Statistical Papers (2020) 61:1347–1349
https://doi.org/10.1007/s00362-020-01199-9

PREFACE

Editorial for the special issue: Change point detection

Georgy Sofronov[1] · Martin Wendler[2] · Volkmar Liebscher[3]

Published online: 11 July 2020

More and more data are collected in order to analyse the time-varying behaviour of an underlying stochastic process. It is of particular interest to be able to detect systematic changes, so-called change points, in the underlying structure despite the random fluctuations and to estimate the locations of these changes. Change point detection is a ubiquitous problem, arising in almost any field including biology, medicine, finance, economics among many others.

For example, change points in DNA sequences can indicate the location of genes and other functional elements. In epidemiology, the timely detections of rising numbers of infections is a very important issue. In intensive care, change point detection methods can contribute to better alarm systems.

This special issue includes a selection of papers presented at the Workshop on Change Point Detection Limit Theorems, Algorithms, and Applications in Life Sciences held in Alfried Krupp Wissenschaftskolleg Greifswald from 8 to 10 July 2019. The aim of this workshop was to give an overview of the new statistical methods, to discuss the strength and weaknesses of different approaches and to see how they can deal with challenging data situations, which occur in up to date experiments from the life sciences. There were 29 talks and 5 posters presented by researchers from different countries around the world. The workshop gave the participants (experienced academics, early career researchers and PhD students) an excellent opportunity not only to present their new developments, but also to engage in discussion to get a better understanding of their different views on change point problems.

Depending on a way of how data are collected, change point problems may be divided into two large classes: posteriori (also called retrospective or off-line) and sequential (also quickest or on-line). The former class of change point problems arises when the data have been collected already, whereas in the latter class of problems, the data are collected sequentially over time and future observations are not known yet. Articles (Gapeev 2020; Szajowski 2020) study sequential change point problems while

✉ Georgy Sofronov
georgy.sofronov@mq.edu.au

1 Macquarie University, Sydney, Australia

2 Otto-von-Guericke-Universität, Magdeburg, Germany

3 Universität Greifswald, Greifswald, Germany

all other articles in the special issue consider posteriori change point problems. The results of article (Hlávka et al. 2020) can be applied to both on-line and retrospective problems.

There are different approaches to deal with change point problems in multivariate data. Article (Hlávka et al. 2020) proposes paired and two-sample break-detection procedures for vectorial observations and multivariate time series. The new methods involve L^2-type criteria based on empirical characteristic functions and are easy to compute regardless of dimension. Article (Maciak et al. 2020) introduces detection procedures for a change in means of panel data are proposed, where the considered model allows for mutually dependent and generally nonstationary panels with an extremely short follow-up period. Two competitive self-normalized test statistics are employed and their asymptotic properties are derived for a large number of available panels.

Nonparametric change point models have become increasingly popular over the last two decades. Very often, the models give rise to challenging mathematical research questions, for example, concerning asymptotic results. Article (Wendler and Račkauskas 2020) presents a robust test statistic to detect a changed segment (so called epidemic changes) based on the Wilcoxon statistic. To obtain asymptotic results, functional limit theorems for U-processes in Hölder spaces are proved. Article (Mohr and Selk 2020) considers a regression model that allows for time series covariates as well as heteroscedasticity with a regression function that is modelled nonparametrically. Consistency of the estimator based on a Kolmogorov–Smirnov functional of the marked empirical process of residuals is shown. The case of lagged dependent covariates is also considered. Article (Yang et al. 2020) considers the estimation of the structural change point in a nonparametric model with dependent observations. It is shown that the statistic, which is based on a maximum-CUSUM-estimation procedure, tends to zero almost surely if there is no change, and is larger than a threshold asymptotically almost surely otherwise. The strong consistency of the change point estimator is also proved.

Multiple change point detection is a challenging problem, demanding sophisticated techniques to validate the identified change points. In article (Ma et al. 2020), data are modelled assuming each segment is an autoregressive time series with possibly different autoregressive parameters. Once these potential change points are identified using a likelihood ratio scan based estimation technique, modified parametric spectral discrimination tests are used to validate the proposed segments.

Article (Gapeev 2020) studies the sequential hypothesis testing and sequential (or quickest) change point detection problems with linear delay penalty costs for observable Wiener processes under constantly delayed detection times. Closed-form expressions for the Bayesian risk functions and optimal stopping boundaries for the associated weighted likelihood ratio processes in the original problems of sequential analysis are derived. Article (Szajowski 2020) considers sequential change point detection with applications to the game theory. Article (Bandt 2020) studies order patterns and permutation entropy with applications to change point problems in medical and financial time series data.

References

Bandt C (2020) Order patterns, their variation and change points in financial time series and Brownian motion. Stat Pap. https://doi.org/10.1007/s00362-020-01171-7

Gapeev PV (2020) On the problems of sequential statistical inference for wiener processes with delayed observations. Stat Pap. https://doi.org/10.1007/s00362-020-01178-0

Hlávka Z, Hušková M, Meintanis SG (2020) Change-point methods for multivariate time-series: paired vectorial observations. Stat Pap. https://doi.org/10.1007/s00362-020-01175-3

Ma L, Grant AJ, Sofronov G (2020) Multiple change point detection and validation in autoregressive time series data. Stat Pap. https://doi.org/10.1007/s00362-020-01198-w

Maciak M, Pešta M, Peštová B (2020) Changepoint in dependent and non-stationary panels. Stat Pap. https://doi.org/10.1007/s00362-020-01180-6

Mohr M, Selk L (2020) Estimating change points in nonparametric time series regression models. Stat Pap. https://doi.org/10.1007/s00362-020-01162-8

Szajowski KJ (2020) Rationalization of detection of the multiple disorders. Stat Pap. https://doi.org/10.1007/s00362-020-01168-2

Wendler M, Račkauskas A (2020) Convergence of U-processes in Hölder spaces with application to robust detection of a changed segment. Stat Pap. https://doi.org/10.1007/s00362-020-01161-9

Yang Q, Li Y-N, Zhang Y (2020) Change-point detection for nonparametric regression under strongly mixing data. Stat Pap. https://doi.org/10.1007/s00362-020-01196-y