Extremes (2025) 28:557–578
https://doi.org/10.1007/s10687-025-00513-8

Statistical inference on the local dependence condition of extreme values in a stationary sequence

Jan Holešovský[1] · Michal Fusek[2]

Received: 8 July 2024 / Revised: 24 April 2025 / Accepted: 15 May 2025 /
Published online: 23 May 2025

Abstract
The extremal index is an important characteristic measuring dependence of extreme values in a stationary series. Several new estimators that are mostly based on interexceedance times within the Peaks-over-Threshold model have been recently published. Nevertheless, in many cases these estimators rely on suitable choice of auxiliary parameters and/or are derived under assumptions that are related to validity of the local dependence condition $D^{(k)}(u_n)$. Although the determination of the correct order k in the $D^{(k)}(u_n)$ condition can have major effect on the extremal index estimates, there are not many reliable methods available for this task. In this paper, we present various approaches to assessing validity of the $D^{(k)}(u_n)$ condition including a graphical diagnostics and propose several statistical tests. A simulation study is carried out to determine performance of the statistical tests, particularly the type I and type II errors.

Keywords Local dependence · Extremal index · Extreme value theory · Clusters

AMS 2000 Subject Classifications 62G32 · 62N01 · 60G70 · 62M09 · 62N02

1 Introduction

Statistical inference for extreme events is usually based on extreme value theory. A typical issue is to determine the frequency of occurrence of extremes. The analysis is being conducted with one of two possible approaches, the block maxima model based on the generalized extreme value distribution, or the Peaks-over-Threshold model

✉ Jan Holešovský
jan.holesovsky@vut.cz

Michal Fusek
fusekmi@vut.cz

1 Institute of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Brno University of Technology, Brno, Czech Republic

2 Department of Mathematics, Faculty of Electrical Engineering and Communication, Brno University of Technology, Brno, Czech Republic

(POT) based on the generalized Pareto (GP) distribution. The latter method is usually preferable as it allows to conduct the analysis on samples of larger size than in the first case, which is a typical problem of the extreme value inference.

It is known from the theory that when analysing extreme events of a stationary series, the extremes tend to cluster. When the block maxima approach is used, the dependence in the series can be under certain circumstances overcome. However, in the POT model, presence of the dependence can cause serious troubles. If it is neglected, then the marginal quantiles like the return level tend to be underestimated. Solving this issue requires either (i) application of a declustering scheme leading to further reduction of the sample size, or (ii) estimation of the dependence structure in the series. Under some additional restrictions on the dependence of distant observations, the cluster size can be described with a single parameter, the so-called extremal index. The latter point then corresponds to finding a suitable estimator of this parameter.

Many extremal index estimators have been discussed in the literature. For example the runs and blocks estimators (Smith and Weissman 1994), the maxima estimators (Ancona-Navarrete and Tawn 2000; Gomes 1993; Northrop 2015), and the sliding block estimator of Robert et al. (2009). Recently, attention has especially been paid to estimators based on interexceedance times within the framework of the POT model. Most of these estimators require, however, the selection of auxiliary parameters that are related to the local dependence restriction, particularly given by the $D^{(k)}(u_n)$ condition of Chernick et al. (1991) (cf. Section 2). Any unsuitable choice of auxiliary parameters stemming from a false assumption on the dependence condition can have a serious effect on the estimators quality. At the same time, there has recently been a lack of reliable methods to determine the validity of the condition. In this paper, we turn back to the studies of Holešovský and Fusek (2022), Holešovský and Fusek (2020) and Süveges and Davison (2010), and discuss properties of their proposed extremal index estimators under model misspecification. On this basis, we propose graphical diagnostics and several statistical tests for the $D^{(k)}(u_n)$ condition.

Ferro and Segers (2003) derived the limit distribution of interexceedance times and proposed the intervals estimator for extremal index, a moment estimator with improvement on the finite-sample bias. They also showed, that direct application of the likelihood method to the limit distribution is not suitable as it overestimates the extremal index towards independence. Other advanced approaches followed, suggesting different ways to solve this issue. Süveges and Davison (2010) proposed the K-gaps estimator, Holešovský and Fusek (2020) introduced the censored estimator, and the truncated estimator was derived in Holešovský and Fusek (2022). However, these advanced estimators require the selection of auxiliary parameters K or D. They are of similar nature and are related to the clustering tendency of the process given by the $D^{(k)}(u_n)$ condition in Section 2. For the estimators to work properly, the choice of K and D is of utmost importance.

When analyzing real data, it can be difficult to prove validity of the $D^{(k)}(u_n)$ condition. Several approaches have been proposed, for example, the graphical diagnostics of anti-$D^{(k)}(u_n)$ events (Süveges 2007; Ferreira and Ferreira 2018). Nevertheless, the graphical approach generally leads to subjective conclusions. Other two methods are the information matrix test and its extensions (Süveges and Davison 2010; Fukutome et al. 2015, 2019; Ferreira 2018), and the stability analysis of the runs estimator (Cai 2023). Both of them are, however, based on estimators that are rather sensitive to the selection of the auxiliary parameters, including the threshold itself.

In this paper, we derive the information matrix test for the censored estimator and compare it with the test introduced in Süveges and Davison (2010). In addition, we apply a graphical approach and propose statistical tests, that make use of the censored and/or the truncated estimators. Both of these estimators have shown respectable properties and are in some sense more flexible that the K-gaps or the runs estimators.

The paper is organized as follows. In Section 2 we introduce the basic concepts. Section 3 is devoted to the information matrix test. We modify it in the context of the censored estimator for the purpose of testing the local dependence condition of a general order. In Section 4 we derive the distribution of interexceedance times under model misspecification and on this basis we propose graphical diagnostics on the bias of extremal index estimators. Several goodness-of-fit tests for the misspecified model are applied in Section 5 to assess validity of the local dependence condition. We compared the tests in simulation study in terms of type I and II errors, the results of which we summarize in final Section 6.

2 Preliminaries

Let $X_1, X_2, \dots$ be a strictly stationary sequence with marginal cumulative distribution function (cdf) F, tail function $\overline{F} = 1-F$, and right end-point $x^* = \sup\{x : F(x) < 1\}$. Let $M_{i,j} = \max\{X_{i+1}, \dots, X_j\}$ where for $i \geq j$ we set $M_{i,j} = -\infty$, and put $M_n = M_{0,n}$. We say $X_1, X_2, \dots$ has extremal index $\theta \in (0, 1]$ if, for all $\tau > 0$, there is a sequence $\{u_n\}_{n\geq 1}$ such that, as $n \to \infty$,

$$n\overline{F}(u_n) \to \tau$$

and

$$P(M_n \leq u_n) \to \exp(-\theta\tau),$$

see Leadbetter et al. (1983). When $\theta = 1$, exceedances of a high threshold u_n occur in an isolated way, similarly as if $X_1, X_2, \dots$ were independent. In the case $\theta < 1$, the extremes tend to cluster in the limit. Moreover, the extremal index can be interpreted as the reciprocal of the mean cluster size and it is a key parameter for assessing risks resulting from extreme events in many application areas.

The assumptions in the previous paragraph are too general, and it is necessary to restrict the dependence in the sequence. Leadbetter et al. (1983) assumed validity of the $D(u_n)$ condition which limits the long-range dependence at extreme levels. It is a standard mixing condition that implies asymptotic independence of sufficiently distant extreme observations (see Leadbetter et al. 1983, for more details). Nevertheless, it is often required to restrict the dependence more locally. The following $D^{(k)}(u_n)$ condition of Chernick et al. (1991) is usually considered.

Condition 1 *$D^{(k)}(u_n)$ condition is said to be satisfied for some positive integer k, if a stationary series $X_1, \ldots, X_n$ under the $D(u_n)$ condition of Leadbetter et al. (1983) also satisfies*

$$nP(X_1 > u_n, M_{1,k} \leq u_n < M_{k,r_n}) \to 0 \tag{1}$$

as $n \to \infty$ with $r_n = o(n)$ and u_n as in $D(u_n)$.

Clearly, if $D^{(k_0)}(u_n)$ holds, then $D^{(k)}(u_n)$ also holds for all $k \geq k_0$. It means that within a cluster, high-threshold exceedances are asymptotically almost surely separated by at most $k-1$ observations. The $D^{(k)}(u_n)$ condition plays a crucial role in estimation of the extremal index θ. Most of available threshold-based estimators of θ are derived under the assumption of validity of a particular $D^{(k)}(u_n)$ condition or, alternatively, the order of the condition or some related quantity appear therein as auxiliary parameters. From here on, let $k_0 \geq 1$ denote an integer such that $X_1, \ldots, X_n$ satisfies the $D^{(k_0)}(u_n)$ condition, but the condition $D^{(k_0-1)}(u_n)$ is not fulfilled.

Recently, attention has been paid to estimators that are based on times between two successive threshold exceedances. Denote

$$T(u_n) = \min\{j \geq 1 : X_{j+1} > u_n \mid X_1 > u_n\} \tag{2}$$

a random variable corresponding to interexceedance times in the underlying series. For threshold u_n increasing with n, it holds for $x > 0$ that

$$P(\overline{F}(u_n)T(u_n) \leq x) \to 1 - \theta\exp(-\theta x) =: F_\theta(x) \tag{3}$$

as $n \to \infty$ (see Theorem 1 in Ferro and Segers 2003). It means that as $u_n \to x^*$ for $n \to \infty$,

$$\overline{F}(u_n)T(u_n) \xrightarrow{d} T_\theta, \tag{4}$$

where $\xrightarrow{d}$ denotes the convergence in distribution and T_θ follows the mixture of the degenerate distribution at 0 and the exponential distribution with expected value $1/\theta$. The extremal index θ then determines both the proportion of intra- and inter-cluster times, as well as the expected value of the normalized inter-cluster times.

3 Information matrix test for the censored estimator

One of available statistical tests for the $D^{(k)}(u_n)$ condition is the information matrix test (IMT) proposed in Süveges and Davison (2010). The authors applied the test for identification of a proper K in context of the K-gaps estimator of θ proposed in the same paper. Later study has revealed (Holešovský and Fusek 2020), that such test is appropriate for testing k_0, but it is not suited for testing the $D^{(k)}(u_n)$ condition in general. In this section we modify the IMT statistics in context of the censored estimator that overcomes this limitation.

The K-gaps estimator $\widehat{\theta}_{\mathrm{SD}}$ of Süveges and Davison (2010) is based on the sample of gaps $S^{(K)}(u_n) = \max\{T(u_n) - K; 0\}$ for some $K \geq 0$ (see Appendix B for a brief overview). Their likelihood model for the gaps requires, however, a proper choice K_0 of the parameter K. The IMT was originally developed by White (1982). It utilizes the fact that for a well-specified model with a log-likelihood $\ell(\theta)$ the Fisher's information matrix $I(\theta) = -\mathrm{E}\left(\frac{\partial^2 \ell}{\partial \theta^2}\right)$ equals the variance of the score vector $J(\theta) = \mathrm{Var}\left(\frac{\partial \ell}{\partial \theta}\right)$. In case the true model $\ell_0(\theta)$ is not contained in the assumed model $\ell(\theta)$, i.e. there is no θ_0 such that $\ell_0 = \ell(\theta_0)$, the estimator of θ is consistent for the parameter value $\tilde{\theta}$ minimizing the Kullback-Leibler divergence with the true distribution. In such case its asymptotic variance is in the form $I(\tilde{\theta})^{-1} J(\tilde{\theta}) I(\tilde{\theta})^{-1}$. Let $\Delta(\theta) = J(\theta) - I(\theta)$, then the null hypothesis $H_0 : \Delta(\theta) = 0$ is set against the alternative $H_1 : \Delta(\theta) \neq 0$.

Let us consider a stationary series $X_1, \ldots, X_n$ with N observations exceeding a sufficiently high threshold u, i.e. there are $N-1$ interexceedance times $T_1, \ldots, T_{N-1}$. The IMT statistics is of the form

$$\mathrm{IMT}(\theta) = (N-1)\Delta^2(\theta)\, V^{-1}(\theta)\,, \tag{5}$$

where $V(\theta)$ denotes the variance of $\Delta(\theta)$, and θ should be replaced by its estimate $\widehat{\theta}$. Assuming independence of the interexceedance times, and under regularity conditions from Theorem 4.1 of White (1982), the test statistic (5) has approximately χ^2 distribution with one degree of freedom. Süveges and Davison (2010) derived the form of $\Delta(\theta)$ and $V(\theta)$ in the context of their K-gaps estimator $\widehat{\theta}_{\mathrm{SD}}$. In addition, Fukutome et al. (2015, 2019) proposed an automated procedure for simultaneous selection of the threshold and the parameter K based on the IMT and also provide corrections of the original formulas.

From the discussion in Holešovský and Fusek (2020) it turns out that the proper choice K_0 is related to validity of the $D^{(k_0)}(u_n)$ condition with $K_0 = k_0 - 1$. Hence, the IMT can also be applied to assess validity of the $D^{(k_0)}(u_n)$ condition. On the other hand, the model of Süveges and Davison (2010) is misspecified for $K > K_0$ and as such it is inappropriate for testing any condition of order $k > k_0$. We modify the test in the context of the censored estimator $\widehat{\theta}_{\mathrm{C}}$ of Holešovský and Fusek (2020) (see Appendix B for a brief overview). This estimator is derived on the basis of likelihood model that is well specified for all $D > D_0 = k_0 - 1$. Moreover, the censored estimator often exhibits more desirable properties than the K-gaps estimator. In case of the censored estimator, the statistic (5) remains in the same form, where $\Delta(\theta)$, $V(\theta)$ and details of the derivation can be found in Appendix A.

Süveges and Davison (2010) also discussed the regularity conditions of White (1982) in the context of their likelihood model based on Eq. 3. It follows directly that the same theory applies to the log-likelihood (B2) from Appendix B related to the estimator $\widehat{\theta}_{\mathrm{C}}$. The test is based on assumptions of mutual independence of interexceedance times. As it was noted, for example, in Ferro and Segers (2003), the first one is typically violated for the intra-cluster times, although different sets of clusters are asymptotically independent. It was already discussed in Holešovský and Fusek (2020) that the requirement of independence may be disregarded under validity of suitable $D^{(k)}(u_n)$ condition. Moreover, the performance of the IMT can be harmed by poor limit approximation of the distribution of times by F_θ from Eq. 3, usually at low threshold levels. This generally applies also to alternative approaches we discuss in Sections 4 and 5.

4 Graphical diagnostics on the bias of the estimators

Another approach to assess validity of the $D^{(k)}(u_n)$ condition is based on graphical diagnostics. Some suggestions were made in Süveges (2007) and Ferreira and Ferreira (2018) using plots of anti-$D^{(k)}$ events, but the optimal settings of parameters in the plot remains unclear, making it difficult to deal with. It is more convenient to build such a diagnostics that is linked to the behaviour of the extremal index estimators. In this section, we derive the distribution of interexceedance times truncated by $D < D_0$, that can be understood as a misspecification of a model on which the truncated estimator is based. Using this result, we propose graphical inspection of estimates of θ to identify k_0 at which the model is well specified.

First proposal of the anti-$D^{(2)}$ plot to verify the $D^{(2)}(u_n)$ condition appeared in Süveges (2007) and was later generalized for an arbitrary $D^{(k)}(u_n)$ in Ferreira and Ferreira (2018). Given a threshold u and a block size r, the proportion $p_k(u, r)$ of anti-$D^{(k)}(u_n)$ events is given by

$$p_k(u, r) = \frac{\sum_{j=1}^{n-r+1} \mathbf{1}_{[X_j > u, X_{j+1} \leq u, \ldots, X_{j+k-1} \leq u, M_{j+k-1, r+j-1} > u]}}{\sum_{j=1}^{n} \mathbf{1}_{[X_j > u]}}. \tag{6}$$

Süveges (2007) suggests to evaluate the proportion for a range of thresholds and block sizes, whereas validity of the condition should result in a path (u_i, r_j) with $u_i \to x^*$ and $r_j \to \infty$ such that $p_k(u_i, r_j) \to 0$.

This approach was extended in Ferreira and Ferreira (2018) by considering both u and r depending on a sequence length n, specifically $u = u(n)$ corresponding to the sample quantile $1 - \tau/n$ and $r = r(n) = [n/(\log n)^s]$, $s > 0$. Moreover, for a given sequence of thresholds and block sizes, they suggest to plot rather the difference

$$d_k(u(n), r(n)) = \sum_{j=1}^{n-r(n)+1} \mathbf{1}_{[X_j > u(n), M_{j, j+k-1} \leq u(n)]}.$$

A suitable choice of k_0 is accompanied by distant trajectories of d_{k_0-1} and d_{k_0} and close trajectories of d_{k_0} and d_{k_0+1}. The idea is based on the fact that a series $X_1, \ldots, X_n$ under the $D^{(k_0)}(u_n)$ condition satisfies also the $D^{(k_0+1)}(u_n)$ condition but not $D^{(k_0-1)}(u_n)$ or any lower order condition. Nevertheless, the graphical diagnostics can lead to subjective conclusions.

Holešovský and Fusek (2022) proposed the truncated estimator $\widehat{\theta}_{\mathrm{T}} = \widehat{\theta}_{\mathrm{T}}(D)$ that is based on the distribution of truncated interexceedance times. For brief details on derivation of $\widehat{\theta}_{\mathrm{T}}$ see Appendix B. Such an estimator overcomes the uncertainty of the degenerate part in left tail of F_θ and is concentrated purely on the right tail. They derived the relation

$$P\left(\overline{F}(u_n)(T(u_n) - D) > x \mid T(u_n) > D\right) \to \exp(-\theta x), \tag{7}$$

for $x > 0$, $u_n \to x^*$ for $n \to \infty$, and some $D \geq 0$. However, the convergence in Eq. 7 requires validity of the $D^{(D+1)}(u_n)$ condition or, alternatively, $D \geq D_0 = k_0 - 1$.

Specifically, in case $D = 0$, the relation (7) is valid under the $D^{(1)}(u_n)$ condition. Since this is implied by the $D'(u_n)$ condition of Leadbetter et al. (1983), it is necessarily related to the case $\theta = 1$.

Suppose the case $D_0 \geq 1$ such that there exists a $D < D_0$. We may discuss the convergence in Eq. 7 in the misspecified model where the parameter D is not chosen adequately. We have

$$P\left(\overline{F}(u_n)(T(u_n) - D) > x \mid T(u_n) > D\right) = \frac{P\left(\overline{F}(u_n)(T(u_n) - D) > x\right)}{P(T(u_n) > D)}. \quad (8)$$

Holešovský and Fusek (2022) have shown that for a shifted time $T(u_n) - D$ and $x > 0$, it holds that

$$P\left(\overline{F}(u_n)(T(u_n) - D) > x\right) \to \theta \exp(-\theta x).$$

The probability $P(T(u_n) > D)$ can be alternatively rewritten using

$$P(T(u_n) > D) = P(M_{1,D+1} \leq u_n \mid X_1 > u_n). \quad (9)$$

From Corollary 1.3 of Chernick et al. (1991) it follows that if $D \geq D_0$, the limit of Eq. 9 for $n \to \infty$ is θ, which yields (7). However, for $D < D_0$ this result is not applicable. Instead, we may follow the argumentation in Cai (2023). The function

$$\lambda(s) = \lim_{n \to \infty} P(M_{1,s} \leq u_n \mid X_1 > u_n) \quad (10)$$

is a non-increasing function in s, i.e. for $D < D_0$ it holds

$$\begin{aligned} \lambda(D+1) \geq \lambda(D+2) \geq \cdots \geq \lambda(D_0) > \lambda(D_0+1) \\ = \lim_{n\to\infty} P(M_{1,r_n} \leq u_n \mid X_1 > u_n) = \theta \quad (11) \end{aligned}$$

with $r_n = o(n)$. The strict inequality between $\lambda(D_0)$ and $\lambda(D_0 + 1)$ comes from the fact that the underlying series satisfies the $D^{(k_0)}(u_n)$ condition but not the $D^{(k_0-1)}(u_n)$ condition. Specifically, if $D < D_0$ we have $\theta < \lim_{n\to\infty} P(M_{1,D+1} \leq u_n \mid X_1 > u_n) \leq 1$. Notice that the relation (11), which was derived assuming $D_0 \geq 1$ (and hence $k_0 \geq 2$), also implies $\theta < 1$. This has already been mentioned in Cai (2023).

If we put

$$\theta^*(D) = \frac{\theta}{\lim\limits_{n\to\infty} P(M_{1,D+1} \leq u_n \mid X_1 > u_n)} = \frac{\theta}{\lambda(D+1)},$$

it is easy to see that $\theta \leq \theta^*(D) < 1$ for $D < D_0$ and $\theta = \theta^*(0) \leq \theta^*(1) \leq \cdots < \theta^*(D_0) = 1$. If we combine the results above into the relation (8), we obtain

$$P\left(\overline{F}(u_n)(T(u_n) - D) > x \mid T(u_n) > D\right) \to \theta^*(D) \exp(-\theta x), \quad (12)$$

for $x > 0$. This means that for $D < D_0$ the limiting distribution of the truncated and normalized time is again a mixture of degenerate and exponential distributions whereby the proportion of the intra-cluster times is determined by the value $1-\theta^*(D)$. On the other hand, in case the $D^{(D+1)}(u_n)$ condition is satisfied (i.e. if $D \geq D_0$), $\theta^*(D) = 1$ and the degenerate part is suppressed.

A natural requirement of an estimator of the extremal index is its stability with respect to auxiliary parameters. This is, for example, usual technique applied for threshold selection in the POT model (see Scarrott and MacDonald 2012). The relation (12) offers possibility for graphical assessment of the $D^{(k)}(u_n)$ condition based on the truncated estimator. Moreover, similar approach can be also utilized for the time censor D which is being used by the censored estimator $\widehat{\theta}_{\mathrm{C}}$. The simulation study in Holešovský and Fusek (2020) confirmed that if $D < D_0$ the censored estimator is burdened with a significant bias, while for $D \geq D_0$ the bias is reduced.

5 Testing for presence of the degenerate part

A statistical test for the $D^{(k)}(u_n)$ condition can be based on identification of the degenerate part in the sample distribution of truncated or censored interexceedance times. In this section we discuss corresponding modifications of classical goodness-of-fit tests, and we derive forms of their statistics.

As before, consider a series $X_1, \ldots, X_n$ satisfying the $D^{(k_0)}(u_n)$ condition and related times $T_1, \ldots, T_{N-1}$ corresponding to a given high threshold u, where $T_{(1)} \leq \cdots \leq T_{(N-1)}$ denote the order statistics of the times. From Eq. 7 it follows that in case $D \geq D_0 = k_0 - 1$, we can approximate the distribution of $\overline{F}(u)(T_i - D)$ conditioned by $T_i > D$ by the limit exponential distribution. If $D < D_0$, the limit distribution would be in the form of Eq. 12 including some degenerate part. Hence, the inference on suitable D can be based on testing a conformity between the empirical and the theoretical exponential distribution which offers a possibility to test the null hypothesis $H_0: D \geq D_0$ against the alternative $H: D < D_0$. In order to do so, we can use goodness-of-fit tests based on the empirical distribution function.

Let $\{S_1, \ldots, S_{N_D}\} = \{T_{(N-N_D)} - D, \ldots, T_{(N-1)} - D\}$ be the set of truncated times T_i for which it holds $T_i > D$. Let $F_{\mathrm{E}}(x;\theta) = 1 - \exp(-\theta x)$ for $x > 0$ denote the cdf of the exponential distribution and $\widehat{F}_{\mathrm{T}}(x)$ denote the empirical cdf (ecdf) of the times S_i normalized by $\overline{F}(u)$, i.e.

$$\widehat{F}_{\mathrm{T}}(x) = \frac{1}{N_D}\sum_{i=1}^{N_D} \mathbf{1}_{[\overline{F}(u)S_i \leq x]},$$

where $\overline{F}(u)$ can be replaced by its estimator N/n. We apply two commonly used test statistics based on the ecdf, specifically the Kolmogorov–Smirnov (KS) statistic

$$\mathrm{KS}_{\mathrm{T}} = \sup_x \left| \widehat{F}_{\mathrm{T}}(x) - F_{\mathrm{E}}(x;\theta) \right|, \tag{13}$$

and the Anderson–Darling (AD) statistic

$$\mathrm{AD_T} = N_D \int_{-\infty}^{\infty} \frac{[\widehat{F}_\mathrm{T}(x) - F_\mathrm{E}(x;\theta)]^2}{F_\mathrm{E}(x;\theta)[1 - F_\mathrm{E}(x;\theta)]} \, \mathrm{d}F_\mathrm{E}(x;\theta), \tag{14}$$

where θ can be replaced by its ML estimate $\widehat{\theta}$. Since θ is the scale parameter, distributions of the ecdf test statistics do not depend on the true values of the unknown parameter, and depend only on the tested distribution and on the sample size. On that account, critical values of the test statistics are implemented in many softwares.

The ecdf statistics can be modified within the framework of censored interexceedance times. Let us have a censor $D \geq D_0$ and assume there are N_C largest observed times $D < T_{(N-N_C)} \leq \cdots \leq T_{(N-1)}$ for a given D, while the rest of times are taken as censored. From the likelihood (B2) in Appendix B we see, that the observed times $\overline{F}(u)T_i > \overline{F}(u)D$ should follow the exponential distribution, while the times $\overline{F}(u)T_i \leq \overline{F}(u)D$ should correspond to mixture of degenerate and continuous part of the distribution. The KS statistics for the censored sample takes the following form

$$\mathrm{KS_C} = \sup_{x \geq \overline{F}(u)D} \left| \widehat{F}_\mathrm{C}(x) - F_\theta(x;\theta) \right|, \tag{15}$$

where

$$\widehat{F}_\mathrm{C} = \frac{1}{N-1} \left((N - 1 - N_C) + \sum_{i=1}^{N_C} \mathbf{1}_{[\overline{F}(u)T_{(N-i)} \leq x]} \right)$$

is the ecdf corresponding to the censored sample, $F_\theta(x;\theta)$ is the limit cdf (3) and θ should be replaced by the censored estimator $\widehat{\theta}_\mathrm{C}$ for the given D. The AD statistics for the censored sample takes the following form

$$\mathrm{AD_C} = N_D \int_{-\infty}^{\infty} \frac{[\widehat{F}_\mathrm{C}(x) - F_\theta(x;\theta)]^2}{F_\theta(x;\theta)[1 - F_\theta(x;\theta)]} \, \mathrm{d}F_\theta(x;\theta), \tag{16}$$

where θ should again be replaced by the censored estimator $\widehat{\theta}_\mathrm{C}$ for the given D. Alternative forms of statistics $\mathrm{KS_C}$ and $\widehat{F}_\mathrm{C}$, which are more suitable for computational purposes, can be found in Fusek (2023). Critical values of the ecdf statistics can be obtained by means of Monte Carlo simulations as quantiles of the test statistics calculated from a large number of repetitions.

Here above we assume that F_θ is appropriate model for some possibly large D. However, this approach can also result in rejection of the model due to other reasons, particularly because of slow convergence of the cdf of $\overline{F}(u)T_i$ to F_θ.

6 Simulations

To assess the performance of methods discussed in previous sections, we apply them to four processes, three of them with known order k_0 of the $D^{(k_0)}(u_n)$ condition and one for which, as far as we know, the theoretical result has not been derived yet. Specifically, we consider:

1. Max-autoregressive (maxAR) process $X_i = \max\{\alpha X_{i-1}, (1-\alpha)Z_i\}$, $i = 2, \dots, n$, $X_1 = Z_1$, with $0 \le \alpha < 1$ and Z_i independent standard Fréchet, with $\theta = 1-\alpha$ satisfying the $D^{(2)}(u_n)$ condition (Robinson and Tawn 2000).
2. Moving maxima (movMax) process $X_i = \max_{0\le j\le m} \alpha_j Z_{i-j}$, $i = 1, \dots, n$, where $\alpha_j \ge 0$ satisfy $\sum_{j=0}^{m} \alpha_j = 1$, and $\theta = \max_{0\le j\le m} \alpha_j$. We take into account two variants of the movMax process. First, the coefficients are set to form a non-increasing sequence leading to a process satisfying the $D^{(2)}(u_n)$ condition (movMax-$D^{(2)}$). In the second variant of the movMax process, three subsequent zero elements are incorporated into the sequence of coefficients from the first variant leading to a process satisfying the $D^{(5)}(u_n)$ condition (movMax-$D^{(5)}$). See Ferreira and Ferreira (2018) for general conditions on the parameters under the validity of a $D^{(k)}(u_n)$ condition.
3. AR(1) process with Gaussian marginals $X_i = \phi X_{i-1} + \varepsilon_i$, where ε_i are independent $\mathrm{N}(0, 1-\phi^2)$ variables for $i \ge 2$, and X_1 has the standard Gaussian distribution, with $\phi = 0.5$ and $\theta = 1$ satisfying the $D^{(1)}(u_n)$ condition (Ancona-Navarrete and Tawn 2000).
4. Markov chain (MC) process with Gumbel margins and symmetric logistic dependence structure, i.e. $P(X_i \le x, X_{i+1} \le y) = \exp\left[-(\mathrm{e}^{-x/\alpha} + \mathrm{e}^{-y/\alpha})^{\alpha}\right]$, with $\alpha = 0.5$ and $\theta \approx 0.331$ (Fawcett and Walshaw 2012). The order of the dependence condition is unknown.

For all these processes, we take the estimators of θ discussed in previous sections and explore the existing and newly proposed methods for assessment of the local dependence condition $D^{(k_0)}(u_n)$ and estimation of the order k_0. Specifically, we consider the censored estimator $\widehat{\theta}_{\mathrm{C}}$, the truncated estimator $\widehat{\theta}_{\mathrm{T}}$ and the K-gaps estimator $\widehat{\theta}_{\mathrm{SD}}$. Results will be graphically illustrated only for $\theta = 0.4$ in case of maxAR process, and for $\theta = 0.8$ ($\alpha_0 = 0.8, \alpha_1 = \alpha_2 = \alpha_3 = 0, \alpha_4 = \alpha_5 = 0.1$) in case of movMax-$D^{(5)}$. More results on other variants of maxAR or movMax processes can be found in supplementary material provided to this paper (Holešovský and Fusek 2024).

Firstly, we may assess validity of the $D^{(k)}(u_n)$ condition by graphical inspection of the anti-$D^{(k)}(u_n)$ events and stability of the extremal index estimators. In Fig. 1, there are typical trajectories of the maxAR, movMax and AR(1) processes, where the k_0 values are known. In addition, there are the proportions $p_{k_0}(u_i, r_j)$ of anti-$D^{(k_0)}(u_n)$ from Eq. 6 as it was proposed by Süveges (2007). These illustrative plots are based on single realizations of the processes and have been obtained for samples of size $n = 10\,000$. The values of $p_{k_0}(u_i, r_j)$ coloured in gray indicate the proportions below given significance level 0.05. Based on our experience with the plots, we conclude that Fig. 1 demonstrates very typical paths for the processes of interest.

For maxAR and movMax processes a path (u_i, r_j) can be possibly identified such that $p_{k_0}(u_i, r_j) \to 0$. In case of the AR(1), however, such path is only apparent for thresholds above 99% quantile (not visible in Fig. 1). Hence, it may require a large sample size to be identified. For the MC process in Fig. 2 we show typical plots of anti-$D^{(k)}(u_n)$ events with $k = 3, 4$ and 5, whereas the value $k = 4$ was presumed valid by Ferreira and Ferreira (2018). This is in accordance with the middle and right

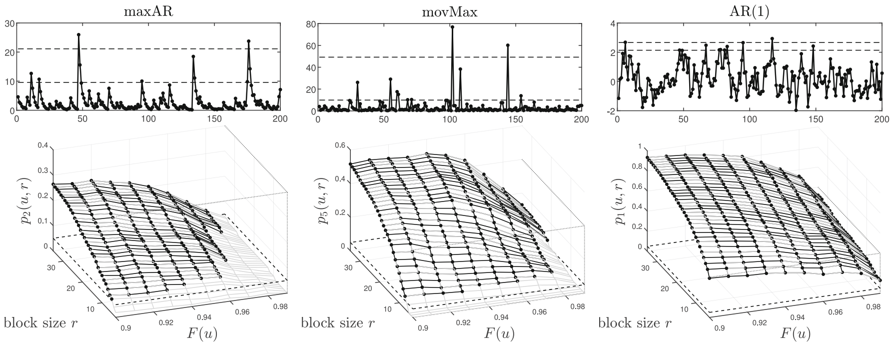

Fig. 1 Upper plots: example trajectories of the maxAR ($\theta = 0.4$), movMax-$D^{(5)}$ ($\theta = 0.8$), and AR(1) processes. Dashed lines show 95% and 99% sample quantiles. Lower plots: proportions $p_{k_0}(u, r)$ of the anti-$D^{(k_0)}(u_n)$ events evaluated from a single trajectory. Dashed line around the box indicates the 0.05 level, gray parts of $p_{k_0}(u_i, r_j)$ indicate values below this level

plots in Fig. 2. In contrast to that, the left plot also indicates some decrease of $p_2(u, r)$ for large enough thresholds.

It is evident, that the inspection of the anti-$D^{(k)}(u_n)$ events can lead to unclear and rather subjective results, since the conclusions made from these plots strongly depend on the range of considered block sizes and thresholds. On the other hand, simplicity of this approach makes it suitable for initial evaluation. The situation gets more difficult if the goal is not the assessment of a given condition, but determination of the minimal order k_0. This requires comparisons of multiple plots, such as in Fig. 2, and it turns out that an inspection of plots of differences discussed in Section 4 can be more convenient. In this case, however, the threshold and the block size are connected and their effects cannot be investigated separately.

We build our graphical diagnosis on stability of the estimators $\widehat{\theta}_{\mathrm{C}}$ and $\widehat{\theta}_{\mathrm{T}}$ with respect to the choice of auxiliary parameter D. Suitable $D_0 = k_0 - 1$ should be chosen

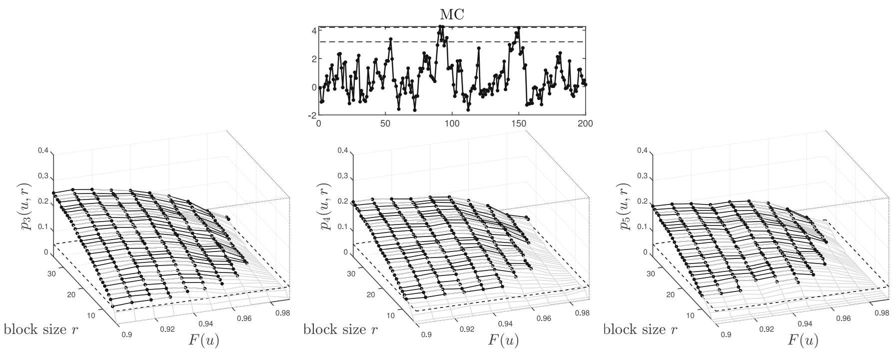

Fig. 2 Upper plot: example trajectory of the MC process, dashed lines show 95% and 99% sample quantiles. Lower plots: proportions $p_k(u, r)$ of the anti-$D^{(k)}(u_n)$ events evaluated from a single trajectory for $k = 3$ (left), $k = 4$ (middle) and $k = 5$ (right). Dashed line around the box indicates the 0.05 level, gray parts of $p_k(u_i, r_j)$ indicate values below this level

such that a certain stability in estimation of θ for values $D \geq D_0$ is reached. The paths of the estimators obtained for single realizations of considered processes are shown in Fig. 3. We also include confidence intervals for $\widehat{\theta}_{\mathrm{C}}$ based on asymptotic normality of this ML estimator using the asymptotic variance derived in Holešovský and Fusek (2020). The paths are in accordance with the conclusions made in Holešovský and Fusek (2020, 2022), who have shown through simulation that the estimators $\widehat{\theta}_{\mathrm{C}}$ and $\widehat{\theta}_{\mathrm{T}}$ are burdened with significant bias in case $D < D_0$, while for $D \geq D_0$ the bias is significantly reduced. The values of D, for which a stability of the estimators is achieved and/or the confidence interval for $\widehat{\theta}_{\mathrm{C}}$ contains the true value of θ, satisfy the theoretical k_0 for maxAR and movMax processes. In case of the AR(1) process, the stability plot leads to $D = 1$ instead of $D = 0$. For the MC process, the bias is rather reduced for $D = 1$ and vanishes for $D = 3$. This could indicate that the process does not exhibit significant differences from a process satisfying $D^{(2)}(u_n)$ or $D^{(4)}(u_n)$ condition.

In Fig. 3, paths of the estimator $\widehat{\theta}_{\mathrm{SD}}$ are also visualized. As it was already addressed in Section 3, it can be observed that any deviation of the auxiliary parameter K from proper $K_0 = k_0 - 1$ could result in increase of $\widehat{\theta}_{\mathrm{SD}}$ bias. This is mostly apparent for the movMax-$D^{(5)}$ process in Fig. 3 (top right), where increasing the auxiliary parameter above K_0 results in a significant underestimation of θ. We may conclude, based on our experience, that this instability is mainly emphasized for moderate to large values of θ irrespective of any specific process.

Besides the graphical diagnostics, we deal with statistical tests presented in Sections 3 and 5 and study their properties via simulations. We consider the following statistics: IMT for the censored sample (Eq. A1 in Appendix A), $\mathrm{KS_T}$ (13) and $\mathrm{AD_T}$ (14) tests for truncated times, and $\mathrm{KS_C}$ (15) and $\mathrm{AD_C}$ (16) for censored times. We draw 1000 samples of size $n = 10\,000$ from each process and perform all the tests with a given significance level $\alpha = 0.05$. Critical values for $\mathrm{KS_C}$ and $\mathrm{AD_C}$ tests are obtained

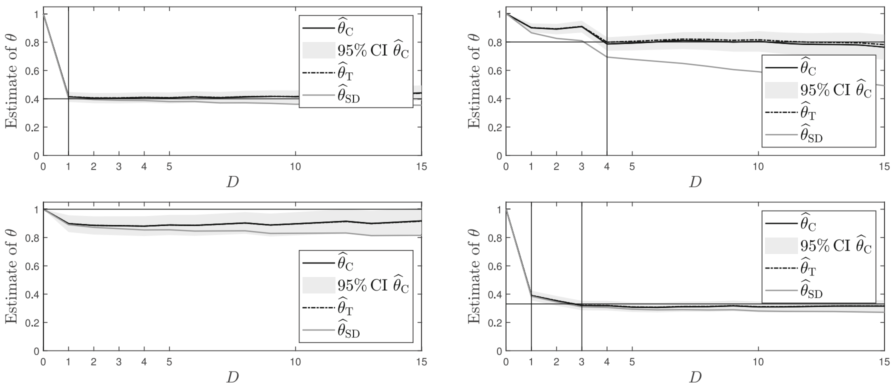

Fig. 3 Examples of stability plots of the extremal index estimators for maxAR (top left), movMax-$D^{(5)}$ (top right), AR(1) (bottom left), and MC process (bottom right). The estimates were obtained for u corresponding to 95% sample quantile, or 99% quantile for AR(1). Confidence interval (CI) for $\widehat{\theta}_{\mathrm{C}}$ is based on the asymptotic normality of the ML estimator. True value of θ is indicated by the horizontal line. Vertical lines indicate expected values D_0 or K_0 (zero for AR(1))

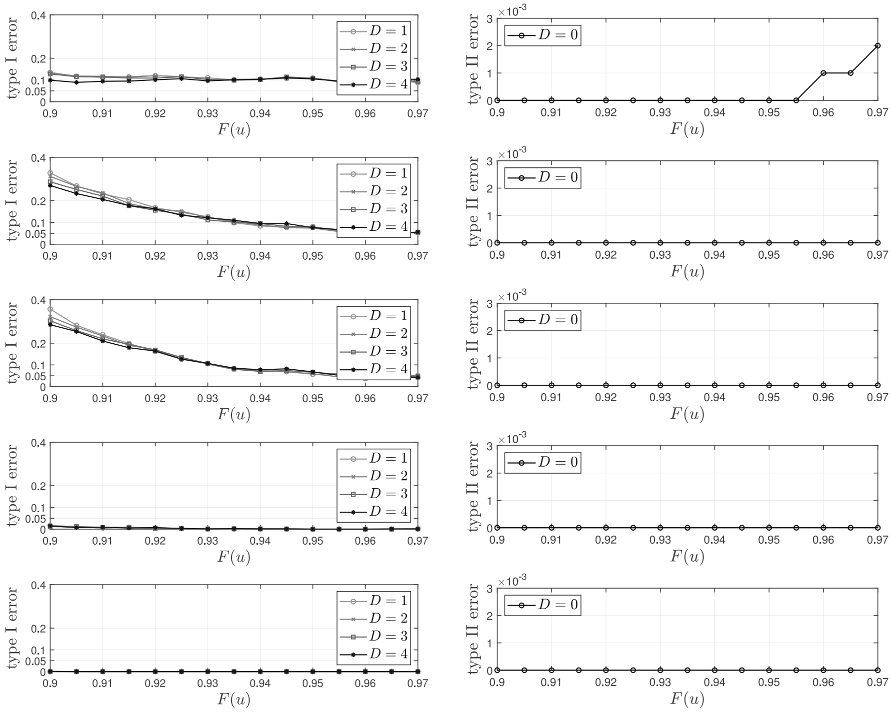

Fig. 4 Estimated probabilities of type I (left) and II (right) errors obtained for maxAR process ($\theta = 0.4$) at a significance level of $\alpha = 0.05$. From top to bottom: IMT, KS_T, AD_T, KS_C and AD_C test

by means of Monte Carlo method based on 10 000 samples. The interexceedance times are determined by thresholds corresponding to sample quantiles ranging from 0.9 to 0.97, i.e. from 1000 to 300 threshold exceedances. For AR(1) we choose the quantiles up to 0.99 as discussed below. Effect of the threshold selection is of particular interest. The auxiliary parameter D is selected in the range from 0 up to $D_0 + 3$, some of them being unsuitable with respect to models (7) or (B2) (in Appendix B). The probabilities of type I and type II errors are estimated as the proportions of cases in which the null hypothesis of D being suitable is or is not rejected, respectively.

For a reliable assessment of the local dependence condition $D^{(k)}(u_n)$, or its minimal order k_0, it is desirable to reject the null hypothesis for all $k < k_0$ and not to reject it for $k \geq k_0$. This means to obtain small probabilities of the type I error (below the significance level) for $D \geq D_0$ and of the type II error for $D < D_0$, respectively.

The estimated probabilities of errors for all processes under consideration with the exception of movMax-$D^{(2)}$ process are shown in Figs. 4–7. Results for the movMax-$D^{(2)}$ process are similar to those for the maxAR process and can be found in the supplementary material. In case of the maxAR process, all the tests are capable to identify the closest inappropriate value $D = D_0 - 1$ with a high probability (small estimated probability of type II error) for all values of θ (see Fig. 4 for the case $\theta = 0.4$) with the exception of IMT, which shows a rapid increase of probability of type II error

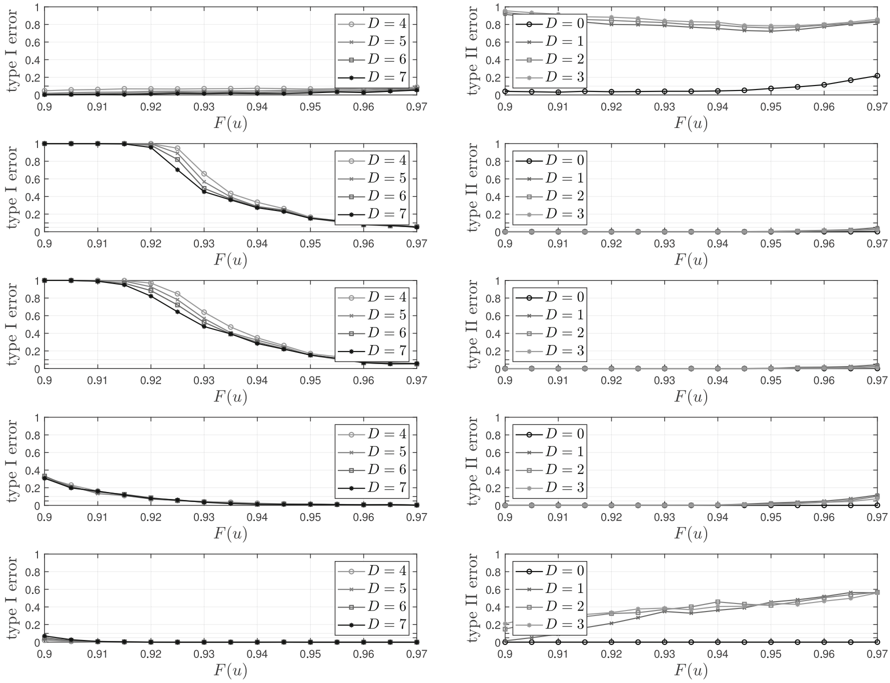

Fig. 5 Estimated probabilities of type I (left) and II (right) errors obtained for movMax-$D^{(5)}$ process ($\theta = 0.8$) at a significance level of $\alpha = 0.05$. From top to bottom: IMT, KS_T, AD_T, KS_C and AD_C test

for $\theta = 0.9$ (not shown in figures). Moreover, KS_C and AD_C tests outperform the other alternatives with respect to the probabilities of type I errors which fall below the given significance level for all θ and high thresholds. On the other hand, the probabilities of

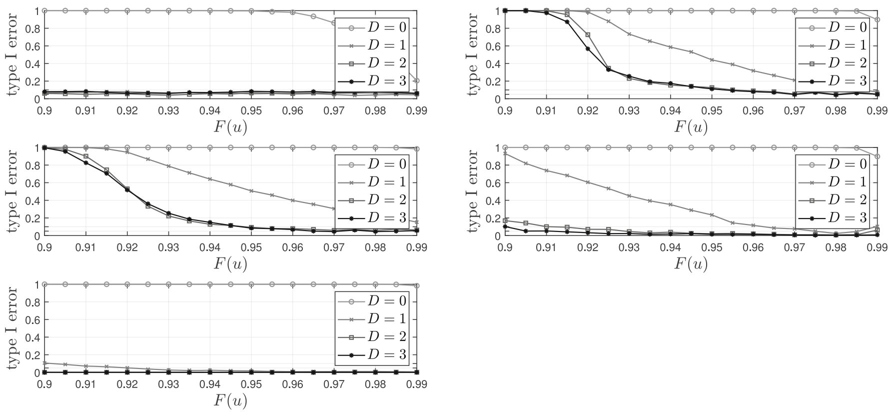

Fig. 6 Estimated probabilities of type I errors (any D being suitable) obtained for AR(1) process at a significance level of $\alpha = 0.05$: IMT (top left), KS_T (top right), AD_T (middle left), KS_C (middle right) and AD_C test (bottom left)

the type I error are well below the 0.05 level, which indicates that the tests could be better calibrated. In general, the AD_C test performs the best. The KS_T and AD_T tests perform reasonably well for high thresholds, especially if $\theta \leq 0.4$.

In case of the movMax-$D^{(5)}$ process, the estimated probability of type II error starts to grow for IMT and AD_C tests and $\theta \geq 0.6$. For example, Fig. 5 shows for $\theta = 0.8$ that the IMT is reasonably able to detect the invalidity of the local dependence condition only when the difference between D and D_0 is large. All the other tests show a slow increase of the probability of type II error for high thresholds and high values of θ. On that account, neither the IMT nor the AD_C test could be recommended. Regarding the probability of type I error, KS_C and AD_C tests outperform the other alternatives as the significance level of 0.05 is met for all θ and high thresholds. Nevertheless, the probabilities of type I error are again well below the required limit for both tests. This implies the poor performance of KS_C and AD_C in identifying any unsuitable D, especially when it is close to D_0. For a sufficiently large sample size, the KS_T and AD_T represent a better option at high threshold levels.

A special case is the AR(1) process that satisfies the $D^{(1)}(u_n)$ condition, hence any choice of parameter D should be suitable. The value $D = 0$ is, however, mostly rejected by all tests unless the threshold is set enormously high. Similarly, the probabilities of type I errors of KS_T and AD_T tests exhibit very slow decrease for any other D, see Fig. 6. It follows from Leadbetter et al. (1983) that validity of the $D^{(1)}(u_n)$ condition

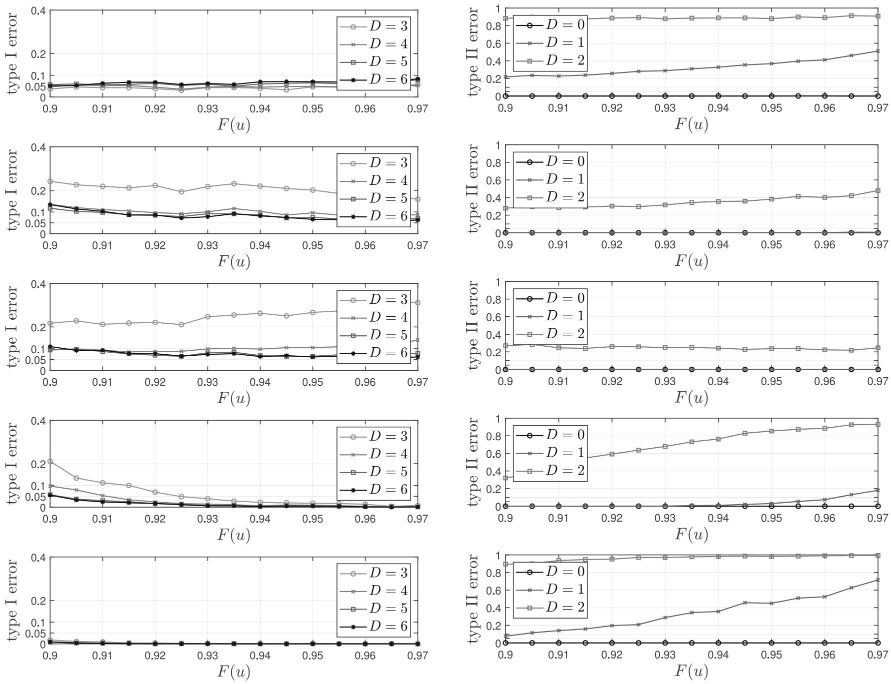

Fig. 7 Estimated probabilities of type I (left) and II (right) errors obtained for MC process at a significance level of $\alpha = 0.05$ under assumption $k_0 = 4$. From top to bottom: IMT, KS_T, AD_T, KS_C and AD_C test

necessarily implies $\theta = 1$. The difficulty of estimating θ at the boundary of its domain is a well-known problem of all current estimators, and it also affects the performance of the tests, which is possibly supported by a slower convergence of the interexceedance times in Eq. 3.

In case of the MC process, the proper order k_0 of the dependence condition is unknown. From the stability plot in Fig. 3 we observed that the process exhibits only a small deviation from both $k_0 = 2$ and $k_0 = 4$. We assume $k_0 = 4$ (and hence $D_0 = 3$), which is also based on results from Ferreira and Ferreira (2018). Probabilities of type I and II errors under such assumption are plotted in Fig. 7. Because $D \geq D_0$ does not violate any of the considered models (B2) or (12), possible misspecification in selected $k_0 = 4$ against the alternative $k_0 < 4$ does not harm the probabilities of type I errors (left figures). It may be, however, misleading in terms of probability of the type II error, since some of the values of parameter D are incorrectly classified as inappropriate. For such D the tests would indicate large proportion of cases in which the null hypothesis is not rejected. This is evidently the case of $D = 2$ in Fig. 7, while the probabilities of type II errors for $D = 1$ correspond rather to the results obtained for maxAR and movMax processes.

7 Conclusion

We proposed several approaches for assessment of the local dependence condition $D^{(k)}(u_n)$ and identification of its minimal order k_0. The first approach was based on assessing stability of the extremal index estimators with respect to choice of the auxiliary parameter D (K respectively). Our experience with commonly considered processes indicates that this approach may work for estimators $\widehat{\theta}_{\mathrm{C}}$ and $\widehat{\theta}_{\mathrm{T}}$, which show certain stability with respect to an appropriate choice of auxiliary parameters. Nevertheless, the stability plot is not suited for estimators which are heavily dependent on the choice of auxiliary parameters including choice of the threshold, for example $\widehat{\theta}_{\mathrm{SD}}$. It is also necessary to distinguish whether the aim is the determination of a suitable D for the purpose of extremal index estimation (by $\widehat{\theta}_{\mathrm{C}}$ or $\widehat{\theta}_{\mathrm{T}}$, for example), or the estimation of the minimal order k_0 itself. A wrong determination of k_0 or D_0 does not have to necessarily lead to a significant deterioration of the estimate of θ.

The second approach was based on five statistical tests. It was found out that in case of maxAR and movMax processes, the $\mathrm{KS_C}$ and $\mathrm{AD_C}$ tests outperform the other alternatives with respect to the probabilities of type I errors which fall below the given significance level for all θ and high thresholds. The $\mathrm{AD_C}$ test performs the best for maxAR and movMax-$D^{(2)}$ (see the supplementary material) processes, and the $\mathrm{KS_C}$ test performs the best for the movMax-$D^{(5)}$ process. Nevertheless, probabilities of the type I error well below the significance level imply a poor performance of $\mathrm{KS_C}$ and $\mathrm{AD_C}$ in identifying any unsuitable D. For large sample sizes and high thresholds, the $\mathrm{KS_T}$ and $\mathrm{AD_T}$ represent a better option, because their type I error probabilities approach the given significance level and their type II error probabilities remain low.

In case of the AR(1) process, the proper $k_0 = 1$ in the $D^{(k_0)}(u_n)$ condition is mostly rejected by all tests unless the threshold is set enormously high, which could be caused by a slower convergence of the interexceedance times to the limiting distribution. In this case, the test results are also affected by the quality of the $\theta = 1$ estimate at the boundary of its domain, which is problematic for all estimators. If we assume that the MC process satisfies the $D^{(4)}(u_n)$ condition, the best probability of type I error is obtained in case of the $\mathrm{AD_C}$ test. Nevertheless, all tests can be misleading in terms of type II error.

Supplementary information

We provide supplementary material to this paper in Holešovský and Fusek (2024). In this supplement we give additional results of estimated probabilities of type I and II errors obtained for statistical tests proposed in the paper on validity of the $D^{(k)}(u_n)$ condition of Chernick et al. (1991). Particularly, we provide the results for various setting of parameters for maxAR and movMax processes.

Appendix A: Information matrix test for the censored estimator

Assume that $X_1, \ldots, X_n$ is a stationary series satisfying the $D(u_n)$ condition of Leadbetter et al. (1983). For a high enough threshold u, consider that there are interexceedance times $T_1, \ldots, T_{N-1}$, where $N = \sum_{i=1}^{n} \mathbf{1}_{[X_i > u]}$ is the number of exceedances. Assume the normalized times $\overline{F}(u)T_i$, $i = 1, \ldots, N-1$, are i.i.d. variables drawn from the distribution with cdf F_θ in Eq. 3.

For a given time censor $D \geq 0$ treat the time $T_i \leq D$ as censored while the time $T_i > D$ is observed. For the censored estimator of the extremal index we have the log-likelihood function

$$\ell(\theta) = \sum_{i=1}^{N-1} \ell_i(\theta),$$

where

$$\ell_i(\theta) = \log\kappa + \mathbf{1}_{[\overline{F}(u)T_i \leq d]} \log F_\theta(d) + \mathbf{1}_{[\overline{F}(u)T_i > d]} \left(2\log\theta - \theta\overline{F}(u)T_i\right)$$

is a single i-th observation contribution, $\kappa = (N-1)!/(N-1-N_C)!$ and $d = \overline{F}(u)D$. We derive the following derivatives

$$\ell_i'(\theta) = \mathbf{1}_{[\overline{F}(u)T_i \leq d]} \frac{-e^{-\theta d} + d\theta e^{-\theta d}}{1 - \theta e^{-\theta d}} + \mathbf{1}_{[\overline{F}(u)T_i > d]} \left(\frac{2}{\theta} - \overline{F}(u)T_i\right),$$

$$\ell_i''(\theta) = \mathbf{1}_{[\overline{F}(u)T_i \leq d]} \frac{\left(2d - d^2\theta\right) e^{-\theta d} \left(1 - \theta e^{-\theta d}\right) - (d\theta - 1)^2 e^{-2\theta d}}{\left(1 - \theta e^{-\theta d}\right)^2} - \mathbf{1}_{[\overline{F}(u)T_i > d]} \frac{2}{\theta^2}.$$

The score function for the i-th observation is

$$j_i(\theta) = \left[\ell_i'(\theta)\right]^2 = \mathbf{1}_{[\overline{F}(u)T_i \leq d]} \frac{e^{-2\theta d}(d\theta - 1)^2}{\left(1 - \theta e^{-\theta d}\right)^2} + \mathbf{1}_{[\overline{F}(u)T_i > d]} \left(\frac{2}{\theta} - \overline{F}(u)T_i\right)^2,$$

and the Fisher's information is

$$
\begin{aligned}
i_i(\theta) &= -\ell_i''(\theta) \\
&= \mathbf{1}_{[\overline{F}(u)T_i>d]}\frac{2}{\theta^2} - \mathbf{1}_{[\overline{F}(u)T_i\le d]}\frac{\left(2d-d^2\theta\right)e^{-\theta d}\left(1-\theta e^{-\theta d}\right)-(d\theta-1)^2e^{-2\theta d}}{\left(1-\theta e^{-\theta d}\right)^2}.
\end{aligned}
$$

The difference between the score function and the Fisher's information together with its derivative is

$$
\begin{aligned}
\delta_i(\theta) =& j_i(\theta) - i_i(\theta) \\
=& \mathbf{1}_{[\overline{F}(u)T_i>d]}\left(\frac{2}{\theta^2}-\frac{4\overline{F}(u)T_i}{\theta}+(\overline{F}(u)T_i)^2\right)+\mathbf{1}_{[\overline{F}(u)T_i\le d]}\frac{\left(2d-d^2\theta\right)e^{-\theta d}}{\left(1-\theta e^{-\theta d}\right)}, \\
\delta_i'(\theta) =& \mathbf{1}_{[\overline{F}(u)T_i>d]}\left(\frac{4\overline{F}(u)T_i}{\theta^2}-\frac{4}{\theta^3}\right) \\
&+\mathbf{1}_{[\overline{F}(u)T_i\le d]}\frac{\left(\theta d^3-3d^2\right)e^{-\theta d}\left(1-\theta e^{-\theta d}\right)+\left(\theta^2d^3-3\theta d^2+2d\right)e^{-2\theta d}}{\left(1-\theta e^{-\theta d}\right)^2}.
\end{aligned}
$$

Denote the sample means

$$
\Delta(\theta)=\frac{1}{N-1}\sum_{i=1}^{N-1}\delta_i(\theta),\quad \Delta'(\theta)=\frac{1}{N-1}\sum_{i=1}^{N-1}\delta_i'(\theta),\quad I(\theta)=\frac{1}{N-1}\sum_{i=1}^{N-1}i_i(\theta).
$$

The sample variance of $\Delta(\theta)$ is of the form

$$
V(\theta)=\frac{1}{N-1}\sum_{k=1}^{N-1}\left[\delta_k(\theta)-\Delta'(\theta)I^{-1}(\theta)\ell_k'(\theta)\right]^2,
$$

and the information matrix test statistic is

$$
\mathrm{IMT}(\theta)=(N-1)\Delta^2(\theta)V^{-1}(\theta), \tag{A1}
$$

where θ should be replaced by its estimate $\widehat{\theta}$.

Appendix B: Extremal index estimators based on interexceedance times

Here below we provide a brief overview of extremal index estimators based on interexceedance times that are discussed in Sections 3, 4, and 5. The methods are covered for completeness, and more details can be found in the papers mentioned below.

Consider a sequence of interexceedance times $T_1, \dots, T_{N-1}$ and assume the times are independent. Denote $T_{(1)} \leq \cdots \leq T_{(N-1)}$ the ordered statistics of the times.

K-gaps estimator

The K-gaps estimator $\widehat{\theta}_{\mathrm{SD}}$ of Süveges and Davison (2010) is based on the sample of gaps $S_i^{(K)} = \max\{T_i - K; 0\}$, $i = 1, \dots, N-1$, for some $K \geq 0$. Süveges and Davison (2010) have shown that the limit distribution of the gaps remains in the form of Eq. 3 with $T(u_n)$ replaced by $S^{(K)}(u_n)$. Under some additional assumptions on the local dependence, the corresponding log-likelihood function is of the form

$$\ell_{\mathrm{SD}}(\theta) = (N - 1 - N_C)\log(1-\theta) + 2N_C \log\theta - \theta \sum_{i=1}^{N-1} \overline{F}(u) S_i^{(K)},$$

where $N_C = \sum_{i=1}^{N-1} \mathbf{1}_{[T_i > K]}$ is the number of interexceedance times exceeding the value of K. The K-gaps estimator is obtained by maximizing the log-likelihood function, i.e.

$$\widehat{\theta}_{\mathrm{SD}} = \underset{0\leq\theta\leq 1}{\arg\max}\, \ell_{\mathrm{SD}}(\theta)\,.$$

The adequateness of the likelihood model, however, requires a suitable value K_0 of the parameter K. Süveges (2007) assumes $K_0 = 1$ in case the $D^{(2)}(u_n)$ condition holds. The discussion in Holešovský and Fusek (2020) leads to $K_0 = k_0 - 1$ for the $D^{(k_0)}(u_n)$ condition. A deviation of K from K_0 brings a bias into the estimator $\widehat{\theta}_{\mathrm{SD}}$ as the times corresponding to the degenerate part of F_θ are assigned to the exponential part or vice versa.

Censored estimator

Consider a time censor $D \geq 0$. The censored estimator $\widehat{\theta}_{\mathrm{C}}$ of Holešovský and Fusek (2020) is based on division of $T_1, \dots, T_{N-1}$ into two sets. The N_C largest times greater than D are considered observed (uncensored), while the times less than or equal to D are treated as censored. Assuming independence of the interexceedance times, the log-likelihood function of the censored sample can be written in the form of

$$\ell(\theta) = \sum_{i=1}^{N-1} \left\{ \log\kappa + \mathbf{1}_{[\overline{F}(u)T_i \leq d]} \log F_\theta(d) \; + \mathbf{1}_{[\overline{F}(u)T_i > d]} \log\left[\theta^2 e^{-\theta\overline{F}(u)T_i}\right] \right\}, \quad \text{(B2)}$$

where $\mathbf{1}_{[\,.\,]}$ is the indicator function and $\kappa = (N-1)!/(N-1-N_C)!$ is a constant with $N_C = \sum_{i=1}^{N-1} \mathbf{1}_{[T_i > D]}$. The normalized time $\overline{F}(u)T_i$ is considered to be a random variable drawn from F_θ and $d = \overline{F}(u)D$ is the normalized censor, whereas the value $\overline{F}(u)$ can be estimated by N/n. The censored estimator $\widehat{\theta}_{\mathrm{C}} = \widehat{\theta}_{\mathrm{C}}(D)$ of the extremal index is obtained by maximizing of the log-likelihood function (B2). The observed times $T_i > D$ in Eq. B2 are assigned purely to the non-degenerate part of F_θ, from which follows the restriction $D \geq D_0 = k_0 - 1$ for the model to be valid (see Holešovský and Fusek 2020, for more details).

Truncated estimator

The truncated estimator $\widehat{\theta}_{\mathrm{T}}$ of Holešovský and Fusek (2022) is based on limit distribution (7) of truncated interexceedance times that exceed some suitable truncation point D. Given interexceedance times, let $\{S_1, \ldots, S_{N_D}\} = \{T_{(N-N_D)} - D, \ldots, T_{(N-1)} - D\}$ be the set of times above the truncation point D, i.e. with $T_{(N-N_D-1)} \le D$. A simple ML estimator $\widehat{\theta}$ of θ could be derived from Eq. 7 as the reciprocal of the sample mean

$$\widehat{\theta} = \frac{N_D}{\sum_{i=1}^{N_D} \overline{F}(u) S_i}. \tag{B3}$$

The application of a bias correction and a penultimate approximation $P(T > n) = \theta p^{n\theta}$, $n = 1, 2, \ldots$, of the limiting distribution leads to the so-called truncated estimator

$$\widehat{\theta}_{\mathrm{T}} = \widehat{\theta}^{\mathrm{BC}} - \frac{\overline{F}(u)}{2(N-1)} \left[1 + \widehat{\theta}^{\mathrm{BC}}(N-4) - \left(\widehat{\theta}^{\mathrm{BC}}\right)^2 (N-1) \right],$$

where $\overline{F}(u)$ should be replaced by its estimator N/n,

$$\widehat{\theta}^{\mathrm{BC}} = \frac{(N-1)\widehat{\theta} - 1}{N - 1 + \overline{F}(u) D},$$

and $\widehat{\theta}$ is given in Eq. B3.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10687-025-00513-8.

Acknowledgements The authors would like to thank the editor and the reviewers for a number of useful suggestions which helped improve the manuscript.

Author Contributions J.H. provided theoretical framework of the misspecified models and its derivation, M.F. derived and implemented statistical tests for identification of degenerate part in the distribution mixture. Both authors contributed equally on the simulation study, the manuscript text and its review.

Funding The paper was supported by specific research project No. FAST-S-22-7867 at Brno University of Technology.

Data Availability A Matlab live script demonstrating graphical methods to assessing validity of the $D^{(k)}(u_n)$ condition is included as a supplementary material. All generated datasets and the rest of methods implemented in Matlab are available from the corresponding author.

Declarations

Conflict of interests Authors have no conflict of interest to declare that are relevant to the content of this article.

Ethics approval Not applicable.

Competing interests The authors declare no competing interests.

References

Ancona-Navarrete, M.A., Tawn, J.A.: A comparison of methods for estimating the extremal index. Extremes **3**(1), 5–38 (2000)

Cai, J.J.: Statistical inference on $D^{(d)}(u_n)$ condition and estimation of the extremal index. Accessed 5 May 2024. (2023). https://arxiv.org/abs/1911.06674v2

Chernick, M.R., Hsing, T., McCormick, W.P.: Calculating the extremal index for a class of stationary sequences. Adv. Appl. Probab. **23**, 835–850 (1991)

Fawcett, L., Walshaw, D.: Estimating return levels from serially dependent extremes. Environmetrics **23**, 272–283 (2012)

Ferreira, H., Ferreira, M.: Estimating the extremal index through local dependence. Annales de l'Institute Henri Poincare (B) Probability and Statistics **54**(2), 587–605 (2018)

Ferreira, M.: Analysis of estimation methods for the extremal index. Electron. J. App. Stat. Anal. **11**(1), 296–306 (2018)

Ferro, C.A.T., Segers, J.: Inference for clusters of extreme values. J. R. Stat. Soc. B Met. **65**(2), 545–556 (2003)

Fukutome, S., Liniger, M.A., Süveges, M.: Automatic threshold and run parameter selection: a climatology for extreme hourly precipitation in switzerland. Theor. Appl. Climatol. **120**(3–4), 403–416 (2015)

Fukutome, S., Liniger, M.A., Süveges, M.: Correction to: Automatic threshold and run parameter selection: a climatology for extreme hourly precipitation in switzerland. Theor. Appl. Climatol. **137**(3–4), 3215 (2019)

Fusek, M.: Statistical power of goodness-of-fit tests for type i left-censored data. Austrian J. Stat. **52**(1), 51–61 (2023)

Gomes, M.I.: On the estimation of parameter of rare events in environmental time series. In: Statistics for the Environment. Water Related Issues, vol. 2, pp. 225–241. Wiley, New York (1993)

Holešovský, J., Fusek, M.: Supplement to 'Statistical inference on the local dependence condition of extreme values in a stationary sequence'. (2024)

Holešovský, J., Fusek, M.: Estimation of the extremal index using censored distributions. Extremes **23**(2), 197–213 (2020)

Holešovský, J., Fusek, M.: Improved interexceedance-times-based estimator of the extremal index using truncated distribution. Extremes **25**(4), 695–720 (2022)

Leadbetter, M.R., Lindgren, G., Rootzén, H.: Extremes and Related Properties of Random Sequences and Series. Springer, New York (1983)

Northrop, P.: An efficient semiparametric maxima estimator of the extremal index. Extremes **18**(4), 585–603 (2015)

Robert, C.Y., Segers, J., Ferro, C.A.: A sliding blocks estimator for the extremal index. Electron. J. Stat. **3**, 993–1020 (2009)

Robinson, M.E., Tawn, J.A.: Extremal analysis of processes sampled at different frequencies. J. R. Stat. Soc. B. **62**(1), 117–135 (2000)

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

Smith, R.L., Weissman, I.: Estimating the extremal index. J. R. Stat. Soc., Ser. B. **50**, 151–528 (1994)

Süveges, M.: Likelihood estimation of the extremal index. Extremes **10**, 41–55 (2007)

Süveges, M., Davison, A.C.: Model misspecification in peaks over threshold analysis. Ann. Appl. Stat. **4**(1), 203–221 (2010)
White, H.: Maximum likelihood estimation of misspecified models. Econometrica **50**, 1–25 (1982)