Extremes
https://doi.org/10.1007/s10687-026-00534-x

RESEARCH

MOPED: A moving sum method for change point detection in pairwise extremal dependence

Euan T. McGonigle[1] · Matthew Pawley[2] · Jordan Richards[3] · Christian Rohrbeck[4]

Received: 30 August 2025 / Revised: 16 February 2026 / Accepted: 4 March 2026

Abstract

It is increasingly the case with modern time series that many data sets of practical interest contain abrupt changes in structure. These changes may occur in complex characteristics such as the extremal dependence structure, and identifying such structural breaks remains a challenging problem. Many existing change point detection algorithms focus on changes in dependence across the entire distribution, rather than the tails, and approaches that are tailored to extremes typically make strict parametric assumptions or they are only applicable to bivariate data. We propose a nonparametric MOving sum based approach for detecting multiple changes in the Pairwise Extremal Dependence (MOPED) of multivariate regularly varying data. To avoid the classical problem of threshold selection in the study of multivariate extremes, we further propose a multiscale, multi-threshold variant of MOPED that pools change point estimates across choices of the threshold and the bandwidth used in local estimation. Good performance of MOPED is illustrated in a simulation study, and we showcase its ability to identify subtle changes in tail dependence class in the absence of correlation changes. We further demonstrate the usefulness of MOPED by identifying changes in the extremal connectivity of electroencephalogram (EEG) signals of seizure-prone neonates.

Keywords Electroencephalograms · Extreme value theory · Multivariate regular variation · Seizures · Structural breaks

AMS classiication62G3262G32 Statistics of extreme values · Tail inference

✉ Jordan Richards
jordan.richards@ed.ac.uk

[1] School of Mathematical Sciences, University of Southampton, Southampton, UK

[2] ATASS Sports, Exeter, UK

[3] School of Mathematics and Maxwell Institute for Mathematical Sciences, University of Edinburgh, Edinburgh, UK

[4] Department of Mathematical Sciences, University of Bath, Bath, UK

Published online: 16 March 2026

1 Introduction

In many applications, interest lies in analysing the tail behaviour of a random vector (Coles et al. 1999; Cooley and Thibaud 2019; Simpson et al. 2020; Rohrbeck and Cooley 2023; Murphy-Barltrop et al. 2024; De Monte et al. 2025). This has given rise to a wide range of methods in multivariate extreme value theory; we refer the reader to Engelke and Ivanovs (2021) for a review article on classical methods and a summary of recent developments. Due to the scarcity of extreme events, one assumption commonly made in practice is that of stationarity in the tail dependence, with any change in the distribution of extremes being captured by a nonstationary model for the marginal distributions (see, e.g., Chavez-Demoulin and Davison 2005; Youngman 2019; Richards and Huser 2026). For instance, in environmental studies, such a framework allows extreme events at spatial sites to become more severe and frequent over time, while the overall spatial structure of extreme events is assumed to remain unchanged. However, it is increasingly commonplace for data to exhibit changing tail dependence structure: Poon et al. (2003) identified complex nonstationarity in the tail dependence of stock market indices; flood events in the UK tend to be more widespread in winter than in summer (Rohrbeck and Tawn 2021); Zhong et al. (2022) modelled temporal trends in the spatial extent of European heatwaves; Kakampakou et al. (2024) found that the spatial dependence in sea surface temperature extremes across the Red Sea has changed over the past decades; Murphy-Barltrop and Wadsworth (2024) identified increasing trends in the strength of extremal dependence between UK temperatures and dryness.

There has been a growing number of methods in extreme value theory which aim to detect nonstationarity in the tail dependence of data. An exploratory approach is to estimate and compare the coefficient of tail dependence χ (Coles et al. 1999) for non-overlapping time windows (see, e.g., Poon et al. 2003; Kakampakou et al. 2024). Formally, for random variables X and Y with cumulative distribution functions F_X and F_Y respectively,

$$\chi = \lim_{u \to 1} \Pr\{F_X(X) > u \mid F_Y(Y) > u\}. \tag{1}$$

We say that X and Y are asymptotically independent when $\chi = 0$, and asymptotically dependent when $\chi > 0$. Beyond this diagnostic tool, several statistical tests for detecting structural changes in extremal dependence are available. Bücher et al. (2015) introduced a test for a change in tail dependence of a sequence of independent bivariate random vectors, and this framework was generalized to β-mixing time series by Hoga (2018). Hazra and Bose (2025) used a parametric extremal dependence model and likelihood ratio tests to identify a single change point in bivariate extremal dependence, whilst Hoga (2022) introduce a modelling framework which allows for a transition between asymptotic dependence and asymptotic independence. Drees (2023) introduced another testing procedure for a structural change in the tail dependence which is, however, limited to lower dimensions due to its computational burden. An alternative has been proposed by Pawley (2025) who focus on testing for changes in summaries of pairwise extremal dependence, which decreases the computational burden and permits application in higher dimensions at the cost of

reduced power in certain cases. Finally, we note that changes in tail dependence may also refer to a shift in temporal dependence. For instance, Dupuis and Trapin (2019) considered testing for a change point in serial dependence of a single time series of extremes.

Whilst the above methods directly consider nonstationary tail dependence, another class of methods – so-called nonparametric change point detection approaches – may also be used. These approaches can discriminate between general distributional changes and can be applied to the problem of detecting change points, or structural breaks, in extremal dependence. For example, there are approaches based on density functions (Padilla et al. 2022) and kernel transforms of the data (Celisse et al. 2018; Arlot et al. 2019). Another common technique relies on energy-based notions of distance (Matteson and James 2014; Mcgonigle and Cho 2025), whilst graph-based methods applicable to non-Euclidean data may also be employed (Chen and Zhang 2015). However, nonparametric methods are less informative in practical situations, where it is often the aim to study a particular type of change point present in the data based on the application of interest.

Working in a framework similar to Bücher et al. (2015) and Drees (2023), we consider the problem of detecting abrupt changes in the tail dependence structure amongst a subset, or all, of the variables of a serially independent multivariate time series. In this work, we constrain our focus to identifying structural changes in only the tail dependence, and hereafter assume that the marginal distributions are stationary. For identifying structural changes in marginal tail behaviour, we instead direct the reader to, for example, Dierckx and Teugels (2010), Dupuis et al. (2015), Hoga (2017), Kojadinovic and Naveau (2017), e Silva et al. (2020), Castillo-Mateo (2022), and Girard et al. (2025). The key contribution of this paper is to introduce a computationally efficient algorithm which identifies structural changes in tail dependence. We summarise tail dependence in the vector over a time period using the tail pairwise dependence matrix (TPDM) of Cooley and Thibaud (2019) which makes the assumption of asymptotic dependence between variables. The TPDM acts analogously to the covariance matrix in classical Gaussian statistics, which has made it an increasingly popular tool in the study of multivariate extremes; examples of its usage include clustering (Elsom and Pawley 2024; Richards et al. 2025), causal inference (Jiang et al. 2025b), asymmetric dependence modelling (Jiang et al. 2025a), graph learning (Lee and Cooley 2022; Gong et al. 2024), principal component analysis (Jiang et al. 2020), linear prediction (Lee and Cooley 2021), and time series modelling (Mhatre and Cooley 2024).

We propose a moving sum (MOSUM; see, e.g., Eichinger and Kirch 2018) procedure for multiple change point detection in the extremal dependence of a multivariate time series, using a detector statistic that is carefully devised to detect changes in the TPDM; we refer to this method as MOPED (MOving sum method for changes in Pairwise Extremal Dependence). We showcase the efficacy of MOPED, relative to an existing state-of-the-art nonparametric change point detection algorithm, via a simulation study and real data application. In our application to multivariate EEG signals from seizure-prone neonatal subjects, we find that MOPED identifies significant structural changes in the extremal dependence of the signals when the subjects undergo seizures, and these changes typically occur close to the point of seizure.

The remainder of the paper is organised as follows. In Section 2, we provide background on multivariate regular variation and the tail pairwise dependence matrix, which we then use to define our extremal dependence change point algorithm, MOPED. In Section 3, we provide a discussion of the numerical implementation of our change point algorithm and a simulation study highlighting its efficacy and empirical properties. In Section 4, we apply our algorithm to identify change points and structural changes in the extremal dependence of multivariate EEG signals for seizure-prone neonates. We provide concluding remarks and a discussion of extensions in Section 5. Accompanying software implementing the method is available as the R package !moped! at https://github.com/EuanMcGonigle/moped.

2 Methodology

In this section, we introduce the necessary background concepts, define the multiple change point model for the TPDM, and discuss our proposed change point detection methodology.

2.1 Multivariate regular variation

A d-dimensional random vector $\mathrm{X} \in \mathbb{R}_+^d$ is regularly varying with tail index $\alpha > 0$, denoted by $\mathrm{X} \in \mathrm{RV}_+^d(\alpha)$, if there exists a sequence $b_n \to \infty$ such that $n \Pr(b_n^{-1}\mathrm{X} \in \cdot) \xrightarrow{v} v(\cdot)$ as $n \to \infty$, where the limiting Radon measure $v(\cdot)$ is supported on the cone $\mathbb{E}_+^d := [0, \infty]^d \setminus \{0\}$ and where $\xrightarrow{v}$ denotes vague convergence; see, e.g., Resnick (2007). As a result of this vague convergence statement, the limit measure $v(\cdot)$ satisfies $(-\alpha)$-homogeneity, i.e., $v(r\mathcal{B}) = r^{-\alpha}v(\mathcal{B})$ for any $r > 0$ and any Borel subset $\mathcal{B} \subset \mathbb{E}_+^d$. Courtesy of the $(-\alpha)$-homogeneity of $v(\cdot)$, we can make $v(\cdot)$ amenable to modelling via a radial-angular decomposition of the limit measure into a product measure, defined on $\mathbb{R}_+ \times \mathcal{S}_+^{d-1}$, where $\mathcal{S}_+^{d-1} := \{\mathrm{x} \in \mathbb{R}_+^d : ||\mathrm{x}||_2 = 1\}$ is the positive part of the unit $(d-1)$-sphere: for $r > 0$, we consider

$$v(\{\mathrm{x} \in \mathbb{E}_+^d : ||\mathrm{x}||_2 \geq r, \mathrm{x}/||\mathrm{x}||_2 \in \mathcal{B}_H\}) = r^{-\alpha}H(\mathcal{B}_H),$$

where $\mathcal{B}_H \subset \mathcal{S}_+^{d-1}$ is a Borel subset and where $H(\cdot)$ is an angular mass measure supported on $\mathcal{S}_+^{d-1}$. Loosely speaking, the angular measure $H(\cdot)$ determines the direction of a multivariate extreme event, while the measure $r^{-\alpha}$ determines the relative magnitude.

Note that $H(\cdot)$ is not a valid probability measure, but can be normalised by dividing by its total mass, $H(\mathcal{S}_+^{d-1})$, to give a valid probability measure, denoted by $N(\cdot) := H(\cdot)/H(\mathcal{S}_+^{d-1})$. In the case where $\alpha = 2$ and X has standardised margins such that $X_i \in \mathrm{RV}_+^1(\alpha), i = 1, \ldots, d$, then we have $H(\mathcal{S}_+^{d-1}) = d$; hereafter we assume that both of these assumptions hold and note that, in practice, this is without loss of generality (with respect to marginal transformations; see, e.g., Jiang

et al. 2025a, for discussion). An example of a marginal distribution that satisfies $X \in \mathrm{RV}_+^1(\alpha)$ is the Pareto(2) distribution with unit scale and shape two, and distribution function $F(x) = 1 - x^{-2}, x \geq 1$.

2.2 Tail pairwise dependence matrix

The tail pairwise dependence matrix (TPDM; Cooley and Thibaud 2019) is a matrix of pairwise summaries of the extremal dependence in $\mathrm{X} = (X_1, \ldots, X_d) \in \mathrm{RV}_+^d(\alpha)$. The TPDM is denoted by $\Sigma := (\sigma_{ij})_{i,j=1,\ldots,d}$, with pairwise entries

$$\sigma_{ij} := \int_{\mathcal{S}_+^{d-1}} \theta_i \theta_j \mathrm{d}H(\boldsymbol{\theta}) = d \int_{\mathcal{S}_+^{d-1}} \theta_i \theta_j \mathrm{d}N(\boldsymbol{\theta}) = d \lim_{r \to \infty} \mathbb{E}\left[\frac{X_i X_j}{R^2} \mid R > r \right], \quad (2)$$

with $R = ||\mathrm{X}||_2$. The strength of asymptotic dependence between X_i and X_j increases with σ_{ij}, with $\sigma_{ij} = 0$ if and only if X_i and X_j are asymptotically independent. When $d = 2$, the pairwise σ_{ij} is known as the Extremal Dependence Measure (EDM; Resnick 2004; Larsson and Resnick 2012). Note that $\mathrm{X} = (X_1, \ldots, X_d) \in \mathrm{RV}_+^d(\alpha) \Rightarrow X_i \in \mathrm{RV}_+^1(\alpha), i = 1, \ldots, d$, and that the TPDM is invariant under marginal transformations of X which conserve $\mathrm{RV}_+^1(\alpha)$. We can estimate σ_{ij} empirically using n samples $\left\{ \mathrm{X}_t = (X_{1,t}, \ldots, X_{d,t})^{\mathrm{T}} \right\}_{t=1}^n$. Assuming that the limit in Eq. 2 holds exactly for sufficiently large R, the empirical estimator of σ_{ij} is

$$\widehat{\sigma}_{ij} := \frac{d}{k} \sum_{t=1}^{n} \frac{X_{i,t} X_{j,t}}{R_t^2} \mathbb{I}(R_t > r_0), \quad (3)$$

where $r_0 > 0$ is a suitably-chosen high radial threshold, $R_t := ||\mathrm{X}_t||_2$, and $k = \sum_{t=1}^n \mathbb{I}(R_t > r_0)$ is the number of radial threshold exceedances, where $\mathbb{I}(\cdot)$ is the indicator function. Choice of the exceedance threshold r_0 is an ubiquitous problem in extreme value analysis and involves the classical bias-variance trade-off: larger choices of r_0 reduce bias in estimation of σ_{ij} at the cost of reduced threshold exceedances and higher estimation uncertainty. Following, e.g., Cooley and Thibaud (2019), we take r_0 to be the k-th largest order statistic of $\{R_t\}_{t=1}^n$. Similarly to the bandwidth in moving sum based change point detection methods (to be discussed in Section 2.4), the order k is a hyper-parameter that must be specified. In Section 2.6, we discuss a technique for pooling over a sequence of k values, in order to improve our change point detection algorithm and to reduce sensitivity to the selected threshold.

2.3 Change point model

We consider a model where the tail dependence in a multivariate time series $\{\mathrm{X}_t : t \geq 1\}$ undergoes abrupt changes. In particular, we consider the setting

$$\mathrm{X}_t = \sum_{l=1}^{q+1} \mathrm{X}_t^{(l)} \mathbb{I}(\tau_{l-1} < t \leq \tau_l), \qquad 1 \leq t \leq n, \tag{4}$$

where each of the sequences $\left\{ \mathrm{X}_t^{(l)} = \left(X_{1,t}^{(l)}, \ldots, X_{d,t}^{(l)} \right)^{\mathrm{T}} : t \geq 1 \right\}, l = 1, \ldots, q+1,$ are independent, identically distributed random vectors in $\mathrm{RV}_+^d(2)$ (with standardised $\mathrm{RV}_+^1(2)$ margins and) with corresponding angular measure $H^{(l)}(\cdot)$, such that consecutive sequences differ in their extremal dependence measure characterised by their TPDM $\Sigma^{(l)} := (\sigma_{ij}^{(l)})_{i,j=1,\ldots,d}$:

$$\sigma_{ij}^{(l)} := \int_{\mathcal{S}_+^{d-1}} \theta_i \theta_j \mathrm{d}H^{(l)}(\boldsymbol{\theta}) = d \lim_{r \to \infty} \mathbb{E} \left[\frac{X_{i,1}^{(l)} X_{j,1}^{(l)}}{R^{(l)^2}} \mid R^{(l)} > r \right], \tag{5}$$

where $R^{(l)} := ||\mathrm{X}_1^{(l)}||_2$. That is, we consider a d-variate time series model containing q change points (structural breaks), with the TPDM of the observed time series, $\{\mathrm{X}_t : t \geq 1\}$, differing across the change points. The change points are collected in $\mathcal{C} := \{\tau_l\}_{l=0}^{q+1}$ with the convention that $\tau_0 := 0$ and $\tau_{q+1} := n$, and neighbouring segments differ in their TPDM such that $\Sigma^{(l)} \neq \Sigma^{(l-1)}$ for all $l = 2, \ldots, q+1$.

To further clarify our proposed change point model, we reiterate our assumptions on the time series $\{\mathrm{X}_t\}_{t=1}^n$:

- **Assumption 1:** Marginally, $X_{i,t} \in \mathrm{RV}_+^1(2)$ for all $i = 1, \ldots, d, t = 1, \ldots, n$. [Margins]
- **Assumption 2:** $\mathrm{X}_t \in \mathrm{RV}_+^d(2), t = 1, \ldots, n$, are independent. [Serial independence]
- **Assumption 3:** There exist a collection of $q+2$ (with $q \geq 0$) change points $\mathcal{C} = \{\tau_l\}_{l=0}^{q+1}$, with $\tau_0 = 0$ and $\tau_{q+1} = n$, such that the TPDM of the regularly varying X_t is $\Sigma^{(l)}$ for $\tau_{l-1} < t \leq \tau_l, l = 1, \ldots, q$, with $\Sigma^{(l)} \neq \Sigma^{(l')}$ for all $l \neq l'$. [Change points]*Remark*: As an implication of **Assumption 1**, the β-th moment of $X_{i,t}$ is finite only for $\beta < 2$. Note that $\mathrm{X}_t \in \mathrm{RV}_+^d(2)$ is an asymptotic statement, and we make no assumptions on the sub-asymptotic joint distribution of $\mathrm{X}_t \in \mathrm{RV}_+^d(2)$. We assume that only the corresponding TPDM, and thus the pairwise extremal dependence, differs across change points.

2.4 Moving sum based method for detecting changes in the TPDM

The moving sum (MOSUM) methodology has successfully been applied to a variety of change point testing (Chu et al. 1995) and estimation problems (Eichinger and Kirch 2018). We here define a MOSUM detector statistic for estimating multiple change points in the TPDM of the multivariate time series $\{\mathrm{X}_t\}_{t=1}^n$; we refer to this method as MOPED (MOving sum method for changes in Pairwise Extremal Depen-

dence). For a fixed bandwidth $G \in \mathbb{N}$, the method scans the data using a detector statistic, $T(G, t)$, computed on neighbouring moving windows of length G located either side of a candidate change point location t. The detector statistics measures the discrepancy between local estimates of the TPDM of the corresponding windows measured via their matrix norm. In particular, denoting the Frobenius norm by $||\cdot||_F$, we use the detector statistic

$$T(G,t) = ||D(G,t)||_F, \qquad G \leq t \leq n-G, \tag{6}$$

where $D(G, t)$ is the $d \times d$ matrix whose (i, j)-th entry is given by

$$D(G,t)_{i,j} = \frac{d}{k} \sum_{s=t-G+1}^{t} \frac{X_{i,s}X_{j,s}}{R_s^2} \mathbb{I}(R_s > r_{0,t}^{(-)}) - \frac{d}{k} \sum_{s=t+1}^{t+G} \frac{X_{i,s}X_{j,s}}{R_s^2} \mathbb{I}(R_s > r_{0,t}^{(+)}), \tag{7}$$

for $t \in [G, n-G]$, and where $r_{0,t}^{(-)}$ and $r_{0,t}^{(+)}$ are radial exceedance thresholds for the left and right G-windows, respectively, of the candidate change point t, such that

$$\sum_{s=t-G+1}^{t} \mathbb{I}(R_s > r_{0,t}^{(-)}) = k = \sum_{s=t+1}^{t+G} \mathbb{I}(R_s > r_{0,t}^{(+)}).$$

The entry $D(G,t)_{i,j}$ measures the local difference, in a G-window to the left and right of the candidate change point location t, in the pairwise tail dependence between the i-th and j-th components of the random vector X_t. Then, $T(G, t)$ is the aggregated measure of discrepancy across all components of X_t. Note that, theoretically, $\sigma_{i,j}$ for $i = j$ is equal to one. Therefore, in practice, we set $D(G,t)_{ii} = 0$ for $1 \leq i \leq d$ when computing $T(G, t)$.

When the data in the left and right G-windows, i.e., $\{\mathrm{X}_s : |t-s| \leq G\}$, does not undergo a change point, the local estimates of the TPDM to the left and right of t are expected to be similar and thus the detector statistic $T(G, t)$ is expected to be close to zero. On the other hand, if $|t - \tau_j| < G$ where τ_j is a true change point, then $T(G, t)$ is expected to increase and then decrease around τ_j with a local maximum at the change point $t = \tau_j$.

We illustrate this behaviour using the following example. A bivariate time series of length $n = 7000$ is generated with two change points in the tail dependence at locations $\tau_1 = 2000$ and $\tau_2 = 5000$ (for more details on the data generating process, see Scenario 1 in Section 3). The top panel of Fig. 1 shows a realisation of this process (on Pareto(2) margins) whilst the bottom panel gives the detector statistic $T(G, t)$, $G \leq t \leq n-G$, computed with $G = 1000$ and $k = G/10$. The detector statistic forms prominent peaks around the two change points, whilst it is otherwise generally flat.

Based on these observations, we detect and locate the change points in the TPDM by using local maximisers of the detector series $\{T(G,t)\}_{t=G}^{n-G}$ that are significantly large. We adopt the so-called η-criterion, first considered by Eichinger and Kirch (2018), for simultaneous estimation of multiple change points. For some fixed con-

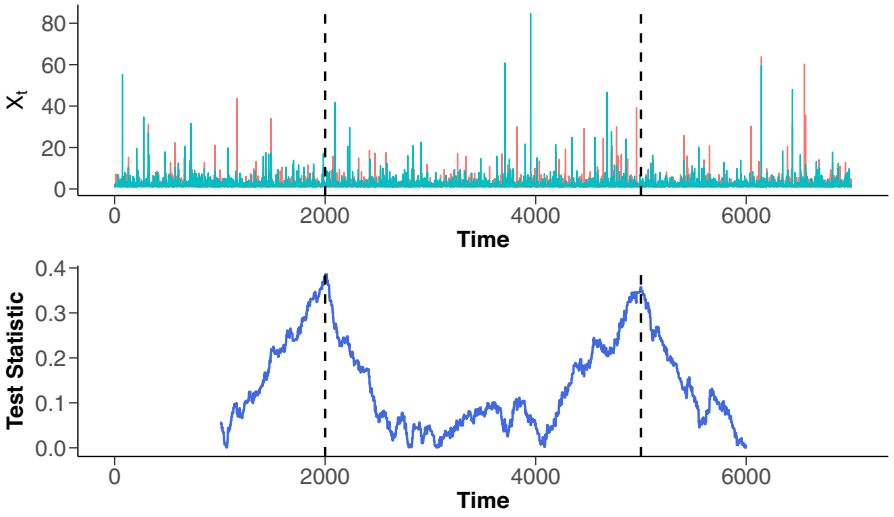

Fig. 1 Top: bivariate time series of length $n = 7000$ with change points $\tau_1 = 2000$ and $\tau_2 = 5000$ (vertical dashed lines). Bottom: corresponding detector statistic $T(G, t)$

stant $\eta \in (0, 1)$ and a threshold $C > 0$ (to be specified in Section 2.5), we identify any local maximiser(s) of $T(G, t)$, denoted $\widehat{\tau}$, which satisfies

$$T(G, \widehat{\tau}) > C \quad \text{and} \quad \widehat{\tau} = \underset{\{t:\, |t-\widehat{\tau}| \leq \eta G\}}{\arg\max} \; T(G, t). \tag{8}$$

That is, $\widehat{\tau}$ is declared a change point if it is a local maximiser of $T(G, t)$ over an ηG -neighbourhood, at which the threshold C is exceeded. In practice, local maximisers satisfying Eq. 8 may not be unique, in which case we use the mid-point of those such local maximisers instead. We denote the set of such estimators fulfilling Eq. 8 by $\widehat{\mathcal{C}} = \{\widehat{\tau}_1, \dots\}$ with total number of estimated change points $\widehat{q} := |\widehat{\mathcal{C}}|$.

We remark that, whilst here we use the MOSUM-based approach, alternatives based on cumulative sums (CUSUMs) such as the wild binary segmentation algorithm (Fryzlewicz 2014) are also possible. The MOSUM approach, however, offers several advantages relevant to our application: the use of two symmetric windows of size G in the computation of the test statistic allows for a natural use of order statistics to set the exceedance threshold (see discussion on extremal threshold selection in Section 3.1). This ensures that the same number of exceedances contribute in both windows, which helps with the empirical behaviour of the test statistic. Secondly, the MOSUM approach offers superior computational efficiency when combined with resampling methods (see Section 2.5), since the resampling procedure only need be performed once for a given bandwidth G.

2.5 Threshold selection via permutation testing

The choice of threshold C influences the finite sample performance of the method, which will depend on unknown quantities involved in the underlying extremal

dependence structure of the time series. Resampling is a widely-used method for the calibration of change point detection algorithms, including threshold selection (Matteson and James 2014; Bücher and Kojadinovic 2016). Under the assumption of independence of the observations, we utilise a permutation testing procedure in order to approximate the quantiles of $\max_{G \leq t \leq n-G}\{T(G,t)\}$ in the absence of any change point, from which we can then select the threshold C in a data-driven manner.

We perform M permutations of the times series. For the m-th permutation, we first randomly permute the time series $\{\mathrm{X}_t\}_{t=1}^n$ to form a new set of observations $\{\mathrm{X}_t^{[m]}\}_{t=1}^n$. Using these permuted observations, the MOPED detector statistics, $\{T^{[m]}(G,t)\}_{t=G}^{n-G}$, are obtained using Eq. 6. Finally, we record the maximum value of the MOPED detector statistics as $T^{[m]} := \max_{G \leq t \leq n-G}\{T^{[m]}(G,t)\}$.

The permutation test would be exact if all $n!$ possible permutations are computed. Since this is not computationally feasible, we instead perform an approximate test by performing the above procedure using $M \ll n!$ random permutations. From the M permutation test statistic values $\{T^{[m]}\}_{m=1}^M$, we set the threshold $C = Q_{1-\alpha}(\{T^{[m]}\}_{m=1}^M)$ as the empirical $(1-\alpha)$-th quantile of $\{T^{[m]}\}_{m=1}^M$ for the chosen approximate significance level $\alpha \in (0,1)$. Having selected the set of change points $\widehat{\mathcal{C}}$ using the rule in Eq. 8, we can compute an approximate p-value for each $\widehat{\tau} \in \widehat{\mathcal{C}}$ as

$$p(\widehat{\tau}) := \frac{1 + \left|\{1 \leq m \leq M : T^{[m]} \geq T(G,\widehat{\tau})\}\right|}{M+1}. \tag{9}$$

Remark: if independence of observations cannot be reasonably assumed, then alternative resampling methods that are applicable to the times series setting could be employed instead; for example the block bootstrapping approach of Ng et al. (2022), or the wild bootstrap of Shao (2010).

Combining the procedures outlined in Sections 2.4 and 2.5, we have the following full algorithmic description of MOPED in Algorithm 1.

Algorithm 1 MOPED algorithm.

Input: Multivariate time series $\{\mathbf{X}_t\}_{t=1}^n$, bandwidth G, order k, selection parameter η, permutation parameters α and M

for $t \in \{G, \ldots, n-G\}$ **do**
  Compute MOSUM statistic $T(G,t)$ according to Eq. 6

for $m \in \{1, \ldots, M\}$ **do**
  Compute permutation statistic $T^{[m]}$

Set threshold $C \leftarrow Q_{1-\alpha}(\{T^{[m]}\}_{m=1}^M)$

$\widehat{\mathcal{C}} \leftarrow$ Set of change point estimators obtained with bandwidth G and threshold C according to Eq. 8

for $\widehat{\tau} \in \widehat{\mathcal{C}}$ **do**
  Compute p-values $p(\widehat{\tau})$ according to Eq. 9

Output: Estimated change locations $\widehat{\mathcal{C}}$, approximate p-values $\{p(\widehat{\tau}) : \widehat{\tau} \in \widehat{\mathcal{C}}\}$

2.6 Multiscale, multiple threshold change point detection

In moving window-based change point detection methods, one of the key considerations is the choice of detecting bandwidth G, whilst in extreme value statistics, a crucial aspect is the choice of order k used in the extremal threshold. In this section, we extend the single-scale, single extremal threshold MOPED algorithm described in Section 2.4 to incorporate multiple bandwidths and multiple extremal thresholds.

Inspired by methods proposed by Messer et al. (2014) and Mcgonigle and Cho (2025), we first describe an approach that allows us to perform the MOPED procedure for a fixed G but multiple values of k, and combine the resulting change point estimates into a single set. We then describe a similar method to combine these results across multiple values of G.

First, we denote a set of $J > 1$ ranks $\mathcal{K} := \{k_j,\ 1 \leq j \leq J : k_1 < \ldots < k_J\}$. We run MOPED for each $k \in \mathcal{K}$, denoting the corresponding change point estimates as $\widehat{\mathcal{C}}(k)$, and then combine the change point estimates across different values of k using the 'bottom-up' merging method proposed by Messer et al. (2014). We first add all estimates in $\widehat{\mathcal{C}}(k_1)$, for k_1 the smallest rank, to the set of final estimates $\widehat{\mathcal{C}}$. Then, sequentially for $j = 2, \ldots, J$, we accept $\widehat{\tau} \in \widehat{\mathcal{C}}(k_j)$ as a final change point estimate if and only if $\min_{t \in \widehat{\mathcal{C}}} |t - \widehat{\tau}| \geq \eta G$ with η as before. That is, we only accept the estimates that are not close to previously detected change point estimates using a smaller rank/larger extremal threshold; we expect that close change point estimates will correspond to the same true change point, and so enforce that the final change point estimates $\widehat{\mathcal{C}}$ are sufficiently far apart. The algorithmic description of the multiple threshold MOPED procedure is summarised in Algorithm 2 below.

Algorithm 2 Multi-threshold MOPED algorithm.

Input: Multivariate time series $\{X_t\}_{t=1}^n$, bandwidth G, set of ranks $\mathcal{K}$, selection parameter η, permutation parameters α and M

Initialise $\widehat{\mathcal{C}} \leftarrow \mathcal{C}(\mathcal{K}) \leftarrow \emptyset$
for $k \in \mathcal{K}$ **do**
  $\{\widehat{\mathcal{C}}(k), \mathcal{P}(k)\} \leftarrow \mathsf{MOPED}(\{X_t\}_{t=1}^n, G, k, \eta, \alpha, M)$
  $\mathcal{C}(\mathcal{K}) \leftarrow \mathcal{C}(\mathcal{K}) \cup \widehat{\mathcal{C}}(k)$

for $\widehat{\tau} \in \mathcal{C}(\mathcal{K})$ *in increasing order with respect to* k **do**
  if $\min_{t \in \widehat{\mathcal{C}}} |t - \widehat{\tau}| \geq \eta G$ **then** Add $\widehat{\tau}$ to $\widehat{\mathcal{C}}$

Output: Estimated change locations $\widehat{\mathcal{C}}$, estimated number of changes $\widehat{q} = |\widehat{\mathcal{C}}|$

The above approach outlines how to combine MOPED estimates with multiple threshold orders k into a single set of change point estimates, for a single fixed bandwidth G. Appropriate choice of this G is necessary to achieve good practical performance: a larger G is suitable for detecting smaller changes over large time intervals, whilst a smaller G can detect larger changes over shorter time periods. A multiple-bandwidth moving sum procedure can be particularly well-suited to so-called multiscale change points where both of the change point types described above

occur within the same time series. Using multiple bandwidths, and then merging the results, can therefore improve the adaptivity of our proposed moving window-based procedure.

Here, we again combine the MOPED method with the bottom-up merging procedure, but now with respect to the different bandwidths, G. This is motivated by the observation that, with larger bandwidths, the MOPED detector statistic may be contaminated by several changes, resulting in the corresponding estimates being potentially unreliable. Therefore, it is reasonable to keep all change point estimates from the smallest bandwidth and, moving on to the next bandwidth in an iterative manner, only keep new estimates that cannot be accounted for by the previously accepted estimators. Denote by $\mathcal{G} = \{G_h, 1 \leq h \leq H : G_1 < \ldots < G_H\}$ the set of $H > 1$ bandwidths, and let $\widehat{\mathcal{C}}(G)$ denote the set of change point estimates detected using the multi-threshold MOPED procedure, as in Algorithm 2, with single bandwidth G and ranks $\mathcal{K}$. Then, as with the multiple threshold merging, we add all estimates in $\widehat{\mathcal{C}}(G_1)$ returned with the first bandwidth G_1 to the set of final estimates $\widehat{\mathcal{C}}$ and, sequentially for $h = 2, \ldots, H$, accept $\widehat{\tau} \in \widehat{\mathcal{C}}(G_h)$ as a final change point estimate if and only if $\min_{t \in \widehat{\mathcal{C}}} |t - \widehat{\tau}| \geq \eta G_h$ with η as before. That is, we only accept the estimates that do not correspond with the change points which have previously been detected at a finer scale. The full multiscale, multi-threshold version of MOPED is presented in Algorithm 3 for completeness.

Algorithm 3 Multiscale, multi-threshold MOPED algorithm.

Input: Multivariate time series $\{X_t\}_{t=1}^n$, set of bandwidths $\mathcal{G}$, set of ranks $\mathcal{K}$, selection parameter η, permutation parameters α and M

Initialise $\widehat{\mathcal{C}} \leftarrow \mathcal{C}(\mathcal{G}) \leftarrow \emptyset$

for $G \in \mathcal{G}$ **do**

  $\widehat{\mathcal{C}}(G) \leftarrow$ MULTI-THRESH-MOPED$(\{X_t\}_{t=1}^n, G, \mathcal{K}, \eta, \alpha, M)$

  for $\widehat{\tau} \in \widehat{\mathcal{C}}(G)$ **do** Add $(\widehat{\tau}, G)$ to $\mathcal{C}(\mathcal{G})$

for $\widehat{\tau} \in \mathcal{C}(\mathcal{G})$ *in increasing order with respect to* G **do**

  if $\min_{t \in \widehat{\mathcal{C}}} |t - \widehat{\tau}| \geq \eta G$ **then** Add $\widehat{\tau}$ to $\widehat{\mathcal{C}}$

Output: Estimated change point locations $\widehat{\mathcal{C}}$, estimated number of changes $\widehat{q} = |\widehat{\mathcal{C}}|$

3 Numerical results

In this section, we provide practical recommendations for the implementation of our approach, and conduct numerical studies to investigate the efficacy of the method.

3.1 Practical implementation

Computational complexity Fast evaluation of the test statistic is paramount for reduced computational complexity of MOPED. Note that, for a single bandwidth

G and order k, a naive approach to computing the test statistic by recalculating the sums in Eq. 6 would yield a method with computational complexity $O(MnG)$. Owing to the MOSUM-based approach, we can instead sequentially update some of the quantities involved in the computation of Eq. 6, reducing the computational cost to $O(Mn\log(G))$ as follows.

The test statistic $D(G,t)_{i,j}$ in Eq. 7 can be updated in a sequential fashion by first computing the time series of products $\{X_{i,s}X_{j,s}/R_s\}_{s=1}^{n}$, as well as the rolling sorted values $\{R^{(s)}\}_{s=t-G+1}^{t}$ of $\{R_s\}_{s=t-G+1}^{t}$ for $G \leq t \leq n$. Then, since only the largest k values of $\{R^{(s)}\}_{s=t-G+1}^{t}$ make a contribution in the first term in Eq. 7 (and likewise for $\{R^{(s)}\}_{s=t+1}^{t+G}$ and the second term), the summation terms for computing $D(G,t+1)_{i,j}$ can be updated using the associated rolling sorted values, by keeping track of deletions and insertions that occur in updating the rolling sorts from time t to time $t+1$, which can be achieved in $O(\log(G))$ complexity (see e.g., Vanegas et al. 2022). Therefore we require $O(n\log(G))$ total computations to compute the detector statistic, and hence $O(Mn\log(G))$ when including the permutation testing.

Bandwidth Due to the nonparametric nature of MOPED, in combination with the difficulty of detecting changes in the extremal dependence, it is advisable to use larger bandwidths than that shown to work well for the MOSUM procedure for univariate mean change detection (Eichinger and Kirch 2018). In practice, the practitioner may have prior knowledge that aids the choice of G. In our simulation studies and data applications, we consider G in multiples of 500.

Parameters for change point estimation and detection We set $\eta = 0.4$ in Eq. 8 following the recommendation in Meier et al. (2021). For the permutation testing, we use $M = 200$ permutations with approximate significance levels of $\alpha = 0.05$ or $\alpha = 0.1$.

Extremal threshold The choice of threshold level r_0, and the corresponding order k that defines the k-th order statistic, depend on the practitioner's interest and domain knowledge. This is a problem commonly faced both in extreme value theory (see, e.g., Scarrott and MacDonald 2012; Murphy et al. 2025) and by general-purpose change point detection methods, such as the choice of the quantile level in Vanegas et al. (2022) or the choice of lag in Mcgonigle and Cho (2025). In the absence of any information, we recommend setting the threshold as $k = 0.05G$, such that r_0 corresponds to the 95% empirical quantile.

Multiscale, multi-threshold MOPED As discussed in Section 2.6, the multiscale, multi-threshold variant of MOPED pools change point estimates over values of the bandwidth G and order k. We show in the following simulation study, Section 3.2, that this approach can improve the accuracy of change point estimation relative to MOPED with fixed hyper-parameters (G and k). However, the multiscale, multi-threshold version is less conservative than the "fixed" version of MOPED, in the sense that it may produce spurious change point estimates. As observed by McGonigle and Cho (2023), bottom-up merging has a propensity to produce false positives as

it accepts all estimates from the finest bandwidth. Thus, we advocate use of the multiscale, multi-threshold procedure in applications where the accuracy of change point estimates is more important than testing for structural breaks in the extremal dependence of multivariate time series (and vice versa for the fixed variant of MOPED).

3.2 Simulation study

Our MOSUM-based extremal dependence change point algorithm is compared against two competing approaches: the state-of-the-art nonparametric method, E-divisive (Matteson and James 2014), and the parametric extreme value method of Hazra and Bose (2025). The latter identifies a single change point in bivariate data using a likelihood ratio test between two candidate Hüsler-Reiss max-stable distributions (Hüsler and Reiss 1989).

Three data generating scenarios are considered. In each, data are first generated on a uniform scale before the margins are transformed to one appropriate for the applied change point method: for MOPED, E-divisive, and the method of Hazra and Bose (2025), these are Pareto(2), standard Gaussian, and standard Gumbel, respectively. In Scenarios 1 and 2, we generate serially-independent data that satisfy Assumption 2 of our change point model (see Section 2.3). In Scenario 3, we consider a misspecified setting with mild autocorrelation in the multivariate time series data.

For Scenario 1, we generate data from a d-variate Student's t-copula with $\nu = 3$ degrees of freedom and with correlation matrix, Ω, varying across the change point; we switch between i) independence (i.e., $\Omega = I_{d\times d}$) and ii) all pairwise correlations equal to $\rho > 0$ (i.e., where all off-diagonal elements of Ω are equal to ρ). Note that, in cases where we have $q = 0$ change points, the second correlation matrix is used to generate the data, i.e., we consider non-zero pairwise ρ.

For Scenario 2, we switch between a d-variate Student's t-copula ($\nu = 3$) and a Gaussian copula, both with positive-definite correlation matrix Ω. The matrix Ω is generated randomly and kept equal across the change point. In this way, we design a data generating model with no change point(s) in the correlation structure, but where the theoretical value of the TPDM Σ varies across the change point; note that the theoretical TPDM for a Gaussian copula is the identity matrix, whilst the TPDM for the Student's t-copula has non-zero diagonal entries (if the corresponding entry in Ω is strictly positive). In other words, the extremal dependence class of the data changes from asymptotic independence to asymptotic dependence; see discussion in Section 1. Three random seeds are used for generating Ω and results are reported for each.

For Scenario 3, we generate data from a first-order vector autoregressive model:

$$\mathrm{X}_t = \Phi \mathrm{X}_{t-1} + \boldsymbol{\epsilon}_t, \quad t = 1, \ldots, n,$$

where $\Phi = \phi I_{d\times d}$ for autoregressive (AR) parameter $\phi \in [0, 1)$ and where $\boldsymbol{\epsilon}_t$ is d-variate Gaussian with zero mean, unit variance, and correlation matrix Ω. The correlation matrix Ω of $\boldsymbol{\epsilon}_t$ differs across the change points, and it is parametrised as described above for Scenario 1. We consider sample size $n = 5000$, and two values of ρ: a weak and strong dependence case with $\rho = 0.2$ and $\rho = 0.6$, respectively. We

vary dimension $d \in \{2, 8, 15\}$ between scenarios, and consider $q \in \{0, 1, 2\}$ equally-spaced change points. For MOPED, we consider two versions: a "fixed" version of MOPED (see Algorithm 1), where the tuning parameters are $G = 1500$ and $k = 0.1G$, and the multiscale, multi-threshold MOPED (hereafter referred to as MMMOPED; see Algorithm 3). For MMMOPED, we merge over the sequence of bandwidths $\mathcal{H} = \{500, 1000, 1500\}$ and ranks $\mathcal{K} = \{0.2G, 0.1G, 0.05G\}$. In all cases, we use a significance level of $\alpha = 0.1$. For brevity, we do not consider MMMOPED for Scenario 3, but we do consider three values of the AR parameter ϕ: $\phi = 0$, $\phi = 0.3$, and $\phi = 0.6$, corresponding to zero, mild, and strong autocorrelation.

Note that the method of Hazra and Bose (2025) is only applicable for dimension $d = 2$. Moreover, it cannot test for the significance of a change point, and rather it returns exactly the most likely change point (according to a parametric likelihood ratio test). Thus, we only report the results of the Hazra and Bose (2025) method when $d = 2$ and $q = 1$.

Tables 1, 7, and 10–12 (Appendix A) provide the results of the simulation study when $d = 2$ for Scenarios 1, 2, and 3, respectively; similar tables for $d = 8$ and $d = 15$ are provided in Appendix A. Across all tables, we report the distribution of $\widehat{q} - q$, where $\widehat{q}$ is the estimated number of change points and q is the truth, alongside two measures of accuracy for the change point estimates (when $q > 0$) advocated by Van Den Burg and Williams (2020): the average covering metric (CM; Arbelaez et al. 2010) and V-measure (VM; Rosenberg and Hirschberg 2007). Define the partition

Table 1 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 1, $d = 2$, of the simulation study. For the fixed MOPED method, $G = 1500$; for MMMOPED, we consider $\mathcal{H} = \{500, 1000, 1500\}$. The modal value of $\widehat{q} - q$ is provided in bold in each row

			$\widehat{q} - q$						
ρ	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
$\rho = 0.2$	0	MOPED	-	-	**0.897**	0.092	0.011	-	-
		MMMOPED	-	-	**0.568**	0.299	0.133	-	-
		E-divisive	-	-	**0.948**	0.013	0.039	-	-
$\rho = 0.6$		MOPED	-	-	**0.889**	0.100	0.011	-	-
		MMMOPED	-	-	**0.571**	0.305	0.124	-	-
		E-divisive	-	-	**0.953**	0.015	0.032	-	-
$\rho = 0.2$	1	MOPED	-	0.402	**0.565**	0.033	0.000	0.729	0.449
		MMMOPED	-	0.153	**0.542**	0.238	0.067	0.783	0.608
		E-divisive	-	**0.902**	0.059	0.035	0.004	0.527	0.058
		Hazra and Bose (2025)	-	-	1.000	-	-	0.924	0.829
$\rho = 0.6$		MOPED	-	0.000	**0.993**	0.006	0.001	0.973	0.923
		MMMOPED	-	0.000	**0.644**	0.284	0.072	0.928	0.873
		E-divisive	-	0.000	**0.947**	0.018	0.035	0.989	0.975
		Hazra and Bose (2025)	-	-	1.000	-	-	0.995	0.979
$\rho = 0.2$	2	MOPED	**0.365**	0.344	0.291	0.000	0.000	0.604	0.469
		MMMOPED	0.127	0.247	**0.462**	0.146	0.018	0.709	0.637
		E-divisive	**0.932**	0.025	0.039	0.004	0.000	0.347	0.031
$\rho = 0.6$		MOPED	0.000	0.000	**1.000**	0.000	0.000	0.963	0.929
		MMMOPED	0.000	0.000	**0.552**	0.370	0.078	0.918	0.888
		E-divisive	0.000	0.000	**0.947**	0.025	0.028	0.986	0.974

$\mathcal{P} = \{\mathcal{A}_l\}_{l=1}^{q+1}$ of $\{1, \ldots, n\}$ provided by the true change point locations $\{\tau_l\}_{l=1}^{q+1}$, i.e., $\mathcal{A}_j = \{\tau_{j-1} + 1, \ldots, \tau_j\}$, and its corresponding estimate, $\widehat{\mathcal{P}} = \{\widehat{\mathcal{A}}_l\}_{l=1}^{\widehat{q}+1}$. The CM is then

$$\mathrm{CM}(\widehat{\mathcal{P}}, \mathcal{P}) = \frac{1}{n} \sum_{\mathcal{A} \in \mathcal{P}} |\mathcal{A}| \max_{\widehat{\mathcal{A}} \in \widehat{\mathcal{P}}} \left\{ \frac{|\mathcal{A} \cap \widehat{\mathcal{A}}|}{|\mathcal{A} \cup \widehat{\mathcal{A}}|} \right\}.$$

The VM is calculated using the conditional entropy of $\widehat{\mathcal{P}}$; for brevity, we omit its formal definition. Both metrics take values in [0, 1], with larger values indicating better accuracy. Results are averaged over 1000 repeated experiments.

For Scenario 1, Tables 1, 2, and 3 give the simulation study results when $d = 2$, $d = 8$, and $d = 15$, respectively. We observe that MOPED performs competitively with E-divisive across all considered configurations of (d, ρ, q), as both methods attain similar CM and VM scores, and the distributions of $\widehat{q} - q$ are broadly similar. As expected, MOPED performs better when the change in strength of extremal dependence is greater (i.e., when $\rho = 0.6$). In some cases, MOPED performs markedly better than E-divisive; for example, in the case where we expect identification of a change point to be the most difficult, i.e., when $\rho = 0.2$ and $d = 2$, MOPED provides both higher CM/VM scores (when $q \geq 1$) and is more likely to identify the true number of change points. Moreover, Tables 1, 2, and 3 illustrate improvements can be made to MOPED by aggregating over sequences of bandwidths G and ranks k; we note that, in cases where the change in extremal dependence is small, i.e., $\rho = 0.2$ and $q > 0$, multiscale, multi-threshold MOPED provides more accurate estimates of the change point locations (higher CM and VM scores). Finally, we note that the parametric extreme value method of Hazra and Bose (2025) provides the most accurate estimates of the change point when there is known to be only $q = 1$. However, this method is limited to dimension $d = 2$, whereas Tables 2 and 3 illustrate that MOPED performs well for much higher dimensions. We note here that we have not optimised the hyper-parameters, G and k, for MOPED or MMMOPED, and so these results represent a lower-bound on the performance of our change point detection algorithm. For Scenario 1, we repeat the numerical experiment with a lower sample size of $n = 2000$. In this setting, to reflect the lower sample size, we use a smaller bandwidth of $G = 500$ for MOPED, and a sequence of bandwidths $\mathcal{H} = \{300, 400, 500\}$ for MMMOPED. The results are reported in Tables 4–6. We observe generally poorer performance of all methods (relative to $n = 5000$), but MOPED retains its competitiveness with E-divisive and the parametric detection method (Hazra and Bose 2025). We also observe that MOPED exhibits robustness to the sample size in cases where change points are easier to detect, i.e., $\rho = 0.6$. Note that, in our real data application (Section 4), the lowest sample size that we consider is $n \approx 3500$.

For Scenario 2, Tables 7, 8, and 9 give the results for $d = 2$, $d = 8$, and $d = 15$, respectively. Note that, for Scenario 2, we only consider the true number of change points $q > 0$; the case where $q = 0$ is equivalent to the similar case in Scenario 1. In contrast with Scenario 1, the change in the distribution across the change point is more subtle in Scenario 2; the data generating copula changes from one which exhibits asymptotic dependence to one which exhibits asymptotic independence (see

Section 1), but with the (sub-asymptotic) correlation structure remaining constant across the change point. In the case where the change point is particularly difficult to identify, for example, when the dimension is low ($d = 2$), all methods underestimate the true number of change points q. However Tables 7 and 8 illustrate that the fixed-parameter MOPED method generally outperforms both E-divisive and the method of Hazra and Bose (2025), in terms of both accuracy (higher CM/VM scores) and correct determination of q (with more mass at zero for the estimated distribution of $\widehat{q} - q$). We see further improvements on these results when using MMMOPED, especially in the case where the change point is most difficult to identify, i.e., $d = 2$; see Table 7. The performance of MOPED and E-divisive for higher dimension $d = 15$ are broadly similar, and there is no clear dominant approach.

For Scenario 3, Tables 10–12 give the results for $d = 2$ while Tables 13–15 give the results for $d = 8$. As the value of the autocorrelation parameter ϕ increases from 0 to 0.6, we observe the expected degradation in the performance of all considered change point detection methods: MOPED, E-divisive, and the parametric approach of Hazra and Bose (2025). For all methods, we observe a reduction in estimation accuracy for q, the number of change points, as well as the average covering metric and V-measure. However, we note that our extremal change point method, MOPED, exhibits relative robustness to the presence of autocorrelation. When $\phi = 0$, i.e., the data are independent, E-divisive outperforms MOPED in some cases (Tables 10 and 13). However, even under mild autocorrelation (Tables 11 and 14), the performance of E-divisive is markedly worse than MOPED, with E-divisive routinely overestimating the number of change points by at least two. Conversely, MOPED facilitates reasonable estimation accuracy even under strong autocorrelation (Tables 12 and 15), albeit with some evidence that the true number of change points is underestimated. Note that the method of Hazra and Bose (2025) is only designed for estimating exactly one change point, without any form of testing, whereas MOPED and E-divisive estimate the number of change points. As a result, we observe that the performance of the Hazra and Bose (2025) method is relatively unaffected by the presence of autocorrelation.

4 Application

4.1 Overview

We now consider multivariate electroencephalogram (EEG) recordings from 79 human neonates, who were monitored for the occurrence of seizures in the Baby Brain Activity Center, which is a neonatal intensive care unit in Finland (Stevenson et al. 2019). The signals are observed at $d = 19$ channels of the brain, which are illustrated in Fig. 2, and the sample size n differs between subjects. Alongside the observed time series, the dataset includes annotations of the occurrence of seizures provided by three expert clinicians. Extremal dependence of these data has been previously analysed by Talento et al. (2025a, 2025b), who identified changes in the extremal dependence strength and brain connectivity for those neonates who had experienced seizures. Similarly, Guerrero et al. (2023) and Redondo et al. (2026)

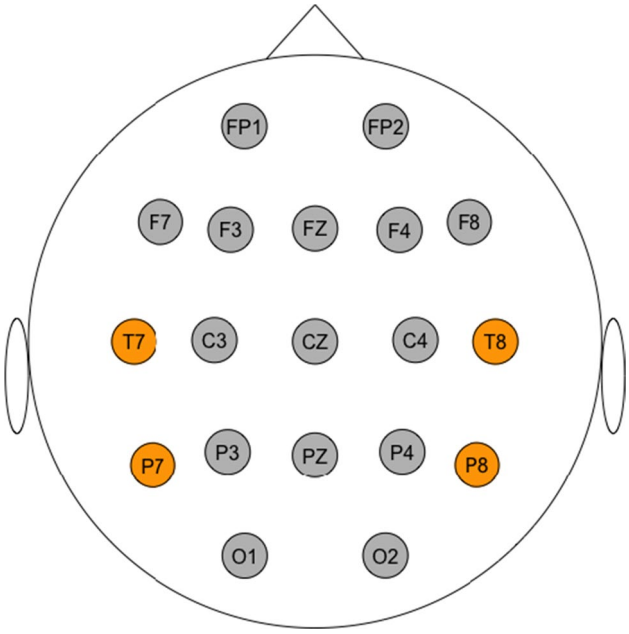

Fig. 2 Standard 10-20 EEG scalp topography with 19 channels. The $d = 4$ channels coloured in orange are those considered in the lower dimensional case study

have previously identified changes in the pre- and post-seizure extremal behaviour of brains, albeit in adults rather than neonates. Given the abrupt nature of seizures (Nih 2023) and their links with changing extremal dependence, MOPED may be able to identify their occurrence or absence.

Following Talento et al. (2025b), we pre-process each time series by applying a Butterworth low-pass filter which retains only the EEG signals comprising oscillations in the Delta frequency band of (0, 4] Hertz (Hz). We focus on the Delta band, commonly associated with sleep patterns, as erratic behaviour in this frequency band is often linked with seizures (Douglass et al. 2002; Muthuswamy et al. 2002). The Butterworth low-pass filter is applied using the `R` package `eegkit` (Helwig 2018). To reduce serial dependence in these data, we sub-sample every 256th value, where 256Hz is the sampling rate, and thus we sub-sample every second. While this does not completely remove the autocorrelation in the data, the results from Scenario 3 of the simulation study (Section 3.2) suggest that MOPED is relatively robust to the presence of mild autocorrelation. We consider only continuous time series, i.e., neonates with no interruption in the Delta band activity across the observation period, and discard all other time series. The remaining data are marginally standardised to Pareto(2) margins by applying an empirical rank transform. As discussed by Guerrero et al. (2023), these data exhibit temporal non-stationarity in their marginal extremal behaviour. For simplicity, we have not explicitly modelled the data margins. As such, the following estimated change points may reflect changes in the marginal extremal behaviour and/or the extremal dependence structure. However, MOPED still provides a quick and convenient exploratory tool for identifying structural changes in the

extremal behaviour of these data and, as we are to show, also provides estimates that are in better agreement, relative to the competitor (E-divisive), with expert opinion.

We compare change point estimates using MOPED and E-divisive. For MOPED, the hyper-parameters are set to $G = 1000$ and $k = 0.1G$; we report estimates for other hyper-parameter configurations in Appendix B. As discussed in Sections 2.6 and 3.2, the multiscale, multi-threshold MOPED (MMMOPED) algorithm is less conservative than the fixed MOPED method, and so it always provides larger estimates $\widehat{q}$ of the total number of change points. Thus, for brevity, we generally omit the results of MMMOPED from the main text, and instead provide these in Figs. 10 and 11 of Appendix B; for MMMOPED, we pool over bandwidth values of $\mathcal{H} = \{500, 1000, 1500\}$ and ranks $\mathcal{K} = \{0.1G, 0.05G, 0.025G\}$. For all three methods, we consider both a low and high dimensional setting, with $d = 4$ and $d = 19$ signals, respectively, and, to account for any autocorrelation in the data, we use a more conservative significance level of $\alpha = 0.05$. When $d = 4$, we follow Talento et al. (2025b) and consider signals observed at channels T7, T8, P7, and P8, which are equally distributed across the right and left hemispheres of the brain (see Fig. 2).

We consider a subset of the 79 neonates and focus on three cases, using the expert clinician annotations: 1) where there are no recorded seizures; 2) where there is one contiguous time period of seizure activity; 3) where there are multiple discontiguous periods of seizure activity.

4.2 Results

For Case 1, we consider $d = 4$ and Subjects 10, 27, and 37 with sample sizes $n = 5427$, $n = 3496$, and $n = 4578$, respectively. We may expect that subjects that are free of seizures experience typical brain activity, and thus the "ground truth" is $q = 0$ change points in the extremal dependence of their EEG time series. Figure 3 provides estimates of the change points from E-divisive and MOPED. Here we

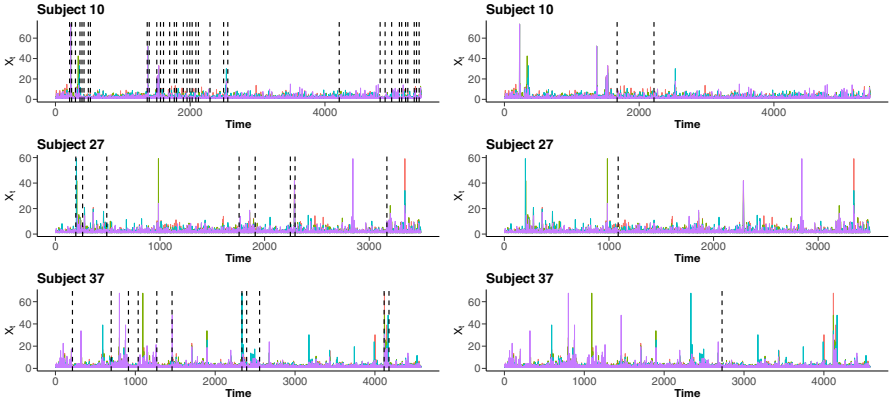

Fig. 3 Change point estimates for Subjects 10, 27, and 37 (top to bottom row). Estimates (vertical dotted lines) are provided by E-divisive (left) and MOPED (right), with $G = 1000$ and $k = 100$. Observations of $\{X_{it}\}_{t=1}^{n}$ are plotted against time t, with the colour corresponding to the i-th channel, $i = 1, \ldots, d$. Here, $d = 4$ and no subjects were identified as having experienced a seizure

observe that E-divisive has a tendency to estimate substantially more change points than MOPED; conversely, MOPED routinely estimates $\widehat{q}$ close to zero significant change points in the TPDM. A similar conclusion is drawn when considering all $d = 19$ channels; see Fig. 7. Figure 8 in Appendix B provides a sensitivity study of the MOPED estimates for Subject 27, when $d = 4$. Across a range of G and k values, MOPED provides at most $\widehat{q} = 2$ significant change points, whilst E-divisive returns $\widehat{q} = 7$. As we discuss later, MOPED tends to provide higher estimates of $\widehat{q}$ for subjects who have experienced seizures.

We now focus on three subjects for Case 2, i.e., they were identified as experiencing a single uninterrupted period of seizure throughout the observation period. Subjects 24, 33, and 75 ($n = 3706$, $n = 3747$, and $n = 3955$, respectively) all experience one contiguous period of interrupted seizure; for Subjects 33 and 75, this occurs midway through the recording period, whereas for Subject 24, the seizure occurs at the end. We note that for Subject 33 there is a small period of seizure activity at the start of the observation period, but this was only identified by one of the clinical experts; hence we choose to ignore this period of activity. Here, we consider all channels ($d = 19$) with G and k as previously, and provide the results in Fig. 4. As in Figs. 3, 4 shows that the E-divisive method is sensitive to changes in the EEG signals and routinely estimates a larger number q of change points compared to MOPED. Although we have no ground truth in this application, it may be reasonable to assume that the true number of change points is close to two, with a change in the extremal behaviour in brain activity occurring both before and after the seizure period. Estimates from MOPED seem to agree with this hypothesis, as the method provides a single change point estimate proceeding the seizure occurrence of Subject 24 (where it would not be feasible to estimate one after the seizure event) and provides two change point estimates for Subject 33, which are either side of the annotated period of seizure activity. For Subject 75, MOPED provides a single change point estimate, at the end

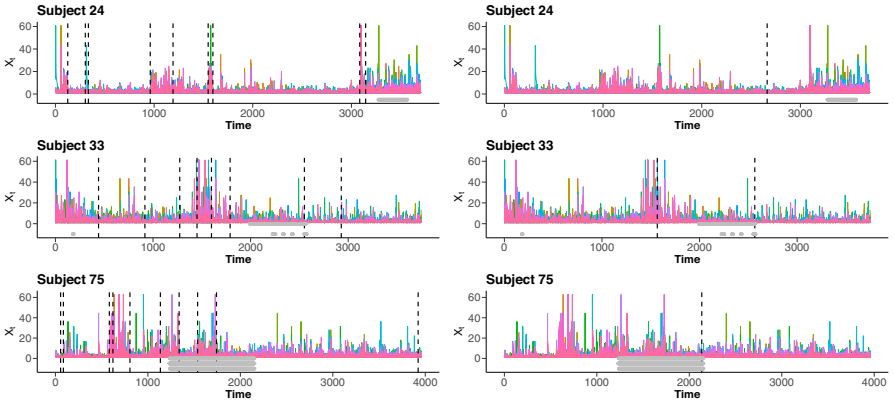

Fig. 4 Change point estimates for Subjects 24, 33, and 75 (top to bottom row). Estimates (vertical dotted lines) are provided by E-divisive (left) and MOPED (right), with $G = 1000$ and $k = 100$. Observations of $\{X_{i,t}\}_{t=1}^{n}$ are plotted against time t, with the colours corresponding to the i-th channel, $i = 1, \ldots, d$. Here $d = 19$ and the grey dots along the x-axis denote time-points identified as seizures by expert clinicians; these points are arranged in three rows, corresponding to the different expert clinicians

of the seizure activity period. Figure 9 in Appendix B provides a sensitivity study of the MOPED estimates for Subject 33; we generally observe one or two change point estimates, that are close to the period of seizure activity.

Figure 5 provides local estimates of the TPDM for the subjects shown in Fig. 4. For each window of data demarcated by the estimated change points provided in the right panels of Fig. 4, we estimate the TPDM using all data and with r_0 corresponding to the 0.95-quantile of the empirical radii. Figure 5 showcases the distinct changes in the TPDM which are being identified by MOPED. Of note are the TPDM estimates for Subject 33, which have a block-diagonal structure that corresponds to strong extremal dependence between only Channels C4–O2 (lower

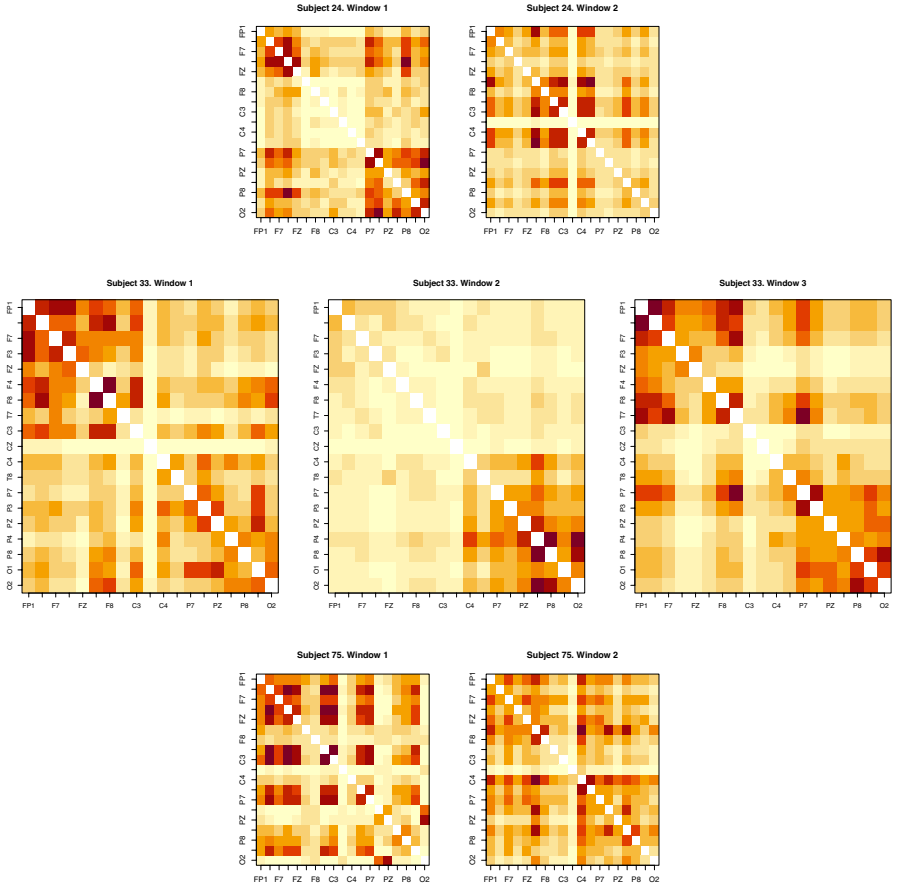

Fig. 5 Heat maps of estimated TPDM for Subjects 24, 33, and 75 for the windows of observations demarcated by the MOPED change point estimates shown in Fig. 4. The windows are ordered from left to right, and the number of windows changes with the subject. The column and row names correspond to the channels in Fig. 2

hemisphere; see Fig. 2) during the period of seizure activity. This localised extremal activity may be indicative of a seizure.

For the final Case 3, we consider Subjects 22, 51, and 74 ($n = 3844$, $n = 4701$, and $n = 4364$), which all experience multiple prolonged periods of seizure activity. Figure 6 provides the E-divisive, MOPED, and MMMOPED estimates of the change points. As we have seen previously for Cases 1 and 2, E-divisive provides a higher estimated number of change points, which occur throughout the times series. For Subject 22, we observe that MOPED provides two change point estimates: one at the end of the first seizure period, and one at the start of the second seizure period. We observe a similar result for Subject 51, except with an additional change point estimate close to the onset of the second seizure period. For Subject 74, MOPED provides two change point estimates, which are located between the first and second, and second and third, periods of seizure activity. As expected, MMMOPED provides more change point estimates. For Subjects 22 and 74, Fig. 6 illustrates that MMMOPED provides a total of $\widehat{q} = 6$ change point estimates, which are clustered around the periods of seizure activity.

5 Discussion

We have proposed a computationally efficient algorithm, MOPED, for detecting structural changes in the tail dependence of sequences of regularly varying random vectors. Our approach uses a moving sum procedure to detect changes in the TPDM of Cooley and Thibaud (2019) in a nonparametric fashion. In Section 2.6, we proposed a multiscale, multi-threshold method for pooling change point estimates across values of the bandwidth G and exceedance threshold required for local estimation of the TPDM. Our simulation studies demonstrated that

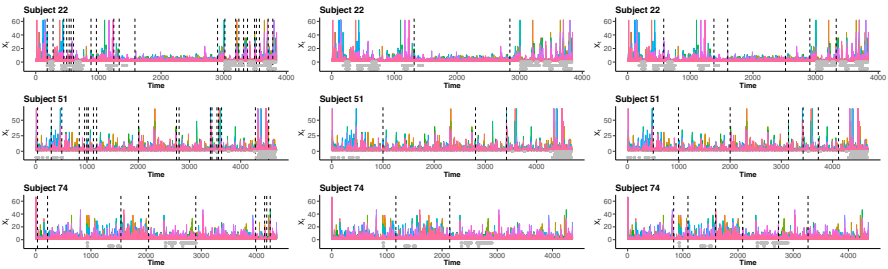

Fig. 6 Change point estimates for Subjects 22, 51, and 74 (top to bottom row). Estimates (vertical dotted lines) are provided by E-divisive (left), MOPED (centre; $G = 1000$ and $k = 100$), and MMMOPED (right). Observations of $\{X_{i,t}\}_{t=1}^{n}$ are plotted against time t, with the colours corresponding to the i-th channel, $i = 1, \ldots, d$. Observations of $\{X_{i,t}\}_{t=1}^{n}$ are plotted against time t, with the colours corresponding to the i-th channel, $i = 1, \ldots, d$. Here $d = 19$ and the grey dots along the x-axis denote time-points identified as seizures by expert clinicians; these points are arranged in three rows, corresponding to the different expert clinicians

MOPED performs competitively with both the current state-of-the-art approach for nonparametric change point detection, E-divisive, and the competing extremal dependence method of Hazra and Bose (2025). We showcased in our simulation study that, where nonparametric methods fail, MOPED is able to identify changes in the extremal dependence class of data, even when the sub-asymptotic dependence (correlation) structure remains unchanged. We also illustrated that our multiscale, multi-threshold MOPED method is able to improve accuracy of change point estimates. In our application to seizure-prone neonatal EEG signals in Section 4, we found that MOPED consistently returned a lower number of change points than the E-divisive approach in cases where practitioners identified no seizure activity in the EEG recordings. Moreover, in cases where seizures were identified by clinicians, we found that the MOPED change point estimates tended to coincide with the onset and end of seizure activity. This may suggest that seizures can be characterised by changes in the extremal dependence of brain signals, and that MOPED may be useful as a data-driven and automatic method for identifying seizures.

We illustrated cases where the MOPED algorithm outperforms the competing parametric extremal dependence change point method of Hazra and Bose (2025). Moreover, due to its computational efficiency, MOPED scales to higher dimensions than previously considered in other applications of extremal dependence change point models. In contrast, E-divisive performs well for high dimensional data. However, it does not enjoy the interpretability of MOPED (and other extremal change point algorithms), in the sense that it is unknown what aspects of the distribution have changed across the change points; MOPED specifically targets changes in the pairwise extremal dependence. The interpretability, efficiency, and good performance of MOPED come at the price of moderate hyperparameter tuning. While we did not optimise the hyperparameters for MOPED, in our application we found only small differences in the change point estimates when applying MOPED with different choices for k and G. A more in-depth analysis may reveal some intricate link between the optimal choice of tuning parameters and the change in tail dependence. One may also consider performing further simulation studies to investigate the similarities and differences of E-divisive and MOPED for a wider range of tail dependence models, in particular in higher dimensions. Our multiscale, multi-threshold MOPED algorithm may alleviate the need for hyperparameter tuning, but further investigation into the optimal choice of ranks $\mathcal{K}$ and bandwidths $\mathcal{H}$, and the manner in which change point estimates are pooled across them, is required.

In this context, it is important to note that our approach is not able to detect all types of changes in extremal dependence, due to the TPDM only considering pairwise tail dependence. Further work may develop change point algorithms designed to detect higher-order changes in extremal dependence by, for example, identifying

changes in the angular measure. An example for when we are likely unable to detect a change can be found in Pawley (2025).

Although we have provided a permutation-based method for testing the existence of structural breaks in the TPDM, one could also consider an approach based on the asymptotic theory of the TPDM estimator; see, e.g., Pawley (2025). The theoretical treatment of the MOPED algorithm, i.e., the asymptotic characterisation of the MOSUM test statistic and convergence rates for the change point estimators, remains an avenue for further work.

Appendix

A Supplementary tables

Table 2 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 1, $d = 8$, of the simulation study. For the fixed MOPED method, $G = 1500$; for MMMOPED, we consider $\mathcal{H} = \{500, 1000, 1500\}$. The modal value of $\widehat{q} - q$ is provided in bold in each row

				$\widehat{q} - q$						
d	ρ	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
8	$\rho = 0.2$	0	MOPED	-	-	**0.881**	0.114	0.005	-	-
			MMMOPED	-	-	**0.500**	0.355	0.145	-	-
			E-divisive	-	-	**0.936**	0.024	0.040	-	-
	$\rho = 0.6$		MOPED	-	-	**0.901**	0.096	0.003	-	-
			MMMOPED	-	-	**0.534**	0.343	0.123	-	-
			E-divisive	-	-	**0.948**	0.013	0.039	-	-
	$\rho = 0.2$	1	MOPED	-	0.166	**0.772**	0.061	0.001	0.847	0.680
			MMMOPED	-	0.016	**0.537**	0.364	0.083	0.859	0.765
			E-divisive	-	0.000	**0.939**	0.026	0.035	0.988	0.972
	$\rho = 0.6$		MOPED	-	0.000	**1.000**	0.000	0.000	0.991	0.965
			MMMOPED	-	0.000	**0.790**	0.188	0.022	0.951	0.913
			E-divisive	-	0.000	**0.947**	0.021	0.032	0.993	0.987
	$\rho = 0.2$	2	MOPED	0.129	0.307	**0.562**	0.002	0.000	0.762	0.697
			MMMOPED	0.009	0.070	**0.617**	0.246	0.058	0.838	0.801
			E-divisive	0.000	0.001	**0.930**	0.037	0.032	0.981	0.966
	$\rho = 0.6$		MOPED	0.000	0.000	**1.000**	0.000	0.000	0.982	0.961
			MMMOPED	0.000	0.000	**0.765**	0.210	0.025	0.951	0.926
			E-divisive	0.000	0.000	**0.937**	0.026	0.037	0.993	0.989

Table 3 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 1, $d = 15$, of the simulation study. For the fixed MOPED method, $G = 1500$; for MMMOPED, we consider $\mathcal{H} = \{500, 1000, 1500\}$. The modal value of $\widehat{q} - q$ is provided in bold in each row

				$\widehat{q} - q$						
d	ρ	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
15	$\rho = 0.2$	0	MOPED	-	-	**0.908**	0.087	0.005	-	-
			MMMOPED	-	-	**0.556**	0.271	0.174	-	-
			E-divisive	-	-	**0.957**	0.009	0.034	-	-
	$\rho = 0.6$		MOPED	-	-	**0.891**	0.104	0.005	-	-
			MMMOPED	-	-	**0.500**	0.322	0.178	-	-
			E-divisive	-	-	**0.958**	0.006	0.036	-	-
	$\rho = 0.2$	1	MOPED	-	0.068	**0.874**	0.057	0.001	0.898	0.778
			MMMOPED	-	0.004	**0.519**	0.393	0.084	0.873	0.794
			E-divisive	-	0.000	**0.945**	0.020	0.035	0.990	0.980
	$\rho = 0.6$		MOPED	-	0.000	**1.000**	0.000	0.000	0.992	0.971
			MMMOPED	-	0.000	**0.768**	0.204	0.028	0.948	0.915
			E-divisive	-	0.000	**0.946**	0.018	0.036	0.992	0.987
	$\rho = 0.2$	2	MOPED	0.042	0.161	**0.796**	0.001	0.000	0.865	0.818
			MMMOPED	0.000	0.007	**0.601**	0.299	0.093	0.875	0.839
			E-divisive	0.000	0.000	**0.950**	0.025	0.025	0.990	0.981
	$\rho = 0.6$		MOPED	0.000	0.000	**1.000**	0.000	0.000	0.986	0.967
			MMMOPED	0.000	0.000	**0.793**	0.175	0.032	0.951	0.929
			E-divisive	0.000	0.000	**0.956**	0.024	0.020	0.995	0.992

Table 4 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 1, $d = 2$, of the simulation study. Here we used a lower sample size of $n = 2000$. For the fixed MOPED method, $G = 500$; for MMMOPED, we consider $\mathcal{H} = \{300, 400, 500\}$. The modal value of $\widehat{q} - q$ is provided in bold in each row

			$\widehat{q} - q$						
ρ	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
$\rho = 0.2$	0	MOPED	-	-	**0.883**	0.103	0.014	-	-
		MMMOPED	-	-	**0.617**	0.278	0.105	-	-
		E-divisive	-	-	**0.948**	0.019	0.033	-	-
$\rho = 0.6$		MOPED	-	-	**0.888**	0.101	0.011	-	-
		MMMOPED	-	-	**0.620**	0.279	0.101	-	-
		E-divisive	-	-	**0.941**	0.026	0.033	-	-
$\rho = 0.2$	1	MOPED	-	**0.737**	0.241	0.021	0.001	0.589	0.181
		MMMOPED	-	**0.451**	0.387	0.135	0.027	0.658	0.353
		E-divisive	-	**0.936**	0.026	0.035	0.003	0.509	0.026
		Hazra and Bose (2025)	-	-	1.000	-	-	0.833	0.677
$\rho = 0.6$		MOPED	-	0.012	**0.923**	0.064	0.001	0.934	0.846
		MMMOPED	-	0.000	**0.554**	0.372	0.074	0.901	0.823
		E-divisive	-	0.000	**0.949**	0.022	0.029	0.979	0.957
		Hazra and Bose (2025)	-	-	1.000	-	-	0.988	0.957
$\rho = 0.2$	2	MOPED	**0.692**	0.218	0.086	0.004	0.000	0.438	0.199
		MMMOPED	**0.391**	0.330	0.210	0.062	0.007	0.539	0.389
		E-divisive	**0.937**	0.027	0.033	0.001	0.002	0.345	0.028
$\rho = 0.6$		MOPED	0.004	0.030	**0.943**	0.023	0.000	0.901	0.853
		MMMOPED	0.000	0.002	**0.498**	0.392	0.108	0.876	0.840
		E-divisive	0.044	0.030	**0.875**	0.037	0.014	0.925	0.887

Table 5 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 1, $d = 8$, of the simulation study. Here we used a lower sample size of $n = 2000$. For the fixed MOPED method, $G = 500$; for MMMOPED, we consider $\mathcal{H} = \{300, 400, 500\}$. The modal value of $\widehat{q} - q$ is provided in bold in each row

				$\widehat{q} - q$						
d	ρ	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
8	$\rho = 0.2$	0	MOPED	-	-	**0.903**	0.020	0.037	-	-
			MMMOPED	-	-	**0.586**	0.303	0.111	-	-
			E-divisive	-	-	**0.943**	0.020	0.037	-	-
	$\rho = 0.6$		MOPED	-	-	**0.920**	0.069	0.011	-	-
			MMMOPED	-	-	**0.565**	0.341	0.094	-	-
			E-divisive	-	-	**0.947**	0.019	0.034	-	-
	$\rho = 0.2$	1	MOPED	-	**0.702**	0.276	0.020	0.002	0.612	0.223
			MMMOPED	-	0.294	**0.469**	0.199	0.038	0.722	0.489
			E-divisive	-	0.078	**0.857**	0.043	0.022	0.927	0.844
	$\rho = 0.6$		MOPED	-	0.000	**0.981**	0.016	0.003	0.973	0.925
			MMMOPED	-	0.000	**0.646**	0.290	0.064	0.915	0.857
			E-divisive	-	0.000	**0.943**	0.025	0.032	0.989	0.981
	$\rho = 0.2$	2	MOPED	**0.602**	0.312	0.081	0.005	0.000	0.467	0.259
			MMMOPED	0.208	**0.363**	0.324	0.090	0.015	0.628	0.535
			E-divisive	**0.708**	0.165	0.110	0.016	0.001	0.431	0.189
	$\rho = 0.6$		MOPED	0.000	0.000	**0.999**	0.001	0.000	0.953	0.916
			MMMOPED	0.000	0.000	**0.626**	0.321	0.053	0.908	0.873
			E-divisive	0.000	0.000	**0.942**	0.032	0.026	0.990	0.983

Table 6 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 1, $d = 15$, of the simulation study. Here we used a lower sample size of $n = 2000$. For the fixed MOPED method, $G = 500$; for MMMOPED, we consider $\mathcal{H} = \{300, 400, 500\}$. The modal value of $\widehat{q} - q$ is provided in bold in each row

				$\widehat{q} - q$						
d	ρ	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
15	$\rho = 0.2$	0	MOPED	-	-	**0.905**	0.088	0.007	-	-
			MMMOPED	-	-	**0.573**	0.312	0.115	-	-
			E-divisive	-	-	**0.950**	0.024	0.026	-	-
	$\rho = 0.6$		MOPED	-	-	**0.892**	0.093	0.015	-	-
			MMMOPED	-	-	**0.588**	0.299	0.113	-	-
			E-divisive	-	-	**0.955**	0.020	0.025	-	-
	$\rho = 0.2$	1	MOPED	-	**0.599**	0.374	0.026	0.001	0.655	0.307
			MMMOPED	-	0.195	**0.503**	0.254	0.049	0.775	0.592
			E-divisive	-	0.000	**0.955**	0.023	0.022	0.985	0.960
	$\rho = 0.6$		MOPED	-	0.000	**0.983**	0.017	0.000	0.979	0.938
			MMMOPED	-	0.000	**0.652**	0.289	0.060	0.914	0.858
			E-divisive	-	0.000	**0.944**	0.033	0.023	0.991	0.984
	$\rho = 0.2$	2	MOPED	**0.489**	0.353	0.147	0.011	0.000	0.517	0.342
			MMMOPED	0.143	0.269	**0.416**	0.144	0.027	0.687	0.613
			E-divisive	0.166	0.228	**0.555**	0.043	0.008	0.765	0.700
	$\rho = 0.6$		MOPED	0.000	0.000	**0.999**	0.001	0.000	0.962	0.930
			MMMOPED	0.000	0.000	**0.655**	0.291	0.054	0.914	0.881
			E-divisive	0.000	0.000	**0.950**	0.033	0.017	0.992	0.985

Table 7 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 2, $d = 2$, of the simulation study. For the fixed MOPED method, $G = 1500$; for MMMOPED, we consider $\mathcal{H} = \{500, 1000, 1500\}$. The modal value of $\widehat{q} - q$ is provided in bold in each row

			$\widehat{q} - q$						
Ω	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
1	1	MOPED	-	**0.582**	0.391	0.026	0.001	0.660	0.316
		MMMOPED	-	0.293	**0.471**	0.197	0.039	0.726	0.489
		E-divisive	-	**0.949**	0.019	0.029	0.003	0.508	0.021
		Hazra and Bose (2025)	-	-	1.000	-	-	0.553	0.195
2		MOPED	-	**0.559**	0.417	0.024	0.000	0.669	0.332
		MMMOPED	-	0.255	**0.505**	0.198	0.042	0.739	0.517
		E-divisive	-	**0.940**	0.021	0.035	0.004	0.509	0.024
		Hazra and Bose (2025)	-	-	1.000	-	-	0.554	0.202
3		MOPED	-	**0.555**	0.424	0.021	0.000	0.670	0.334
		MMMOPED	-	0.272	**0.493**	0.194	0.041	0.734	0.506
		E-divisive	-	**0.940**	0.018	0.038	0.004	0.508	0.024
		Hazra and Bose (2025)	-	-	1.000	-	-	0.554	0.202
1	2	MOPED	**0.536**	0.314	0.150	0.000	0.000	0.511	0.326
		MMMOPED	0.238	**0.351**	0.315	0.085	0.011	0.617	0.517
		E-divisive	**0.938**	0.013	0.049	0.000	0.000	0.345	0.028
2		MOPED	**0.494**	0.344	0.161	0.001	0.000	0.526	0.355
		MMMOPED	0.215	**0.357**	0.330	0.091	0.007	0.626	0.535
		E-divisive	**0.941**	0.013	0.046	0.000	0.000	0.344	0.027
3		MOPED	**0.495**	0.334	0.170	0.001	0.000	0.529	0.357
		MMMOPED	0.219	**0.364**	0.324	0.086	0.007	0.623	0.530
		E-divisive	**0.941**	0.013	0.045	0.000	0.001	0.344	0.027

Table 8 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 2, $d = 8$, of the simulation study. For the fixed MOPED method, $G = 1500$; for MMMOPED, we consider $\mathcal{H} = \{500, 1000, 1500\}$. The modal value of $\widehat{q} - q$ is provided in bold in each row

				$\widehat{q} - q$						
d	Ω	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
8	1	1	MOPED	-	0.056	**0.870**	0.073	0.001	0.909	0.800
			MMMOPED	-	0.008	**0.436**	0.421	0.135	0.847	0.767
			E-divisive	-	0.000	**0.942**	0.025	0.033	0.974	0.935
	2		MOPED	-	0.132	**0.788**	0.077	0.003	0.867	0.721
			MMMOPED	-	0.017	**0.456**	0.379	0.148	0.836	0.746
			E-divisive	-	0.046	**0.899**	0.027	0.028	0.948	0.879
	3		MOPED	-	0.257	**0.691**	0.051	0.001	0.817	0.619
			MMMOPED	-	0.053	**0.529**	0.310	0.108	0.830	0.717
			E-divisive	-	0.023	**0.934**	0.018	0.025	0.959	0.900
	1	2	MOPED	0.012	0.114	**0.872**	0.002	0.000	0.898	0.857
			MMMOPED	0.000	0.015	**0.513**	0.350	0.122	0.846	0.815
			E-divisive	0.096	0.059	**0.785**	0.042	0.018	0.876	0.821
	2		MOPED	0.042	0.221	**0.732**	0.005	0.000	0.845	0.801
			MMMOPED	0.005	0.043	**0.530**	0.322	0.100	0.837	0.805
			E-divisive	**0.431**	0.130	0.413	0.016	0.010	0.632	0.477
	3		MOPED	0.144	0.306	**0.546**	0.004	0.000	0.755	0.687
			MMMOPED	0.023	0.109	**0.600**	0.228	0.039	0.817	0.778
			E-divisive	0.363	0.162	**0.450**	0.015	0.010	0.664	0.531

Table 9 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 2, $d = 15$, of the simulation study. For the fixed MOPED method, $G = 1500$; for MMMOPED, we consider $\mathcal{H} = \{500, 1000, 1500\}$. The modal value of $\widehat{q} - q$ is provided in bold in each row

				$\widehat{q} - q$						
d	Ω	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
15	1	1	MOPED	-	0.000	**0.971**	0.029	0.000	0.966	0.911
			MMMOPED	-	0.000	0.248	**0.406**	0.345	0.813	0.765
			E-divisive	-	0.000	**0.950**	0.024	0.026	0.984	0.958
	2		MOPED	-	0.001	**0.966**	0.033	0.000	0.962	0.903
			MMMOPED	-	0.000	0.363	**0.421**	0.216	0.843	0.785
			E-divisive	-	0.000	**0.954**	0.015	0.031	0.982	0.955
	3		MOPED	-	0.000	**0.959**	0.040	0.001	0.961	0.901
			MMMOPED	-	0.000	0.343	**0.416**	0.241	0.844	0.788
			E-divisive	-	0.000	**0.958**	0.021	0.021	0.984	0.956
	1	2	MOPED	0.000	0.001	**0.999**	0.000	0.000	0.961	0.927
			MMMOPED	0.000	0.000	0.261	**0.406**	0.333	0.851	0.837
			E-divisive	0.000	0.000	**0.936**	0.039	0.025	0.977	0.957
	2		MOPED	0.000	0.001	**0.999**	0.000	0.000	0.959	0.924
			MMMOPED	0.000	0.000	0.310	**0.386**	0.303	0.853	0.838
			E-divisive	0.000	0.000	**0.950**	0.024	0.026	0.973	0.951
	3		MOPED	0.000	0.004	**0.996**	0.000	0.000	0.956	0.919
			MMMOPED	0.000	0.000	0.276	**0.471**	0.253	0.854	0.838
			E-divisive	0.001	0.000	**0.957**	0.025	0.017	0.973	0.951

Table 10 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 3, $d = 2$ and AR parameter $\phi = 0$, of the simulation study. Here we use a sample size of $n = 5000$, with $G = 1500$ for MOPED. The modal value of $\widehat{q} - q$ is provided in bold in each row

			$\widehat{q} - q$						
ρ	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
$\rho = 0.2$	0	MOPED	-	-	**0.913**	0.084	0.003	-	-
		E-divisive	-	-	**0.948**	0.014	0.038	-	-
$\rho = 0.6$		MOPED	-	-	**0.901**	0.093	0.006	-	-
		E-divisive	-	-	**0.945**	0.015	0.040	-	-
$\rho = 0.2$	1	MOPED	-	0.465	**0.508**	0.027	0.000	0.702	0.397
		E-divisive	-	**0.914**	0.040	0.043	0.003	0.519	0.044
		Hazra and Bose (2025)	-	-	1.000	-	-	0.963	0.900
$\rho = 0.6$		MOPED	-	0.000	**0.985**	0.015	0.000	0.975	0.927
		E-divisive	-	0.000	**0.942**	0.019	0.039	0.988	0.973
		Hazra and Bose (2025)	-	-	1.000	-	-	0.997	0.986
$\rho = 0.2$	2	MOPED	**0.411**	0.334	0.254	0.001	0.000	0.578	0.431
		E-divisive	**0.930**	0.021	0.045	0.003	0.001	0.346	0.031
$\rho = 0.6$		MOPED	0.000	0.000	**1.000**	0.000	0.000	0.964	0.930
		E-divisive	0.000	0.000	**0.934**	0.028	0.038	0.987	0.975

Table 11 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 3, $d = 2$ and AR parameter $\phi = 0.3$, of the simulation study. Here we use a sample size of $n = 5000$, with $G = 1500$ for MOPED. The modal value of $\widehat{q} - q$ is provided in bold in each row

			$\widehat{q} - q$						
ρ	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
$\rho = 0.2$	0	MOPED	-	-	**0.854**	0.135	0.011	-	-
		E-divisive	-	-	0.104	0.034	**0.862**	-	-
$\rho = 0.6$		MOPED	-	-	**0.874**	0.116	0.010	-	-
		E-divisive	-	-	0.175	0.044	**0.781**	-	-
$\rho = 0.2$	1	MOPED	-	0.442	**0.517**	0.041	0.000	0.708	0.411
		E-divisive	-	0.072	0.029	0.158	**0.741**	0.593	0.432
		Hazra and Bose (2025)	-	-	1.000	-	-	0.957	0.888
$\rho = 0.6$		MOPED	-	0.000	**0.988**	0.011	0.001	0.972	0.921
		E-divisive	-	0.000	0.161	0.062	**0.777**	0.784	0.747
		Hazra and Bose (2025)	-	-	1.000	-	-	0.997	0.984
$\rho = 0.2$	2	MOPED	**0.398**	0.343	0.258	0.001	0.000	0.584	0.441
		E-divisive	0.085	0.027	0.149	0.104	**0.635**	0.539	0.498
$\rho = 0.6$		MOPED	0.000	0.000	**1.000**	0.000	0.000	0.961	0.926
		E-divisive	0.000	0.000	0.164	0.084	**0.752**	0.837	0.850

Table 12 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 3, $d = 2$ and AR parameter $\phi = 0.6$, of the simulation study. Here we use a sample size of $n = 5000$, with $G = 1500$ for MOPED. The modal value of $\widehat{q} - q$ is provided in bold in each row

			$\widehat{q} - q$						
ρ	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
$\rho = 0.2$	0	MOPED	-	-	**0.664**	0.286	0.050	-	-
		E-divisive	-	-	0.000	0.000	**1.000**	-	-
$\rho = 0.6$		MOPED	-	-	**0.696**	0.252	0.052	-	-
		E-divisive	-	-	0.000	0.000	**1.000**	-	-
$\rho = 0.2$	1	MOPED	-	0.328	**0.576**	0.095	0.001	0.734	0.472
		E-divisive	-	0.000	0.000	0.000	**1.000**	0.149	0.313
		Hazra and Bose (2025)	-	-	1.000	-	-	0.917	0.816
$\rho = 0.6$		MOPED	-	0.000	**0.963**	0.034	0.003	0.955	0.888
		E-divisive	-	0.000	0.000	0.000	**1.000**	0.175	0.331
		Hazra and Bose (2025)	-	-	1.000	-	-	0.995	0.978
$\rho = 0.2$	2	MOPED	0.305	0.345	**0.346**	0.004	0.000	0.627	0.507
		E-divisive	0.000	0.000	0.000	0.000	**1.000**	0.198	0.451
$\rho = 0.6$		MOPED	0.000	0.000	**1.000**	0.000	0.000	0.953	0.914
		E-divisive	0.000	0.000	0.000	0.000	**1.000**	0.234	0.475

Table 13 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 3, $d = 8$ and AR parameter $\phi = 0$, of the simulation study. Here we use a sample size of $n = 5000$, with $G = 1500$ for MOPED. The modal value of $\widehat{q} - q$ is provided in bold in each row

				$\widehat{q} - q$						
d	ρ	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
8	$\rho = 0.2$	0	MOPED	-	-	**0.895**	0.097	0.008	-	-
			E-divisive	-	-	**0.960**	0.004	0.036	-	-
	$\rho = 0.6$		MOPED	-	-	**0.882**	0.107	0.011	-	-
			E-divisive	-	-	**0.954**	0.021	0.025	-	-
	$\rho = 0.2$	1	MOPED	-	0.114	**0.794**	0.091	0.001	0.864	0.715
			E-divisive	-	0.000	**0.955**	0.013	0.032	0.989	0.973
	$\rho = 0.6$		MOPED	-	0.000	**1.000**	0.000	0.000	0.991	0.966
			E-divisive	-	0.000	**0.961**	0.014	0.025	0.994	0.990
	$\rho = 0.2$	2	MOPED	0.083	0.222	**0.695**	0.000	0.000	0.817	0.760
			E-divisive	0.000	0.000	**0.957**	0.022	0.021	0.985	0.971
	$\rho = 0.6$		MOPED	0.000	0.000	**1.000**	0.000	0.000	0.982	0.960
			E-divisive	0.000	0.000	**0.960**	0.013	0.027	0.995	0.993

Table 14 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 3, $d = 8$ and AR parameter $\phi = 0.3$, of the simulation study. Here we use a sample size of $n = 5000$, with $G = 1500$ for MOPED. The modal value of $\widehat{q} - q$ is provided in bold in each row

				$\widehat{q} - q$						
d	ρ	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
8	$\rho = 0.2$	0	MOPED	-	-	0.273	**0.486**	0.241	-	-
			E-divisive	-	-	0.005	0.002	**0.993**	-	-
	$\rho = 0.6$		MOPED	-	-	**0.768**	0.209	0.023	-	-
			E-divisive	-	-	0.083	0.013	**0.904**	-	-
	$\rho = 0.2$	1	MOPED	-	0.079	**0.823**	0.095	0.003	0.873	0.734
			E-divisive	-	0.000	0.003	0.001	**0.996**	0.509	0.524
	$\rho = 0.6$		MOPED	-	0.000	**1.000**	0.000	0.000	0.990	0.964
			E-divisive	-	0.000	0.149	0.050	**0.801**	0.818	0.771
	$\rho = 0.2$	2	MOPED	0.061	0.201	**0.738**	0.000	0.000	0.832	0.778
			E-divisive	0.000	0.000	0.002	0.007	**0.991**	0.616	0.684
	$\rho = 0.6$		MOPED	0.000	0.000	**1.000**	0.000	0.000	0.981	0.958
			E-divisive	0.000	0.000	0.135	0.059	**0.806**	0.837	0.856

Table 15 Distribution of the estimated number of change points and the average CM and VM over 1000 realisations for Scenario 3, $d = 8$ and AR parameter $\phi = 0.6$, of the simulation study. Here we use a sample size of $n = 5000$, with $G = 1500$ for MOPED. The modal value of $\widehat{q} - q$ is provided in bold in each row

				$\widehat{q} - q$						
d	ρ	q	Method	≤ -2	-1	0	1	≥ 2	CM	VM
8	$\rho = 0.2$	0	MOPED	-	-	0.277	**0.488**	0.235	-	-
			E-divisive	-	-	0.000	0.000	**1.000**	-	-
	$\rho = 0.6$		MOPED	-	-	0.209	**0.518**	0.273	-	-
			E-divisive	-	-	0.000	0.000	**1.000**	-	-
8	$\rho = 0.2$	1	MOPED	-	0.021	**0.715**	0.246	0.018	0.860	0.736
			E-divisive	-	0.000	0.000	0.000	**1.000**	0.039	0.257
	$\rho = 0.6$		MOPED	-	0.000	**1.000**	0.000	0.000	0.987	0.954
			E-divisive	-	0.000	0.000	0.000	**1.000**	0.143	0.316
	$\rho = 0.2$	2	MOPED	0.024	0.170	**0.795**	0.011	0.000	0.848	0.799
			E-divisive	0.000	0.000	0.000	0.000	**1.000**	0.057	0.379
	$\rho = 0.6$		MOPED	0.000	0.000	**1.000**	0.000	0.000	0.974	0.947
			E-divisive	0.000	0.000	0.000	0.000	**1.000**	0.198	0.461

B Supplementary figures

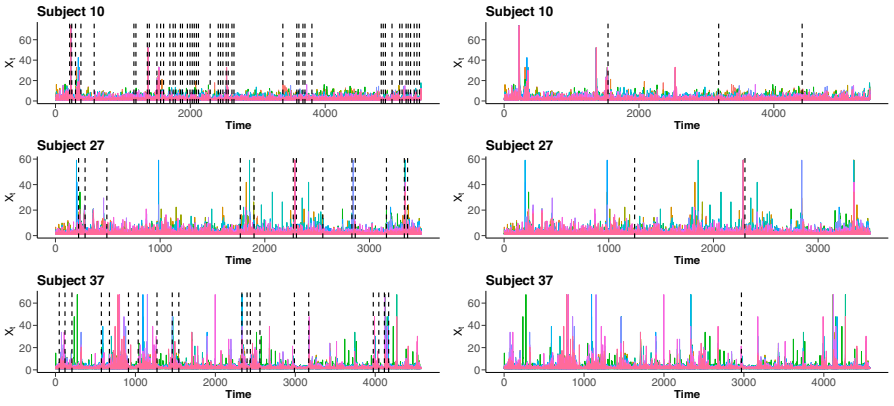

Fig. 7 Change point estimates for Subjects 10, 27, and 37 (top to bottom row) of the case study. Estimates are denoted by the vertical dotted lines; the left and right columns provide estimates from E-divisive and MOPED ($G = 1000$ and $k = 100$), respectively. Observations of $\{X_{i,t}\}_{t=1}^{n}$ are plotted against time t, with the colour corresponding to the i-th channel, $i = 1, \ldots, d$. Here, $d = 19$ and no subjects were identified as having experienced a seizure

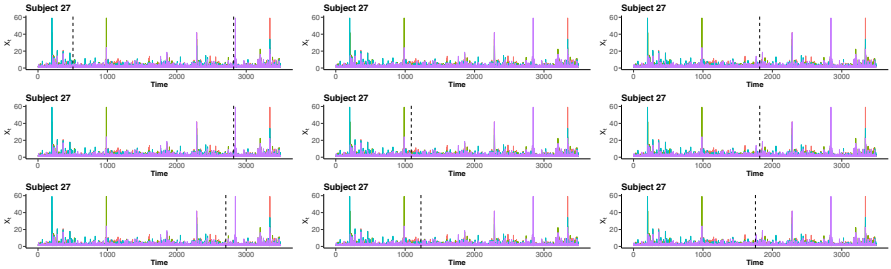

Fig. 8 Change point estimates for Subject 27, using MOPED, are denoted by the vertical dotted lines. The hyper-parameters differ across panels, with $G = 500$, $G = 1000$, and $G = 1500$ for the left, middle, and right columns, respectively, and k set to $0.1G$, $0.05G$, and $0.025G$ for the top, central, and bottom rows, respectively. Observations of $\{X_{i,t}\}_{t=1}^{n}$ are plotted against time t, with the colour corresponding to the i-th channel, $i = 1, \ldots, 4$

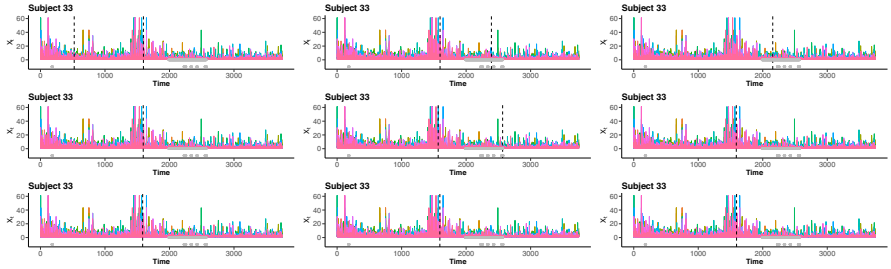

Fig. 9 Change point estimates for Subject 33, using MOPED, are denoted by the vertical dotted lines. The hyper-parameters differ across panels, with $G = 500$, $G = 1000$, and $G = 1500$ for the left, middle, and right columns, respectively, and k set to $0.1G$, $0.05G$, and $0.025G$ for the top, central, and bottom rows, respectively. Observations of $\{X_{i,t}\}_{t=1}^{n}$ are plotted against time t, with the colour corresponding to the i-th channel, $i = 1, \ldots, 19$

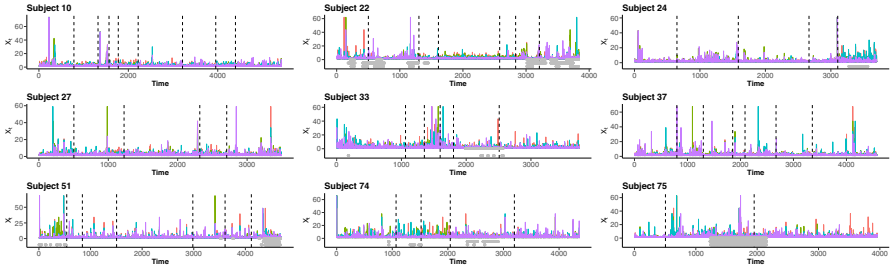

Fig. 10 Change point estimates, using MMMOPED, are denoted by the vertical dotted lines. Observations of $\{X_{i,t}\}_{t=1}^{n}$ are plotted against time t, with the colour corresponding to the i-th channel, $i = 1, \ldots, d$, for $d = 4$

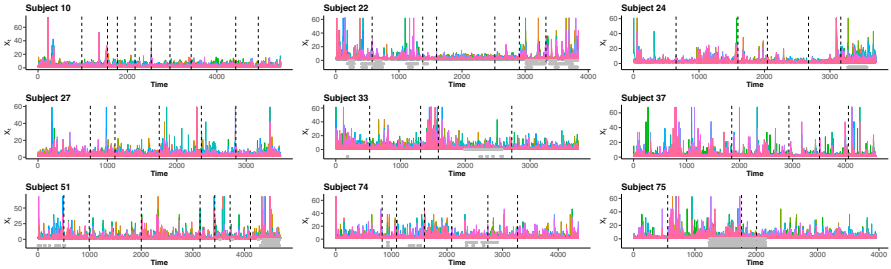

Fig. 11 Change point estimates, using MMMOPED, are denoted by the vertical dotted lines. Observations of $\{X_{i,t}\}_{t=1}^{n}$ are plotted against time t, with the colour corresponding to the i-th channel, $i = 1, \ldots, d$, for $d = 19$

Acknowledgements The authors thank Mara Sherlin Talento for helpful discussions related to the dataset. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (www.ecdf.ed.ac.uk/).

Author Contributions E.T.M. and J.R. wrote the main manuscript text. All authors reviewed the manuscript and contributed to the methodology.

Funding The authors have no relevant funding details to disclose.

Data Availability The open access EEG data and expert annotations of Stevenson et al. (2019) are available at https://doi.org/10.5281/zenodo.4940267 (last accessed 19/08/2025). Accompanying software implementing the method is available as the R package moped at https://github.com/EuanMcGonigle/moped.

Declarations

Competing Interests The authors declare no competing interests.

Ethical Approval Not applicable.

References

Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **33**(5), 898–916 (2010)

Arlot, S., Celisse, A., Harchaoui, Z.: A kernel multiple change-point algorithm via model selection. J. Mach. Learn. Res. **20**(162), 1–56 (2019)

Bücher, A., Jäschke, S., Wied, D.: Nonparametric tests for constant tail dependence with an application to energy and finance. J. Econ. **187**(1), 154–168 (2015)

Bücher, A., Kojadinovic, I.: Dependent multiplier bootstraps for non-degenerate u-statistics under mixing conditions with applications. J. Stat. Plan. Inference **170**, 83–105 (2016)
Castillo-Mateo, J.: Distribution-free changepoint detection tests based on the breaking of records. Environ. Ecol. Stat. **29**(3), 655–676 (2022)
Celisse, A., Marot, G., Pierre-Jean, M., Rigaill, G.: New efficient algorithms for multiple change-point detection with reproducing kernels. Comput. Stat. Data Anal. **128**, 200–220 (2018)
Chavez-Demoulin, V., Davison, A.C.: Generalized additive modelling of sample extremes. J. R. Stat. Soc.: Ser. C: Appl. Stat. **54**(1), 207–222 (2005)
Chen, H., Zhang, N.: Graph-based change-point detection. Ann. Stat. **43**(1), 139–176 (2015)
Chu, C.-S.J., Hornik, K., Kaun, C.-M.: Mosum tests for parameter constancy. Biometrika **82**(3), 603–617 (1995)
Coles, S.G., Heffernan, J.E., Tawn, J.A.: Dependence measures for extreme value analyses. Extremes **2**, 339–365 (1999)
Cooley, D., Thibaud, E.: Decompositions of dependence for high-dimensional extremes. Biometrika **106**(3), 587–604 (2019)
De Monte, L., Huser, R., Papastathopoulos, I., Richards, J.: Generative modelling of multivariate geometric extremes using normalising flows. arXiv:2505.02957 (2025)
Dierckx, G., Teugels, J.L.: Change point analysis of extreme values. Environmetrics **21**(7–8), 661–686 (2010)
Douglass, L.M., Wu, J.Y., Rosman, N.P., Stafstrom, C.E.: Burst suppression electroencephalogram pattern in the newborn: predicting the outcome. J. Child Neurol. **17**(6), 403–408 (2002)
Drees, H.: Statistical inference on a changing extreme value dependence structure. Ann. Stat. **51**(4), 1824–1849 (2023)
Dupuis, D., Sun, Y., Wang, H.J.: Detecting change-points in extremes. Stat. Its Interface **8**(1), 19–31 (2015)
Dupuis, D.J., Trapin, L.: Ground-level ozone: Evidence of increasing serial dependence in the extremes. Ann. Appl. Stat. **13**(1), 34–59 (2019)
e Silva, W. V. M., do Nascimento, F. F., and Bourguignon, M.: A change-point model for the r-largest order statistics with applications to environmental and financial data. Appl. Math. Model. **82**, 666–679 (2020)
Eichinger, B., Kirch, C.: A MOSUM procedure for the estimation of multiple random change points. Bernoulli **24**, 526–564 (2018)
Elsom, H., Pawley, M.: Extreme value statistics for analysing simulated environmental extremes. Extremes, pp. 1–27 (2024)
Engelke, S., Ivanovs, J.: Sparse structures for multivariate extremes. Ann. Rev. Stat. Its Appl. **8**(1), 241–270 (2021)
Fryzlewicz, P.: Wild binary segmentation for multiple change-point detection. Ann. Stat. **42**, 2243–2281 (2014)
Girard, S., Opitz, T., Usseglio-Carleve, A., Yan, C.: Changepoint identification in heavy-tailed distributions. Working paper or preprint (2025)
Gong, Y., Zhong, P., Opitz, T., Huser, R.: Partial tail-correlation coefficient applied to extremal-network learning. Technometrics **66**(3), 331–346 (2024)
Guerrero, M.B., Huser, R., Ombao, H.: Conex-Connect: Learning patterns in extremal brain connectivity from multichannel EEG data. Ann. Appl. Stat. **17**(1), 178–198 (2023)
Hazra, A., Bose, S.: Estimating changepoints in extremal dependence, applied to aviation stock prices during COVID-19 pandemic. J. Appl. Stat. **52**(3), 525–554 (2025)
Helwig, N.: eegkit: Toolkit for electroencephalography data. R package (2018). Available at http://CRAN.R-project.org/package=eegkit. Accessed 28 July 2025
Hoga, Y.: Testing for changes in (extreme) VaR. Economet. J. **20**(1), 23–51 (2017)
Hoga, Y.: A structural break test for extremal dependence in β-mixing random vectors. Biometrika **105**(3), 627–643 (2018)
Hoga, Y.: Modeling time-varying tail dependence, with application to systemic risk forecasting. J. Financ. Economet. **20**(5), 1007–1037 (2022)
Hüsler, J., Reiss, R.-D.: Maxima of normal random vectors: between independence and complete dependence. Stat. Prob. Lett. **7**(4), 283–286 (1989)
Jiang, J., Richards, J., Huser, R., Bolin, D.: The efficient tail hypothesis: An extreme value perspective on market efficiency. J. Bus. Econ. Stat. To appear (2025a)
Jiang, J., Richards, J., Huser, R., Bolin, D.: Separation-based causal discovery for extremes. arXiv:2505.08008 (2025b)

Jiang, Y., Cooley, D., Wehner, M.F.: Principal component analysis for extremes and application to US precipitation. J. Clim. **33**(15), 6441–6451 (2020)
Kakampakou, L., Simpson, E.S., Wadsworth, J.L.: Spatial extremal modelling: A case study on the interplay between margins and dependence. Stat **13**(4), e70021 (2024)
Kojadinovic, I., Naveau, P.: Detecting distributional changes in samples of independent block maxima using probability weighted moments. Extremes **20**(2), 417–450 (2017)
Larsson, M., Resnick, S.I.: Extremal dependence measure and extremogram: the regularly varying case. Extremes **15**(2), 231–256 (2012)
Lee, J., Cooley, D.: Transformed-linear prediction for extremes. arXiv:2111.03754 (2021)
Lee, J., Cooley, D.: Partial tail correlation for extremes. arXiv:2210.02048 (2022)
Matteson, D.S., James, N.A.: A nonparametric approach for multiple change point analysis of multivariate data. J. Am. Stat. Assoc. **109**(505), 334–345 (2014)
McGonigle, E.T., Cho, H.: Nonparametric data segmentation in multivariate time series via joint characteristic functions. Biometrika **112**(2), asaf024 (2025)
McGonigle, E.T., Cho, H.: Robust multiscale estimation of time-average variance for time series segmentation. Comput. Stat. Data Anal. **179**, 107648 (2023)
Meier, A., Kirch, C., Cho, H.: mosum: A package for moving sums in change-point analysis. J. Stat. Softw. **97**, 1–42 (2021)
Messer, M., Kirchner, M., Schiemann, J., Roeper, J., Neininger, R., Schneider, G.: A multiple filter test for the detection of rate changes in renewal processes with varying variance. Ann. Appl. Stat. **8**(4), 2027–2067 (2014)
Mhatre, N., Cooley, D.: Transformed-linear models for time series extremes. J. Time Ser. Anal. **45**(5), 671–690 (2024)
Murphy-Barltrop, C. J., Majumder, R., Richards, J.: Deep learning of multivariate extremes via a geometric representation. arXiv:2406.19936 (2024)
Murphy-Barltrop, C., Wadsworth, J.: Modelling non-stationarity in asymptotically independent extremes. Comput. Stat. Data Anal. **199**, 108025 (2024)
Murphy, C., Tawn, J.A., Varty, Z.: Automated threshold selection and associated inference uncertainty for univariate extremes. Technometrics **67**(2), 215–224 (2025)
Muthuswamy, J., Sherman, D.L., Thakor, N.V.: Higher-order spectral analysis of burst patterns in EEG. IEEE Trans. Biomed. Eng. **46**(1), 92–99 (2002)
Ng, W.L., Pan, S., Yau, C.Y.: Bootstrap inference for multiple change-points in time series. Economet. Theor. **38**(4), 752–792 (2022)
NIH.: Epilepsy and seizures (2023). https://www.ninds.nih.gov/health-information/disorders/epilepsy-and-seizures. Accessed: 02 Jun 2025
Padilla, O.H.M., Yu, Y., Wang, D., Rinaldo, A.: Optimal nonparametric multivariate change point detection and localization. IEEE Trans. Inf. Theory **68**(3), 1922–1944 (2022)
Pawley, M.: Extensions and applications of the tail pairwise dependence matrix. PhD thesis, University of Bath (2025)
Poon, S.-H., Rockinger, M., Tawn, J.: Modelling extreme-value dependence in international stock markets. Stat. Sin. **13**(4), 929–953 (2003)
Redondo, P.V., Guerrero, M.B., Ombao, H., Huser, R.: Statistics of extremes for neuroscience. In: de Carvalho, M., Huser, R., Naveau, P., Reich, B. (eds.) Handbook of Statistics of Extremes. Chapman & Hall/CRC Press, Boca Raton, FL (2026)
Resnick, S. I.: Heavy-tail phenomena: Probabilistic and statistical modeling. Springer Science & Business Media (2007)
Resnick, S.: The extremal dependence measure and asymptotic independence. Stoch. Model. **20**(2), 205–227 (2004)
Richards, J., Alotaibi, N., Cisneros, D., Gong, Y., Guerrero, M.B., Redondo, P.V., Shao, X.: Modern extreme value statistics for Utopian extremes. Extremes **28**(1), 149–171 (2025)
Richards, J., Huser, R.: Extreme Quantile Regression with Deep Learning. In: de Carvalho, M., Huser, R., Naveau, P., Reich, B.J. (eds.) Handbook of Statistics of Extremes. Chapman & Hall/CRC Press, Boca Raton, FL (2026)
Rohrbeck, C., Cooley, D.: Simulating flood event sets using extremal principal components. Ann. Appl. Stat. **17**(2), 1333–1352 (2023)
Rohrbeck, C., Tawn, J.A.: Bayesian spatial clustering of extremal behavior for hydrological variables. J. Comput. Graph. Stat. **30**(1), 91–105 (2021)

Rosenberg, A., Hirschberg, J.: V-measure: A conditional entropy-based external cluster evaluation measure. In Empirical Methods in Natural Language Processing and Computational Natural Language Learning, pp. 410–420 (2007)

Scarrott, C., MacDonald, A.: A review of extreme value threshold estimation and uncertainty quantification. REVSTAT-Stat. J. **10**(1), 33–60 (2012)

Shao, X.: The dependent wild bootstrap. J. Am. Stat. Assoc. **105**(489), 218–235 (2010)

Simpson, E.S., Wadsworth, J.L., Tawn, J.A.: Determining the dependence structure of multivariate extremes. Biometrika **107**(3), 513–532 (2020)

Stevenson, N.J., Tapani, K., Lauronen, L., Vanhatalo, S.: A dataset of neonatal EEG recordings with seizure annotations. Sci. Data **6**(1), 1–8 (2019)

Talento, M. S. D., Richards, J., Huser, R., Ombao, H.: Canonical tail dependence for soft extremal clustering of multichannel brain signals. arXiv:2512.06435 (2025a)

Talento, M. S. D., Richards, J., Pinto-Orellana, M., Huser, R., Ombao, H. C.: Spectral extremal connectivity of two-state seizure brain waves. arXiv:2503.04169 (2025b)

Van den Burg, G. J., Williams, C. K.: An evaluation of change point detection algorithms. arXiv:2003.06222 (2020)

Vanegas, L.J., Behr, M., Munk, A.: Multiscale quantile segmentation. J. Am. Stat. Assoc. **117**(539), 1384–1397 (2022)

Youngman, B.D.: Generalized additive models for exceedances of high thresholds with an application to return level estimation for US wind gusts. J. Am. Stat. Assoc. **114**(528), 1865–1879 (2019)

Zhong, P., Huser, R., Opitz, T.: Modeling nonstationary temperature maxima based on extremal dependence changing with event magnitude. Ann. Appl. Stat. **16**(1), 272–299 (2022)