*Journal of the Korean Statistical Society (1999)*, **28**: *2, pp 235–251*

# The Cluster Damage in a $k$th-Order Stationary Markov Chain[†]

**Seokhoon Yun**[1]

Abstract

In this paper we examine extremal behavior of a $k$th-order stationary Markov chain $\{X_n\}$ by considering excesses over a high level which typically appear in clusters. Excesses over a high level within a cluster define a cluster damage, i.e., a normalized sum of all excesses within a cluster, and all excesses define a damage point process. Under some distributional assumptions for $\{X_n\}$, we prove convergence in distribution of the cluster damage and obtain a representation for the limiting cluster damage distribution which is well suited for simulation. We also derive formulas for the mean and the variance of the limiting cluster damage distribution. These results guarantee a compound Poisson limit for the damage point process, provided that it is strongly mixing.

*Keywords:* Extreme values; Excesses; Cluster damage distributions; Damage point processes; $k$th-order stationary Markov chains.

## 1. INTRODUCTION

It is often the case that we are concerned about exceedance data above a certain level when we handle environmental time series (cf. Smith (1989) and Yun (1996)). For example, to assess the degree of air pollution in a certain area using collected ground-level ozone one of our major concerns is to model how often and how large the ozone concentrations above a high level occur. This kind of problem is deeply related to extreme value analysis of stationary time series.

Extreme value theory for stationary sequences has been well established during last few decades (see, e.g., Leadbetter (1974), Leadbetter, Lindgren and Rootzén (1983), O'Brien (1987) and Hsing, Hüsler and Leadbetter (1988)). Exceedances above a high level in a weakly dependent stationary sequence typically

[†]The author wishes to acknowledge the financial support of the Korea Research Foundation made in the program year of 1997: Grant No. 1997-003-D00044.

[1]Department of Applied Statistics, University of Suwon, Suwon, 445-743, Korea.

appear in approximately independent clusters due to its local dependence, and the average cluster size, i.e., the average number of exceedances within a cluster, can be approximated by the reciprocal of the so-called extremal index (cf. O'Brien (1987), Leadbetter and Rootzén (1988) and Hsing (1993)).

Another important measure in applications is the aggregate excess over a high level, i.e,. the cumulative total of all excesses over a high level (cf. Anderson and Dancy (1992) and Leadbetter (1995)). In the hydrological context, this is a measure of the volume of overflow. Excesses over a high level in a weakly dependent stationary sequence also typically form approximately independent clusters like exceedances.

To give a reasonable basis for modeling the exceedances as well as the excesses of weakly dependent stationary data, it is therefore important to calculate both the (limiting) cluster size distribution (or the extremal index) and the (limiting) cluster aggregate excess distribution when the dependence structure of the stationary sequence is given. Generally, the calculating procedure is intractable to go through with.

However, it turns out that the calculating task works well in stationary Markov chains. Extreme value theory for stationary Markov chains has been actively studied since O'Brien (1987) and Rootzén (1988). In particular, Smith (1992) and Perfekt (1994) obtained representations for the extremal index of a stationary Markov chain which are well suited for Monte Carlo simulation. This was extended to higher-order stationary Markov chains by Perfekt (1997) and Yun (1998). Let $\{X_n\}_{n=1}^{\infty}$ be a real-valued, $k$th-order ($k \geq 1$) stationary Markov chain. Using limiting distributions of some rescaled stationary transition kernels for $\{X_n\}$, Yun (1998) introduced a new $(k-1)$th-order Markov chain $\{Y_n\}_{n=1}^{\infty}$ and defined a $k$th-order Markov chain $\{Z_n\}_{n=1}^{\infty}$ by $Z_n = Y_1 + \cdots + Y_n$. The chain $\{Z_n\}$ was then effectively used to derive a representation for the extremal index of $\{X_n\}$.

In this paper we prove convergence in distribution of a cluster damage, i.e., a normalized cluster aggregate excess over a high level, in the $k$th-order stationary Markov chain $\{X_n\}$ and obtain a representation for the limiting cluster damage distribution using the chain $\{Z_n\}$. This enables us to calculate the (limiting) cluster aggregate excess distribution by simulation. From the representation, we also derive formulas for the mean and the variance of the limiting cluster damage distribution. This result together with the existing extreme value theory for stationary sequences guarantees a compound Poisson limit for the damage point process in $\{X_n\}$, provided that it is strongly mixing (see Section 2 for details).

In Section 2 we discuss the compound Poisson limits for the damage point process and the exceedance point process under a general stationary sequence. In Section 3 we present the basic distributional assumptions for the Markov chain $\{X_n\}$ and prove the weak convergence of the cluster damage distribution in the chain $\{X_n\}$. By $\xrightarrow{w}$ we denote weak convergence. Also we write $x^+ := \max\{x, 0\}$ and $x^- := -\min\{x, 0\}$.

## 2. DAMAGE POINT PROCESS AND EXCEEDANCE MODELING IN A GENERAL STATIONARY SEQUENCE

Let $\{X_n\}_{n=1}^{\infty}$ be a strictly stationary sequence of random variables with marginal distribution function $F$. Let $\{u_n\}_{n=1}^{\infty}$ be a sequence of constants ("levels" or "thresholds") such that $n(1 - F(u_n)) \to \lambda$ as $n \to \infty$ for some $\lambda > 0$. In particular, if $F$ is continuous, such a sequence $\{u_n\}$ always exists (cf. Theorem 1.7.13 of Leadbetter, Lindgren and Rootzén (1983)). If an event $\{X_i > u_n\}$ ("exceedance") occurs, $X_i - u_n$ is referred to as an "excess". We consider a point process $N_n$ on $(0, 1]$ defined by

$$N_n(B) := \sum_{i=1}^{n} \delta_{i/n}(B) a_n (X_i - u_n)^+, \; B\text{: Borel set in } (0,1], \tag{2.1}$$

for suitably chosen constants $a_n > 0$, where $\delta_{i/n}(\cdot)$ denotes the Dirac measure with mass 1 at $i/n$. The $N_n$ may be regarded as the exceedance marked point process with "points" at the normalized exceedances $\{i/n : X_i > u_n\}$ and "marks" given by the "damages" $a_n(X_i - u_n)$, i.e., normalized excesses. Therefore $N_n(0, 1]$ is the total accumulated damage from $X_1, ..., X_n$.

In this section it is assumed that $\{N_n\}_{n=1}^{\infty}$ is strongly mixing in the following sense (cf. Leadbetter (1995)). For $0 \le s < t \le 1$, define $\mathcal{B}_{s,t}^{(n)}$ to be the $\sigma$-field generated by the random variables $N_n(a, b]$, $s \le a < b \le t$. Also, for $0 \le l < 1$, write

$$\alpha_{n,l} = \sup\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{B}_{0,s}^{(n)}, B \in \mathcal{B}_{s+l,1}^{(n)}, 0 < s < 1 - l\}.$$

Then $\{N_n\}$ is called strongly mixing if $\alpha_{n,l_n} \to 0$ as $n \to \infty$ for some sequence $l_n$ with $l_n \to 0$. In particular, if $X_1, X_2, ...$ are independent, we get $\alpha_{n,l_n} = l_n = 0$. The strong mixing for $\{N_n\}$ is much weaker than the usual strong mixing for the original sequence $\{X_n\}$. It is also noted that the strong mixing for $\{N_n\}$ implies the $\Delta(u_n)$ condition for $\{X_n\}$ in Hsing, Hüsler and Leadbetter (1988).

Let $\{k_n\}_{n=1}^{\infty}$ be a sequence of positive integers ("standard sequence") such that

$$k_n = o(n),\ k_n \to \infty \text{ and } k_n(\alpha_{n,l_n} + l_n) \to 0 \text{ as } n \to \infty.$$

The less the long range dependence, the faster the possible rate for $k_n$. Finally, put $r_n = [n/k_n]$, the integer part of $n/k_n$, and divide the integer set $\{1, 2, ..., k_n r_n\}$ ($k_n r_n \sim n$) into $k_n$ consecutive groups ("blocks") $\{(i-1)r_n + 1, (i-1)r_n + 2, ..., ir_n\}$, $i = 1, ..., k_n$, of length $r_n$. Then the exceedances (if any) in such a block are said to form a "cluster". In other words, any two exceedances belonging to the same block are considered to stem from the same cluster. Further if we write $J_i := ((i-1)r_n/n, ir_n/n]$, $i = 1, ..., k_n$, then $\{J_1, ..., J_{k_n}, (k_n r_n/n, 1]\}$ forms a partition of $(0, 1]$. The "cluster damage distribution" $\pi_n$ is now defined by the distribution function

$$\begin{aligned} \pi_n(x) &:= P\{N_n(J_1) \le x | N_n(J_1) > 0\} \\ &= P\{\sum_{i=1}^{r_n} a_n(X_i - u_n)^+ \le x | M_{1,r_n} > u_n\},\ x > 0, \end{aligned}$$

where $M_{i,j} := \max\{X_i, ..., X_j\}$, $1 \le i \le j$.

The following result is a variant of Theorem 4.2 of Leadbetter (1995).

**Theorem 2.1.** *Let $\{X_n\}_{n=1}^{\infty}$ be a strictly stationary sequence of random variables with marginal distribution function $F$ and $\{u_n\}_{n=1}^{\infty}$ a sequence of levels such that $n(1 - F(u_n)) \to \lambda$ as $n \to \infty$ for some $\lambda > 0$. Suppose that*

*(i) $\{N_n\}$ is strongly mixing,*

*(ii) $P\{M_{2,r_n} \le u_n | X_1 > u_n\} \to \theta$ as $n \to \infty$ for some $\theta \in (0, 1]$,*

*(iii) the cluster damage distribution $\pi_n$ converges weakly to some probability distribution $\pi$ on $(0, \infty)$ as $n \to \infty$.*

*Then the damage point process $N_n$ converges in distribution as $n \to \infty$ to a compound Poisson process CP($\theta\lambda, \pi$) with intensity $\theta\lambda$ and multiplicity distribution $\pi$.*

**Proof:** We sketch the proof. Since $\{k_n\}$ is a standard sequence for $N_n$, the sequence $\{r_n\}$ must satisfy

$$r_n \to \infty,\ r_n = o(n) \text{ and } n(\alpha_{n,l_n} + l_n) = o(r_n) \text{ as } n \to \infty.$$

Also, since $n(1 - F(u_n)) \to \lambda$ if and only if $F^n(u_n) \to e^{-\lambda}$, applying Theorem 2.1 of O'Brien (1987) yields that

$$P\{M_{1,n} \le u_n\} - (F(u_n))^{nP\{M_{2,r_n} \le u_n | X_1 > u_n\}} \to 0 \text{ as } n \to \infty.$$

Thus it follows from condition (ii) that

$$P\{M_{1,n} \le u_n\} \to e^{-\theta\lambda} \text{ as } n \to \infty.$$

Now the assertion of the theorem can be proved by the methods similar to Theorem 4.2 of Hsing, Hüsler and Leadbetter (1988). □

**Remark 2.1.** 1. The parameter $\theta$, which has important uses in extreme value theory, is called the "extremal index" of the sequence $\{X_n\}$, i.e.,

$$P\{M_{1,n} \le u_n\} \approx F^{n\theta}(u_n) \text{ as } n \to \infty.$$

The reciprocal of $\theta$ typically agrees with the asymptotic mean cluster size, i.e.,

$$\theta^{-1} = \lim_{n\to\infty} \sum_{j=1}^{\infty} jP\{\sum_{i=1}^{r_n} I\{X_i > u_n\} = j | M_{1,r_n} > u_n\}.$$

2. The theorem indicates that the damages over a high level can be well modeled by (limiting) clusters with independent damage distributions $\pi$, located at the points of a Poisson process with mean $\theta\lambda$ per (normalized) unit time.

In the theorem if we replace $N_n$ by the "exceedance point process"

$$N_n(B) = \sum_{i=1}^{n} \delta_{i/n}(B) I\{X_i > u_n\}, \ B\text{: Borel set in } (0,1],$$

and $\pi_n$ by the "cluster size distribution"

$$\begin{aligned} \pi_n(j) &= P\{N_n(J_1) = j | N_n(J_1) > 0\} \\ &= P\{\sum_{i=1}^{r_n} I\{X_i > u_n\} = j | M_{1,r_n} > u_n\}, \ j = 1, 2, ..., \end{aligned}$$

then we obtain the result that the exceedance point process $N_n$ converges in distribution as $n \to \infty$ to a compound Poisson process CP$(\theta\lambda, \pi)$ with intensity $\theta\lambda$ and multiplicity distribution $\pi$, provided that $\pi_n$ converges weakly to a probability distribution $\pi$ on $\{1, 2, ...\}$. This is a variant of Theorem 4.2 of Hsing, Hüsler and Leadbetter (1988). Thus the exceedances above a high level can be modeled by (limiting) clusters with independent size distributions $\pi$, located at the points of a Poisson process with mean $\theta\lambda$ per (normalized) unit time. This result together with Theorem 2.1 provides statisticians with a full basis for modeling both the exceedances and the excesses over a high level of weakly dependent stationary data.

## 3. CLUSTER DAMAGE DISTRIBUTION IN A MARKOV CHAIN

To apply the compound Poisson limit for the damage point process $N_n$ in practice, we have to prove the underlying conditions of Theorem 2.1 and need to compute the limiting cluster damage distribution $\pi$ as well as the extremal index $\theta$. Though this procedure is potentially very complicated in general, we may have nice and simple representations for $\pi$ and $\theta$ which are well suited for simulation if the original stationary sequence $\{X_n\}$ is allowed to have a Markovian structure.

Let $\{X_n\}_{n=1}^{\infty}$ be a real-valued, $k$th-order ($k \geq 1$) stationary Markov chain. Then the distribution of the whole chain is determined by the joint distribution of any $k+1$ consecutive variables, which is assumed absolutely continuous throughout. Let $F$ and $f$ denote the distribution function and the probability density function of $X_n$, respectively, and for each $j = 1, ..., k$, let $f_{j+1}(\cdot\,|x_1, ..., x_j)$ denote the conditional probability density function of $X_{n+j+1}$ given that $(X_{n+1}, ..., X_{n+j}) = (x_1, ..., x_j)$.

Assume that the well-known von Mises condition holds for $F$ (cf. Resnick (1987)), i.e., there exists a constant $\xi \in \Re$ such that

$$\lim_{u \uparrow x_F} \frac{g(u)f(u)}{1 - F(u)} = 1, \tag{3.1}$$

where $x_F := \sup\{x : F(x) < 1\}$, the right endpoint of $F$, and the function $g$ satisfy

$$\begin{array}{ll} x_F = \infty \text{ and } g(u) = \xi u, & \text{if } \xi > 0, \\ g \text{ is some strictly positive function,} & \text{if } \xi = 0, \\ x_F < \infty \text{ and } g(u) = -\xi(x_F - u), & \text{if } \xi < 0. \end{array}$$

This is a sufficient condition for $F$ to belong to the domain of attraction of the extreme value distribution $\Omega_\xi(x) := \exp(-(1 + \xi x)^{-1/\xi})$, $1 + \xi x > 0$, and is satisfied by a very large class of absolutely continuous distribution functions $F$.

Throughout the case $\xi = 0$ is interpreted as the limit $\xi \to 0$ so that $\Omega_0(x) = \exp(-e^{-x})$, $x \in \Re$. When $\xi = 0$, one appropriate choice of the function $g$ is $g(u) = \int_u^{x_F}(1-F(t))\,dt/(1-F(u))$. Under condition (3.1), it holds that (cf. Yun (1997))

$$\lim_{u \uparrow x_F} \frac{g(u)f(u+g(u)x)}{1-F(u)} = (1+\xi x)^{-1/\xi-1}$$

locally uniformly in $x$ with $1+\xi x > 0$.

For the conditional probability density functions, we assume that, for each $j = 1, ..., k$,

$$\begin{aligned} & g(u) f_{j+1}(u+g(u)x_{j+1}|u+g(u)x_1,...,u+g(u)x_j) \\ & \to \frac{1}{1+\xi x_{j+1}} h_j\left(\frac{1}{\xi}\log\left(\frac{1+\xi x_{j+1}}{1+\xi x_j}\right); \frac{1}{\xi}\log\left(\frac{1+\xi x_2}{1+\xi x_1}\right), ..., \frac{1}{\xi}\log\left(\frac{1+\xi x_j}{1+\xi x_{j-1}}\right)\right), \\ & \quad 1+\xi x_1 > 0, ..., 1+\xi x_{j+1} > 0, \text{ as } u \uparrow x_F, \end{aligned} \tag{3.2}$$

for some function $h_j : \Re^j \to [0, \infty)$ with $\int_{-\infty}^{\infty} h_j(t; y_1, ..., y_{j-1})\,dt \le 1$, $y_1, ...,$ $y_{j-1} \in \Re$, which is typically the case if the joint distribution function of $(X_{n+1},$ $... ,X_{n+k+1})$ belongs to the domain of attraction of a $(k+1)$-dimensional extreme value distribution with equal univariate marginals $\Omega_\xi$ (cf. Yun (1998)). For each $j = 1, ..., k$ and for $y_1, ..., y_{j-1} \in \Re$, define

$$H_j(y; y_1, ..., y_{j-1}) := 1 - \int_y^{\infty} h_j(t; y_1, ..., y_{j-1})\,dt, \; y \in \{-\infty\} \cup \Re,$$

which may then be considered as a distribution function (of $y$) on $\{-\infty\} \cup \Re$ with possibly positive mass $1 - \int_{-\infty}^{\infty} h_j(t; y_1, ..., y_{j-1})\,dt$ at $y = -\infty$. The limiting distributions $H_j$, $j = 1, ..., k$, are used to define a $\{-\infty\} \cup \Re$-valued, $(k-1)$th-order Markov chain $\{Y_n\}_{n=1}^{\infty}$ as follows:

1. $Y_1 \sim H_1(\cdot)$.

2. For $j = 2, ..., k$, $Y_j|(Y_1, ..., Y_{j-1}) \sim H_j(\cdot; Y_1, ..., Y_{j-1})$ if $Y_1, ..., Y_{j-1} > -\infty$; put $Y_j = -\infty$, otherwise.

3. For $j \ge k+1$, $Y_j|(Y_{j-k+1}, ..., Y_{j-1}) \sim H_k(\cdot; Y_{j-k+1}, ..., Y_{j-1})$ if $Y_{j-k+1}, ...,$ $Y_{j-1} > -\infty$; put $Y_j = -\infty$, otherwise.

Finally, define

$$Z_n := \sum_{i=1}^{n} Y_i, \; n = 1, 2, ..., \; (Z_0 \equiv 0) \tag{3.3}$$

so that $\{Z_n\}_{n=1}^{\infty}$ is a $\{-\infty\} \cup \Re$-valued, $k$th-order Markov chain. The state $-\infty$ is an absorbing state of the chain $\{Z_n\}$ if at least one of $H_j(y; y_1, ..., y_{j-1})$, $j = 1, ..., k$, has a positive mass at $y = -\infty$. When $k = 1$, the chain $\{Z_n\}$ is nothing but a random walk in which $Y_1, Y_2, ...$ are independent with distribution function $H_1$. Further if at least one of $H_j(y; y_1, ..., y_{j-1})$, $j = 1, ..., k$, has a positive mass at $y = -\infty$, we assume in addition that

(A1) for each $j = 1, ..., k$ and for every fixed $x_1, ..., x_j$ with $1 + \xi x_i > 0$, there exists a constant $u_j^*(x_1, ..., x_j) < x_F$ such that the class

$$\{g(u)f_{j+1}(u + g(u)x_{j+1}|u + g(u)x_1, ..., u + g(u)x_j) : u_j^*(x_1, ..., x_j) \leq u < x_F\}$$

of functions of $x_{j+1}$ is locally uniformly integrable over $\{x_{j+1} : 1 + \xi x_{j+1} > 0\}$,

(A2) when $k = 1$,

$$\lim_{L\to\infty} \overline{\lim_{u\uparrow x_F}} \sup\{P\{X_2 > u|X_1 = x\} : x \leq u - g(u)(1 - L^{-\xi})/\xi\} = 0;$$

when $k \geq 2$, for each $j = 2, ..., k$,

$$\lim_{L\to\infty} \overline{\lim_{u\uparrow x_F}} \sup\Big\{P\{X_{j+1} > u|(X_1, ..., X_j) = (x_1, ..., x_j)\} : \min_{1\leq i\leq j} x_i \leq u - g(u)(1 - L^{-\xi})/\xi\Big\} = 0.$$

Examples of Markov chains satisfying these conditions can be found in Smith (1992) and Yun (1998).

The $k$th-order Markov chain $\{Z_n\}$ defined by (3.3) will be termed the "tail chain" associated with the original chain $\{X_n\}$. The terminology was used in Perfekt (1994, 1997), but the construction scheme is slightly different from ours (see Remark 3.2 of Yun (1998)). The tail chain in (3.3) has a simple structure so that it is usually easy to generate the chain by simulation. This tail chain was used in Yun (1998) to develop a method of computing the extremal index of $\{X_n\}$. Using the tail chain, we here provide a representation for the limiting cluster damage distribution $\pi$ by proving the weak convergence of the cluster damage distribution $\pi_n$ in Section 2.

We need the following general result to prove Theorem 3.1.

**Lemma 3.1.** *Let $\{X_n\}_{n=1}^{\infty}$ be a strictly stationary sequence of random variables, which is not necessarily Markovian, with marginal distribution function $F$. Let*

*$\{u_n\}$ be a sequence of levels such that $u_n \uparrow x_F$ as $n \to \infty$ and $\{r_n\}$ a sequence of positive integers such that*

$$r_n \to \infty \text{ and } P\{M_{r_n+1,2r_n} > u_n | M_{1,r_n} > u_n\} \to 0 \text{ as } n \to \infty. \tag{3.4}$$

*Then, for any constants $a_n > 0$ and for $x > 0$,*

$$P\{\sum_{i=1}^{r_n} a_n(X_i - u_n)^+ \le x | M_{1,r_n} > u_n\}$$
$$-P\{\sum_{i=1}^{r_n} a_n(X_i - u_n)^+ \le x | X_{r_n} > u_n, M_{r_n+1,2r_n} \le u_n\} \to 0 \text{ as } n \to \infty.$$

**Remark 3.1.** If the damage point process $N_n$ in (2.1) is strongly mixing, then the sequence $\{r_n\}$ with $r_n = [n/k_n]$ ($\{k_n\}$: standard sequence) satisfies (3.4) (e.g., see Lemma 2.3 of Hsing, Hüsler and Leadbetter (1988)).

**Proof of Lemma 3.1:** First, observe that

$$\begin{aligned}
&|P\{M_{1,r_n} > u_n\} - r_n P\{X_{r_n} > u_n, M_{r_n+1,2r_n} \le u_n\}| \\
&\quad = |\sum_{i=1}^{r_n}(P\{X_i > u_n, M_{i+1,r_n} \le u_n\} - P\{X_i > u_n, M_{i+1,r_n+i} \le u_n\})| \\
&\quad \le \sum_{i=1}^{r_n} P\{X_i > u_n, M_{i+1,r_n} \le u_n, M_{r_n+1,2r_n} > u_n\} \\
&\quad = P\{M_{1,r_n} > u_n, M_{r_n+1,2r_n} > u_n\}.
\end{aligned} \tag{3.5}$$

Also, writing $Q_n(i,j) = \sum_{s=i}^{j} a_n(X_s - u_n)^+$, $1 \le i \le j$, we get

$$\begin{aligned}
&|P\{Q_n(1,r_n) \le x, M_{1,r_n} > u_n\} \\
&\quad -r_n P\{Q_n(1,r_n) \le x, X_{r_n} > u_n, M_{r_n+1,2r_n} \le u_n\}| \\
&\quad = |\sum_{i=1}^{r_n}(P\{Q_n(r_n+1,2r_n) \le x, X_{r_n+i} > u_n, M_{r_n+i+1,2r_n} \le u_n\} \\
&\qquad\qquad -P\{Q_n(i+1,r_n+i) \le x, X_{r_n+i} > u_n, M_{r_n+i+1,2r_n+i} \le u_n\})| \\
&\quad \le |\sum_{i=1}^{r_n}(P\{Q_n(r_n+1,r_n+i) \le x, M_{i+1,r_n} > u_n, \\
&\qquad\qquad\qquad X_{r_n+i} > u_n, M_{r_n+i+1,2r_n} \le u_n\}
\end{aligned}$$

$$
\begin{aligned}
&-P\{Q_n(i+1,r_n+i)\le x, M_{i+1,r_n}>u_n,\\
&\qquad\qquad X_{r_n+i}>u_n, M_{r_n+i+1,2r_n}\le u_n\})|\\
&+|\sum_{i=1}^{r_n}(P\{Q_n(i+1,r_n+i)\le x, X_{r_n+i}>u_n, M_{r_n+i+1,2r_n}\le u_n\}\\
&\qquad -P\{Q_n(i+1,r_n+i)\le x, X_{r_n+i}>u_n, M_{r_n+i+1,2r_n+i}\le u_n\})|\\
&\le \sum_{i=1}^{r_n} P\{M_{1,r_n}>u_n, X_{r_n+i}>u_n, M_{r_n+i+1,2r_n}\le u_n\}\\
&\quad+\sum_{i=1}^{r_n} P\{X_{r_n+i}>u_n, M_{r_n+i+1,2r_n}\le u_n, M_{2r_n+1,3r_n}>u_n\}\\
&=2P\{M_{1,r_n}>u_n, M_{r_n+1,2r_n}>u_n\}.
\end{aligned}
\tag{3.6}
$$

Using (3.5) and (3.6), we therefore have

$$
\begin{aligned}
&|P\{Q_n(1,r_n)\le x|M_{1,r_n}>u_n\}-P\{Q_n(1,r_n)\le x|X_{r_n}>u_n, M_{r_n+1,2r_n}\le u_n\}|\\
&\quad\le (P\{M_{1,r_n}>u_n\})^{-1}|P\{Q_n(1,r_n)\le x, M_{1,r_n}>u_n\}\\
&\qquad\qquad -r_nP\{Q_n(1,r_n)\le x, X_{r_n}>u_n, M_{r_n+1,2r_n}\le u_n\}|\\
&\qquad+(P\{M_{1,r_n}>u_n\})^{-1}|r_nP\{X_{r_n}>u_n, M_{r_n+1,2r_n}\le u_n\}-P\{M_{1,r_n}>u_n\}|\\
&\quad\le 3P\{M_{r_n+1,2r_n}>u_n|M_{1,r_n}>u_n\}\to 0 \text{ as } n\to\infty,
\end{aligned}
$$

which completes the proof. □

In the following theorem we show that conditions (ii) and (iii) of Theorem 2.1 hold for a large class of $k$th-order stationary Markov chains $\{X_n\}$ having tail chains. Particularly the representation for $\pi$ enables us to compute the limiting cluster damage distribution easily by simulating the corresponding tail chain. Define $\phi(x):=\xi^{-1}(e^{\xi x}-1)$.

**Theorem 3.1.** *Let $\{X_n\}_{n=1}^{\infty}$ be a $k$th-order stationary Markov chain satisfying (3.1) and (3.2). Let $\{Z_n\}$ be the tail chain associated with $\{X_n\}$ and $T$ an Exp(1)-distributed random variable which is independent of $\{Z_n\}$. Assume that (A1) and (A2) hold if necessary. Define $\tau=\max\{n\ge 0: Z_n>-T\}$. Let $\{u_n\}$ be a sequence of levels such that $u_n\uparrow x_F$ as $n\to\infty$ and $\{r_n\}$ a sequence of positive integers satisfying (3.4). Suppose that*

$$
\lim_{r\to\infty}\overline{\lim_{n\to\infty}}\sum_{i=r}^{r_n} P\{X_i>u_n|X_1>u_n\}=0. \tag{3.7}
$$

*Then,*

$$P\{M_{2,r_n} \le u_n | X_1 > u_n\} \to \theta := P\{\tau = 0\} \text{ as } n \to \infty.$$

*Further if $\theta > 0$, then it holds that for $x > 0$*

$$\pi_n(x) = P\{\sum_{i=1}^{r_n} a_n(X_i - u_n)^+ \le x | M_{1,r_n} > u_n\}$$
$$\to \pi(x) := 1 - \theta^{-1} P\{x - \phi(T) < \sum_{j=1}^{\tau} \phi^+(Z_j + T) \le x\} \text{ as } n \to \infty,$$

*where $a_n = 1/g(u_n)$ and $\phi^+(x) = \max\{\phi(x), 0\}$.*

**Remark 3.2.** 1. In fact, under condition (3.7) it can similarly be shown that

$$\lim_{n\to\infty} P\{M_{r+1,r_n} > u_n | X_1 > u_n\} = P\{\sup_{j\ge r} Z_j > -T\} = P\{\tau \ge r\}$$

for $r = 1, 2, \ldots$, and so

$$P\{\tau = \infty\} = \lim_{r\to\infty} P\{\tau \ge r\} = \lim_{r\to\infty} \lim_{n\to\infty} P\{M_{r+1,r_n} > u_n | X_1 > u_n\} = 0$$

(by (3.7) again) which implies $P\{0 \le \tau < \infty\} = 1$ and hence

$$P\{\sum_{j=1}^{\tau} \phi^+(Z_j + T) < \infty\} = 1.$$

Therefore $\pi$ defines a distribution function on $(0, \infty)$.
2. The random variable $\sum_{j=1}^{\tau} \phi^+(Z_j + T)$ has a probability $\theta$ at 0 and is absolutely continuous on $(0, \infty)$. Note that $\phi^+(x) = \phi(x) I\{x > 0\}$.
3. If we let $W$ be a random variable having distribution function $\pi$, the "cluster aggregate excess" can therefore be approximated as

$$\sum_{i=1}^{r_n} (X_i - u_n)^+ | M_{1,r_n} > u_n \overset{d}{\approx} g(u_n) W \text{ for large } n.$$

**Proof of Theorem 3.1:** The first assertion on the convergence to $\theta$ was given in Theorem 3.1 and Corollary 3.3 of Yun (1998). For the second assertion, it is enough by Lemma 3.1 to show that for $x > 0$

$$\begin{aligned} &P\{\sum_{i=1}^{r_n} a_n(X_i - u_n)^+ \le x | X_{r_n} > u_n, M_{r_n+1,2r_n} \le u_n\} \\ &\quad = (P\{M_{2,r_n+1} \le u_n | X_1 > u_n\})^{-1} A_n / P\{X_1 > u_n\} \\ &\quad \to \pi(x) \text{ as } n \to \infty, \end{aligned} \tag{3.8}$$

where

$$A_n = P\{\sum_{i=1}^{r_n} a_n(X_i - u_n)^+ \le x, X_{r_n} > u_n, M_{r_n+1,2r_n} \le u_n\}.$$

Write $Q_n(i,j) = \sum_{s=i}^{j} a_n(X_s - u_n)^+$, $1 \le i \le j$, as before. Then, since $\{X_{r_n} > u_n\} = \{M_{1,r_n} > u_n, X_{r_n} > u_n\}$, we can rewrite $A_n$ as

$$\begin{aligned} A_n &= \sum_{i=1}^{r_n-r} B_n^{(1)}(i) + \sum_{i=r_n-r+1}^{r_n} B_n^{(2)}(i) - \sum_{i=r_n-r+1}^{r_n} B_n^{(3)}(i) \\ &= B_{n,r}^{(1)} + B_{n,r}^{(2)} - B_{n,r}^{(3)}, \text{ say,} \end{aligned}$$

for any $1 \le r \le r_n$, where

$$\begin{aligned} B_n^{(1)}(i) &= P\{Q_n(1,r_n) \le x, M_{1,i-1} \le u_n, X_i > u_n, X_{r_n} > u_n, M_{r_n+1,2r_n} \le u_n\}, \\ B_n^{(2)}(i) &= P\{Q_n(i,r_n) \le x, X_i > u_n, X_{r_n} > u_n, M_{r_n+1,2r_n} \le u_n\}, \\ B_n^{(3)}(i) &= P\{Q_n(i,r_n) \le x, M_{1,i-1} > u_n, X_i > u_n, X_{r_n} > u_n, M_{r_n+1,2r_n} \le u_n\}. \end{aligned}$$

Here, by (3.7) we have

$$\begin{aligned} \lim_{r\to\infty} \overline{\lim_{n\to\infty}} B_{n,r}^{(1)}/P\{X_1 > u_n\} &\le \lim_{r\to\infty} \overline{\lim_{n\to\infty}} \sum_{i=1}^{r_n-r} P\{X_{r_n-i+1} > u_n | X_1 > u_n\} \\ &= \lim_{r\to\infty} \overline{\lim_{n\to\infty}} \sum_{j=r+1}^{r_n} P\{X_j > u_n | X_1 > u_n\} = 0. \quad (3.9) \end{aligned}$$

For the second term $B_{n,r}^{(2)}$ of $A_n$, consider

$$\begin{aligned} & B_{n,r}^{(2)}/P\{X_1 > u_n\} \\ &= \sum_{i=r_n-r+1}^{r_n} P\{Q_n(1, r_n - i + 1) \le x, \\ &\qquad X_{r_n-i+1} > u_n, M_{r_n-i+2,2r_n-i+1} \le u_n | X_1 > u_n\} \\ &= \sum_{j=1}^{r} P\{Q_n(1,j) \le x, X_j > u_n, M_{j+1,r_n+j} \le u_n | X_1 > u_n\}. \end{aligned}$$

Now, using (3.7) again and applying similar methods as in Lemma 2.2 and Theorem 3.1 of Yun (1998), it can be shown that for any $j = 1, ..., r$,

$$\begin{aligned} & P\{(a_n(X_1 - u_n), ..., a_n(X_j - u_n)) \in \cdot, X_j > u_n, M_{j+1,r_n+j} \le u_n | X_1 > u_n\} \\ & \xrightarrow{w} P\{(\phi(T), \phi(Z_1 + T), ..., \phi(Z_{j-1} + T)) \in \cdot, \tau = j - 1\} \text{ as } n \to \infty. \end{aligned}$$

Since the function $(x_1, ..., x_j) \mapsto \sum_{i=1}^{j} x_i^+$ is continuous, it thus follows that

$$\begin{aligned}\lim_{n\to\infty} B_{n,r}^{(2)}/P\{X_1 > u_n\} &= \sum_{j=1}^{r} P\{R(0, j-1) \leq x, \tau = j-1\} \\ &\to P\{R(0,\tau) \leq x\} \text{ as } r \to \infty, \end{aligned} \tag{3.10}$$

where $R(i,j) = \sum_{s=i}^{j} \phi^+(Z_s + T)$, $0 \leq i \leq j$ (recall $Z_0 \equiv 0$). For the third term $B_{n,r}^{(3)}$ of $A_n$, observe that

$$\begin{aligned} & B_{n,r}^{(3)}/P\{X_1 > u_n\} \\ &= \sum_{i=r_n-r+1}^{r_n} \sum_{j=1}^{i-1} P\{Q_n(2, r_n - j + 1) \leq x, M_{2,i-j} \leq u_n, X_{i-j+1} > u_n, \\ &\qquad X_{r_n-j+1} > u_n, M_{r_n-j+2, 2r_n-j+1} \leq u_n | X_1 > u_n\} \\ &= \sum_{l=1}^{r} \sum_{m=l+1}^{l+s} C_n(l,m) + \sum_{l=1}^{r} \sum_{m=l+s+1}^{r_n} C_n(l,m) \\ &= C_{n,r,s}^{(1)} + C_{n,r,s}^{(2)}, \text{ say,} \end{aligned}$$

for any $1 \leq s \leq r_n - r$, where we put $l = r_n - i + 1$, $m = r_n - j + 1$ and

$$\begin{aligned} C_n(l,m) = P\{Q_n(2,m) \leq x, M_{2,m-l} \leq u_n, X_{m-l+1} > u_n, \\ X_m > u_n, M_{m+1, r_n+m} \leq u_n | X_1 > u_n\}. \end{aligned}$$

Here, using (3.7) it can be again seen that for any $l = 1, ..., r$ and $m = l+1, ..., l+s$,

$$\begin{aligned} & P\{(a_n(X_2 - u_n), ..., a_n(X_m - u_n)) \in \cdot, M_{2,m-l} \leq u_n, \\ &\quad X_{m-l+1} > u_n, X_m > u_n, M_{m+1,r_n+m} \leq u_n | X_1 > u_n\} \\ &\xrightarrow{w} P\{(\phi(Z_1 + T), ..., \phi(Z_{m-1} + T)) \in \cdot, \\ &\qquad \tilde{M}_{1,m-l-1} \leq -T, Z_{m-l} > -T, \tau = m-1\} \text{ as } n \to \infty, \end{aligned}$$

where $\tilde{M}_{i,j} := \max\{Z_i, ..., Z_j\}$, $1 \leq i \leq j$. Thus it follows that

$$\begin{aligned} & \lim_{n\to\infty} C_{n,r,s}^{(1)} \\ &= \sum_{l=1}^{r} \sum_{m=l+1}^{l+s} P\{R(1, m-1) \leq x, \tilde{M}_{1,m-l-1} \leq -T, Z_{m-l} > -T, \tau = m-1\} \\ &= \sum_{m=2}^{r+s} \sum_{l=1}^{m-1} P\{R(1, m-1) \leq x, \tilde{M}_{1,m-l-1} \leq -T, Z_{m-l} > -T, \tau = m-1\} \\ &= \sum_{m=2}^{r+s} P\{R(1, m-1) \leq x, \tau = m-1\} \\ &\to P\{R(1,\tau) \leq x, \tau \geq 1\} \text{ as } s \to \infty \text{ and } r \to \infty. \end{aligned}$$

Also, by (3.7) we have

$$\lim_{r\to\infty}\lim_{s\to\infty}\overline{\lim_{n\to\infty}}\, C^{(2)}_{n,r,s} \le \lim_{r\to\infty}\sum_{l=1}^{r}\Big(\lim_{s\to\infty}\overline{\lim_{n\to\infty}}\sum_{m=s+l+1}^{r_n} P\{X_m > u_n | X_1 > u_n\}\Big) = 0.$$

These two results imply that

$$\lim_{r\to\infty}\lim_{n\to\infty} B^{(3)}_{n,r}/P\{X_1 > u_n\} = P\{R(1,\tau) \le x, \tau \ge 1\}. \tag{3.11}$$

From (3.9), (3.10) and (3.11), we therefore have

$$\begin{aligned}\lim_{n\to\infty} A_n/P\{X_1 > u_n\} &= P\{R(0,\tau) \le x\} - P\{R(1,\tau) \le x, \tau \ge 1\} \\ &= P\{\tau = 0\} - P\{R(1,\tau) \le x, R(0,\tau) > x\} \\ &= \theta - P\{x - \phi(T) < \sum_{j=1}^{\tau} \phi^+(Z_j + T) \le x\},\end{aligned}$$

which proves (3.8). □

The representation for the limiting cluster damage distribution $\pi$ in Theorem 3.1 makes it easy to derive its mean and variance. These quantities can then be used effectively in modeling the clusters above a high level of weakly dependent, $k$th-order stationary Markovian data.

**Corollary 3.1.** *Under the same conditions as in Theorem 3.1, let* $V = \sum_{j=1}^{\tau} \phi^+(Z_j + T)$ *and* $W$ *a random variable having distribution function* $\pi$. *If* $E(V) < \infty$, *then*

$$E(W) = \begin{cases} \dfrac{1}{\theta(1-\xi)}, & \text{if } \xi < 1, \\ \infty, & \text{otherwise.} \end{cases}$$

*If* $E(V^2) < \infty$, *then*

$$Var(W) = \begin{cases} \dfrac{2}{\theta\xi^2}\displaystyle\sum_{j=1}^{\infty}\Big\{\frac{1}{1-2\xi}E(\exp(\xi Z_j^+ - (1-\xi)Z_j^-)) \\ \qquad -\dfrac{1}{1-\xi}E(\exp((\xi-1)Z_j^-)) \\ \qquad -\dfrac{1}{1-\xi}E(\exp(\xi Z_j^+ - Z_j^-)) + E(\exp(Z_j))\Big\} \\ \quad +\dfrac{2}{\theta(1-\xi)(1-2\xi)} - \dfrac{1}{\theta^2(1-\xi)^2}, & \text{if } \xi < 1/2, \\ \infty, & \text{otherwise.} \end{cases}$$

**Proof:** Since $W$ and $V$ are nonnegative random variables, if $E(V) < \infty$, then

$$\begin{aligned} E(W) &= \int_0^\infty P\{W > x\}\,dx = \theta^{-1}\int_0^\infty (P\{V + \phi(T) > x\} - P\{V > x\})\,dx \\ &= \theta^{-1}(E(V + \phi(T)) - E(V)) = (\theta\xi)^{-1}(E(\exp(\xi T)) - 1) \\ &= \begin{cases} \dfrac{1}{\theta(1-\xi)}, & \text{if } \xi < 1, \\ \infty, & \text{otherwise.} \end{cases} \end{aligned}$$

Also, if $E(V^2) < \infty$, then

$$\begin{aligned} E(W^2) &= 2\int_0^\infty xP\{W > x\}\,dx \\ &= 2\theta^{-1}\int_0^\infty x(P\{V + \phi(T) > x\} - P\{V > x\})\,dx \\ &= \theta^{-1}(E(V + \phi(T))^2 - E(V^2)) = \theta^{-1}(2E(V\phi(T)) + E(\phi(T))^2) \\ &= 2\theta^{-1}\xi^{-2}\sum_{j=1}^{\infty} E((\exp(\xi(Z_j + 2T)) - \exp(\xi T) \\ &\qquad - \exp(\xi(Z_j + T)) + 1)I\{Z_j > -T\}) \\ &\quad + \theta^{-1}\xi^{-2}(E(\exp(2\xi T)) - 2E(\exp(\xi T)) + 1) \\ &= \begin{cases} \dfrac{2}{\theta\xi^2}\displaystyle\sum_{j=1}^{\infty}\Big\{\dfrac{1}{1-2\xi}E(\exp(\xi Z_j^+ - (1-\xi)Z_j^-)) \\ \qquad -\dfrac{1}{1-\xi}E(\exp((\xi - 1)Z_j^-)) \\ \qquad -\dfrac{1}{1-\xi}E(\exp(\xi Z_j^+ - Z_j^-)) + E(\exp(Z_j))\Big\} \\ \quad +\dfrac{2}{\theta(1-\xi)(1-2\xi)}, \quad \text{if } \xi < 1/2, \\ \infty, \qquad\qquad \text{otherwise,} \end{cases} \end{aligned}$$

since $\{Z_n\}$ and $T \sim \text{Exp}(1)$ are independent. This completes the proof. □

For examples of Markov chains which satisfy the conditions of Theorem 3.1, the reader is referred to Smith (1992) and Yun (1998).

## References

Anderson, C. W. and Dancy, G. P. (1992). "The severity of extreme events," Research Report 409/92, Dept. Probab. Statist., Univ. Sheffield.

Hsing, T. (1993). "Extremal index estimation for a weakly dependent stationary sequence,"vvvvvv *Ann. Statist.*, **21**, 2043-2071.

Hsing, T., Hüsler, J. and Leadbetter, M. R. (1988). "On the exceedance point process for a stationary sequence," *Probab. Theory Related Fields*, **78**, 97-112.

Leadbetter, M. R. (1974). "On extreme values in stationary sequences," *Z. Wahrsch. verw. Gebiete*, **28**, 289-303.

Leadbetter, M. R. (1995). "On high level exceedance modeling and tail inference," *J. Statist. Plann. Inference*, **45**, 247-260.

Leadbetter, M. R. and Rootzén, H. (1988). "Extremal theory for stochastic processes," *Ann. Probab.*, **16**, 431-478.

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

O'Brien, G. L. (1987). "Extreme values for stationary and Markov sequences," *Ann. Probab.*, **15**, 281-291.

Perfekt, R. (1994). "Extremal behaviour of stationary Markov chains with applications," *Ann. Appl. Probab.*, **4**, 529-548.

Perfekt, R. (1997). "Extreme value theory for a class of Markov chains with values in $\Re^d$," *Adv. in Appl. Probab.*, **29**, 138-164.

Resnick, S. I. (1987). *Extreme Values, Point Processes and Regular Variation*, Springer, New York.

Rootzén, H. (1988). "Maxima and exceedances of stationary Markov chains," *Adv. in Appl. Probab.*, **20**, 371-390.

Smith, R. L. (1989). "Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone," *Statist. Science*, **4**, 367-393.

Smith, R. L. (1992). "The extremal index for a Markov chain," *J. Appl. Probab.*, **29**, 37-45.

Yun, S. (1996). "Modeling extreme values of ground-level ozone based on threshold methods for Markov chains," *Korean Comm. Statist.*, **3**, 249-273.

Yun, S. (1997). "On domains of attraction of multivariate extreme value distributions under absolute continuity," *J. Multivariate Anal.*, **63**, 277-295.

Yun, S. (1998). "The extremal index of a higher-order stationary Markov chain," *Ann. Appl. Probab.*, **8**, 408-437.