 200 \) and the setting is the same as in Fig. 1. Fig. 2(a) shows the power of test (3.24) for the null hypothesis of no relevant change in the mean, where the variances are the same before and after the change point and are given by \( \sigma^2 = 0.2^2, 0.5^2, 1, 2^2, 5^2 \). We observe that the approximation of the nominal level is quite accurate at the point \( \delta = 1 \). Moreover, the rejection probabilities decrease in \( \sigma^2 \). There is essentially no power for \( \sigma^2 = 5^2 \) because in this case the variance is dominating the mean. Moreover, in this case the level of the test is not very well approximated, which is because it is difficult to estimate the change point \( t \) accurately under a large signal-to-noise ratio. In Fig. 2(b) we display the effect of changing variances in the same setting as in Fig. 2(a) where the variance in the first half is equal to 1 and in the second half is given by \( \sigma^2 = 0.2^2, 0.5^2, 1, 2^2, 5^2 \). We do not observe substantial differences with respect to the quality of approximation of the nominal level. Compared with the case of constant variances the power is in general lower for \( \sigma^2 > 1 \) and higher for \( \sigma^2 < 1 \). These empirical findings reflect the asymptotic theory, because the asymptotic variance of the estimator \( \hat{M}_n \) is an increasing function of \( \sigma_1^2 \) and \( \sigma_2^2 \) (see formula (4.1)) and it follows from approximation (5.1) that the power of test (3.24) is decreasing with this variance.

Finally, we investigate the effect of serial dependence on test (3.24) for the null hypothesis of no relevant change in the mean. For this we generate \( n = 200 \) and \( n = 500 \) realizations of an AR(1) process with AR parameter \( \rho = 0, 0.4, 0.8 \), mean 0 and standard-normal-distributed innovations by using the R function \texttt{arima.sim}. Such a process fulfils a strong mixing condition with mixing coefficients that decay exponentially (see for example Doukhan (1994), theorem 6, page 99). After that, we add \( \delta \) to the last 100 realizations. Fig. 3 shows that the serial dependence has an effect on the quality of the approximation of the nominal level if the sample size is \( n = 200 \).
Fig. 1. Empirical rejection probabilities of the test (3.24) for the null hypothesis of no relevant change in the mean, where \( \mu_1 = 0, \mu_2 = 1 \) and \( t = 0.5 \) \( (---, n = 200; \ldots, n = 500; \ldots, n = 1000): \) (a) constant \( \delta = 1, \) varying \( \Delta \); (b) constant \( \Delta = 1, \) varying \( \delta \)

Moreover, the power decreases with increasing correlation. These properties have also been observed by others in the context of CUSUM-type testing procedures for classical hypotheses (see Xiao and Phillips (2002) and Aue, Horváth, Hušková and Ling (2009)). Moreover, using the asymptotic theory from Section 4 we can also give a precise explanation of these observations. For the AR(1) model under consideration the quantities \( V_{\text{mean}} \) in expression (4.1) are given by \( V_{\text{mean}} = V_{\text{mean}}^2 = \rho^2/(1 - \rho^2) \). Consequently the asymptotic variance \( \tau_{F_1,F_2,t}^2 \) is increasing with \( |\rho| \) and by formula (5.1) the power is decreasing.

5.2. Relevant changes in the parameters of a regression

In this section we investigate the finite sample properties of the test for a relevant change in the
Fig. 2. Empirical rejection probabilities of the test (3.24) for the null hypothesis of no relevant change in the mean, where

\[ \Delta = 1, \sigma_1 = \sigma_2 = 0.2, \sigma_1 = 1, \sigma_2 = 0.5, \sigma_1 = 1, \sigma_2 = 1, \sigma_1 = 1, \sigma_2 = 2; \]

\(- - - - , \sigma_1 = 1, \sigma_2 = 5\) (the sample size \(n = 200\) and the horizontal line marks the significance level 0.05): (a) constant variances; (b) different variances before and after the change point.

slope parameter of the regression model

\[ Y_i = \beta X_i + \epsilon_i \quad i = 1, \ldots, n, \]

based on a bivariate sample. In Fig. 4(a) we display the power of test (3.24) for the null hypothesis of no relevant change in the parameter \(\beta\), where \(\beta_1 = 0\) in the first half and \(\beta_2 = \delta \geq 0\) of the sample and the explanatory variables \(X_i\) and errors \(\epsilon_i\) in the linear regression model are independent identically standard normal distributed. The approximation of the nominal level is quite accurate and the power is increasing with the sample size. In contrast, the power of the test for a change in the slope is lower than the power for the test for a change of the same size in the mean as considered in Fig. 1. (Additional simulations show that this power difference still
Empirical rejection probabilities of the test (3.24) for the null hypothesis of no relevant change in the mean under serial dependence, where $\Delta = 1$ (---, $p = 0$; -- - - - , $p = 0.4$; - - - - - - , $p = 0.8$) (the horizontal line marks the significance level 0.05): (a) sample size $n = 200$; (b) sample size $n = 500$

exists if we do not account for serial dependence in the mean test, that means if we consider $V_{\text{mean}}^1 = (1/|n|) \Sigma_{i=1}^{|n|} (X_i - \hat{\mu}_1)^2$ and the analogue formula for $V_{\text{mean}}^2$. This observation can be easily explained by the asymptotic representation of the probability of rejection in expression (5.1) which is a decreasing function of the asymptotic variance $\tau_{F_1,F_2,t}^2$. For the test of the null hypothesis of no relevant change in the mean and slope these variances are given by 307.2 and 576 respectively (see equations (4.1) and (4.6)). In Fig. 4(b) we display the results for heavy-tailed predictors $X_i$, i.e. $X_i \sim \sqrt{(\xi)} t_f$, where $t_f$ denotes a $t$-distribution with $f$ degrees of freedom. Note that the $t$-distribution is standardized with $\text{var}(X_i) = 1$. We observe a less accurate approximation of the nominal level if the sample size is $n = 200$. Moreover, $t_5$-distributed regressors yield also a loss in power. This observation can also be explained by formula (5.1), where the asymptotic variance $\tau_{F_1,F_2,0.5}^2$ is given by 576 and 1024 for the normal and $t_5$-distribution respectively.
Fig. 4. Empirical rejection probabilities of the test (3.24) for the null hypothesis of no relevant change in the parameter of a linear regression model ($\Delta$, $n = 200$; $\cdots$, $n = 500$; $\cdots$, $n = 1000$) (the horizontal line marks the significance level 0.05): (a) normally distributed regressors; (b) $t_5$-distributed regressors

5.3. Relevant changes in the distribution

We continue with a brief finite sample study of the test for the null hypothesis of no relevant change in the distribution function, which was discussed in Section 4. We choose sample sizes $n = 200, 500, 1000$ with serially independent random variables, $\mathcal{N}(0, 1)$ distributed in the first half and $\chi^2$ distributed with different degrees of freedom $f = 0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4$ in the second half of the sample. The $\chi^2$-distributed random variables are standardized such that they have mean 0 and variance 1. The distance $\|F_1 - F_2\|$ for $f = 1$ is approximately equal to 0.2254 and this value was chosen as $\Delta$ in test (3.24). Table 1 shows the rejection probabilities of test (3.24). Because of the different distance measure, they are somewhat difficult to compare with the other figures but, apparently, the test does work well. The power decreases in $f$ as the $\chi^2_f$-distribution, standardized such that it has mean 0 and variance 1, converges to the $\mathcal{N}(0, 1)$ distribution for $f \to \infty$. 
Table 1. Empirical rejection probabilities of test (3.24) for the null hypothesis of no relevant change in the distribution function†

| n   | Results for the following degrees of freedom and values of n/δ: |
|-----|---------------------------------------------------------------|
|     | f = 0.2, f = 0.4, f = 0.6, f = 0.8, f = 1, f = 1.2, f = 1.4, n/δ = 0.3730 n/δ = 0.3154 n/δ = 0.2764 n/δ = 0.2476 n/δ = 0.2254 n/δ = 0.2077 n/δ = 0.1932 |
|-----|---------------------------------------------------------------|
| 200 | 0.995 0.784 0.404 0.174 0.078 0.042 0.021 |
| 500 | 1.000 0.978 0.614 0.221 0.069 0.023 0.006 |
| 1000| 1.000 0.986 0.846 0.313 0.064 0.011 0.001 |

†The first half of the sample is generated from an \( N(0, 1) \) distribution and the second half from a (standardized) \( \chi^2 \)-distribution with various degrees of freedom. The size of a relevant change is defined by \( \Delta = 0.2254 \) and corresponds to \( f = 1 \).

Table 2. Empirical coverage probabilities of the confidence interval (3.28) for the null hypothesis of no relevant change in the mean†

| n   | Results for the following values of n/δ: |
|-----|---------------------------------------------------------------|
|     | 0.4 0.6 0.8 1 1.2 1.4 0.4 0.6 0.8 1 1.2 1.4 |
|     | \( \alpha = 0.05 \) \( \alpha = 0.1 \) |
|-----|---------------------------------------------------------------|
| 200 | 0.973 0.955 0.949 0.946 0.944 0.944 0.939 0.911 |
| 500 | 0.956 0.953 0.951 0.952 0.951 0.951 0.914 0.906 |
| 1000| 0.957 0.954 0.953 0.952 0.952 0.952 0.907 0.902 |

†The first half of the sample is generated from an \( N(0, 1) \) and the second half from an \( N(\delta, 1) \) distribution.

5.4. Confidence intervals

As discussed in Section 3, the method that is proposed in this paper can also be used for constructing confidence intervals for the magnitude of the change between the parameters \( \theta_1 \) and \( \theta_2 \). We illustrate this for a change in the mean considering the setting from Fig. 1(b). In Table 2 we display the simulated coverage probabilities of the confidence interval (3.28), where \( \delta \in \{0.4, 0.6, 0.8, 1, 1.2, 1.4\} \). We observe that the empirical coverage probabilities are close to the theoretical coverage probabilities. If \( \delta \) approaches 0 the corresponding confidence intervals become conservative if the sample size is not too large.

5.5. Empirical illustration

We illustrate the application of the new test in an example from portfolio management. For this we consider continuous returns from the closing prices of BASF and Sanofi. The time period for the prices is from June 1st, 2009, to June 1st, 2012, i.e. \( n = 784 \), and we observe the bivariate vectors \( (X_i, Y_i) (i = 1, \ldots, n) \), where \( X \) represents BASF and \( Y \) Sanofi. We first used the procedure that was described in Section 4 to test the hypothesis of a relevant change point in the variance \( \Sigma \), where \( \Sigma \) is the covariance matrix of \( (X, Y) \). More precisely, we test the series \( \{Z_i = (X_i^2, Y_i^2, X_i Y_i)^T| i = 1, \ldots, n\} \) for a relevant change in the mean. Assuming constant means, which is standard with financial returns, this is equivalent to testing for a relevant change in the variance–covariance matrix (see section C.1 in the on-line supplement for more details). On a
5% level of significance, test (3.24) does not reject the null hypothesis of a relevant change of size $\Delta = 10^{-4}$ in the norm of the variance matrix. The test statistic is given by $M_n^2 = 4.67 \times 10^{-8}$ and the right-hand end of the critical region is given by $4.72 \times 10^{-8}$. Moreover, the 95% confidence interval for the squared norm of the vector of differences in the second-order moments is given by $[-1.9 \times 10^{-9}, 9.5 \times 10^{-8}]$.

Next we consider the classical testing problem $H_0 : \Sigma_1 = \Sigma_2$ to detect changes in the variance structure by using the test proposed in Aue, Hörmann, Horváth and Reimherr (2009). More precisely, we consider the test statistic $\Omega_n$ defined in this paper and kernel or bandwidth chosen as described in Section 4.1, where the AR parameters are estimated from the full sample. The value of the test statistic is 1.41, whereas the critical value is 1.00 (see Page (1959), page 444). Consequently this test rejects the null hypothesis $H_0 : \Sigma_1 = \Sigma_2$. The estimated break point $\hat{t}$ is 0.717.

Finally, we illustrate some consequences of the different decisions. For this we consider two simplified investment strategies which are both based on the global minimum variance portfolio given by $\Sigma^{-1}(1, 1)^T / [(1, 1)^T \Sigma^{-1} (1, 1) ]$. The first strategy is a consequence of the test for the hypothesis of no relevant change points. As pointed out in the previous paragraph the method proposed in this paper does not reject this hypothesis and the matrix $\Sigma$ is estimated by the empirical covariance matrix from the full sample. The second strategy is based on the test of Aue, Hörmann, Horváth and Reimherr (2009), which rejects the classical hypothesis of a change in the variance. Here the empirical covariance matrix is estimated from the sample $(X_{[n\hat{t}]+1}, Y_{[n\hat{t}]+1}), \ldots, (X_n, Y_n)$ (with $\hat{t} = 0.717$). We calculate the daily returns for both strategies and the profit or loss at the end of the time period: with a start amount of £10000, one would have a loss of £388.04 for the first and of only £101.19 for the second strategy. Therefore, in a world with only fixed but no variables transactions the second strategy would be useful if the fixed transaction costs are not higher than £388.04 − £101.19 = £286.85. In other words, if the costs are higher, one should not reject the null hypothesis and the break in the covariance matrix should be regarded as an irrelevant change. These results can be used for the choice of the threshold $\Delta$ in future investment strategies. More precisely, if an investor assumes these retrospective calculations are valid for future decisions and if the transaction costs are higher than £286.85, the new test developed in this paper should be used with threshold $\Delta = 8 \times 10^{-5}$. If the null hypothesis of no relevant change in the variance is not rejected, the investor should not shift his portfolio even if the test by Aue, Hörmann, Horváth and Reimherr (2009) rejects the classical hypothesis $H_0 : \Sigma_1 = \Sigma_2$.

6. Conclusions and future research

In this paper we have investigated the problem of testing for a relevant change in the parameters of a time series. Our work was motivated by the observation that in many cases data analysts are interested only in changes where the difference between the parameters before and after the change point exceeds a minimum threshold, say $\Delta$. An important ingredient in our approach is the appropriate choice of this threshold, which must be defined carefully in every particular application to distinguish between statistical and scientific significance. The classical approach avoids this choice by simply putting $\Delta = 0$, and we strongly argue to choose this threshold thoroughly considering the scientific background of the testing problem. Statistical methodology is developed for testing the hypothesis of no relevant change points. Moreover, if a reasonable choice of the threshold based on the scientific background is not possible, our approach still provides a confidence interval for the norm of the difference between the parameters of the process before and after the change point.
The method proposed in this paper is based on an $L_2$-distance of the classical CUSUM process and can be applied in numerous change point problems. Asymptotic analysis is performed to derive (asymptotic) critical values under the assumption that the process before and after the change point is strictly stationary and a simulation study demonstrates good finite sample properties of the new test. An important and interesting topic for future research is to extend these results to non-stationary processes by using a similar approach to that in Zhou (2013). For example, consider the model $Y_i = \mu_i + e_i$ for $i = 1, \ldots, n$, where $(e_i)_{i \in \mathbb{Z}}$ is a piecewise locally stationary process in the sense of definition 1 of Zhou (2013), and $\mu_i = \mu_1$ if $1 \leq i \leq \lfloor nt \rfloor$ and $\mu_i = \mu_2$ if $\lfloor nt \rfloor + 1 \leq i \leq n$. It then follows by using similar results to those presented in this paper that
\[ \sqrt{n} \left\{ \mathbb{U}_{[s]} - (s + t - st)(\mu_1 - \mu_2) \right\}_{s \in [0, 1]} \overset{D}{\Rightarrow} \left\{ U(s) - sU(1) \right\}_{s \in [0, 1]} \]
where $\{ U(s) \}_{s \in [0, 1]}$ is a centred Gaussian process with covariance kernel
\[ k(s, t) = \sum_{i=0}^{\tau} \int_{b_i}^{b_{i+1}} \sum_{k \in \mathbb{R}} \text{cov}\{ G_i(u, F_k), G_i(u, F_0) \} \, du \]
and the constants $b_i$, the $\sigma$-fields $F_k$ and the non-linear filters $G_i$ appear in the definition of the piecewise locally stationary $(e_i)_{i \in \mathbb{Z}}$ (see Zhou (2013)). From this result an analogue of theorem 1 could be derived. However, the asymptotic variance depends in a complicated way on the covariance kernel $k$ and critical values cannot be derived from the asymptotic theory. For the implementation of a test for the hypothesis of no relevant changes in the mean bootstrap, methods must be developed which address the particular problems appearing with tests of interval hypotheses adequately.

A further challenging direction for future research is the extension of the methodology for detecting relevant changes in processes with multiple break points, which have found considerable attention in the recent literature. One obvious idea is to use binary segmentation methods for this problem (see for example Cho and Fryzlewicz (2015)), but multiscale inference (see Frick et al. (2014)) or methods based on spectral analysis (see for example Preuss et al. (2014)) might yield to other powerful methods.

Acknowledgements

The authors thank Martina Stein, who typed numerous versions of this manuscript with considerable technical expertise, and Axel Bücher and Stanislav Volgushev for helpful discussions. This work has been supported in part by the Collaborative Research Center ‘Statistical modeling of nonlinear dynamic processes’ Sonderforschungsbereich 823, Teilprojekt A1 and C1) of the German Research Foundation. Parts of this paper were written while H. Dette was visiting the Isaac Newton Institute, Cambridge, UK, in 2014 (‘Inference for change-point and related processes’) and the authors thank the Institute for its hospitality. The authors are also grateful to three referees and the Associate Editor. Their constructive comments on an earlier version yielded a substantial improvement of an earlier version of this paper.

References

Andrews, D. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59, 817–858.
Andrews, D. W. K. (1993) Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61, 128–156.
Andrews, D. W. K., Lee, I. and Ploberg, W. (1996) Optimal change-point tests for normal linear regression. *J. Econom.*, 70, 9–36.
Aue, A., Hörmann, S., Horváth, L. and Reimherr, M. (2009) Break detection in the covariance structure of multivariate time series models. *Ann. Statist.*, 37, 4046–4087.

Aue, A. and Horváth, L. (2013) Structural breaks in time series. *J. Tim. Ser. Anal.*, 34, 1–16.

Aue, A., Horváth, L., Huskova, M. and Ling, S. (2009) On distinguishing between random walk and change in the mean alternative. *Econometr. Theor.*, 25, 411–441.

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

Berger, J. O. and Delampady, M. (1987) Testing precise hypotheses. *Statist. Sci.*, 2, 317–335.

Berkes, I., Gombay, E. and Horváth, L. (2009) Testing for changes in the covariance structure of linear processes. *J. Statist. Plann. Inf.*, 139, 2044–2063.

Berkson, J. (1938) Some difficulties of interpretation encountered in the application of the chi-square test. *J. Am. Statist. Ass.*, 33, 526–536.

Brown, R. L., Durbin, J. and Evans, J. M. (1975) Techniques for testing the constancy of regression relationships over time. *J. R. Statist. Soc. B*, 37, 149–163.

Carlsen, E. (1988) Nonparametric change-point estimation. *Ann. Statist.*, 16, 188–197.

Chow, G. (1960) Tests of equality between sets of coefficients in two linear regressions. *Ann. Math. Statist.*, 31, 181–185.

Chow, S.-C. and Liu, P.-J. (1992) *Limit Theorems in Change-point Analysis*. New York: Dekker.

Cho, H. and Fryzlewicz, P. (2015) Multiple change-point detection for high dimensional time series via sparsified binary segmentation. *J. R. Statist. Soc. B*, 77, 475–507.

Cho, G. (1960) Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28, 591–605.

Cho, S.-C. and Liu, P.-J. (1992) *Design and Analysis of Bioavailability and Bioequivalence Studies*. New York: Dekker.

Csörgő, M. and Horváth, L. (1997) *Limit Theorems in Change-point Analysis*. New York: Wiley.

Doukhan, P. (1994) *Mixing Properties and Examples*. Berlin: Springer.

Frick, K., Munk, A. and Sieling, H. (2014) Multiscale change point inference (with discussion). *J. R. Statist. Soc. B*, 76, 495–580.

Hansen, B. E. (1992) Testing for parameter instability in regression with I(1) process. *J. Bus. Econ. Statist.*, 10, 321–335.

Horváth, L., Kokoszka, P. and Steinebach, J. (1999) Testing for changes in multivariate dependent observations with an application to temperature changes. *J. Multiv. Anal.*, 68, 96–119.

Jandhyala, V., Fotopoulos, S., MacNeill, I. and Liu, P. (2013) Inference for single and multiple change-points in time series. *J. Tim. Ser. Anal.*, to be published, doi: 10.1111/jtsa12035.

Kim, H.-J. and Cai, L. (1993) Robustness of the likelihood ratio test for a change in simple linear regression. *J. Am. Statist. Ass.*, 88, 864–871.

Kim, H.-J. and Siegmund, D. (1989) The likelihood ratio test for a change-point simple linear regression. *Biometrika*, 76, 409–423.

Krämer, W., Ploberger, W. and Alt, R. (1988) Testing for structural change in dynamic models. *Econometrica*, 56, 1355–1369.

Lehmann, E. L. (1986) *Testing Statistical Hypotheses*, 2nd edn. New York: Wiley.

Lorden, G. (1971) Procedures for reacting to a change in distribution. *Ann. Math. Statist.*, 42, 1897–1908.

McBride, G. B. (1999) Equivalence tests can enhance environmental science and management. *Aust. New Zeal. J. Statist.*, 41, 19–29.

Moustakides, G. V. (2004) Optimality of the CUSUM procedure in continuous time. *Ann. Statist.*, 32, 302–315.

Nosek, K. and Skuzutnika, Z. (2014) Change-point detection in a shape-restricted regression model. *Statistics*, 48, 641–656.

Page, E. S. (1955) Control charts with warning lines. *Biometrika*, 42, 243–257.

Page, E. S. (1959) K-sample analogues of the Kolmogorov-Smirnov and Cramer-v. Mises tests. *Ann. Math. Statist.*, 30, 420–447.

Preuss, P., Puchstein, R. and Dette, H. (2014) Detection of multiple structural breaks in multivariate time series. *J. Am. Statist. Ass.*, to be published, doi: 10.1080/01621459.2014.920613.

Wied, D. (2013) CUSUM-type testing for changing parameters in a spatial autoregressive model for stock returns. *J. Tim. Ser. Anal.*, 34, 221–229.

Wied, D., Krämer, W. and Dehling, H. (2012) Testing for a change in correlation at an unknown point in time using an extended functional delta method. *Econometr. Theor.*, 28, 570–589.

Xiao, Z. and Philips, P. (2002) A cusum test for cointegration using regression residuals. *J. Econometr.*, 108, 43–61.

Zhou, Z. (2013) Heteroscedasticity and autocorrelation robust structural change detection. *J. Am. Statist. Ass.*, 108, 726–740.

**Supporting information**

Additional ‘supporting information’ may be found in the on-line version of this paper: ‘Online Appendix’.