[Article] Weakest-link scaling and extreme events in finite-sized systems

<table>
<thead>
<tr>
<th>Original Citation:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Availability:</th>
</tr>
</thead>
<tbody>
<tr>
<td>This version is available at: <a href="http://porto.polito.it/2649607/">http://porto.polito.it/2649607/</a> since: September 2016</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Publisher:</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDPI AG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Published version:</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOI:10.3390/e17031103</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Terms of use:</th>
</tr>
</thead>
<tbody>
<tr>
<td>This article is made available under terms and conditions applicable to Open Access Policy Article (&quot;Creative Commons: Attribution 3.0&quot;), as described at <a href="http://porto.polito.it/terms_and_conditions.html">http://porto.polito.it/terms_and_conditions.html</a></td>
</tr>
</tbody>
</table>

Porto, the institutional repository of the Politecnico di Torino, is provided by the University Library and the IT-Services. The aim is to enable open access to all the world. Please share with us how this access benefits you. Your story matters.

(Article begins on next page)
Weakest-Link Scaling and Extreme Events in Finite-Sized Systems

Dionissios T. Hristopulos 1,*, Manolis P. Petrakis 1 and Giorgio Kaniadakis 2

1 School of Mineral Resources Engineering, Technical University of Crete, Chania 73100, Greece; E-Mail: petrakis@mred.tuc.gr
2 Department of Applied Science and Technology, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy; E-Mail: giorgio.kaniadakis@polito.it

* Author to whom correspondence should be addressed; E-Mail: dionis@mred.tuc.gr; Tel.: +30-28210-37688; Fax: +30-28210-37853.

Academic Editor: Kevin H. Knuth

Received: 21 January 2015 / Accepted: 3 March 2015 / Published: 9 March 2015

Abstract: Weakest-link scaling is used in the reliability analysis of complex systems. It is characterized by the extensivity of the hazard function instead of the entropy. The Weibull distribution is the archetypical example of weakest-link scaling, and it describes variables such as the fracture strength of brittle materials, maximal annual rainfall, wind speed and earthquake return times. We investigate two new distributions that exhibit weakest-link scaling, i.e., a Weibull generalization known as the $\kappa$-Weibull and a modified gamma probability function that we propose herein. We show that in contrast with the Weibull and the modified gamma, the hazard function of the $\kappa$-Weibull is non-extensive, which is a signature of inter-dependence between the links. We also investigate the impact of heterogeneous links, modeled by means of a stochastic Weibull scale parameter, on the observed probability distribution.

Keywords: Weibull; gamma distribution; return intervals; superstatistics; fracture strength; stretched exponential; reliability modeling

PACS classifications: 02.50.-r; 02.50.Ey; 02.60.Ed; 89.20.-a; 89.60.-k
1. Introduction

In statistical mechanics, the velocities of non-interacting particles are viewed as random variables that follow the Maxwell–Boltzmann distribution. Quantities of interest are often macroscopic (system) properties, such as the energy or the entropy, that involve averages over many particles. Evaluating these properties involves calculating averages over suitable probability distributions. Randomness is also a factor in the analysis of macroscopic variables. Environmental time series, energy and mineral reserves and earthquakes are just a few examples of phenomena that exhibit stochastic fluctuations. In addition, the lifetimes (times to failure) of various natural and engineered materials and devices are also random variables. In contrast with statistical mechanics, there is a strong interest not only in average quantities, but also in the extreme values of such distributions [1]. The classical extreme value theory studies the probability distributions of extreme events [2,3]. The Fisher–Tippet–Gnedenko theorem shows that the distribution of suitably-scaled transformations of the maxima or minima of a collection of independent and identically distributed (i.i.d.) variables follows the generalized extreme value (GEV) distributions, which include the Gumbel, Weibull and Fréchet distributions. The type of GEV obtained depends on the tail behavior of the probability distribution for the i.i.d. random variables. Extensions that involve stochastic scalings have been recently proposed in [4].

We investigate systems that obey the weakest-link scaling (WLS) principle. In the WLS framework, the system response (e.g., material strength) is controlled by the weakest link (unit), which implies a marked deviation from statistical independence. The Weibull model [3] is the archetypical probability distribution for WLS systems and is widely used in reliability modeling [5]. It finds application as a model of the fracture strength of brittle and quasi-brittle technological materials [6,7] and geologic media [8], waiting times between earthquakes [9–15], wind speed [16] and annual hydrological maxima [17]. We study the impact of the randomness (heterogeneity) of the Weibull link parameters on the system survival function by introducing doubly-stochastic representations analogous to superstatistics [18]. We also investigate weakest-link scaling with non-exponential link survival functions. In this context, we introduce a modified gamma distribution, and we provide theoretical arguments for the \( \kappa \)-Weibull distribution, which exhibits power-law tails in finite-sized systems.

1.1. Mathematical Preliminaries

We will use the capital letter \( X \) to denote a random variable (e.g., tensile strength, wind speed, earthquake return time, maximum annual rainfall or river discharge) and the lowercase \( x \) to denote values of the random variable, where \( x \in [0, \infty) \). We will assume that the random variable represents the property of interest for a particular spatial or temporal window (support), the size of which we will denote by \( \ell \). The probability density function (pdf) will be denoted by \( f_\ell(x) \). It is customary to use \( f_X(x) \) to denote the pdf of \( X \) when more than one random variable is involved. In this case, however, there is no risk of confusion, and thus, we replace \( X \) with the support scale, which is relevant in our analysis. The integral of the pdf, i.e., the cumulative distribution function (cdf), will be denoted by:

\[
F_\ell(x) = \int_{-\infty}^{x} dx' f_\ell(x').
\]
Another useful probability function, which is mostly used in reliability analysis, is the survival function or complementary cumulative distribution function, defined by:

$$R_\ell(x) = 1 - F_\ell(x). \quad (1)$$

The survival function $R_\ell(x)$ gives the probability that the system has not “failed” when the variable $X$ has reached the level $x$. The survival function is equal to one for $x = x_{\text{min}}$ and tends to zero for $x \rightarrow \infty$. A characteristic property of the survival function is that its integral equals the expectation of $X$, i.e.,

$$\mathbb{E}[X] = \int_0^\infty dx' R_\ell(x'). \quad (2)$$

The above is based on $dR_\ell(x)/dx = -f_\ell(x)$, which follows from (1) and integration by parts.

- In the case of fracture strength, $R_\ell(x)$ is the probability that the system has not failed if the external load takes values that do not exceed $x$.
- In the case of earthquake return times, the survival function is the probability that there are no successive earthquakes separated by a time interval less than or equal to $x$.
- In the case of annual rainfall maxima, $R_\ell(x)$ gives the probability that the maximum rainfall over the support area for the duration of one year does not exceed the value $x$.

Finally, the hazard rate or hazard function $h_\ell(x)$ is the conditional probability that the system will fail at time $x^* > x$ within the infinitesimal time window $x < x^* \leq x + \delta x$, given that the system has not failed during the interval $[0, x]$. Let $A$ denote the event that the system has not failed at $x$ and $B$ the event that the system fails at $x < x^* \leq x + \delta x$. Then, $h_\ell(x) = \text{Prob}(B|A) = \text{Prob}(B \cap A)/\text{Prob}(A)$, which leads to the following:

$$h_\ell(x) = \frac{\text{Prob}(x < x^* \leq x + \delta x)}{R_\ell(x) \delta x} = \frac{f_\ell(x)}{R_\ell(x)}. \quad (3)$$

The asymptotic dependence of the hazard rate is a matter of interest. For example, in the case of earthquakes, it is often desired to estimate the probability of the time until the next big earthquake. If the particular hazard rate tends to infinity as $x \rightarrow \infty$, the probability that an earthquake will occur increases with the time elapsed since the last earthquake [19]. In contrast, a decreasing or constant hazard rate indicates that the probability of a new occurrence is reduced or remains constant, respectively, as time elapses, and no event takes place.

2. Weakest-Link Scaling Theory

The weakest-link scaling (WLS) theory has been developed to model the strength of disordered systems based on the strengths of “critical elements” or links [20]. In WLS, the entire system fails if the link with the lowest mechanical strength breaks, hence the term “weakest-link”. The foundations of weakest-link theory are found in the studies of Gumbel [2] and Weibull [3] on the statistics of extremes. The concept of links is straightforward in simple systems, e.g., one-dimensional chains. In higher dimensions, the links correspond to critical subsystems, the failure of which causes an avalanche, leading to the failure of the entire system.
Weakest-link scaling is used in statistical physics to investigate size effects in the strength of heterogeneous materials \cite{entropy2015osci171106,entropy2015osci21347,entropy2015osci23189,entropy2015osci23385}. In reliability modeling, a weakest-link system is known as a series system \cite{entropy2015osci24492}. The same scaling ideas, however, can be used in other systems, as well. For example, in the case of maximum annual rainfall, the scale links could correspond to single days and the system scale to an entire year. Then, if "one link fails," i.e., if the rainfall on a specific day exceeds the specified threshold $x$, then the "entire system fails," i.e., the maximum annual daily rainfall also exceeds the threshold.

Let us consider a system with support $\ell_N$ consisting of $N$ links with support $\ell_0$. The WLS premise is then expressed as follows:

$$R_{\ell_N}(x) = \prod_{i=1}^{N} R_{\ell_0}(x),$$  \hspace{1cm} (4)

where $R_{\ell_0}(x)$ refers to the survival function of the $i$-th link. If we assume that all of the links have a common survival function, then (4) can be expressed as:

$$R_{\ell_N}(x) = [R_{\ell_0}(x)]^N.$$  \hspace{1cm} (5)

The above leads to an additive relation for the logarithms of the survival function, i.e.,

$$\ln R_{\ell_N}(x) = N \ln R_{\ell_0}(x).$$  \hspace{1cm} (6)

Now, if we take into account that:

$$\frac{d \ln R_{\ell_N}(x)}{dx} = \frac{1}{R_{\ell_N}(x)} \frac{dR_{\ell_N}(x)}{dx} = -\frac{f_{\ell_N}(x)}{R_{\ell_N}(x)} = -h_{\ell_N}(x),$$  \hspace{1cm} (7)

it follows from (6) that the WLS hazard function satisfies the additive property:

$$h_{\ell_N}(x) = N h_{\ell_0}(x).$$  \hspace{1cm} (8)

For a system of $N$ independent and identically distributed random variables, it holds that:

$$f_{\ell_N}(x) = \prod_{i=1}^{N} f_{\ell_0}(x).$$

Equation (5) implies that under WLS conditions, the link independence does not hold, since:

$$f_{\ell_N}(x) = -\frac{dR_{\ell_N}(x)}{dx} = N \left[R_{\ell_0}(x)\right]^{N-1} f_{\ell_0}(x),$$  \hspace{1cm} (9)

which is markedly different from the link independence condition.

To compare general WLS distribution functions with the Weibull, we use the double logarithm of the reciprocal survival probability,

$$\Phi(x; \ell_N) := \ln \left[\ln \left(\frac{1}{R_{\ell_N}(x)}\right)\right].$$  \hspace{1cm} (10)

If $F_{\ell_N}(x)$ follows the Weibull dependence, then $\Phi(x; \ell_N) = m \ln \frac{x}{x_s}$.

For a general WLS distribution, we obtain the following relation based on (5):

$$\Phi(x; \ell_N) = \ln N + \ln \left[-\ln R_{\ell_0}(x)\right].$$  \hspace{1cm} (11)

The above scaling relation in combination with (10), replacing $\ell_N$ in the latter with $\ell_0$, implies the following scaling property:

$$\Phi(x; \ell_N) = \ln N + \Phi(x; \ell_0).$$  \hspace{1cm} (12)
2.1. Systems Composed of Interacting Links

Above, we assumed an identical survival function for all of the links, a condition that is plausible in a mean-field framework. Herein, we consider systems in which the link properties vary. This is an interesting possibility, which can lead to new probability models for reliability engineering. In general, if \( R_{\ell_0}(x) = \text{Prob}(\Lambda_i) \) is the probability that the \( i \)-th link survives, the system’s survival probability under WLS conditions is given by:

\[
R_{\ell_N}(x) = \text{Prob}(\Lambda_1 \cap \Lambda_2 \ldots \cap \Lambda_N).
\]

Using the law of joint probability it follows that:

\[
R_{\ell_N}(x) = \text{Prob}(\Lambda_1) \text{Prob}(\Lambda_2|\Lambda_1) \ldots \text{Prob}(\Lambda_{\ell_N}|\Lambda_1, \ldots, \Lambda_{N-1}).
\]

The conditional probability \( \text{Prob}(\Lambda_2|\Lambda_1) \) may in general depend on the state of \( \Lambda_1 \) and the degree of interaction between the two links. This agrees with the observation of complex interaction patterns in geological fault systems, if the links are identified with single faults or fault groups [25]. One possibility for expressing this dependence is by means of the following ansatz:

\[
\text{Prob}(\Lambda_2|\Lambda_1) = R_{\ell_0}(\lambda_1 x),
\]

where \( \lambda_1 > 0 \) is a scaling factor. Since \( R_{\ell_0}(\cdot) \) is a decreasing function, \( \lambda_1 > 1 \) implies that the conditional survival probability of \( \Lambda_2 \) is reduced compared to the probability \( R_{\ell_0}(x) \).

More generally, for an \( N \)-link system, the survival function of the link \( \Lambda_i \) can be expressed as:

\[
\text{Prob}(\Lambda_{\ell_i}|\Lambda_1, \ldots, \Lambda_{i-1}) = R_{\ell_0}(\lambda_{i-1} x), \quad i = 2, \ldots, N.
\]

**Figure 1.** Survival function \( R_{\ell_2}(x, \lambda) \) for a system comprising two links that follow the Weibull survival function \( R_{\ell_0}(x) \) with \( x_s = 1 \) and \( m = 2.5 \). Continuous surface: \( R_{\ell_2}(x, \lambda) \) for different values of \( \lambda \). Mesh: \( R_{\ell_2}(x, 1) \).
In light of the above, the system survival function is given by the following equation:

$$R_{\ell_N}(x, \vec{\lambda}) = R_{\ell_0}(x) \prod_{i=2}^{N} R_{\ell_0}(\lambda_{i-1} x),$$

(13)

which is determined by the coefficient vector $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_{N-1})^T$. The above system survival function is given by a product rule in agreement with the survival function (4) of the homogeneous-link system. In addition, all terms of the product involve the same function $R_{\ell_0}(\cdot)$, albeit the argument of the latter is modified by the respective coefficients $\lambda_i$, $i = 1, \ldots, N - 1$. Hence, (13) maintains the product rule and the same functional form of the link survival function, which are central features of WLS, but it also captures interactions between the links via the coefficients $\vec{\lambda}$. We illustrate the behavior of the system survival function for a fictitious two-link system in Figure 1, where $R_{\ell_2}(x, \lambda)$ is compared with $R_{\ell_2}(x, 1) = [R_{\ell_0}(x)]^2$: the former is the survival function of the correlated system, whereas the latter is the standard WLS survival function. For the Weibull survival function, $\lambda$ effectively renormalizes the scale parameter $x_s$ to $x'_s = [2/(1 + \lambda m)]^{1/m} x_s$. Hence, the ansatz (13) essentially accounts for the correlations by renormalizing a parameter of the “free distribution”, which is an idea with deep roots in statistical physics. Figure 1 demonstrates that values $\lambda > 1$ ($\lambda < 1$) lead to lower (higher) survival probability for the same $x$ than the uniform-link WLS function.

The proposed system survival function (13) is quite general, since it allows for different functional forms of the link survival function. In (13), it is possible to consider correlated random variables $\lambda_i$ as coefficients. In this case, the random variable $X$ becomes doubly stochastic and is described by superstatistics [18,26]. We expect that the variability of link parameters is a plausible scenario in heterogeneous physical or engineered systems. For example, a theoretical framework has been proposed that links the statistics of earthquake recurrence times to fracture strength in the Earth’s crust [15,27]. Since the crust is a heterogeneous material decorated with flaws (faults) of widely varying sizes and orientations, it is reasonable to expect that the fracture strength parameters exhibit local random fluctuations in addition to systematic changes with depth [25].

2.2. Hazard Function for Links with Variable Parameters

There is a distinct possibility that links have different statistical properties, even in the case of engineered materials. Thus, a link parameter $\theta$ can become a random variable with pdf $g(\theta)$. In light of the additivity property (4), the hazard rate of the system is given by the following sum:

$$h_{\ell_N}(x) = \sum_{i=1}^{N} h_{\ell_i \ell_0}(x; \theta).$$

(14)

We can define the average hazard function over the links:

$$\bar{h}_{\ell_0}(x) = \frac{1}{N} \sum_{i=1}^{N} h_{\ell_i \ell_0}(x; \theta),$$

so that $h_{\ell_N}(x) = N \bar{h}_{\ell_0}(x)$. If $\theta$ follows the pdf $g(\theta)$, and the support of $\theta$ is the finite interval $[\theta_{\min}, \theta_{\max}]$, we can define an effective link hazard rate $h^*_{\ell_0}(x)$ by:

$$h^*_{\ell_0}(x) = \int_{\theta_{\min}}^{\theta_{\max}} d\theta \ h_{\ell_0}(x, \theta) g(\theta).$$

(15)
If \( h_{\ell_0}(x, \theta) \) and \( g(\theta) \) are continuous functions of \( \theta \), the requirements of the mean value theorem are fulfilled; thus, there exists a \( \theta^* \in [\theta_{\min}, \theta_{\max}] \), such that:

\[
 h_{\ell_0}^*(x) = h_{\ell_0}(x|\theta^*) \int_{\theta_{\min}}^{\theta_{\max}} d\theta \ g(\theta) = h_{\ell_0}(x|\theta^*). \tag{16}
\]

Equation (16) states that the effective link hazard rate maintains the original functional form and that the impact of random \( \theta \) can be described by means of an “effective parameter” \( \theta^* \), provided that the conditions of the mean value theorem are upheld.

If ergodic conditions are satisfied so that the events on the links sufficiently sample the space of probable states, one can assume that \( h_{\ell_0}(x) = h_{\ell_0}^*(x) \). The observed survival function then is \( R_{\ell_0}(x) = R_{\ell_0}(x|\theta^*) \), and:

\[
 \Phi(x; \ell_0) = \ln [-\ln(R_{\ell_0}(x|\theta^*))], \tag{17}
\]

which implies that:

\[
 \Phi(x; \ell_0) = \Phi(x; \ell_0|\theta^*). \tag{18}
\]

### 2.3. Observed Distributions for Non-ergodic Conditions

In the case of natural phenomena, however, the available sample is often finite, and the ergodic condition may not be justifiable. For example, an earthquake catalog for a particular area covers only a small portion of the area’s geological history. This means that the full ensemble, involving both the distribution of the scale parameters and the distribution of \( x \) values for fixed parameter values, is not adequately sampled during the observation process. If \( g(\theta) \) has a narrow mode (e.g., Weibull with \( m > 1 \)), the estimated pdf \( \hat{g}(\theta) \) will have a peak in the interval \([\theta_{\min}, \theta_{\max}]\) near the mode of \( g(\theta) \). For disperse parameter distributions (e.g., Weibull with \( m \leq 1 \)), \( \hat{g}(\theta) \) may exhibit more complicated behavior due to insufficient sampling.

In order to illustrate the effect of the sample size and the link parameter’s variability on the observed link probability distribution, we conduct a numerical experiment. We simulate a random variable \( X \) that represents individual links, assuming that \( X \) follows the Weibull distribution \( f_{\ell_0}(x; m, x_s) \) with a fixed shape parameter \( m \) and uniformly distributed scale parameter \( x_s \). The distribution of \( x_s \) is \( g(x_s) = (x_{s,\max} - x_{s,\min})^{-1} \) if \( x_s \in [x_{s,\min}, x_{s,\max}] \) and \( g(x_s) = 0 \) otherwise. We generate a large set containing \( n_{\text{total}} = 10^6 \) random numbers drawn from Weibull distributions with different values of \( x_s \). The latter correspond to different links drawn from the uniform distribution \( g(x_s) \). We define the mixing ratio as the number of sampled links over the number of samples per link, i.e.,

\[
 \rho = \frac{N}{n_{s}}, \quad N = \# \text{number of links}, \quad n_s = \# \text{samples per link}
\]

The ratio \( \rho \) determines the mixing of random values from distributions with different scale parameters. Small values of \( \rho \) denote that a small number of links are sampled compared to the samples taken from each link, whereas high values of \( \rho \) denote the opposite.

In order to investigate different mixing scenarios, we consider two values of \( \rho \) equal to 0.1 and \( 10^{-4} \). The lower limit \( x_{s,\min} \) of the scale parameter distribution is set to one, whereas the width \( \delta = x_{s,\max} - x_{s,\min} \) of the uniform distribution is set to \( \delta = 1 \) and 100. From the initial set of \( n_{\text{total}} \)
random numbers, we randomly select $n = 10^3$ and $10^5$ samples, for each of the four scenarios obtained by the different combinations of $\rho$ and $\delta$. The different values of $n$ allow investigating the sample size effect. For each sample, we calculate the maximum likelihood estimates of the optimal-fit Weibull parameters, which are listed in Table 1. For $m = 0.7$, the sample-based estimates of the Weibull modulus are close to the true value, especially in the case of the narrower distribution of the scale parameter. For $m = 2.5$, the estimates of the Weibull modulus are close to the true value only for the narrow scale parameter distribution.

Figure 2. Sample estimates $\hat{\Phi}(x; \ell)$ of data sets with sizes $n = 10^3$ (top) and $n = 10^5$ (bottom). The samples are randomly drawn from a total of $n_{\text{total}} = 10^6$ simulated values generated from Weibull distributions with fixed $m = 0.7$ (left column) and $m = 2.5$ (right column). The $x_s$ are uniformly distributed in the interval $[1, 1 + \delta]$ with $\delta = 1, 100$. The mixing ratio that determines the number of $x_s$ values (links) over the number of random values per link is $\rho = 0.1, 10^{-4}$. 
We also show the estimated \( \hat{\Phi}(x; \ell) \) derived from the empirical (sample-based) cumulative distribution function in Figure 2. For reference purposes, if the data follow the Weibull distribution, \( \hat{\Phi}(x; \ell) \) is a straight line with slope \( \hat{m} \) and intercept \( \hat{m} \ln \hat{x}_s \). As shown in Figure 2(c), for \( m < 1 \) and large \( n \), the Weibull distribution is obtained. For smaller \( n \) (Figure 2(a)), deviations from the Weibull emerge. For \( m > 1 \) and small \( n \), the data cannot be described with a Weibull distribution (Figure 2(b)). For a larger sample size (Figure 2(d)), \( \hat{\Phi}(x; \ell) \) demonstrates Weibull tails with different slopes. Overall, for \( m = 0.7 \) (in general, for \( m < 1 \)), the distribution of sampled values maintains, albeit approximately, the Weibull characteristic shape and the original modulus. For \( m > 1 \), the Weibull scaling is maintained only for a narrow scale parameter distribution.

Table 1. Estimates of Weibull distribution parameters for the simulated data of Figure 2.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \delta )</th>
<th>( \rho )</th>
<th>( \hat{x}_s )</th>
<th>( \hat{m} )</th>
<th>( \hat{x}_s )</th>
<th>( \hat{m} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 10^3 )</td>
<td>1</td>
<td>( 10^{-4} )</td>
<td>1.39</td>
<td>0.688</td>
<td>1.42</td>
<td>2.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( 10^{-1} )</td>
<td>1.55</td>
<td>0.694</td>
<td>1.54</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>( 10^{-4} )</td>
<td>45.6</td>
<td>0.592</td>
<td>51.8</td>
<td>1.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( 10^{-1} )</td>
<td>43.6</td>
<td>0.634</td>
<td>49.6</td>
<td>1.38</td>
</tr>
<tr>
<td>( 10^5 )</td>
<td>1</td>
<td>( 10^{-4} )</td>
<td>1.40</td>
<td>0.689</td>
<td>1.42</td>
<td>2.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( 10^{-1} )</td>
<td>1.50</td>
<td>0.691</td>
<td>1.52</td>
<td>2.23</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>( 10^{-4} )</td>
<td>57.0</td>
<td>0.679</td>
<td>59.3</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( 10^{-1} )</td>
<td>42.9</td>
<td>0.611</td>
<td>48.4</td>
<td>1.30</td>
</tr>
</tbody>
</table>

3. Exponential and Non-Exponential Link Survival Functions

Whereas the Weibull distribution has proven quite successful in the analysis of extreme value data, in certain cases, deviations from the Weibull scaling have been noticed in the tails of empirical distributions [27,28]. As shown above, such deviations could occur due to fluctuations of the link hazard rate parameters. Below, we investigate a different possibility, which is based on non-exponential link survival functions with deterministic parameters. Ideally, such functions should be obtained from microscopic theories that describe the dynamics of the relevant physical processes inside the links. Nevertheless, the definition of links and their respective probability functions is not always possible in terms of such detailed analysis [6,28]. Hence, for practical purposes, it is useful to consider phenomenological definitions of the survival function.

3.1. Weibull Link Ansatz

In his original paper, Weibull assumed that \( R_{\ell_0}(x) = e^{-\phi(x)} \), where \( \phi(x) \geq 0 \) is a monotonically increasing function [3]. In particular, he proposed that \( \phi(x) = (x/x_s)^m \), where \( x_s \) is the scale parameter and \( m \) is the shape parameter or Weibull modulus. More generally, the exponential dependence \( R_{\ell_n}(x) = e^{-n \phi(x)} \), implies the self-similarity relation:

\[
R_{\ell_N}(x) = R_{\ell_{N/2}}(x) R_{\ell_{N/2}}(x) = \ldots = [R_{\ell_0}(x)]^{2^m}, \quad N = 2^m.
\] (19)
This self-similarity property implies that the survival function for any cluster comprising several links has the same functional form, except for a trivial size scaling, as the single-link survival function.

The Weibull distribution can also be derived without invoking the exponential dependence of the link survival function. This approach is based on the additivity of the logarithmic survival functions, i.e.,

\[
\ln R_{\ell_N}(x) = \sum_{i=1}^{N} \ln R_{\ell_0}(x) = \sum_{i=1}^{N} \ln [1 - F_{\ell_0}(x)].
\]

Since the system’s failure is caused by the weakest link, we can assume that for \(N \gg 1\), this occurs at \(x\), such that \(F_{\ell_0}(x) \ll 1\). Thus, we can use the approximation:

\[
\ln R_{\ell_N}(x) \approx -NF_{\ell_0}(x) \Rightarrow R_{\ell_N}(x) \approx e^{-NF_{\ell_0}(x)}.
\]

Next, one assumes that \(F_{\ell_0}(x) \approx (x/x_0)^m\), \(x < x_c\), where \(x_c\) is a threshold above which the algebraic dependence of \(F_{\ell_0}(x)\) ceases. This argument leads to the Weibull expression; however, it does not clarify the tail behavior of the distribution function for \(x \gg x_c\).

Based on the self-similarity relation (19) and the pioneering works of Daniels and Smith [29,30] that use a Gaussian strength distribution for the links, Curtin showed that the strength distribution of heterogeneous materials tends asymptotically to the Weibull form and that the system’s strength parameters depend on both the system and the link sizes [21].

3.2. Non-Exponential Link Ansatz

An acceptable link survival function should satisfy the following properties: (i) it should take real values in the interval \([0, 1]\); (ii) it should be a monotonically decreasing function of \(x\). Let us introduce the following simple functional form:

\[
R_{\ell_0}(x) = \sqrt{1 + \phi^2(x)} - \phi(x),
\]

which gives admissible survival functions if \(\phi(x)\) satisfies (i) \(\phi(0) = 0\), (ii) \(\lim_{x \to \infty} \phi(x) \to \infty\) and (iii) \(\phi(x)\) is a monotonically increasing function of \(x\). Condition (iii) is necessary to ensure that \(dR_{\ell_0}(x)/dx \leq 0\), because the latter is given by:

\[
\frac{dR_{\ell_0}(x)}{dx} = \frac{d\phi(x)}{dx} \left[ \frac{\phi(x)}{\sqrt{1 + \phi^2(x)}} - 1 \right].
\]

The second term in the product on the right-hand side of the above is negative, and thus, \(dR_{\ell_0}(x)/dx \leq 0\) only if \(d\phi(x)/dx \geq 0\).

In light of (7), one can then show that the hazard function is given by:

\[
h(x) = \frac{d\phi(x)}{dx} \frac{1}{\sqrt{1 + \phi^2(x)}}.
\]

Hence, for \(x \gg 1\), i.e., \(\phi(x) \to \infty\), it holds that \(h(x) \approx d\ln \phi(x)/dx\). The main appeal of (20) is its simplicity. On the other hand, the condition (19) requires solving the following set of nonlinear equations:

\[
\sqrt{1 + \phi_{\ell_N/2}^2(x)} - \phi_{\ell_N/2}(x) = \left[ \sqrt{1 + \phi^2(x)} - \phi(x) \right]^{N/2}, \quad \text{for all } N = 2^m, \ m = 1, 2, \ldots,
\]
which is satisfied by:
\[
\phi_{N/2}(x) = \frac{1 - \left[ \sqrt{1 + \phi^2(x)} - \phi(x) \right]^N}{2 \left[ \sqrt{1 + \phi^2(x)} - \phi(x) \right]^{N/2}}.
\]

The above, however, does not lead to a self-similar expression for \(\phi_{N/2}(x)\).

In the following sections, we investigate certain probability distributions for systems with a finite number of links that are based on non-exponential link survival functions.

4. Finite-Size System with Gamma Link Distribution

4.1. The Gamma Distribution

First, we consider a slight modification of the exponential link survival function, which is based on the gamma distribution.

\[
F_{\ell_0}(x) = F_\gamma(x) := \frac{1}{\Gamma(m)} \gamma(x/x_s, m),
\]

where \(x_s\) and \(m\) are respectively the scale and shape parameters, \(\gamma(z, m)\) is the lower incomplete gamma function defined by means of the integral:

\[
\gamma(z, m) = \int_0^z du u^{m-1} e^{-u},
\]

and \(\Gamma(m)\) is the gamma function that is obtained from \(\gamma(z, m)\) at the asymptotic limit, \(\Gamma(m) = \lim_{z \to \infty} \gamma(z, m)\). The pdf of the gamma distribution is given by:

\[
f_\gamma(x) = \frac{x^{m-1} e^{-x/x_s}}{\Gamma(m) x_s^m}.
\]

The hazard rate of the gamma distribution is respectively given by:

\[
h_{\ell_0}(x) = \frac{e^{-x/x_s}}{\gamma(x/x_s, m) x_s} \left( \frac{x}{x_s} \right)^{m-1}.
\]

4.2. The Modified Gamma Distribution

We use the WLS scaling relation (5) and a gamma link probability function to define the modified gamma distribution, which applies to the entire system as follows:

\[
F_{\ell_N}(x) = 1 - [1 - F_\gamma(x)]^N = 1 - [R_{\ell_0}(x)]^N.
\]

The quantile function (inverse survival function) of the modified gamma distribution is:

\[
T_{\ell_N}(r) = x_s \gamma^{-1} \left[ \Gamma(m) \left( 1 - r^{1/N} \right), m \right],
\]

where \(\gamma^{-1}(\cdot)\) is the inverse of the lower incomplete gamma function. The quantile function of the modified gamma distribution is plotted in Figure 3. The median of the modified gamma distribution is \(x_{med} = T_{\ell_N}(1/2, m)\), i.e.,

\[
x_{med} = x_s \gamma^{-1} \left[ \Gamma(m) \left( 1 - \frac{1}{\sqrt{2}} \right), m \right].
\]
Based on (5) and assuming uniform link parameters, Equation (9) leads to the following modified gamma pdf for the $N$-link system:

$$f_{\ell_N}(x) = N \frac{e^{-x/x_s}}{\Gamma(m)} \left( \frac{x}{x_s} \right)^{m-1} \left[ 1 - \frac{\gamma(x/x_s, m)}{\Gamma(m)} \right]^{N-1}.$$  \hspace{1cm} (29)

Equation (29) can also be expressed as:

$$f_{\ell_N}(x) = f_{\gamma}(x) g_{\ell_N}(x),$$

where $f_{\gamma}(x)$ is given by (24) and $g_{\ell_N}(x)$ is a modulation function given by:

$$g_{\ell_N}(x) = N \left[ 1 - F_{\gamma}(x) \right]^{N-1} = N \left[ R_{\ell_0}(x) \right]^{N-1}. \hspace{1cm} (30)$$

The hazard rate of the modified gamma distribution is given according to (8) by:

$$h_{\ell_N}(x) = \frac{f_{\ell_N}(x)}{R_{\ell_N}(x)} = N h_{\ell_0}(x), \hspace{1cm} (31)$$

where $h_{\ell_0}(x)$ is given by (25).

![Figure 3](image_url)

*Figure 3.* Quantile function for the modified gamma distribution (27) with $m = 2.5$ (a) and $m = 0.7$ (b) for different link sizes $N$ (shown with different line types).

### 4.3. Motivation and Properties

The gamma distribution is used in reliability engineering to model extreme events, such as the lifetime of devices, the intensity of rainfall and earthquake return times [1]. The gamma distribution is also known in statistical physics as the stretched exponential distribution for $m < 1$, whereas it defaults to the exponential distribution for $m = 1$.

The modified gamma survival function is constructed assuming an WLS system that comprises $N$ links that have an identical link failure probability that follows the gamma cdf. The modulation function
$g_{\ell N}(x)$ given by (30) takes values between $N$ (at $x = 0$) and zero (at $x \to \infty$). For $m > 1$ and increasing $N$, $g_{\ell N}(x)$ shifts the mode of the pdf to lower values and reduces the width of the peak, as shown in Figure 4(a). Increasing $N$ also suppresses the right tail of the pdf. The gamma distribution is obtained for $N = 1$, as it is easily verified from (29). For $m < 1$, as $N$ increases, the $g_{\ell N}(x)$ increases the values of $f_{\ell N}(x)$ as $x \to 0$ and reduces the right tail, as shown in Figure 4(b). For $m < 1$, $f_{\ell N}(x) \to \infty$ as $x \to 0$.

The gamma hazard rate is plotted in Figure 5; it is an increasing function for $x < x_s$, whereas at $x \gg 1$, it is dominated by the exponential dependence, which leads to an asymptotically constant hazard function. In contrast, the Weibull distribution for $m > 1$ has a hazard rate that increases with $x$ [13,19].

![Figure 4](image_url)

**Figure 4.** Modified gamma distribution pdf $f_{\ell N}(x)$ with $m = 2.5$ (a) and $m = 0.7$ (b) for different link sizes $N$ (shown with different line types).

![Figure 5](image_url)

**Figure 5.** Hazard rate $h_{\ell 0}(x)$ of the gamma probability distribution for $m = 2.5$ (a) and $m = 0.7$ (b).
In Figure 6 we plot the cdf and $\Phi(x; \ell_N)$ of the modified gamma distribution for $m = 0.7$, $m = 2.5$ and for different values of $N$. We compare these functions to the respective functions for the Weibull distribution. For $m > 1$, the $\Phi(x; \ell_N)$ curves are straight lines on the Weibull plot for $\ln(x/x_s) \leq a_c$, where $a_c \approx -1$, but become concave as $\ln(x/x_s)$ increases further. This means that the right tail of the modified gamma distribution is heavier than the respective Weibull tail. In contrast, for $m < 1$, the $\Phi(x; \ell_N)$ curve becomes convex upwards for $\ln(x/x_s) > a_c$, and the right tail of the modified gamma distribution is lighter than the respective Weibull tail. The $\Phi(x; \ell_N)$ curves of the modified gamma distribution for different $N$ with fixed $m$ collapse onto a single curve by parallel shifting, in agreement with the scaling relation (12).

**Figure 6.** Cumulative distribution function $F_{\ell_N}(x)$ (left) and associated $\Phi(x; \ell_N)$ (right) for the modified gamma distribution with $m = 2.5$ (top) and $m = 0.7$ (bottom) and different values of $N$ (shown by different line types), as well as the respective Weibull functions with the corresponding Weibull modulus.
5. Finite-Sized System with $\kappa$-Weibull Distribution

5.1. The $\kappa$-exponential and $\kappa$-logarithm Functions

The definition of the $\kappa$-exponential function for $\kappa \geq 0$ is given by [31–35]:

$$\exp_\kappa(z) = \left(\sqrt{1 + z^2 \kappa^2 + z \kappa}\right)^{1/\kappa}.$$ (32)

The $\kappa$-exponential tends asymptotically to the standard exponential at the limits $z \to 0$ or $\kappa \to 0$, i.e.,

$$\exp_\kappa(z) \sim \exp(z), \quad z \to 0,$$ (33a)

$$\exp_\kappa(z) \sim \exp(z), \quad \kappa \to 0.$$ (33b)

The inverse function of the $\kappa$-exponential is the $\kappa$-logarithm, which is defined by:

$$\ln_\kappa(z) = \frac{z^\kappa - z^{-\kappa}}{2\kappa}.$$ (34)

5.2. The $\kappa$-Weibull Function

If $\kappa = 1/N$, where $N$ is the number of links in a system, we can define a system survival function based on:

$$R_{\ell N}(x) = \exp_\kappa\left[-\left(\frac{x}{x_s}\right)^m\right].$$ (35)

The $\kappa$-Weibull admits explicit expressions for the cdf, the survival function and the pdf, which are given respectively by the following expressions:

$$F_\kappa(x) = 1 - \exp_\kappa(-[x/x_s]^m),$$ (36)

$$R_\kappa(x) = \exp_\kappa(-[x/x_s]^m),$$ (37)

$$f_\kappa(x) = \frac{m}{x_s} \left(\frac{x}{x_s}\right)^{m-1} \exp_\kappa(-[x/x_s]^m) \sqrt{1 + \kappa^2(x/x_s)^{2m}}.$$ (38)

The $\kappa$-Weibull hazard function is obtained from (31) using (37) and (38) as follows:

$$h_\kappa(x) = \frac{m}{x_s} \left(\frac{x}{x_s}\right)^{m-1} \frac{1}{\sqrt{1 + \kappa^2(x/x_s)^{2m}}}. $$ (39)

The $\kappa$-Weibull quantile function is defined by means of the survival function as follows:

$$T_\kappa(R_\kappa) = \frac{1}{x_s} \left[\ln_\kappa\left(\frac{1}{R_\kappa}\right)\right]^{1/m}. $$ (40)
5.3. Motivation and Properties

The $\kappa$-Weibull is motivated by the non-exponential link ansatz discussed in Section 3.2. In particular, we assume that $\phi(x)$ is the following monotonically increasing function with link-size dependence:

$$\phi(x) = \frac{1}{N} \left( \frac{x}{x_s} \right)^m.$$ 

Then, the link survival function is given according to (20) by the function:

$$R_{\ell_0}(x) = \sqrt{1 + \frac{1}{N^2} \left( \frac{x}{x_s} \right)^{2m} - \frac{1}{N} \left( \frac{x}{x_s} \right)^m}.$$ 

The $\kappa$-Weibull system survival function (35) is then obtained from the above link survival function and the WLS principle (5).

Based on the properties of the $\kappa$-exponential, the survival function (35) corresponds to an extension of the Weibull survival function. The term $\kappa$-Weibull is due to the presence of the $\kappa$-exponential. This distribution has been used in economics to model the distribution of income [32,34] and in statistical physics as a model of earthquake recurrence times [27]. Other potential applications include the fracture strength of quasi-brittle materials, which exhibit deviations from Weibull scaling [7,28], and reliability models that involve the Pareto distribution.

The $\kappa$-exponential function can describe heavy-tailed processes, because for $\kappa > 0$, it decays asymptotically as a power-law function, i.e.,

$$\exp_\kappa(z) \sim |2\kappa z|^{\pm 1/\kappa}. \quad (41)$$

The $\kappa$-Weibull pdf also exhibits a power-law tail, which is inherited by the $\kappa$-exponential and is relevant in applications where such heavy tails are observed. In addition, if we define the $\kappa$-Weibull plot by means of the function $\Phi'_\kappa(x) = \ln \ln (1/R_\kappa(x))$, it follows that $\Phi'_\kappa(x) = m \ln \left( \frac{x}{x_s} \right)$. Hence, $\Phi'_\kappa(x)$ is independent of $\kappa$ and regains the logarithmic scaling of the double logarithm of the inverse survival function. This means that the $\kappa$-Weibull plot, in which $\Phi'_\kappa(x)$ is plotted instead of $\Phi_\kappa(x)$, can be used to visually detect the $\kappa$-Weibull scaling.

Plots of the $\kappa$-Weibull pdf for $x_s = 1$ and two values of $m$ ($m < 1$ and $m > 1$) are shown in Figure 7. The plots also include the Weibull pdf ($\kappa = 0$) for comparison. For both $m$, lower $N$ lead to a heavier right tail. For $m = 2.5$ the mode of the pdf moves to the right as $N$ increases (for $m = 0.7$ the mode is at zero independently of $N$, since the distribution is zero-modal for $m \leq 1$) [34]. To the right of the mode, higher $N$ correspond, at first, to higher pdf values. This is reversed at a crossover point beyond which the lower-$N$ pdfs exhibit slower power-law decay, i.e., $\lim_{x \to \infty} f_\kappa(x) \propto x^{-\alpha}$, where $\alpha = 1 + m/\kappa$ [34]. The crossover point occurs at $\approx 1.5x_s$ for $m = 2.5$, whereas for $m = 0.7$ at $\approx 5x_s$.  

6. Conclusions

We investigated the statistical implications of the principle of weakest-link scaling. This type of scaling is used to model the probability distributions of extreme events, including the failure strength of various heterogeneous materials, wind speed, earthquake recurrence times and hydrological maxima. We have focused on various statistical mechanisms leading to deviations from the Weibull distribution, which is commonly used in weakest-link scaling systems. Our motivation stems from empirical distributions that deviate from the Weibull in physical systems expected to follow weakest-link scaling, e.g., [7,15,27,28,36]. Our results include a modified gamma distribution, the $\kappa$-Weibull distribution and weakest-link scaling distributions with non-homogeneous links that have random (possibly correlated) parameters. The distributions proposed above modify the tail behavior in comparison with the classical Weibull dependence. We thus believe that they are of interest to both statistical physicists and reliability or systems engineers.

The modified gamma distribution, which we introduce herein, is derived by applying weakest-link scaling to systems with a finite number of links. The modified gamma distribution has a heavier tail than the respective Weibull for $m > 1$ and a lighter tail for $m < 1$. The $\kappa$-Weibull distribution is also shown to follow from weakest-link scaling. A notable property of the $\kappa$-Weibull distribution is its ability to exhibit a power-law tail, whereas the modified gamma distribution is dominated by exponential decay at large values. The power-law dependence of the $\kappa$-Weibull is derived from the interaction between the system and the links, expressed in terms of a finite-valued scaling factor $N$ in the link survival function. For $N \to \infty$, the $\kappa$-Weibull tends to the Weibull model, and the power-law decay at $x \gg 1$ is replaced by the exponential Weibull dependence. We have argued in [27] that the $\kappa$-Weibull distribution can be observed in the case of earthquake recurrence times if the monitored area comprises a finite number of faults. We have also linked the number of effective units in this system to the inverse of $\kappa$. This type of behavior can also be observed if the system under study does not include all of the units (e.g., faults).
of an interdependent system. In this case, events that are outside the boundaries of the observed system are not taken into account, thus leading to longer recurrence times, which generate the power-law tail. The heavy tail of the \( \kappa \)-Weibull may also find applications in biology, e.g., in modeling the inter-spike interval distributions of cortical neurons, which is known to follow a power law [37].

We have also proposed a different approach for inserting correlations in weakest-link scaling systems by means of a variable scaling factor in the link survival function. If this factor is random (but possibly correlated between links), the observed random variable \( X \) is described by an appropriate superstatistics formulation. In the case of Weibull link scaling, we have shown that if the scale parameter is uniformly distributed and uncorrelated across links, the observable probability distribution is controlled by three factors: the width of the scale parameter distribution, the Weibull modulus and the mixing ratio, which determines the balance between the number of links and the values contributed by each link to the observed sample. In the limit of a very large sample, it is also shown that the system’s survival function regains the Weibull form, albeit with a renormalized scale, provided that the distribution of the scale parameter has a compact support. In further research, we will investigate the impact of correlations between link distribution parameters on the observed system survival function, as well as applications of this formalism to environmental extreme events that exhibit weakest-link scaling.

Another extreme value distribution that fits within the weak-scaling formalism is the Gumbel distribution [38]. The Gumbel distribution is of the general form \( F(x) = \exp(-\exp(-x)) \) and belongs to the family of generalized extreme value distributions. The Gumbel is characterized by a different mathematical expression and size scaling than the Weibull distribution. There is both theoretical and experimental evidence (e.g., in the fracture of brittle ceramics) that in certain cases, the Gumbel is a more suitable asymptotic form than the Weibull [22,38]. One may therefore wonder if the same arguments that led to the \( \kappa \)-Weibull distribution also apply to the Gumbel distribution. At this point, however, we lack insight into how this construction should be implemented. Hence, we will leave this as an open subject for future research.

Author Contributions

All authors contributed to the research and the writing. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References


© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).