Noname manuscript No.
(will be inserted by the editor)

Analysis of crack growth with robust, distribution-free estimators and tests for nonstationary autoregressive processes

Kustosz, Ch.P. · Müller, Ch.H.

Received: date / Accepted: date

Abstract This article investigates the application of depth estimators to crack growth models in construction engineering. Many crack growth models are based on the Paris-Erdogan equation which describes crack growth by a deterministic differential equation. By introducing a stochastic error term, crack growth can be modeled by a nonstationary autoregressive process with Lévy-type errors. A regression depth approach is presented to estimate the drift parameter of the process. We then prove the consistency of the estimator under quite general assumptions on the error distribution. By an extension of the depth notion to simplical depth it is possible to use a degenerated U statistic and to establish tests for general hypotheses about the drift parameter. Since the statistic asymptotically has a transformed χ^2_1 distribution, simple confidence intervals for the drift parameter can be obtained. In the second part, simulations of AR(1) processes with different error distributions are used to examine the quality of the constructed test. Finally we apply the presented method to crack growth experiments. We compare two datasets from independent experiments under different conditions but with the same material. We show that the parameter estimates differ significantly in this case.

Keywords crack growth · stochastic differential equation · autoregressive process · data depth · robustness · maximum depth estimator · simplical depth · tests

Mathematics Subject Classification (2000) 62G10 · 62G35 · 62G20 · 62G05 · 62G51 · 62M10

Christoph P. Kustosz
TU Dortmund University
Tel.: +49(0)231-755-3114
E-mail: kustosz@statistik.tu-dortmund.de

Christine H. Müller
TU Dortmund University

1 Introduction

The understanding of crack growth is very important for predicting the life time of products as wheels of trains, bridges or hip replacement. The crack growth is the faster the longer the crack is. This rule is expressed in the well-known Paris-Erdogan equation by

$$\frac{d\,a}{d\,N} = C\,(\Delta K)^m \quad \text{with} \quad \Delta K = \Delta\sigma\,\sqrt{\pi\,a}\,G \tag{1}$$

(see e.g. Pook 2000). Thereby a denotes the crack length or width, N is the number of load cycles of a cyclic loading, C and m are constants depending on the material. ΔK is the stress intensity factor, where $\Delta\sigma = \sigma_{max} - \sigma_{min}$ is the amplitude of the cyclic stress and G is a geometrical factor. However, the lifetime of material is not deterministic and cracks are influenced by atomic structures. Therefore it is obvious that crack growth is a nondeterministic process. Hence a stochastic process given by the stochastic differential equation

$$\frac{dA(t)}{dt} = \alpha_0 + \alpha_1\,A(t)^{m/2} + \frac{dV(t)}{dt}, \tag{2}$$
$$\text{where } V(\cdot) \text{ is a stochastic error process, } \alpha_1 > 0,\ \alpha_0 \in I\!R,$$

is a more realistic description of the random crack growth process $A(\cdot)$ depending on the time t, which is in line with the Paris-Erdogan equation. The parameter m is given in many applications (often set as $m = 2$) so that α_0 and α_1 are the only unknown parameters.

Most results for stochastic differential equations deal with errors defined by a Wiener process. But for crack growth, the Wiener process is not appropriate since the process should be increasing. Moreover crack growth is often not continuous since jumps appear in the growth process. These jumps are caused by joining of different cracks or by the failure of substructures. A Lévy process with nonnegative increments $V(t+h) - V(t)$ is then much more appropriate. Here in particular prestressed concrete is considered as an example. In this example, the breaking of the ropes for reinforcing the prestressed concrete causes jumps in the width of the crack. These jumps can for example be modeled by a Poisson-Gamma process or more simply by outliers in a process with continuous distribution. Here we will treat these jumps as outliers.

Under these assumptions we consider the problem of estimating and testing the drift parameter α_1 from observations $y_0, \ldots, y_N$ of crack length or crack width at time points $t_0, \ldots, t_N$. For the statistical inference, the Euler-Maruyama approximation of the stochastic process $A(\cdot)$ can be used (see Kloden et al. 2000 or Iacus 2008). It provides

$$\frac{A(t+h) - A(t)}{h} \approx \alpha_0 + \alpha_1\,A(t)^{m/2} + \frac{V(t+h) - V(t)}{h}$$

or

$$\begin{aligned} &A(t+h) - A(t) \approx h\alpha_0 + h\alpha_1 A(t)^{m/2} + V(t+h) - V(t) \\ &\iff A(t+h) \approx A(t) + h\alpha_0 + h\alpha_1 A(t)^{m/2} + V(t+h) - V(t). \end{aligned}$$

If $V(\cdot)$ is a Lévy process then the increments $V(t+h)-V(t)$ are independent and identically distributed. If h is fixed and the process is observed at time points $t_n = t_0 + n\,h$ for $n \in I\!N$, then we obtain an autoregressive process of the form

$$Y_n = Y_{n-1} + \theta\, Y_{n-1}^k + E_n, \tag{3}$$

where $Y_n = A(t_n)$, $k = \frac{m}{2}$ is given, and $\theta = h\alpha_1$ is the unknown parameter. The errors $E_n = V(t_n) - V(t_{n-1}) + h\alpha_0$, $n \in I\!N$, are independent and identically distributed. The speciality of this autoregressive process is that it is nonstationary since $\alpha_1 > 0$ defines a growing process which implies $\theta > 0$. Moreover, we have $Y_0 = a_0 \geq 0$, where a_0 is the nonrandom initial crack length or crack width. Note that $(Y_n)_{n\in I\!N}$ is an AR(1) process if $k = 1$.

Stationary autoregressive and in particular stationary AR(1) processes were extensively discussed in the literature. The case $|1+\theta| < 1$ is analytically well tractable and a wide range of estimation and forecasting methods exist. The first discussion refers to Mann and Wald (1943). The case of explosive processes with $|1+\theta| > 1$ is of lower attention and appears less frequently in applications. Anderson (1959) analysed the asymptotic behaviour of estimates for AR processes in the explosive case. When the errors are assumed to be independent and normally distributed the asymptotic distribution can be derived. In presence of nonnormal errors Anderson pointed out, that asymptotic distributions are hard to derive, since they depend on the assumptions on the error distribution. Basawa et al. (1989) introduced a method to handle nonnormal errors by bootstrapping. Here the errors are assumed to have a zero mean and a finite variance. Another bootstrapping method was proposed by Stute and Gründer (1993). This approach is a nonparametric method to construct prediction intervals for AR(1) processes. The method requires the errors to be defined by a differentiable, bounded, continuous, symmetric and unimodal distribution. Paulaauskas and Rachev (2003) proposed maximum likelihood estimators for AR(1) processes with nonnormal errors. Their approach allows errors with independent and identical symmetric stable distributions where the characteristic exponent is known.
Since the assumed errors in our case are asymmetric and posses heavy tails resulting from jumps in the data, a general idea to derive estimators is to consider a robust approach based on data depth. Data depth for linear regression problems was at first considered by Rousseeuw and Hubert (1999). This approach was extended to polynomial regression by Wellmann et. al (2009), to multiple regression and orthogonal regression by Wellmann and Müller (2010a, b) and for some generalized linear models by Lin and Chen (2006).

Since all of the proposed models are not sufficient to describe the experimental data of crack growth discussed here, in Section 2 we use data depth to introduce a new estimator and a confidence interval for α_1 as well as a test for hypotheses on α_1. Thereby we only assume that the median of the errors E_n is zero. Under this simple assumption, the asymptotic distribution of the test statistic and the consistency of the estimator is derived. Crucial for the consistency is that the process $A(\cdot)$ and thus Y_n is nonnegative. Due to $\theta > 0$

this can not be satisfied by stationary processes. Probably therefore, Huggins (1989) combined the sign test for stationary AR processes with the least squares estimator. In Section 3, the behavior of the new estimator and the new test are studied by simulation. It is shown that the new test is much more robust then the test of Anderson (1959) and has better power than the simple sign test for small parameter values. The application on data in a fatigue experiment of concrete carriers in Section 4 shows that confidence intervals for α_1 are different in an early stage of the fatigue process if different stress levels are used. All proofs are given in the Appendix.

2 Estimators, tests and confidence intervals based on depth

2.1 Depth estimator for AR processes

The idea of data depth is to define a measure which quantifies how deep a parameter lies within the given dataset. Using a quality measure $Q(\theta, y_n)$, Mizera (2002) defines global data depth of a parameter θ within a data set $y_1, \ldots, y_N$ as

$$d_G(\theta, y_1, \ldots, y_N) = \frac{1}{N} \min \{M;\ \exists n_1, \ldots, n_M \text{ and } \theta' \text{ with}$$
$$Q(\theta', y_n) < Q(\theta, y_n) \quad \forall n \in \{1, \ldots, N\} \setminus \{n_1, \ldots, n_M\}\}.$$

Since global data depth is difficult to calculate in most cases, Mizera (2002) also defines tangential data depth as

$$d_T(\theta, y_1, \ldots, y_N) = \frac{1}{N} \min_{0 \neq u \in \mathbb{R}^q} \sharp \left\{ n;\ \frac{\partial}{\partial \theta} Q(\theta, y_n)^\top u \leq 0 \right\}.$$

For the AR process given by (3) a natural quality function is

$$Q(\theta, y_n, y_{n-1}) = (y_n - y_{n-1} - \theta y_{n-1}^k)^2 \tag{4}$$

with derivative

$$\frac{\partial}{\partial \theta} Q(\theta, y_n, y_{n-1}) = -2(y_n - y_{n-1} - \theta y_{n-1}^k) y_{n-1}^k.$$

As soon as $y_n > 0$ is always satisfied, we have

$$(y_n - y_{n-1} - \theta y_{n-1}^k) y_{n-1}^k \lesseqgtr 0 \Longleftrightarrow (y_n - y_{n-1} - \theta y_{n-1}^k) \lesseqgtr 0 \tag{5}$$

so that the tangential data depth for the AR process (3) is

$$\begin{aligned} &d_T^{AR}(\theta, y_0, \ldots, y_N) \\ &= \frac{1}{N} \min \left\{ \sharp \left\{ n;\ y_n - y_{n-1} - \theta y_{n-1}^k \leq 0 \right\}, \right. \\ &\qquad \left. \sharp \left\{ n;\ y_n - y_{n-1} - \theta y_{n-1}^k \geq 0 \right\} \right\}. \end{aligned} \tag{6}$$

A maximum depth estimator is the parameter with maximum depth in the data set $y_0, \ldots, y_N$.

Definition 1 *If $Y_n > 0$ is always satisfied, then the maximum depth estimator for θ of the AR process (3) is*

$$\widehat{\theta}(y_0, \ldots, y_N) = \arg\max_{\theta > 0} d_T^{AR}(\theta, y_0, \ldots, y_N). \tag{7}$$

Since the error E_n is given by $E_n = V(t_n) - V(t_{n-1}) + h\alpha_0$, we have $Y_n \geq Y_{n-1} \geq a_0 \geq 0$ for all $n \in I\!N$ if

$$\begin{aligned} &V(\cdot) \text{ is a process with nonnegative increments,} \\ &Y_0 = A(t_0) = a_0, \\ &-h\alpha_0 \leq \theta a_0^k = h\alpha_1 a_0^k \Longleftrightarrow \alpha_0 \geq -\alpha_1 a_0^k. \end{aligned} \tag{8}$$

Besides the assumption (8), the condition that the median of E_n is zero will be used here. This is in particular satisfied if the initial crack length a_0 is greater than zero so that α_0 in condition (8) can be negative since $\theta = h\alpha_1 > 0$ holds for a growth process. But a median of zero holds also if the distribution of the increments $V(t_n) - V(t_{n-1})$ has zero median which is satisfied for example by a Poisson process with small intensity parameter. Note that the step length h should be small for a good Euler-Maruyama approximation of the process $A(\cdot)$.

Theorem 1 *If zero is the unique median of E_n and assumption (8) holds, then the maximum depth estimator given in (7) is a consistent estimator of θ.*

The proof is given in the Appendix.

2.2 Tests and conficence intervals

The distribution of tangential depth as well as for global depth is difficult to derive. Since the simplicial depth is a U statistic, its asymptotic distribution is in principal known. By extending the first definition by Liu (1988,1990) for multivariate data, simplicial depth can be defined for any depth notion d as follows:

$$d_S(\theta, y_1, \ldots, y_N) = \frac{1}{\binom{N}{q+1}} \sum_{1 \leq n_1 < \ldots < n_{q+1} \leq N} d(\theta, y_{n_1}, \ldots, y_{n_{q+1}})$$

if the parameter θ is q dimensional (see Müller 2005). If the parameter θ is one-dimensional and the depth notion d is the tangential depth d_T given by a

quality function Q it reduces to

$$\begin{aligned} d_S(\theta, y_1, \ldots, y_N) = \frac{1}{\binom{N}{2}} \sharp \Big\{ & (n_1, n_2);\ n_1 < n_2 \text{ with} \\ & \frac{\partial}{\partial\theta} Q(\theta, y_{n_1}) \le 0,\ \frac{\partial}{\partial\theta} Q(\theta, y_{n_2}) \ge 0 \\ & \text{or } \frac{\partial}{\partial\theta} Q(\theta, y_{n_1}) \ge 0,\ \frac{\partial}{\partial\theta} Q(\theta, y_{n_2}) \le 0 \Big\} \\ = \frac{1}{\binom{N}{2}} \sum_{n_1=1}^{N} \sum_{n_2=n_1+1}^{N} & \left(1\left\{ \frac{\partial}{\partial\theta} Q(\theta, y_{n_1}) \ge 0 \right\} \cdot 1\left\{ \frac{\partial}{\partial\theta} Q(\theta, y_{n_2}) \le 0 \right\} \right. \\ & \left. + 1\left\{ \frac{\partial}{\partial\theta} Q(\theta, y_{n_1}) \le 0 \right\} \cdot 1\left\{ \frac{\partial}{\partial\theta} Q(\theta, y_{n_2}) \ge 0 \right\} \right). \end{aligned}$$

This means for the quality function given by (4) using (5) for the AR process given by (3)

$$\begin{aligned} & d_S^{AR}(\theta, y_0, \ldots, y_N) \\ & = \frac{1}{\binom{N}{2}} \sum_{n_1=1}^{N} \sum_{n_2=n_1+1}^{N} \Big(1\{y_{n_1} - y_{n_1-1} - \theta y_{n_1-1}^k \le 0\} \\ & \qquad \cdot 1\{y_{n_2} - y_{n_2-1} - \theta y_{n_2-1}^k \ge 0\} \\ & \qquad + 1\{y_{n_1} - y_{n_1-1} - \theta y_{n_1-1}^k \ge 0\} \cdot 1\{y_{n_2} - y_{n_2-1} - \theta y_{n_2-1}^k \le 0\} \Big). \end{aligned}$$

The following theorem is analogous to the asymptotic distributions of the simplicial depth for regression with independent errors given by Müller (2005), Wellmann et al. (2009), Wellmann and Müller (2010a,b).

Theorem 2 *If the assumption (8) and* $P(E_n \ge 0) = \frac{1}{2} = P(E_n \le 0)$ *hold and* θ_0 *is the true parameter then*

$$N\left(d_S^{AR}(\theta_0, Y_0, \ldots, Y_N) - \frac{1}{2} \right) \overset{N\to\infty}{\longrightarrow} -\frac{1}{2}(X^2 - 1) = \frac{1}{2} - \frac{1}{2}X^2$$

in distribution where X *has a standard normal distribution.*

As pointed out in Müller (2005), very general hypotheses concerning arbitrary subsets Θ_0 of the parameter space can be tested with the simplicial depth.

Corollary 1 *A test which rejects* $H_0 : \theta \in \Theta_0$ *if*

$$\sup_{\theta\in\Theta_0} N\left(d_S^{AR}(\theta, y_0, \ldots, y_N) - \frac{1}{2} \right) < \frac{1}{2} - \frac{1}{2}\chi_1^2(1-\alpha),$$

where $\chi_m^2(\alpha)$ *is the* α *quantile of the* χ^2 *distribution with* m *degrees of freedom, is asymptotically an* α*-level test for* $H_0 : \theta \in \Theta_0$ *against* $H_1 : \theta \notin \Theta_0$.

Applying the general test on $\Theta_0 = \{\theta_0\}$ leads to confidence intervals.

Corollary 2 *$\widehat{\Theta}$ given by*

$$\widehat{\Theta}(y_0, \ldots, y_N) = \left\{ \theta > 0;\ N\left(d_S^{AR}(\theta, y_0, \ldots, y_N) - \frac{1}{2} \right) \geq \frac{1}{2} - \frac{1}{2}\chi_1^2(1-\alpha) \right\}$$

is an asymptotic confidence interval for level $1 - \alpha$.

3 Simulations

3.1 Simulated processes

We consider here only the case $k = 1$, i.e. AR(1) processes defined by

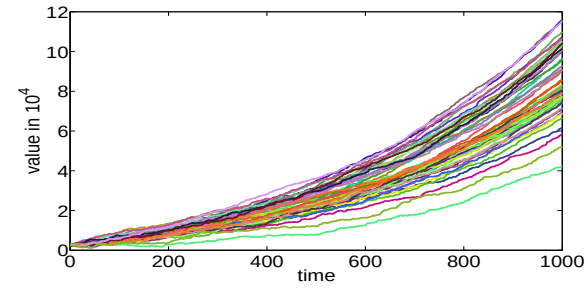
$$Y_n = (1 + \theta)Y_{n-1} + E_{i,n}.$$

The following error distributions are used:

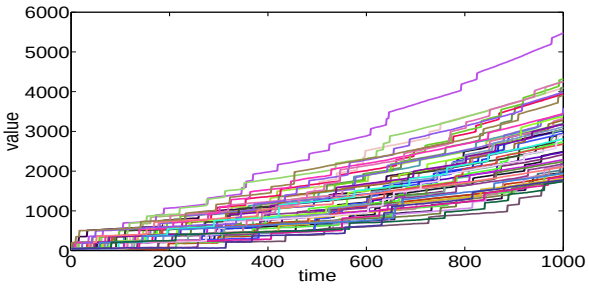
$$\begin{aligned}
&E_{1,n} \sim Gumbel(-36.6513, 100),\\
&E_{2,n} \sim Poisson - Normal - Mixture(\mu = 0, \sigma = 1, \lambda = 2/200),\\
&E_{3,n} \sim Normal(0, 1),\\
&E_{4,n} \sim Fréchet(10, 1.928, -2).
\end{aligned}$$

The choice of the parameters of the acentric Gumbel distribution ensures that the error median is equal to 0 and the expected values are greater than 0. For that we used that the $Gumbel(\alpha, \beta)$ distribution has median $\alpha - \beta ln(ln(2))$ and expected value $\alpha + \beta\xi$, where ξ is the Euler-Mascheroni constant. The same argumentation holds for the Fréchet distribution with median $\xi + \frac{\beta}{ln(2)^{\frac{1}{\alpha}}}$ and expected value $\xi + \beta\Gamma(1 - \frac{1}{\alpha})$. The Poisson-Normal mixture is defined by $X_n + N_n \cdot s$ for $X_n \sim N(\mu, \sigma^2)$, $N_n \sim Pois(\lambda)$ and s a fixed jump size. This model simulates independent and normally distributed errors contaminated by jumps. Due to the small intensity of jumps, the median remains close to 0 in this case. All true parameters θ are set small enough to realize feasible increases for long simulations. Since the results for Fréchet distributed errors were similar to the Gumbel case we do not present the results of the simulation with Fréchet distribution here.

In Figure 1, random paths relating to some processes are depicted. Although theoretically a path of the used processes can be decreasing, this does not really happen if the starting value is large enough as can be seen from Figure 1.



(a) Gumbel errors, $\theta = 0.002$, $a_0 = 2500$



(b) Poisson-normal errors, $\theta = 0.001$, $a_0 = 50$

Fig. 1 Simulated paths of AR(1) processes with predefined error distributions

3.2 Test for explosive AR processes

The proposed test can be addressed to test for explosive AR(1) processes. For this purpose we simulate AR(1) processes with fixed error distributions and fixed growth parameters. The processes are then simulated $M = 1000$ times with different series length $N_k \in \{100, 200, ..., 1900, 2000\}$. Then we apply the derived test of $H_0 : \theta = 0$ with level 5 %. Even if convergency was only proven for parameters larger than zero, we evaluate the boundary case here because simulations indicate that the level is still kept. Figure 2 shows that the relative number of simulations for which the parameter 0 is rejected increases by an increasing series length. Hence the test offers the ability to test for explosive processes by rejecting a parameter of 0 for large sample sizes. The rate of rejection depends on the error distribution and the distance of the true parameter to the boundary value 0. In our cases the Gumbel distribution implies the slowest convergency of the rejection rate to 1. The Normal-Poisson distribution implies the fastest convergency. In both cases we can conclude that the test is consistent for $\theta > 0$ in the sense that the power of the test converges to 1 for growing sample size.

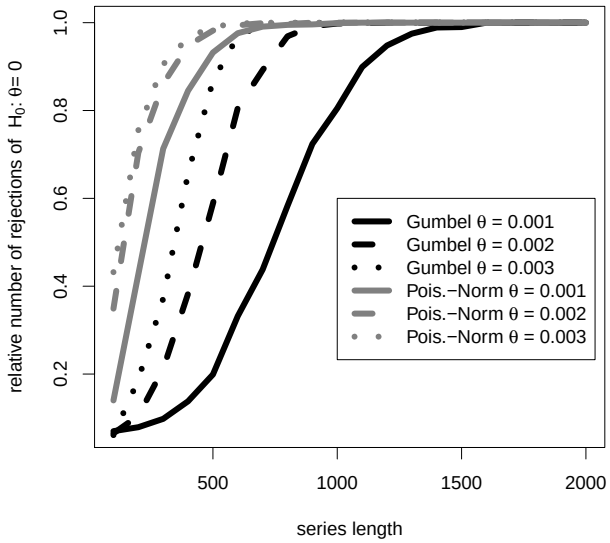


Fig. 2 Rejections of $H_0 : \theta = 0$

3.3 Power functions

To illustrate the power of the proposed test a comparison with alternative methods is used. For normally distributed errors and $H_0 : \theta = \theta_0$ Anderson (1959) proposed the following test statistic in the explosive case (i.e. $\theta_0 > 0$),

$$T_N(\theta_0) = (\hat{\theta} - \theta_0) \sqrt{\sum_{n=1}^{N} y_{n-1}^2},$$

where θ is estimated by

$$\hat{\theta} = \frac{\sum_{n=1}^{N} y_n y_{n-1}}{\sum_{n=1}^{N} y_{n-1}^2} - 1.$$

This test statistic has an approximate normal distribution with mean 0 and variance σ^2 under H_0.

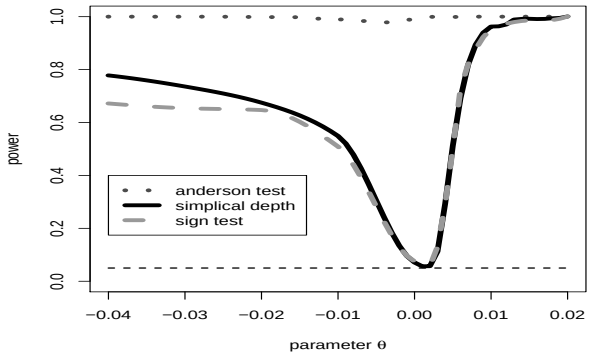
A simple alternative is the sign test studied by Huggins (1989) for AR processes. Here it has the form

$$Q_N(\theta_0) = \sum_{n=1}^{N} sign(Y_n - (1 + \theta_0)Y_{n-1}) = 2(B_N - \frac{N}{2}),$$

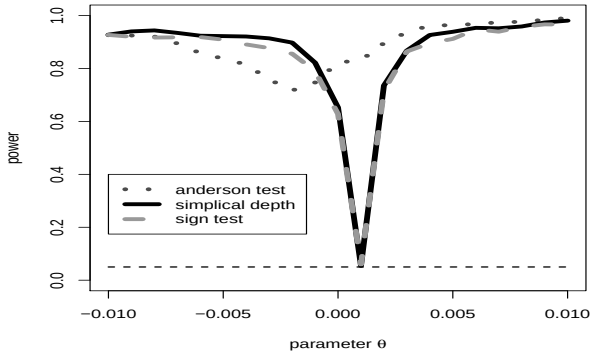
where B_N has a binomial distribution with parameters N and $\frac{1}{2}$ under $H_0 : \theta = \theta_0$. Hence the sign test is an exact test.

In Figure 3 the comparison for a series length of 300 based on 1000 simulated paths is presented. For all simulations with not normally distributed errors, the starting value is set to $a_0 = 1.5$. For the simulation with independent and identical normally distributed errors, a starting value of $a_0 = 100$ was selected. The reason is that in this case the simulated series are not necessarily increasing due to negative errors and the estimates and tests therefore do not reject parameters lower than 0 if the series cross the zero line. With a high starting value this effect can be neglected. The following hypotheses were used: $H_0 : \theta = 0.002$ for the Gumbel errors and $H_0 : \theta = 0.001$ for the Poisson-normal errors and the normal errors. For all simulations, a level of 5% was selected.

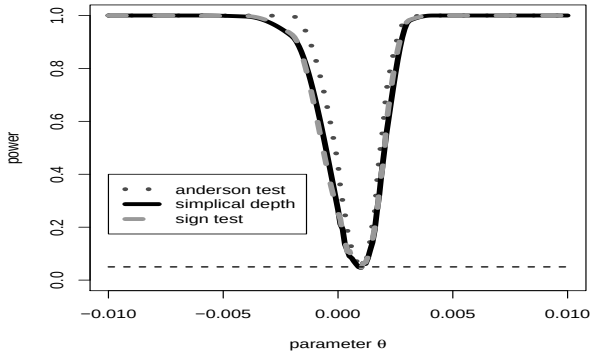
The power clearly depends on the error distribution. With normally distributed errors, the test from Anderson performs slightly better than the other two tests (see Figure 3(c)). This is not surprising because the known variance is used. When the error distribution is not normal, this test fails completely, while this is not the case for the simplical depth test and the sign test. The simplical depth behaves similar to the sign test for normal errors (see Figure 3(c)), slightly better for Poisson-Normal errors (see Figure 3(b)) and quite better than the sign test for $\theta < 0$ with Gumbel errors (see Figure 3(a)). With the selected starting values, the result for Gumbel errors is surprising because the condition of strictly positive values does not hold for all simulated paths in this case.



(a) Gumbel errors



(b) Poisson-normal errors

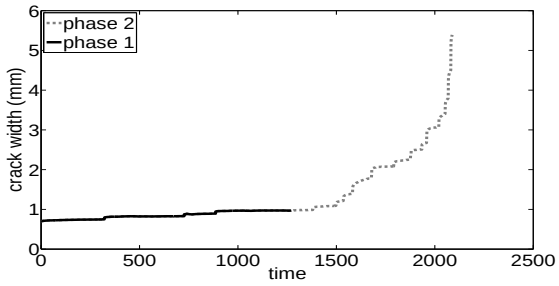


(c) Normal errors

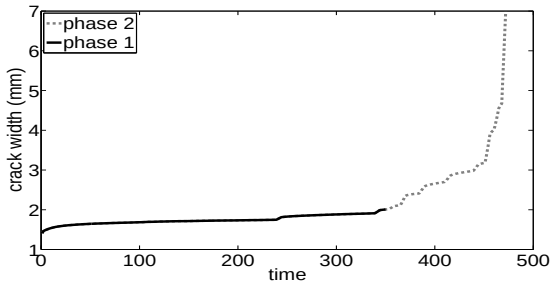
Fig. 3 Power function for sample size 300

4 Application on crack propagation data

Now the estimator is applied to data produced in fatigue experiments conducted by Maurer and Heeke (2010). We consider the experiments TR01 and TR02. Two identical concrete carriers were used for the experiments. Under cyclical loading the crack width of a initial crack was observed. Figure 4 shows the observed crack width in both experiments. Experiment TR01 differs from experiment TR02 due to a higher maximum load applied in TR02. The depth



(a) TR01



(b) TR02

Fig. 4 Experimental crack width data

estimation approach can be used to estimate a stable growth parameter of the crack width series without influence from jumps in the data. Furthermore confidence intervals for the parameter can be calculated.
The model for the crack width Y_n, observed in discrete and equidistant time $n \in I\!N$, is assumed to be

$$Y_n = (1 + \theta)Y_{n-1} + E_n.$$

The errors E_n are assumed to hold the conditions in Sections 1 and 2. Applied to TR01 this assumption is not problematic. The residuals can be assumed to be almost independent with respect to the autocorrelation function. However the assumption of independent residuals can not be assured for TR02. This also

holds if a robust version of the correlation coefficient is used. The robust autocorrelation function surprisingly shows even more correlation. Nevertheless,

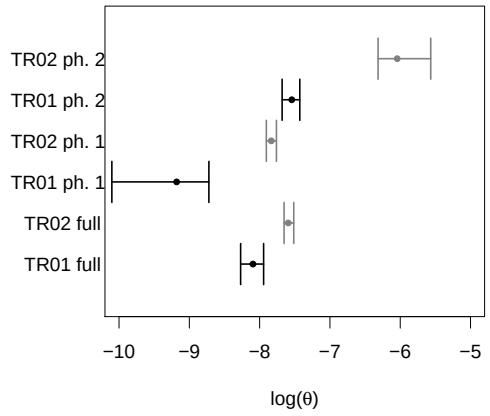


Fig. 5 Comparison of estimated parameters (•) and confidence intervals on a logarithmic scale

the estimation reveals important properties of the series. We consider estimates for the complete series and estimates for phases which visually imply a different crack width progress, indicated in Figure 4 by the solid and dotted lines. In Figure 5, the results of point estimation and confidence interval estimation are presented.

From the confidence intervals, it is possible to conclude that the different experimental conditions result in significantly differing growth parameters. The growth parameter is bigger when a higher load is applied. Furthermore it is possible to identify different parameter values among the phases with a low and high jump frequency. The parameters significantly increase in the phase of high jump activity.

5 Appendix

Proof of Theorem 1
Let be $\epsilon > 0$ arbitrary and let be θ_0 the underlying parameter. Let be B the set of all ω so that the strong law of large numbers holds for
$\frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \geq 0\}$, $\frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \leq 0\}$, $\frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \geq a_0^k\epsilon\}$, and $\frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \leq -a_0^k\epsilon\}$. Then it holds $P(B) = 1$. Now consider an arbitrary $\omega \in B$.

We have to show that there exists $N_* \in I\!N$ such that
$|\widehat{\theta}(Y_0(\omega), \ldots, Y_N(\omega)) - \theta_0| < \epsilon$ for all $N \geq N_*$.
Since $\text{med}(E_1) = 0$ is unique, there exists $\delta > 0$ such that

$$P(E_1 \leq -a_0^k \epsilon) < \frac{1}{2} - 2\,\delta, \quad P(E_1 \geq a_0^k \epsilon) < \frac{1}{2} - 2\,\delta.$$

Because of the strong law of large numbers there exists $N_1 \in I\!N$ such that

$$\begin{aligned}
&\frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \geq 0\} > P(E_1 \geq 0) - \delta \geq \frac{1}{2} - \delta,\\
&\frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \leq 0\} > P(E_1 \geq 0) - \delta \geq \frac{1}{2} - \delta,\\
&\frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \geq a_0^k \epsilon\} < P(E_1 \geq a_0^k \epsilon) + \delta < \frac{1}{2} - \delta,\\
&\frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \leq -a_0^k \epsilon\} < P(E_1 \leq -a_0^k \epsilon) + \delta < \frac{1}{2} - \delta
\end{aligned}$$

for all $N \geq N_1$. Then we have for all $N \geq N_1$

$$\frac{1}{N}\sum_{n=1}^{N} 1\{Y_n(\omega) - Y_{n-1}(\omega) - \theta_0 Y_{n-1}^k \geq 0\} = \frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \geq 0\} > \frac{1}{2} - \delta$$

and

$$\frac{1}{N}\sum_{n=1}^{N} 1\{Y_n(\omega) - Y_{n-1}(\omega) - \theta_0 Y_{n-1}^k \leq 0\} = \frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \leq 0\} > \frac{1}{2} - \delta$$

so that

$$d_T^{AR}(\theta_0, Y_0(\omega), \ldots, Y_N(\omega)) > \frac{1}{2} - \delta \text{ for all } N \geq N_1. \tag{9}$$

Now consider $\theta \geq \theta_0 + \epsilon$. Then we have for all $N \geq N_1$ with $Y_n(\omega) \geq a_0$

$$\begin{aligned}
&\frac{1}{N}\sum_{n=1}^{N} 1\{Y_n(\omega) - Y_{n-1}(\omega) - \theta\, Y_{n-1}^k \geq 0\}\\
&= \frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) + (\theta_0 - \theta)\, Y_{n-1}^k \geq 0\} = \frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \geq (\theta - \theta_0)\, Y_{n-1}^k\}\\
&\leq \frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \geq (\theta - \theta_0)\, a_0^k\} \leq \frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \geq a_0^k\, \epsilon\} < \frac{1}{2} - \delta.
\end{aligned}$$

For $\theta \le \theta_0 - \epsilon$ holds analogously for all $N \ge N_1$

$$\frac{1}{N}\sum_{n=1}^{N} 1\{Y_n(\omega) - Y_{n-1}(\omega) - \theta\, Y_{n-1}^k \le 0\} = \frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \le (\theta - \theta_0)\, Y_{n-1}^k\}$$
$$\le \frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \le (\theta - \theta_0)\, a_0^k\} \le \frac{1}{N}\sum_{n=1}^{N} 1\{E_n(\omega) \le -a_0^k\, \epsilon\} < \frac{1}{2} - \delta.$$

Hence for θ with $|\theta - \theta_0| \ge \epsilon$ we have

$$d_T^{AR}(\theta, Y_0(\omega), \ldots, Y_N(\omega)) < \frac{1}{2} - \delta \text{ for all } N \ge N_1.$$

This means with (9) that only a θ with $|\theta - \theta_0| < \epsilon$ can maximize $d_T^{AR}(\cdot, Y_0(\omega), \ldots, Y_N(\omega))$, i.e. $|\widehat{\theta}(Y_0(\omega), \ldots, Y_N(\omega)) - \theta_0| < \epsilon$ for all $N \ge N_1$. □

Proof of Theorem 2.
If θ_0 is the true parameter then

$$Y_n - Y_{n-1} - \theta_0 Y_{n-1}^k = E_n$$

for all n and the simplicial depth d_S reduces to

$$d_S^{AR}(\theta_0, Y_0, \ldots, Y_N)$$
$$= \frac{1}{\binom{N}{2}} \sum_{n_1=1}^{N} \sum_{n_2=n_1+1}^{N} \left(1\{E_{n_1} \le 0\} \cdot 1\{E_{n_2} \ge 0\} + 1\{E_{n_1} \ge 0\} \cdot 1\{E_{n_2} \le 0\}\right).$$

It is U statistic with the kernel function

$$\psi(e_1, e_2) = 1\{e_1 \le 0\} \cdot 1\{e_2 \ge 0\} + 1\{e_1 \ge 0\} \cdot 1\{e_2 \le 0\}.$$

$P(E_n \ge 0) = \frac{1}{2} = P(E_n \le 0)$ implies $P(E_n = 0) = 0$. With the independence of $E_1, E_2 \ldots$, we obtain

$$\begin{aligned} &\mathrm{E}(\psi(E_1, E_2)|E_1 = e_1) \\ &= 1\{e_1 \le 0\} \cdot P(E_2 \ge 0) + 1\{e_1 \ge 0\} \cdot P(E_2 \le 0) \\ &= 1\{e_1 \le 0\} \cdot \frac{1}{2} + 1\{e_1 \ge 0\} \cdot \frac{1}{2} = \frac{1}{2} \end{aligned}$$

almost surely. Hence the simplicial depth d_S^{AR} is a degenerated U statistic. The spectral decomposition of $\psi(e_1, e_2) - \frac{1}{2}$ is (see also Müller 2005)

$$\psi(e_1, e_2) - \frac{1}{2} = -\frac{1}{2}\, \varphi(e_1)\, \varphi(e_2)$$

with $\varphi(e) = 1\{e \le 0\} - 1\{e \ge 0\}$. The theorem of Hoeffding (see Witting/Müller-Funk 1995, Satz 7.183, p. 650 as well as p. 155) provides the assertion. □

Proof of Corollary 1.
For $\theta_0 \in \Theta_0$ we have

$$\begin{aligned}
&\lim_{N\to\infty} P_{\theta_0}\left(\sup_{\theta\in\Theta_0} N\left(d_S^{AR}(\theta, Y_0, \ldots, Y_N) - \frac{1}{2}\right) < \frac{1}{2} - \frac{1}{2}\chi_1^2(1-\alpha)\right) \\
&\leq \lim_{N\to\infty} P_{\theta_0}\left(N\left(d_S^{AR}(\theta_0, Y_0, \ldots, Y_N) - \frac{1}{2}\right) < \frac{1}{2} - \frac{1}{2}\chi_1^2(1-\alpha)\right) \\
&= P\left(-\frac{1}{2}(X^2-1) < \frac{1}{2} - \frac{1}{2}\chi_1^2(1-\alpha)\right) = P(-X^2+1 < 1-\chi_1^2(1-\alpha)) \\
&= P(-X^2 < -\chi_1^2(1-\alpha)) = P(X^2 > \chi_1^2(1-\alpha)) = \alpha. \ \square
\end{aligned}$$

References

1. Anderson, T. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations. *Ann. Math. Stat.* **30**, 676-687.
2. Basawa, I., Mallik, A., McCormick, W., Taylor, R. (1989). Bootstrapping explosive autoregressive processes. *Ann. Stat.* **17**, 1479-1486.
3. Huggins, R. (1989). The sign test for stochastic processes. *Aust. J. Stat.* **31**, 153-165.
4. Iacus, S.M. (2008). *Simulation and Inference for Stochastic Differential Equations. With R Examples.* Springer, New York.
5. Lin, L. and Chen, M. (2006). Robust estimating equation based on statistical depth. *Stat. Pap.* **47**, 263-278.
6. Liu, R.Y. (1988). On a notion of simplicial depth. *Proc. Nat. Acad. Sci. USA* **85**, 1732-1734.
7. Liu, R.Y. (1990). On a notion of data depth based on random simplices. *Ann. Stat.* **18**, 405-414.
8. Mann, H. and Wald, A. (1943). On the statistical treatment of linear stochastic difference equations. *Econometrica* **11**, 173-220.
9. Maurer, R. and Heeke, G. (2010) Ermüdungsfestigkeit von Spannbeton aus einer älteren Spannbetonbrücke. TU Dortmund, University. Technical report.
10. Mizera, I. (2002). On depth and deep points: a calculus. *Ann. Stat.* **30**, 1681-1736.
11. Müller, Ch.H. (2005). Depth estimators and tests based on the likelihood principle with application to regression. *J. of Multivar. Anal.* **95**, 153-181.
12. Paulaauskas, V. and Rachev, T. (2003). Maximum likelihood estimators in regression models with infinite variance innovations. *Stat. Pap.* **44**, 47-65.
13. Pook, L. (2000). *Linear Elastic Fracture Mechanics for Engineers: Theory and Application.* WIT Press, Southhampton.
14. Rousseeuw, P.J. and Hubert, M. (1999). Regression depth. *J. of Amer. Stat. Assoc.* **94**, 388-402.
15. Stute, W. and Gründer, B. (1993). Nonparametric prediction intervals for explosive AR(1)-processes. *Nonparametr. Stat.* **2**, 155-167.
16. Witting, H. and Müller-Funk, U. (1995). *Mathematische Statistik II.* Teubner, Stuttgart.
17. Wellmann, R., Harmand, P., Müller, Ch.H. (2009). Distribution-free tests for polynomial regression based on simplicial depth. *J. of Multivar. Anal.* **100**, 622-635.
18. Wellmann, R. and Müller, Ch.H. (2010a). Tests for multiple regression based on simplicial depth. *J. of Multivar. Anal.* **101**, 824-838.
19. Wellmann, R. and Müller, Ch.H. (2010b). Depth notions for orthogonal regression. *J. of Multivar. Anal.* **101**, 2358-2371.